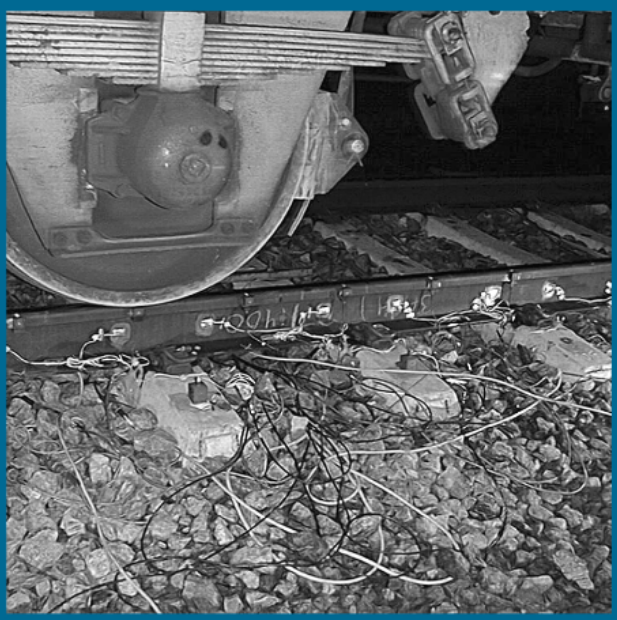
WOODHEAD PUBLISHING
IN MECHANICAL ENGINEERING

Wheel–rail interface handbook

Edited by R. Lewis and U. Olofsson



CRC

WP

Copyrighted Materials
Copyright © 2009 Woodhead Publishing Retrieved from www.knovel.com

Wheel–rail interface handbook

Edited by

R. Lewis and U. Olofsson

CRC Press
Boca Raton Boston New York Washington, DC

WOODHEAD PUBLISHING LIMITED
Oxford Cambridge New Delhi

Published by Woodhead Publishing Limited, Abington Hall, Granta Park, Great Abington, Cambridge CB21 6AH, UK
www.woodheadpublishing.com

Woodhead Publishing India Private Limited, G-2, Vardaan House, 7/28 Ansari Road, Daryaganj, New Delhi – 110002, India
www.woodheadpublishingindia.com

Published in North America by CRC Press LLC, 6000 Broken Sound Parkway, NW, Suite 300, Boca Raton, FL 33487, USA

First published 2009, Woodhead Publishing Limited and CRC Press LLC
© 2009, Woodhead Publishing Limited except Chapter 12 © 2009, M.J.M.M. Steenbergen and Chapter 13 © 2009, Z. Li
The authors have asserted their moral rights.

This book contains information obtained from authentic and highly regarded sources. Reprinted material is quoted with permission, and sources are indicated. Reasonable efforts have been made to publish reliable data and information, but the authors and the publishers cannot assume responsibility for the validity of all materials. Neither the authors nor the publishers, nor anyone else associated with this publication, shall be liable for any loss, damage or liability directly or indirectly caused or alleged to be caused by this book.

Neither this book nor any part may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, microfilming and recording, or by any information storage or retrieval system, without permission in writing from Woodhead Publishing Limited.

The consent of Woodhead Publishing Limited does not extend to copying for general distribution, for promotion, for creating new works, or for resale. Specific permission must be obtained in writing from Woodhead Publishing Limited for such copying.

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

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library.

Library of Congress Cataloging in Publication Data
A catalog record for this book is available from the Library of Congress.

Woodhead Publishing ISBN 978-1-84569-412-8 (book)
Woodhead Publishing ISBN 978-1-84569-678-8 (e-book)
CRC Press ISBN 978-1-4398-0146-8
CRC Press order number: N10033

The publishers' policy is to use permanent paper from mills that operate a sustainable forestry policy, and which has been manufactured from pulp which is processed using acid-free and elemental chlorine-free practices. Furthermore, the publishers ensure that the text paper and cover board used have met acceptable environmental accreditation standards.

Typeset by Replika Press Pvt Ltd, India
Printed by TJ International Limited, Padstow, Cornwall, UK

Copyrighted Materials
Copyright © 2009 Woodhead Publishing Retrieved from www.knovel.com

Contents

Contributor contact details xv

Foreword xxi

Part I State-of-the-art research

1 Introduction to wheel–rail interface research 3
R. Lundén, Chalmers/CHARMEC, Sweden; B. Paulsson, Banverket, Sweden

1.1 History and present situation 3
1.2 Phenomena in the wheel-rail interface 6
1.3 Research fields 14
1.4 Applications 16
1.5 Ongoing research, development and standardization efforts 18
1.6 System aspects and optimization 24
1.7 Future trends 26
1.8 Sources of further information and advice 28
1.9 Acknowledgements 30
1.10 References 30

2 Basic tribology of the wheel–rail contact 34
R. Lewis, University of Sheffield, UK; U. Olofsson, Royal Institute of Technology (KTH), Sweden

2.1 Introduction 34
2.2 Contact mechanics 35
2.3 Wear 40
2.4 Fatigue 49
2.5 Adhesion 53
2.6 References 56

3 Wheel–rail contact mechanics 58
S. Iwnicki, Manchester Metropolitan University, UK; S. Björklund and R. Enblom, Royal Institute of Technology (KTH), Sweden

3.1 Introduction 58
3.2 General contact modelling 62
3.3 Wheel–rail contact analysis 74
3.4 Computer simulation tools for railway vehicle dynamics 84
3.5 Future trends 89
3.6 Sources of further information and advice 89
3.7 References 90

4 Friction and wear simulation of the wheel–rail interface 93
S. Andersson, Royal Institute of Technology (KTH), Sweden

4.1 Introduction 93
4.2 Single-point observation method 95
4.3 Wear maps and transition diagrams 96
4.4 Friction models 99
4.5 Wear simulation 109
4.6 References 123

5 Rail materials 125
J. E. Garnham and C. L. Davis, University of Birmingham, UK

5.1 Introduction 125
5.2 Pearlitic rails 131
5.3 Austenitic rails for switches and crossings 143
5.4 Welding rail 143
5.5 Wear and rolling contact fatigue of pearlitic rail 145
5.6 Bainitic rail 151
5.7 Recent rail material developments 164
5.8 Conclusions 166
5.9 References 166

6 Railway wheel wear 172
F. Braghin and S. Bruni, Politecnico di Milano, Italy; R. Lewis, University of Sheffield, UK

6.1 Introduction 172
6.2 The wheel wear process 173
6.3 Tribological issues in wheel wear 175
6.4 Wheel–rail contact mechanics and its effect on wheel wear 189
6.5 State-of-the-art of uniform wheel wear modelling 195
6.6 Means to reduce uniform wheel wear 201
6.7 Conclusions 206
6.8 References 206

7 Fatigue of railway wheels 211
A. Ekberg, Chalmers University of Technology, Sweden

7.1 Introduction 211
7.2 Appearance and mechanisms of wheel fatigue 213
7.3 Prediction of wheel fatigue 223
7.4 Wheel fatigue put in context 230
7.5 Conclusions 239
7.6 Sources of further information and advice 241
7.7 Acknowledgements 242
7.8 References 242

8 Out-of-round railway wheels 245
J. Nielsen, Chalmers University of Technology, Sweden

8.1 Introduction 245
8.2 Classification and quantification of wheel out-of-roundness 246
8.3 Discrete wheel tread defects 252
8.4 Wheel roughness induced by tread braking 261
8.5 Simulation of consequences of out-of-round wheels 267
8.6 Sources of further information and advice 274
8.7 Acknowledgements 275
8.8 References 275

9 Rail surface fatigue and wear 280
D. I. Fletcher, University of Sheffield, UK; F. J. Franklin, Newcastle University UK; A. Kapoor, Swinburne University of Technology, Australia

9.1 Introduction 280
9.2 Rail rolling contact fatigue 282
9.3 Experimental investigations 290
9.4 Calculating crack growth rate 293
9.5 Crack branching predictions 299
9.6 Rail wear 301
9.7 References 305

10 The evolution and failure of pearlitic microstructure in rail steel: observations and modelling 311
F. J. Franklin, University of Newcastle, UK; J. E. Garnham and C. L. Davis, University of Birmingham, UK; D. I. Fletcher, University of Sheffield, UK; A. Kapoor, Swinburne University of Technology, Australia

10.1 Introduction 311
10.2 Observations of microstructural evolution and failure 312
10.3 Modelling 333

10.4 Conclusions 342
10.5 Acknowledgements 344
10.6 Nomenclature 344
10.7 References 345

11 Rail corrugation 349
S. L. Grassie, Stuart Grassie Engineering Ltd, Germany
11.1 Introduction 349
11.2 Classification of corrugation 350
11.3 Heavy-haul corrugation 353
11.4 Light-rail corrugation 356
11.5 Other P2 resonance corrugation 358
11.6 Rutting 360
11.7 Roaring rails/'pinned-pinned resonance' corrugation 364
11.8 Trackform-specific corrugation 368
11.9 Conclusions and recommendations 370
11.10 Acknowledgements 372
11.11 References 373

12 Rail welds 377
M. J. M. M. Steenbergen, Delft University of Technology, The Netherlands; R. W. van Bezooijen, Id² Consultancy, The Netherlands
12.1 Introduction 377
12.2 Rail welding processes 378
12.3 Rail welds and damage formation 380
12.4 Rail welding irregularities and dynamic effects in the wheel-rail interface 384
12.5 Rail weld geometry assessment; the Dutch rail welding regulations (2005) 395
12.6 Welding irregularities, energy considerations and deterioration 402
12.7 References 406

13 Squats on railway rails 409
Z. Li, Delft University of Technology, The Netherlands
13.1 Introduction 409
13.2 Review of past research 411
13.3 Correlation of squats with track parameters 413
13.4 Characteristics of squats 418
13.5 Three-dimensional dynamic rolling contact solutions in elastoplasticity 421
13.6 Squats initiation due to differential wear and differential plastic deformation 423

13.7 Squats growth process 427
13.8 Detection of squats 433
13.9 Counter measures 433
13.10 Further research 434
13.11 References 435

14 Effect of contaminants on wear, fatigue and traction 437
S. R. Lewis and R. S. Dwyer-Joyce, University of Sheffield, UK

14.1 Introduction 437
14.2 Contaminants 438
14.3 Friction modifiers 448
14.4 Discussion 450
14.5 Conclusions 452
14.6 References 453

15 Effect of damage on vehicle dynamics 456
S. Bruni and F. Braghin, Politecnico di Milano, Italy

15.1 Classification of damage that affects vehicle dynamics 456
15.2 Effects of transversal profile wear 457
15.3 Effects of rail corrugation 467
15.4 Effects of wheel out-of-roundness 469
15.5 Effects of localised damage on wheel and rail profiles 471
15.6 Conclusions 474
15.7 References 474

16 Noise and vibration from the wheel–rail interface 477
D. Thompson and C. Jones, University of Southampton, UK

16.1 Introduction 477
16.2 Basics of noise and vibration 478
16.3 Rolling noise 481
16.4 Reduction of rolling noise 487
16.5 Impact noise 493
16.6 Curve squeal 494
16.7 Ground vibration and ground-borne noise 497
16.8 Conclusions and future trends 503
16.9 Sources of further information and advice 506
16.10 References 506

17 Adhesion and friction modification 510
U. Olofsson, Royal Institute of Technology (KTH), Sweden

17.1 Introduction 510
17.2 The coefficient of friction, adhesion and braking distance 513
17.3 Friction modification 519

17.4 Possible models for low friction at the wheel–rail contact 523
17.5 Future trends 524
17.6 References 524

18 Wheel–rail isolation 528
R. Lewis, University of Sheffield, UK

18.1 Introduction 528
18.2 Third bodies in the wheel–rail contact 530
18.3 Testing for isolation 532
18.4 Effects of contaminants on isolation 536
18.5 Modelling approaches 545
18.6 Conclusions 548
18.7 References 548

19 Air-borne particles from the wheel-rail contact 550
M. Gustafsson, VTI – Swedish National Road and Transport Research Institute, Sweden

19.1 Introduction 550
19.2 Background 551
19.3 Concentrations 551
19.4 Particle properties and sources 555
19.5 Dispersion 564
19.6 Health effects 565
19.7 Measures 567
19.8 Future trends 572
19.9 References 573

20 Maintenance of the wheel–rail interface 576
S. L. Grassie, Stuart Grassie Engineering Ltd, Germany

20.1 Introduction 576
20.2 The importance of friction 578
20.3 Routine metal removal and maintenance of transverse profile 579
20.4 Routine maintenance of longitudinal/circumferential profile 594
20.5 'Corrective' maintenance 599
20.6 Future trends in maintenance of the wheel-rail interface 603
20.7 Sources of further information and advice 605
20.8 References 606

21 Models for infrastructure costs related to the wheel–rail interface 608
E. Andersson, Royal Institute of Technology (KTH), Sweden; J. Öberg, Banverket, Sweden

21.1 Introduction 608

21.2 Track deterioration model 609
21.3 Computational tools and input data 616
21.4 Calibration of model and cost 620
21.5 Examples of results 622
21.6 Conclusions and future trends 624
21.7 References 625
21.8 Appendix 1: Notation 626
21.9 Appendix 2: Vertical dynamic wheel load model 628

Part II Industrial context – managing the wheel–rail interface

22 Managing the wheel–rail interface: Railway infrastructure maintenance in a severe environment; The Swedish experience 633
P.-O. Larsson-Kråik, Banverket, Sweden

22.1 Introduction 633
22.2 General description of Malmbanan (Swedish ore line) 634
22.3 Locomotives and cars 636
22.4 Train control system 642
22.5 Infrastructure configuration 642
22.6 Electrical power system 649
22.7 Track maintenance practices 649
22.8 Maintenance, wheel–rail interaction 651

23 Managing the wheel-rail interface: Europe Metro experience on the London Underground Victoria line 653
D. Scott, London Underground, UK

23.1 Introduction to the Victoria Line and historic wheel–rail interface issues 653
23.2 The Victoria Line upgrade 654
23.3 Wheel–rail interface monitoring 654
23.4 Lubrication management 659
23.5 Identified wheel–rail interface problems 662
23.6 Ongoing work and future plans 666
23.7 Conclusions 668
23.8 References 668

24 Managing the wheel–rail interface: the Canadian experience 669
E. Magel and P. Sroba, National Research Council, Canada

24.1 Introduction 669

24.2 Canadian Pacific Railway – wheel–rail management experience 670
24.3 Cartier Railway Company – optimising the captive railroad 676
24.4 Several wheel–rail problems on Vancouver's Skytrain 678
24.5 Wheel life and ride quality on Edmonton Transit 679
24.6 Wheel shelling on Canada's freight railroads 680
24.7 Canadian National Railway – rail grinding needs are site-specific 682
24.8 Summary and conclusions 684
24.9 References 685

25 Managing the wheel–rail interface: the US experience 686
J. TUNNA, Transportation Technology Center Inc, USA

25.1 Introduction 686
25.2 Wheel and rail materials 688
25.3 Wheel and rail profiles 690
25.4 Wheel and rail surface damage 692
25.5 Lubrication and friction modification 695
25.6 Condition monitoring 696
25.7 Conclusions 697
25.8 References 698

26 Managing the wheel–rail interface: the Japanese experience 701
M. ISHIDA, Railway Technical Research Institute, Japan

26.1 Introduction 701
26.2 Rolling contact fatigue 704
26.3 Wear 717
26.4 Corrugation 725
26.5 Adhesion 737
26.6 Lubrication 745
26.7 References 753

27 Managing the wheel–rail interface: the Australian experience 759
S. MARICH, Marich Consulting Services, Australia

27.1 Introduction 759
27.2 Rail–wheel wear and lubrication 760
27.3 Rail corrugations 764
27.4 Rolling contact and thermal/traction defects 767
27.5 Control of wheel–rail interaction through profiling 774
27.6 Rail grinding 776

27.7 Friction management 780
27.8 Rail and wheel materials 782
27.9 Conclusions 788
27.10 References 789

28 Managing the wheel–rail interface: the Dutch experience 792
A. Zoeteman, R. Dollevoet, R. Fischer and J.-W. Lammers, ProRail, EIM and Delft University of Technology, The Netherlands

28.1 Introduction 792
28.2 Optimising rail maintenance 793
28.3 Optimising wheel maintenance 802
28.4 Special aspects in optimising wheel–rail interface: riding comfort and noise 806
28.5 Monitoring traffic movement: Gotcha/Quo Vadis 812
28.6 Conclusions 815
28.7 References 817

Index *819*

Copyrighted Materials
Copyright © 2009 Woodhead Publishing Retrieved from www.knovel.com

Contributor contact details

(*= main contact)

Chapter 1

Professor Roger Lundén*
Chalmers/CHARMEC
SE-412 96 Gothenburg
Sweden

E-mail: roger.lunden@chalmers.se

Björn Paulsson
Banverket
SE-781 85 Borlänge
Sweden

E-mail: bjorn.paulsson@banverket.se

Chapter 2

Dr Roger Lewis*
Department of Mechanical Engineering
University of Sheffield
Mappin Street
S1 3JD
UK

Email: roger.lewis@sheffield.ac.uk

Professor Ulf Olofsson
Royal Institute of Technology (KTH)
SE-100 44 Stockholm
Sweden

Email: ulf.olofsson@itm.kth.se

Chapter 3

Professor Simon Iwnicki*
Manchester Metropolitan University
Manchester
M1 5GD
UK

E-mail: s.d.iwnicki@mmu.ac.uk

Dr Stefan Björklund
Department of Machine Design
Brinellvägen 83
Royal Institute of Technology (KTH)
SE-100 44
Stockholm
Sweden

Dr Roger Enblom
Aeronautical and Vehicle Engineering
Royal Institute of Technology (KTH)
SE-100 44
Stockholm
Sweden

Chapter 4

Professor Soren Andersson
Royal Institute of Technology (KTH)
SE-100 44
Stockholm
Sweden

E-mail: soren@md.kth.se

Chapter 5

Dr John. E. Garnham and Dr Claire L. Davis*
Department of Metallurgy and Materials
University of Birmingham
Edgbaston
Birmingham
B15 2TT
UK

Email: c.l.davis@bham.ac.uk

Chapter 6

Dr Francesco Braghin* and Professor Stefano Bruni,
Department of Mechanical Engineering
Politecnico di Milano,
Via La Masa 1
20156 Milano
Italy

Email: francesco.braghin@polimi.it; stefano.bruni@polimi.it

Dr Roger Lewis
Department of Mechanical Engineering
The University of Sheffield
Mappin Street
Sheffield
S1 3JD
UK

Email: roger.lewis@sheffield.ac.uk

Chapter 7

Dr Anders Ekberg
Department of Applied Mechanics/ CHARMEC
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden

Email: anders.ekberg@chalmers.se

Chapter 8

Professor J. Nielsen
Department of Applied Mechanics/ CHARMEC
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden

Email: jens.nielsen@chalmers.se

Chapter 9

Dr David I. Fletcher*
Department of Mechanical Engineering
University of Sheffield
Mappin Street
Sheffield
S1 3JD
UK

Email: d.i.fletcher@sheffield.ac.uk;

Dr Francis J. Franklin
NewRail
Department of Mechanical & Systems Engineering
Newcastle University
Newcastle-upon-Tyne
NE1 7RU
United Kingdom

Professor Ajay Kapoor
Faculty of Engineering and Industrial Sciences
Swinburne University of Technology
PO Box 218
Hawthorn, VIC, 3122
Australia

Chapter 10

Dr Francis J. Franklin
School of Mechanical and Systems Engineering
University of Newcastle-upon-tyne
NEI MRU
Newcastle upon tyne
UK

Dr John E. Garnham
Department of Metallurgy and Materials
University of Birmingham
Edgbaston
Birmingham
B15 2TT
UK

Dr David I. Fletcher*
Department of Mechanical Engineering
University of Sheffield
Mappin Street
Sheffield
S1 3JD
UK

Email: d.i.fletcher@sheffield.ac.uk;

Dr Claire L. Davis
Department of Metallurgy and Materials
University of Birmingham
Edgbaston
Birmingham
B15 2TT
UK

Email: c.l.davis@bham.ac.uk

Professor Ajay Kapoor
Faculty of Engineering and Industrial Sciences
Swinburne University of Technology
PO Box 218
Hawthorn, VIC, 3122
Australia

Chapter 11

Dr Stuart Grassie
Stuart Grassie Engineering Ltd
Abbauernring 1
30900 Wedmark
Germany

Email: stuart@stuartgrassie.com

Chapter 12

Dr Michaël J.M.M. Steenbergen*
Delft University of Technology
Faculty of Civil Engineering and Geosciences
Railway Engineering Group
Stevinweg 1
2628 CN Delft
The Netherlands

Email:
M.J.M.M.Steenbergen@tudelft.nl

Ruud W. van Bezooijen
Id² Consultancy
Zwaansweg 9
NL 4247 EX Kedichem
The Netherlands

Chapter 13

Professor Z. Li
Section of Road and Railway Engineering
Faculty of Civil Engineering and Geosciences
Delft University of Technology
Stevinweg 1
2628 CN Delft
The Netherlands

Email: Z.Li@tudelft.nl

Chapter 14

Stephen R. Lewis and Professor Rob S. Dwyer-Joyce*
Department of Mechanical Engineering
University of Sheffield
Mappin Street
Sheffield
S1 3JD
UK

Email Stephen.lewis@sheffield.ac.uk

Chapter 15

Professor Stefano Bruni* and Dr Francesco Braghin
Department of Mechanical Engineering
Politecnico di Milano,
Via La Masa 1
20156 Milano
Italy

Email: stefano.bruni@polimi.it;
Francesco.braghin@polimi.it

Chapter 16

Professor David Thompson* and C. Jones
Institute of Sound and Vibration Research
University of Southampton
Southampton
SO17 1BJ
UK

Email: djt@isvr.soton.ac.uk

Chapter 17

Professor Ulf Olofsson
Royal Institute of Technology (KTH)
Stockholm
SE-100 44
Sweden

Email: ulf.olofsson@itm.kth.se

Chapter 18

Dr Roger Lewis
Department of Mechanical Engineering
University of Sheffield
Mappin Street
S1 3JD
UK

Email: roger.lewis@sheffield.ac.uk

Chapter 19

Dr Mats Gustafsson
VTI – Swedish National Road and Transport Research Institute
SE-58195 Linköping
Sweden

Email address: mats.gustafsson@vti.se

Chapter 20

Stuart Grassie
Stuart Grassie Engineering Ltd
Abbauernring 1
30900 Wedemark
Germany

Email: stuart@stuartgrassie.com

Chapter 21

Professor Evert Andersson*
Professor Railway Technology
Royal Institute of Technology (KTH)
SE-100 44
Stockholm
Sweden

Email: everta@.kth.se

Johan Öberg
M.Sc. Vehicle Engineering
Banverket Operations Division
Track and Civil Engineering
Stockholm
Sweden

Chapter 22

P.-O. Larsson-Kråik
Banverket
Box 43
SE-971 02 Luleå
Sweden

Email: per-olof.larsson-kraik@banverket.se

Chapter 23

Daniel Scott
London Underground
Templar House
81-87 High Holborn
London
WC1V 6NU
UK

Email: Daniel.Scott@metronetrail.com

Chapter 24

Eric Magel*
Centre for Surface Transportation Technology
National Research Council, Canada
2320 Lester Road
Ottawa, Ontario, K1V 1S2
Canada

Email: eric.magel@nrc-cnrc.gc.ca

P. Sroba
Sroba Rail Services Pty Ltd
6/113 King Street
Newcastle
NSW 2300
Australia

Email: peter.sroba@bigpond.com

Chapter 25

John Tunna
Transportation Technology Center Inc.
55500 D.O.T. Road
Pueblo, CO 81001
USA

Email: john_tunna@aar.com

Chapter 26

M. Ishida
Railway Technical Research Institute
Tokyo
Japan

Email: mako@rtri.or.jp

Chapter 27

Dr Stephen Marich
Marich Consulting Services
12 Heatherlea Drive
Wheelers Hill
Victoria 3150
Australia

Email: stephenmarich@bigpond.com

Chapter 28

Arjen Zoeteman*, R. Dollevoet, R. Fischer and J.-W. Lammers
European Rail Infrastructure Managers (EIM)
ProRail and Delft University of Technology
Section of Transport Policy and Logistics
Faculty of Technology
Policy and Management
The Netherlands

Email: Arjen.Zoeteman@eimrail.org

Copyrighted Materials
Copyright © 2009 Woodhead Publishing Retrieved from www.knovel.com

Foreword

The contact between the wheel and rail is the defining feature of the railway. For many people it comes as a surprise to learn that the size of a typical contact is in the same order as a small finger nail and that each passage of each wheel is an irreversible event, resulting in either or both the detachment of a wear particle or the advancement of a fatigue crack. So much for the 'permanent way': the damage caused by the passage of traffic is the financial Achilles Heel of the railway, resulting in major maintenance costs. The damage extends from the site of the contact into both the wheel and vehicle structure and below the rail into the supports and down to the track foundation.

The early railways suffered from many technical problems, including broken rails, wheels and axles. Many investigations were performed and, by a combination of experimentation, experience, empiricism and sensible conservatism, a *modus operandi* was arrived at which served the railway well for the first, say, 150 years of its existence. The mechanical engineering discipline traditionally dealt with the vehicle whilst the rail and foundations were the domain of the civil engineers. This unfortunate demarcation persists, to an extent, today, but increasingly there is an awareness that a systems approach is needed, nowhere more than at the extremely complex interface between the wheel and rail. Furthermore, developments in metallurgy and tribology have led to advances in understanding which need to be accounted for at the interface, whilst increasing pressures on financial performance have necessitated better understanding of 'what is going on'.

Generally, problems at the wheel–rail interface have come into focus as speeds have increased, as trains have become heavier and as the level of traffic has grown. Speeds of up to 350 kph are now commonplace on dedicated high-speed lines, 200 kph on 'ordinary' lines, and for freight trains axle loads of around 30 tonnes are commonplace in many parts of the world. All these developments have seen a huge worldwide increase in interest in the wheel–rail interface for the last 25 years. In the UK, a relatively small accident in terms of injuries and fatalities, at Hatfield in 2000, generated a large programme of technical investigation into contact fatigue. The author of this foreword was invited by Railtrack, the then owners and maintainers

of the rail infrastructure in the UK, to chair their investigations. Much useful work has grown out of this activity, and considerable quantities of the contents of this book stem from this root. It is worth noting here that the privatisation of the UK railway had left Railtrack short on technical know how, lacking in research capability and almost completely bereft of corporate memory, factors which were major contributors to the accident and led to the demise of Railtrack and its replacement by Network Rail.

This book is a most welcome attempt to draw together some disparate strands of relatively recent research and practical experience. The advances made possible by hugely increased computing capabilities, allied to improved techniques for in-service measurements, are well represented. The largest part of the book is devoted to a collection of chapters outlining the state-of-the-art of scientific and engineering knowledge of the problems associated with the wheel–rail interface and the deterioration mechanisms of both wheels and rails. The latter part outlines practical experience from countries around the world, comprising the UK, Sweden, Canada, the USA, the Netherlands, Australia and Japan. Together these represent a valuable collection of knowledge which will become a key reference for both researchers and practitioners for years to come. I am sure that readers will enjoy the contents as much as I have done and will find a great deal to assist their own work.

Professor Roderick A Smith
Future Rail Research Centre, Imperial College London

Copyrighted Materials
Copyright © 2009 Woodhead Publishing Retrieved from www.knovel.com

Part I

State-of-the-art research

1

Introduction to wheel–rail interface research

R. LUNDÉN, Chalmers/CHARMEC, Sweden;
B. PAULSSON, Banverket, Sweden

Abstract: For successful railway transportation, good performance of the wheel–rail interface is of the utmost importance and is, therefore, the subject of worldwide interdisciplinary research efforts. The present chapter should serve as an introduction and background to the following 20 chapters. A brief history of the wheel–rail interface research is given together with an account of the present situation. The main phenomena treated are contact stresses, contact friction and deterioration mechanisms which are related to cost, safety, maintenance, reliability, environment and energy consumption. Major fields of recent research are described followed by four examples of application. Ongoing projects, standardization efforts and legislation activities are reported. Some future trends in research are listed and sources of information are provided. It is emphasized that the wheel–rail interface must be treated as part of the larger train–track system, which means that a full optimization should be carried out on a systems level.

Key words: railway research, wheel–rail interface, train–track interaction, contact mechanics, information sources.

1.1 History and present situation

1.1.1 History of railway operation

Flanged wheels running on a cast iron rail have been in use since the 18th century, and they made it possible to design a railway track system with turnouts. Coned wheels with a flangeway clearance, enabling the wheels to run on a straight track without flange–rail contact, were well established in the 1830s. The British railway pioneers George Stephenson and Isambard Kingdom Brunel both explained the guiding mechanism introduced by the coning.

The first stretch of a railway line with locomotives and regular traffic was opened in England in 1825. The next few decades saw the launch of massive railway projects in many countries around the world. Sparsely populated Sweden may serve as an example of this development. Here, the first long-distance railway across the country was inaugurated in 1862 between Stockholm and Gothenburg. During the 1800s speed was normally not higher than 60 km/h, the axle load was 10 tonnes and 30 kg/m rails were used. Traffic with a speed as 90 km/h was introduced in 1905–1915. The maximum speed was increased to 120 km/h in 1946, 130 km/h in 1948,

160 km/h in 1983 and 200 km/h in 1990 (X2 tilting train). Axle load of 13–14 tonnes was introduced in 1905, 17 tonnes in the 1940s and 22 tonnes in the 1960s.[1]

The iron ore line in northern Sweden and Norway was inaugurated in 1899. The axle load then was 11 tonnes, and it was gradually increased to 18 tonnes in 1950, 25 tonnes in 1965 and 30 tonnes in 2000.[2–4] Other major developments in Sweden were vacuum brakes in the 1880s, electrification from 1915 onwards and compressed air brakes and axle rolling bearings in the1920s. The first wheels had wooden centres, later replaced by wheel centres with forged steel spokes. In the 1920s, wheel rims mounted on solid forged wheel centres were introduced. Forged solid wheels were introduced for freight in the 1960s. Rails in Sweden have gradually been upgraded to 41, 43, 50 and 60 kg/m. More history on wheel and rail developments can be found in works by Wise, Koerfer and Profillidis.[5–7]

The first high-speed train in the world was the Japanese Shinkansen in 1964, then travelling at a maximum speed of 210 km/h. In 1981 the French TGV was introduced at 260 km/h and now runs at 300–320 km/h in regular passenger traffic. Heavy haul operations nowadays run at 30, 35 and even 40 tonne axle load in Australia and the USA. The present speed record on high-speed track is 574.8 km/h set by the French TGV in April 2007.

1.1.2 History of theory for wheel–rail contacts

For the wheel–rail interface, the pioneering scientific breakthroughs occurred in the 1880s. Heinrich Hertz is widely known for his work on elastic contacts. In his paper of 1881 he mentions the wheel–rail contact problem, and his theory soon found application in railway engineering. Assumptions on Hertzian contact are still very common in analysis of the wheel–rail interface and of vehicle dynamics. The work in 1885 by Joseph Valentin Boussinesq laid the foundation for a mathematical analysis of more complex elastic contacts.[8] The first mathematical analysis of the hunting behaviour of a wheelset was given by Johann Klingel in 1883. Frederick Carter formulated the concept of wheel–rail creep in 1926. Fatigue was first studied by August Wöhler in the 1850s in the case of railway axles. The fundamental theory for rolling contact fatigue was published in 1947 by Gustaf Lundberg and Arvid Palmgren. Their application was rolling bearings.[9] Major scientific contributions enabling modern analysis of vehicle dynamics and contact mechanics were carried out by Joost Kalker on general elastic contact mechanics and by Ken L Johnson on elastoplastic rolling contact mechanics. See further the history written by Alan H Wickens in Chapter 2 of the work edited by Iwnicki[10] and the recent paper by Klaus Knothe.[11] Tribology and wear mechanisms are other important fields with major contributions during the 1950s to 1970s by Tabor, Archard and others.[12–13]

1.1.3 Material development

From a small-scale output of pig-iron, major developments in steel production took place during the second half of the 19th century.[14] The Bessemer process (1855) made it possible to produce steel as ingots. The Martin processes (1869 and 1890) rendered it possible to use scrap iron and to drastically reduce the contents of sulphur and phosphorus. Major developments followed during the 1960s and 1970s. Secondary metallurgy processing was introduced, enabling a further reduction of sulphur and phosphorus and other contaminants. Low hydrogen content through vacuum degassing and fine-graining by use of small additives of vanadium or aluminium were introduced. Ingot casting has now been replaced to a large extent by continuous casting. Increased knowledge of non-metallic inclusions and their influence on fatigue and fracture properties,[15] together with modern material analysis techniques and non-destructive testing methods, has raised the material quality in today's wheels and rails to a very high level.

1.1.4 Present situation

Today it is not only the case that higher speeds and greater axle loads are creating more demanding operating conditions for the wheel–rail interface; in a parallel development, computer-based numerical simulations, communication systems, control systems and data collection systems have all invaded railways since the 1980s and have led to a refined adhesion control of traction and braking, which has enabled better utilization of the performance of the wheel–rail interface but, unfortunately, often followed by new and different problems. In simulations, more advanced models can be handled numerically and integration of several modelling fields and also optimizations are now possible. However, the general analysis of problems related to wheel–rail rolling contact is still waiting for a major further increase in computer capacity and speed. For condition monitoring, new possibilities for data collection and computerized decision making are within reach.

In 2002 the European Rail Research Advisory Council (ERRAC, www.errac.org) published a Strategic Rail Research Agenda 2020 (SRRA 2020). The agenda was updated in 2007, and it includes climate change and global warming as new challenges making railway an even more attractive and urgent alternative means for transportation. The railways should meet the market demand which 'in Europe' is estimated by the SRRA to be a 100 % increase in passenger-kilometres and a 200 % increase in tonne-kilometre freight between the years 2000 and 2020. The SRRA highlights critical enabling technologies which need to be developed in five fundamental areas: (i) railway interoperability, (ii) intelligent mobility, (iii) safety and security, (iv) environment and (v) innovative materials and production methods. The

train–track interface is mentioned as an important area of research. ERRAC is an initiative by the European railway sector, the European Commission and the EU member states to revitalize the European rail sector and make it more competitive through increased innovation and research efforts at the European level.

In conclusion, the present situation encompasses higher speeds and axle loads, rapid technological development and a high demand for both quality and quantity of railway transportation. This constitutes a great challenge for continued in-depth research on the wheel–rail interface. There are unsolved problems and potential for improvement and innovation that can contribute to achieving lower production, maintenance, operating and environmental costs and to an overall improvement of the safety and quality of railway transportation. Since railways have a very long operational life, the future technical demands as anticipated in 10, 20 or even 30 years' time must be considered now! Key to success will be collaboration between railways, industries, consultancies, universities and governments in defining, carrying out and implementing research and in the training of a new generation of engineers and researchers.

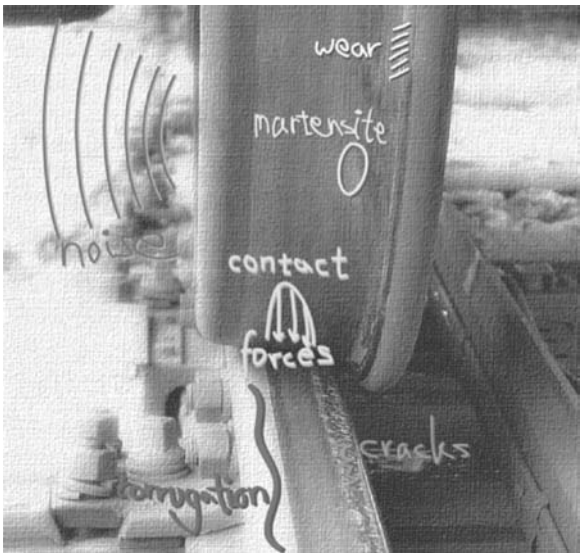
1.2 Phenomena in the wheel–rail interface

1.2.1 Introduction

The high energy efficiency of railway transportation is made possible by the favourably low losses in the rolling contact between the hard surfaces of the wheel and rail, which meet only in a very small contact patch. However, several undesired phenomena may occur in this contact, see Fig. 1.1. High vertical contact forces, but also lateral and longitudinal forces, induce stresses that may cause material yielding and fatigue. Rolling contact forces combined with friction induce wear. Traction and braking may lead to wheel sliding, resulting in rail burns and wheelflats, unfavourable material phase transformations and thermal cracks. These phenomena may create irregularities and/or worn profile geometries of the wheel and rail, resulting in poor vehicle dynamics and a further increase in contact forces and in vibrations and noise. The consequence may be discomfort and disturbance for passengers and the surroundings and also increased maintenance costs for wheels and rails and other components. Severe cases can even result in derailment induced by wheel or rail fracture or by the wheel flange climbing on the rail.

1.2.2 The wheel–rail interface

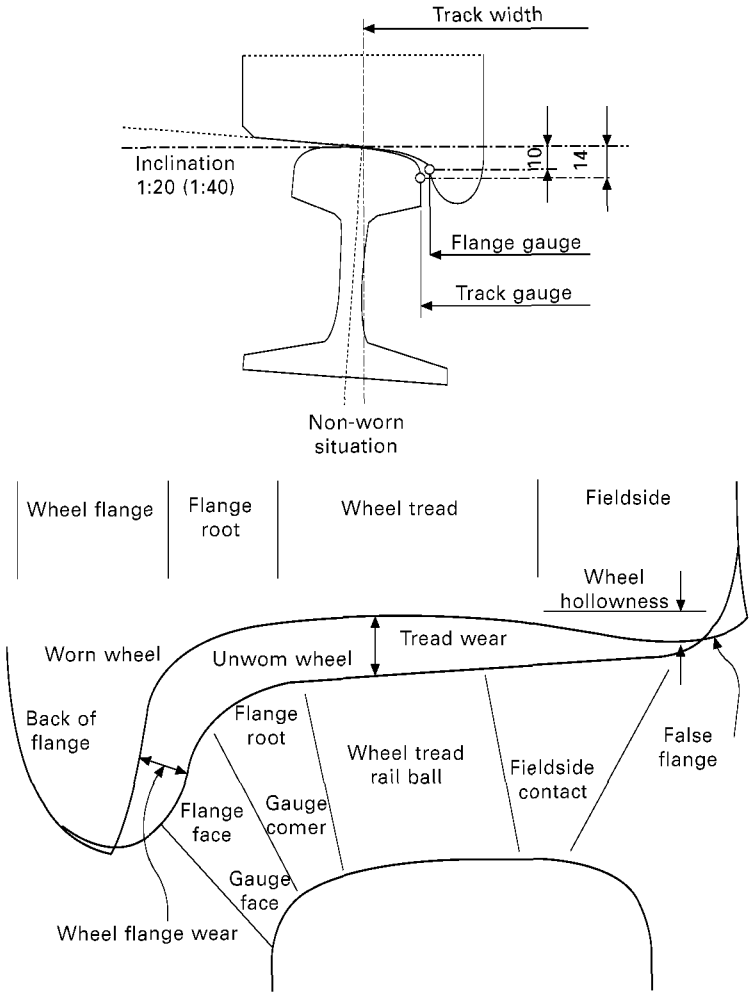
At typical wheel–rail interface in new or newly maintained condition and a wheel profile in worn condition are shown in Fig. 1.1. The original profiles



1.1 Several undesired phenomena may occur in the wheel–rail interface. (Picture from Ekberg *et al.* [16])

often follow some standard or practice, see Section 1.5, depending on the application. A typical contact patch, for standard wheel and rail profiles in new condition and a vertical contact force of 11 tonnes, is elliptic with size 18 × 11 mm, longitudinally and laterally, respectively. For the worn profile, the contact patch will be more circular. The conicity of the wheel tread provides the steering capacity of the wheelset and is therefore important for the running stability of the wheelset and bogie. The rail inclination may be 1.5–3°, dependent on national practices. Optimization of wheel tread and railhead profiles is discussed in Section 1.5.3 and in several other chapters in this book.

The frictional properties of the wheel–rail interface are difficult to characterize and quantify since they are strongly dependent on weather and contamination conditions. The friction coefficient could vary between 0.08 and 0.50. Naturally, this puts a limit on possible traction and braking efforts. Low friction may result in sliding of wheels and damage in the form of wheelflats for vehicles without, or with malfunctioning, slide protection. High friction may lead to increased energy consumption and excessive wear of wheel and rail. Traditional methods to control friction include sanding of the rail and lubrication of the wheel flanges or rail gauge corners in curves. In recent years so-called friction modifiers have been introduced, with the idea that not only should the friction coefficient be altered but also the adhesion curve (friction as function of wheel creep) should be tailored to



1.2 Upper: Wheel profile in contact with inclined rail. Flangeway clearance is seen. Lower: Wheel in new and worn states together with new rail. (Figures from Shevtsov[17])

solve a specific problem. Several chapters in this book are devoted to the wheel–rail friction problem; see Chapters 2, 4, 14 and 17.

In Europe, forged solid wheels, or rolled tyres mounted on a wheel centre, are used almost exclusively. The material is carbon steel, alloyed with some manganese, which is normally obtained by processing and continuous casting of scrap steel. Wheel materials are standardized, see Section 1.5.4. The most common wheel material in Europe is named ER7 and has max 0.52 % carbon and max 0.80 % manganese and a tensile ultimate stress of 820–940 MPa. For certain applications, low-alloyed steels are used.

Manufacturing of a forged railway wheel includes forging, rolling, heat treatment, hardness measurements, machining and ultrasonic testing. Heat treatment in the form of rim chilling results in a microstructure of lamellar pearlite with a favourable combination of toughness, strength and resistance to wear. This heat treatment also creates circumferential compressive residual stresses in the wheel rim which suppress propagation of transverse cracks. Wheels manufactured by casting are used in North America, South Africa and Australia, mainly for freight applications. The material has a carbon content around 0.70 %, following AAR (Association of American Railroads) standards. A higher carbon content elevates the material strength but also reduces the material toughness and makes the material more sensitive to thermal impacts. No chapter in this book considers wheelset technology. For further information, see Section 1.8.2, including the International Wheelset Congresses reference 18 and the partly historic document in reference 5.

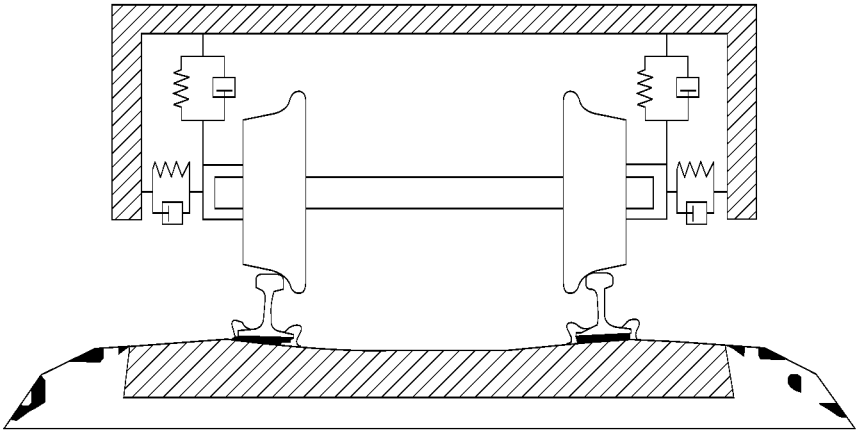
Rails are manufactured from continuously cast blooms. The most common rail material in Europe is R260 (previously named UIC 900A) which has 0.62–0.82 % carbon and 0.70–1.20 % manganese, and a tensile ultimate stress of minimum 880 MPa, see the standard in Section 1.5.4. There is an increased use of head-hardened rails (R350HT and R350LHT). The materials mentioned have a pearlitic microstructure. Critical parts in switches and crossings are often made from manganese steel (13 % manganese). At manufacturing, rails are checked for defects by continuous and automatic ultrasonic testing. See further references 18 and 19, and Chapter 5.

Modern developments in both wheels and rails include cleaner material, better process control and improved and automatic non-destructive testing. Materials having bainitic microstructure can provide certain advantages and are being tested both for wheel and rail applications.

1.2.3 Train–track interaction

The dynamic interaction between the train and the track determines the wheel–rail contact forces and the relative motion between wheelset and track, see Fig. 1.3. Many parameters influence the behaviour: the mechanical properties of the track (including rails, rail pads, fasteners, sleepers, ballast and underground), the mechanical properties of the vehicle (including wheelsets, bogies and wagon primary and secondary suspensions), the conditions in the wheel–rail interface (including its geometrical, frictional and deformation properties) and also track alignment, curves, axle load and train speed. Stability against lateral oscillations (hunting) and curving performance are important aspects.

In recent years, simulation of train–track interaction has become an important part of the train design process. In common simulations of vehicle dynamics, a relatively complex vehicle model is combined with a simple



1.3 Sketch of bogie with primary suspension and wheelset in contact with rails on railpads, sleepers and ballast. (Figure from Andersson *et al.* [26])

track model, see Fig. 1.3. Such a simulation is satisfactory for analysis of vertical and horizontal dynamics at low frequencies, less than, say, 20 Hz. However, for the vertical dynamics a simple vehicle model combined with a more complex track model is often more suitable, and frequencies up to 2000 Hz at conventional speeds may be important for studying forces induced by short-wave irregularities on the wheel and rail surfaces. The dynamic simulations provide data on forces and creep, which can be used as input to an evaluation of stresses, temperatures, wear, material deterioration, noise and ground vibrations. Also the vehicle stability aspects can be assessed. There are several modern textbooks on rail–vehicle dynamics,[10,20–22] and references on vertical dynamics.[23–25] See also Chapters 3 and 15.

Simulations should be verified by field measurements. The wheel–rail contact forces may be indirectly measured by use of instrumented wheels or rails. There is no known operational method for direct measurement in the contact patch of the wheel–rail contact forces. Traction and braking impose additional shear forces, creep, sliding and thermal loading that are often crucial for the wheel–rail interface. Further, tread braking of the wheels may impose excessive heating, thermal cracking and corrugation of the wheel rim that will indirectly affect train–track interaction.[27] Most of these aspects will be dealt with in the following chapters.

1.2.4 Deterioration mechanisms

The wheel–rail contact forces and relative motions will successively induce damage to both wheels and rails. A typical wheel life can vary between

300 000 and 2 500 000 km, including two to five reprofilings. Rails may last for 100–2500 MGT (million gross tonnes on two rails), much dependent on curve radii. The dominating causes of the deterioration differ from one application to another and between wheel and rail.

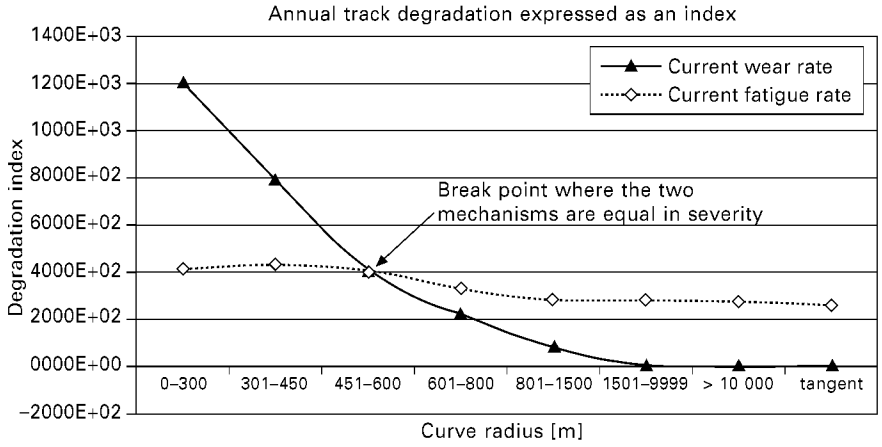
In wheels, the contact stresses may create rolling contact fatigue (RCF) in the surface layer or at subsurface positions (typically 5–25 mm below the surface), the latter with a potential for loss of part of the wheel rim, involving a risk of derailment. Plastification of the tread may change its cross-sectional profile or create wheel out-of-roundness. Wear mechanisms may result in a hollow tread or thin flange, see Fig. 1.2, or in an out-of-roundness. Even moderate wheel slippage may thermally affect the wheel material, creating surface cracks which may grow when exposed to contact forces. Sliding of the wheels may lead to phase transformation in the remaining surface material and to wheelflats that create very high contact forces. Block braked wheels can become severely damaged by the block–wheel interaction, causing thermal cracks in the tread or out-of-roundness of the wheel. The braking may cause an overheating of the entire wheel rim and, after cooling, create tensile residual stresses in the circumferential direction that induce global fracture of the wheel. Different types of wheel tread damage have been the subject of recent overviews.[28,29]

Rails are affected by a combination of local wheel–rail contact stresses and 'global' bending stresses. In continuously welded rails there are also thermal tensile or compressive stresses with a magnitude depending on the ambient temperature. Further, residual stresses from manufacturing may play a role in rail deterioration. Surface-initiated RCF may manifest itself as head checks, gauge corner cracks or squats. RCF can lead to loss of parts of the running surface layer or result in a full rail breakage both involving derailment risks. Gauge corner wear is a common problem in curves. Corrugation of different wavelengths may lead to increased contact forces.[30] Discontinuities such as welds, insulation joints and switches cause increased forces. An overview is given by Cannon *et al.*[19] A recent investigation on rail breaks and insulated joints has also been reported[31].

An interesting phenomenon, occurring on both wheels and rails, is that a certain amount of wear may be beneficial since it removes shallow surface cracks and thereby reduces the risk of larger RCF cracks. The wear rate should then match the crack propagation rate, see Fig. 1.4. Most of the aspects mentioned in this section are addressed in Chapters 3 and 6–13.

1.2.5 Safety issues

As mentioned above, the behaviour of the wheel–rail interface may pose safety problems with failure modes that could result in derailment. Flange climbing, material breakage and wheel and rail fracture are such failure modes.



1.4 Wear and fatigue mechanisms as functions of curve radius. The degradation index corresponds to a relative degradation rate. (Figure from Larsson[32])

Naturally, inspection and maintenance practices focus on preventing these. Two recent major accidents involving human and material losses occurred at Eschede in Germany in 1998 and at Hatfield in UK in 2000. They were initiated by wheel failure and rail failure, respectively.[33,34] The knowledge gained from these accidents is now being used to reduce the risk of similar accidents in the future. Wheel–rail safety aspects with focus on the rail are reported in the special issue edited by Lundén.[35]

1.2.6 Environmental issues

The wheel–rail interface is a major source of the train noise which disturbs both people in the surroundings and passengers onboard the train. Irregularities on the wheel tread and the railhead excite vibrations of the wheels and rails followed by emission of noise. This is called rolling noise and is a major problem for freight trains which often operate at night. Tread braked wheels with cast iron blocks become corrugated which leads to approximately 10 dB more rolling noise compared to disk braked trains at the same speed. New brake blocks are now being implemented to mitigate this problem. Squeal noise is generated by stick–slip in the wheel–rail interface involving double contact between wheel (both by tread and flange) and rail, typically for trams in sharp curves. Railway noise has been a major research issue for more than 15 years, and legislation work on railway noise is ongoing, see Section 1.5.

Structure-borne vibrations are generated at the wheel–rail interface and are transmitted via the bogie structure into the car body.[36] Sources of excitation

are periodic sleeper passages and irregularities on the running surfaces of the wheels and rails. Frequencies in the range 0–80 Hz are of interest for passenger comfort. Vibrations are also transmitted via railpads, sleepers and ballast to form ground vibrations that may propagate into surrounding buildings.[37,38] The frequency range 0–30 Hz is important here. See further Chapter 16.

Particle emissions from railway traffic may involve comfort disturbances and potential health risks. In tunnels and at underground platforms in particular, high particle concentrations may occur. The wheel–rail interface emits worn particles to the air. This problem has previously attracted only limited attention but is now becoming a subject for research, see Section 1.5.2 and Chapter 19.

Railway transportation is energy efficient when utilized well. A main contributing factor is the low rolling resistance in the wheel–rail interface. However, the friction coefficient, which is weather-dependent, the curve radii, the maintenance status of the wheel–rail interface and other parameters affecting vehicle performance may strongly affect energy consumption. Increased energy consumption is coupled to increased wear in the wheel–rail interface, the management of which has a clear cost saving potential.

1.2.7 Operational control, condition monitoring and maintenance

The wheel–rail interface is used to electrically short-circuit the two rails of the track, thereby indicating the presence of a train to the signalling system. However, different contaminants, lubricants, friction modifiers, residuals from brake blocks, etc. may have an insulating effect that disturbs the short-circuiting ability of the wheel–rail interface. The problem is addressed in Chapter 18. See also Section 1.6.3.

Condition monitoring of the wheel–rail interface means regular data collection and processing to plan optimal maintenance activities. For this purpose, implementation of modern technology with sophisticated sensors, computers and data transmission systems has begun. Important areas are effective systems for measuring profiles and detecting surface damage and also ultrasonic testing for internal defects in both wheels and rails. A problem here is to identify those defects and cracks which are potentially dangerous. Continued research efforts into these problems and new innovative products to serve the approximately 1 000 000 km of railway tracks and 25–50 million railway wheels in the world are needed. See Chapter 25 in Part II.

Maintenance of the rail means grinding, rail repair and rail replacement. Grinding is increasingly used by railways in order to improve track quality, reduce track forces, remove defects, modify contact geometry and reduce rolling noise. Wheel tread maintenance means machining-off of material in

a workshop lathe. Aspects of these matters are considered in the EU project INNOTRACK, see Section 1.5.1, and in Chapter 20.

1.2.8 Management aspects and costs

The condition of the wheel–rail interface is an important part of the responsibilities of managers of the infrastructure and rolling stock. An obvious problem for many railways is that the rails and wheels are managed by different organizations. Also, laws and regulations may limit the scope for optimization. Nevertheless, the potential for cost savings should be exploited at all levels. New ideas for implementation of RAMS (reliability, availability, maintainability and safety) and LCC (life cycle cost) have been put forward in the two recent theses,[39,40] and are considered in the INNOTRACK project, see Section 1.5.1 Management of the wheel–rail interface has been addressed in several studies.[41–43] Cost models are presented in Chapter 21.

1.3 Research fields

The wheel–rail interface involves a number of research fields. The 'traditional' fields deal with the physical phenomena in the interface itself. Additional fields are indirectly related since they strongly affect the conditions in the wheel–rail interface or concern a maintenance strategy, etc. Some important research fields related to the wheel–rail interface are listed in the following together with comments on importance, ongoing research etc. Naturally, the areas listed closely connect to the description given in Section 1.2.

A main traditional field is, of course, *contact mechanics* involving forces and relative motion in the wheel–rail rolling contact. The field is well defined by the contents of the well-known textbook by Ken L Johnson.[44] Several aspects of rolling contact are treated such as Hertzian and non-Hertzian contact stresses, surface and subsurface stresses in the contacting bodies, elastic and plastic deformations, material shakedown and ratcheting and formation of residual stresses, together with thermal distributions and thermoelastic phenomena. Microslip, creep, sliding and tractive rolling of the bodies in contact are central. A challenge in this area is to develop more reliable and efficient theoretical models and numerical methods for elastoplastic rolling contacts.

Advanced *material mechanics* is central to modelling and understanding material behaviour under contacting surfaces with severe plastic strains often combined with strong anisotropy.[45]

Rolling contact fatigue manifests itself in crack formation and crack growth in the material close to, or at, the wheel–rail interface. The cracks form at the mesolevel among the grains (their typical size is 0.02 mm). Crack initiation can occur in the surface, normally due to severe plastification, and at a subsurface level, often starting at some microstructural defect.

To predict RCF the first requirement is to translate the uniaxial material behaviour, known from material tests, to the complex rotating triaxial stress state present under the rolling contacts. Second, the risk of RCF or, in cases where RCF is present, the operational life needs to be estimated. One approach is to employ a continuum approach where RCF is related to the (time-dependent) state of stress and strain in the component. A second approach is an assessment based on fracture mechanics as described below. Ongoing RCF research includes the search both for more advanced models capable of capturing more influencing factors and for faster models suitable for engineering applications.

Fracture mechanics deals with the strength of cracked components. It can be employed to predict final fracture, e.g. of a rail. Further, fracture mechanics-based models can be employed to predict the rate of crack growth. Adapting fracture mechanics to the study of RCF (i.e. cracks in the wheel–rail interface) is far from straightforward. Complicating factors as compared to 'plain' fatigue crack growth analysis include complex (time-dependent) states of stress and strain, plastification, crack face friction and anisotropic material. Further, fluid entrapment in the crack may play an important role. Both fundamental and applied research with bearing on the wheel–rail interface is ongoing in this area.

Research on *materials* concerns both the refinement of existing materials and the development of new types of materials. Ideally, the 'best' combinations of material parameters are sought. Today, the specified mechanical properties include ultimate strength, ductility, fatigue strength, hardness, wear resistance, impact strength, fracture toughness, crack threshold value and crack propagation parameters. The debate on the importance of these parameters for good wheel and rail performance is ongoing. Certain aspects may not be covered in the present specifications, for instance, resistance to heat shocks. Special applications, such as heavy duty components in switches and crossings, may give rise to unconventional solutions with high-alloyed materials also requiring special forging or casting operations, heat treatment and machining. Two-layer materials with a cheap base material together with a surface layer with tailored properties, e.g. for high resistance to wear, corrugation and RCF, is a current research area and may be a common solution in the future.

Research on *tribology* is central to understanding and optimizing the wheel–rail interface. Both dry and lubricated friction are important in the understanding of vehicle dynamics and traction and at braking. Ongoing research concerns modelling of wear mechanisms and development and application of friction modifiers to combat corrugation and noise and to improve traction characteristics.

Research in *dynamics* is exploited in the study of vehicle–track interaction. The results affect the behaviour of the wheel–rail interface and vice versa. Important areas of current research are vertical dynamics and also

vehicle–track behaviour in switches and crossings.[46] Often, non-linearities in the wheel–rail interface, originating both in the track and in the vehicle, cannot be disregarded and will call for numerically demanding calculations. Dynamic calculations combined with studies of corrugation, wear, fatigue, crack propagation and optimization are under development.

Research on *railway noise* includes its generation and radiation and also its propagation to the surroundings and into passenger compartments. Rolling noise, induced by wheel and rail corrugation of wavelengths in the order 20–200 mm, continues to be a dominating problem. However, aerodynamic noise is of growing importance since trains with very high speeds are becoming more common. Development and implementation of measures to effectively and economically reduce noise at the source and at different locations along its propagation path is ongoing. The parallel area of *vibration* has similar features although the frequencies considered are much lower than for noise. Of special interest are phenomena occurring at high speeds in combination with soft clay in the underground where excessive vibrations can be encountered.

Research on *particle emissions* has recently found application in railway transportation. Work is ongoing here on health hazards related to the chemical composition and the size distribution of the particles. In addition aerodynamics may find application here to map the transmission paths of the particles.

Some other research areas worth mentioning are as follows. *Optimization* is useful for finding the best solutions when several parameters have an influence, e.g. for tread profiles. *Mechatronics* and *active control* may create new systems and solutions that will affect the wheel–rail interface. In the area of *non-destructive material testing*, new methods are being developed and implemented to systematically and effectively verify the integrity of the wheel–rail interface.

The wheel–rail interface is of such importance for the total safety and economy of railway transportation that special tools are being developed to support management activities. Tools for improving RAMS and lowering LCC, see above, are subjects of current research. Finally, *maintenance* is an area that is studied from both economical and technical aspects.

1.4 Applications

Four application examples are given here to illustrate the different types of challenges that may occur when dealing with the wheel–rail interface.

1.4.1 High axle loads

Competition is forcing railway companies to reduce costs. Increased axle loads can be an effective way to reduce freight costs by providing an increased

capacity. Upgrading of tracks for higher axle loads is quite common for specific transportation routes but not for general lines. Such upgrading can sometimes be made with only minor modifications to track and rolling stock. In other cases, major investment is necessary. For rails, an increased axle load may induce a higher occurrence of gauge corner cracking and wear. For wheels, the result may be increased wear and RCF.[47] Improvements in bogie performance and rail grinding practice as well as lubrication are necessary. Also higher carbon content in the wheel material could be considered. However, when tread braking is used in combination with relatively short braking distances (high retardation), wheel sliding may create brittle martensite and wheelflats in the tread. Wheel–rail interface issues for heavy haul are treated in the IHHA Guidelines[18] and in a paper by Fröhling.[41]

1.4.2 High train speeds

High-speed passenger trains give new possibilities for competitive medium- and long-distance travel in a more sustainable, safer and convenient way than road and air transportation can offer. Between large population centres, new lines dedicated to high-speed operations may be economically feasible. A cheaper alternative is to increase speed by upgrading existing tracks and by introducing tilting trains which make it possible to negotiate curves at higher speeds while preserving passenger comfort. Special attention should then be paid to vertical dynamics, corrugation and RCF phenomena in the wheel–rail interface. Higher speeds influence the statistical scatter of the vertical loads. This means that the occurrence of very high contact forces increases with speed. Further, new mechanisms for corrugation development may follow with higher speed, which leads to a further augmentation of contact forces. The result can be increased wheel and rail RCF which may cause hazardous breakaway of wheel rim material. The described phenomena are often hard to predict. Rail corrugation can be removed by grinding or could possibly be counteracted once the mechanism behind corrugation development has been understood. Wheels for high-speed operations should be carefully handled, with respect both to choice of material and to regular inspections for cracks in the wheel rims.

1.4.3 Metro traffic

Metro applications mean a number of special challenges for the performance of the wheel–rail interface. The traffic is intense with repeated starts and stops. At rush hours there are many passengers in a confined environment. Reliability is of the highest importance here, since failures may cause not only a stop for a single train, but also a general delay or traffic chaos. Normally, disk braking is used, and the main problem for the wheel–rail interface is

high tractive forces and wheel slide that may cause RCF, thermal damage and wheelflats. Corrugation phenomena may occur for wheels and rails, especially on curves. Small curve radii may create problems with excessive flange wear, gauge corner wear and also squeal noise that can be disturbing for passengers both in the trains and at platforms. Mitigation can sometimes be successful through modifications in the wheel–rail interface itself, but in most cases this is achieved by improving the traction or braking system or the manner the train is being driven.

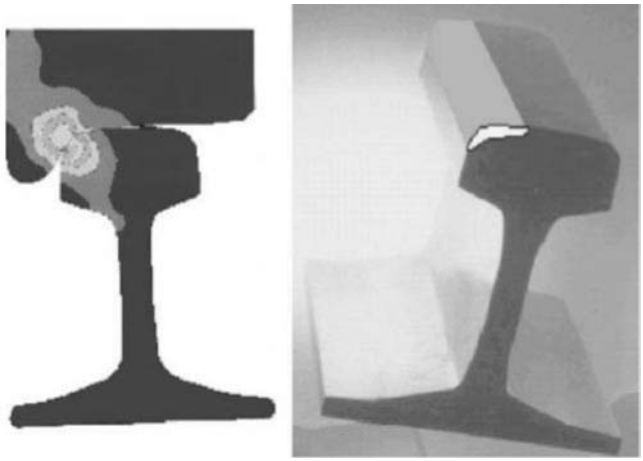
Particle emissions have only recently been recognized as a problem, not only for passengers and staff in the trains but also for people on or near the platforms. The confined environment, together with the wind created by the trains, contributes strongly to the problem. The emissions mainly come from the wheel–rail interface and the braking system but possibly also from the electrical supply system. A present challenge is to understand the magnitude and severity of this problem. The relative contribution from different sources should be found and effective solutions developed and implemented.

1.4.4 Two-layer materials

Two-layer materials may provide a potential solution to problems with RCF, wear and noise. The project INFRASTAR – 'Improving railway infrastructure productivity by sustainable two-material rail development' in the EU Fifth Framework Programme was run during 2000–2003. The aim was to increase the operational life and reduce the emitted noise of particularly exposed sections of railway track, such as small-radius curves subject to large traffic volumes and high axle loads.[48] The application of an extra surface layer to the railhead was investigated, see Fig. 1.5. Two different technologies were considered: the melting of powder onto the surface by means of a laser beam, and the rolling-in of an extra layer of material on the bloom during manufacture. A predictive model that can assist in the design of a fatigue-resistant two-material rail was developed and validated. From field tests of laser-cladded material it was concluded that the method can provide a solution for RCF of rails and that it also has a potential to reduce wear and noise problems. Further work will be needed to find suitable solutions for critical parts such as in switches and crossings. Despite the relatively positive results of INFRASTAR, two-layer materials are still waiting for a commercial breakthrough.

1.5 Ongoing research, development and standardization efforts

Some examples of ongoing research, development, standardization and legislation, all from the European scene, are presented. The projects may



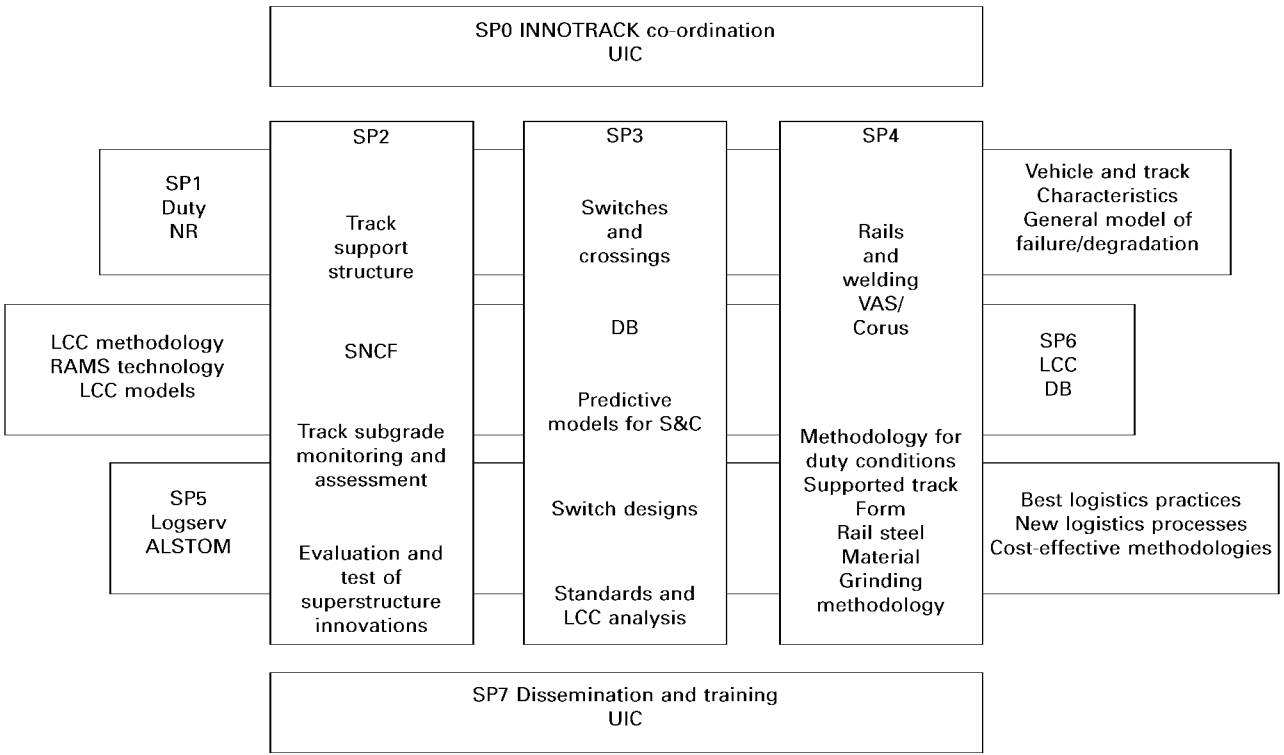
1.5 A surface layer with a selected alloy is applied to achieve required properties. (Figure from Hiensch *et al.*[48])

have a more general scope, but the present text concentrates on wheel–rail interface issues. Naturally, there are many other ongoing projects, both in Europe and elsewhere, that would be worth mentioning. Some further projects are mentioned in the following chapters of this book.

1.5.1 INNOTRACK

A substantial part of the ongoing (40 months from September 2006) project INNOTRACK (Innovative track systems, see www.innotrack.eu) involves research related to the wheel–rail interface. INNOTRACK is an integrated project (IP) within the EU's Sixth Framework Programme that aims to deliver new products, processes and methodologies in order to achieve the ERRAC targets on increased quantities and qualities of rail transport on conventional lines with mixed traffic. INNOTRACK is the first European project with comprehensive co-operation between infrastructure managers, supply industry and academia regarding the complete track construction with the objective of reducing the rate of track degradation and maintenance intervention. Lower LCC and improved RAMS characteristics are the aim. INNOTRACK has 36 partners from railways, supply industry, consultancies and universities in 10 European countries. INNOTRACK has a budget of MEUR 18.6 and comprises 1226 man-months of work. The project is co-ordinated by UIC (Union Internationale des Chemins de Fer, see uic.org). An overview of the subprojects in INNOTRACK is given in Fig. 1.6. The subprojects SP3 and SP4 directly involve the wheel–rail interface issues and are commented upon in the following paragraphs.

Wheel–rail interface issues in the INNOTRACK subproject 'Switches and



1.6 INNOTRACK work breakdown at subproject level.

Crossings' include an innovative prediction model for switch dynamics and deterioration. Predictions are made of the effect of increased axle loads and other changes of traffic in the form of increased dynamic load magnitudes, increased wear and increased RCF. The work will result in guidelines for the selection of rail steel grades and types of crossings and associated installations, and for maintenance procedures and technologies for certain duty conditions.

Advanced wheel–rail contact simulations are performed to identify sections in the switch and crossing panels of a turnout having critical loading. For these sections a local elastoplastic finite element model is employed which can handle both complex contact conditions and advanced material behaviour. The model predicts plastification, crack initiation and wear. A parametric study of the influence of different combinations of wheel profiles, axle loads, train speeds and frictional conditions will provide data for statistical distributions of the loading.

In addition the subproject 'Rails and welding' includes several wheel–rail interface issues. The project aims to deliver guidelines for the use of the full range of rail steel grades based on technical performance and LCC considerations. Relevant tests to quantify material properties of the rail steel that reflect operational strength more closely will be derived. This should make possible a mechanistic understanding of rail degradation. 'Minimum action' rules, based on problem-oriented application of fracture mechanics principles, will be established to enable a move towards predictive maintenance for rail defect management. Further, predictive methodologies will be developed for determining the best frequency of grinding, minimum amount of metal to be removed to maintain a 'crack-free' rail, increased track availability and reduced LCC. Work will be performed on flash-butt rail welds in order to narrow the heat-affected zone, which should reduce differential wear between the weld and the parent rail and the associated consequences for track quality degradation. Field measurements and numerical simulations are combined to study the evolution of squats. The influence of rail corrugation on wheel and rail deterioration is studied with the aim of identifying intervention criteria for rail grinding. To minimize the number of rail replacements, tolerances and minimum action rules for defects in rails and joints are being defined. These rules will depend on defect location and size, but also on operational conditions and rail geometry. Laboratory tests are performed to evaluate the ability of new technologies to improve detection and monitoring.

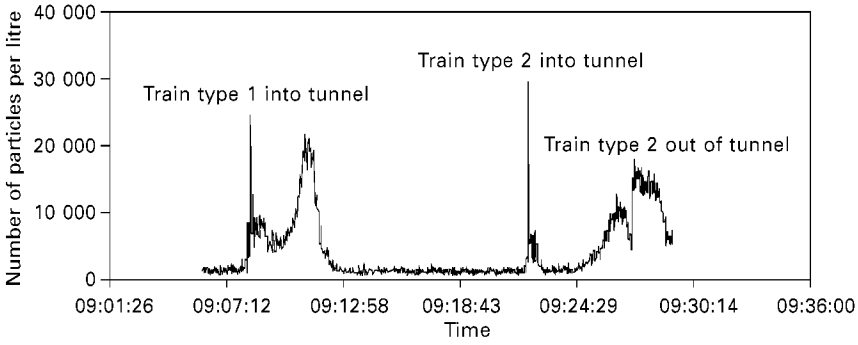
1.5.2 Particle emissions

As mentioned above, particle emissions from railway traffic have previously attracted only limited attention but are now becoming a focus for research.[49,50] Examples of emission sources are diesel engines, brake systems, current

collector systems and the wheel–rail interface. Particle concentrations can become high especially in tunnels and at underground platforms, and can expose passengers and staff to levels that may be hazardous to health. The health risks are related to both size and chemical composition of the emitted particles. IVL Swedish Environmental Research Institute and Chalmers/CHARMEC are presently running a research project to quantify the emissions of wear particles from railways. The emission data will be used in models for estimating the air quality at surrounding areas. Particle emissions are studied in tunnels, see Fig. 1.7, and onboard trains using optical particle counters. For the onboard measurements, the instrument is placed under the train and collects particles from wheels and brakes. The tunnel measurements aim at obtaining comparable data for different train types. One objective of the project is to develop methods to measure, as a function of particle emissions, instantaneous wear rates on locomotives, wagons and rails for different train types, train accelerations/retardations, train speeds, curve radii, weather, etc. Such methods can be very useful for both operational purposes and research projects on wear mechanisms. Simultaneous measurements of emitted noise are planned with the aim of correlating wear rates and noise levels. Banverket (The Swedish National Rail Administration, www.banverket.se) supports the present project and also a parallel project at Karolinska Institutet (www.ki.se) in Stockholm on health aspects of particle emissions.

1.5.3 Wheel–rail profile optimization

Optimization of the geometrical shapes of the interacting wheel and rail is not the subject of a specific chapter of the present book. However, the subject



1.7 Example of measured particle emission in a tunnel at Hindås (close to Gothenburg in Sweden) in August 2007. An optical particle counter was placed at the tunnel entrance approximately 0.5 m from the train and 2.0 m above the ground. Note the secondary peaks several minutes after the train enters the tunnel which is due to wind conditions.

is directly or indirectly addressed in almost every chapter. The geometry influences most critical train performance parameters such as vehicle stability and curving performance, derailment safety, wear, rolling contact fatigue, etc. At TU Delft a project has recently resulted in a doctoral dissertation 'Wheel–rail optimization' which addresses this problem in applications to tram, metro and conventional passenger traffic and which contains some 160 references in the field.[17] Some of the results are published in works by Shevtsov *et al.*[51,52] Examples of other work addressing this problem are found in papers by Fröhling and by Persson and Iwnicki.[41,53]

1.5.4 Rail and wheel standards in Europe

A massive European standardization effort is carried out by CEN (Comité Européen de Normalisation/the European Committee for Standardization, see www.cen.eu) which 'contributes to the objectives of the European Union and European Economic Area with voluntary technical standards which promote free trade, the safety of workers and consumers, interoperability of networks, environmental protection, exploitation of research and development programmes, and public procurement'. CEN organizes some 400 technical committees, working groups and other bodies. Technical Committee 256 (TC256) concerns railway applications, with subcommittees SC1, SC2 and SC3 on track, wheelsets–bogies and braking, respectively. Proposals for standards are being developed in working groups. The proposals are processed in the committees and then subject to review, revision and voting in the EU member states before they are published.

Working groups having a clear connection to wheel–rail interface issues are CEN/TC256/SC1/WG 4: Rails, CEN/TC 256/SC 1/WG 21: Acceptance of trackwork after renewal and/or maintenance work, CEN/TC 256/SC 2/WG 11: Wheels–Wheelsets, CEN/TC256/WG 3: Noise emission, CEN/TC 256/WG 10: Vehicle/Track interaction and CEN/TC 256/WG 38: Flange lubrication. The standard EN 13674-1 Part 1: Vignole railway rails 46 kg/m (on 'ordinary rails') was published in 2003 and is now under revision. It contains specifications on all technical aspects of the rail including the full geometry. It is partly based on the document UIC 860, but the present standard has new regulations on qualifying tests and more detailed requirements on acceptance tests (e.g. microstructure, cleanness, ultrasonics) and geometry. The procurement of rail in Europe today is based on this standard. The standard also covers rails for switches and check rails. Work is ongoing on standards for maintenance (including, e.g. grinding of rails). For wheels there are EN 13262 Wheels – Product requirement and EN 13979-1 Technical approval procedure – Part 1: Forged and rolled wheels. These two standards contain full specifications on all technical aspects of the wheels except for the tread profile that will have its own standard and which is under development.

Also, a standard on wheel maintenance is under development. Again the EN standards on wheels are based on documents from UIC and its previous research organization ERRI (European Rail Research Institute, earlier ORE). In addition to the CEN standardization work, parallel work has been carried out by UIC to develop its standard documents.

1.5.5 Technical specifications for interoperability

TSIs (technical specifications for interoperability) are being developed and implemented to meet requirements in EU directives 96/48/EC and 2001/16/EC. The TSIs are created to co-ordinate the railway systems within the EU, enabling more traffic and higher efficiency. The borders shall be opened for all operators and technical and administrative obstacles removed. Rolling stock and infrastructure operators shall be approved within all EU countries in a common process. The goals for traffic safety shall be the same for all EU countries. The TSIs cover different subsystems, such as infrastructure, rolling stock and energy. There is a separate TSI on noise. The first part of each TSI reproduces the legislative wording of the Commission's decision. The second part lays down the technical specifications. The TSIs are being prepared by ERA (European Railway Agency, see www.era.europa.eu). There are TSIs for conventional railways (maximum speed 190 km/h) and for high-speed railways (maximum speed above 190 km/h). Some TSIs for high speed were implemented in 2002 and for conventional railways during 2006. During 2008 several TSIs were implemented. TSIs relevant for the wheel–rail interface are, for instance, TSI Freight wagon (implemented in 2006) and TSI Infrastructure (not yet implemented). In their technical contents these TSIs follow the EN standards relatively closely for both wheel and rail specifications.

1.6 System aspects and optimization

Technical, economical and managerial aspects of the wheel–rail interface have been described above. It has been emphasized that the wheel–rail interface has a strong interaction with the total train–track system. Phenomena occurring in the wheel–rail interface will affect the train–track system, and vice versa. Vehicle dynamics and contact mechanics are the two most important aspects that will affect wear, corrugation, mechanical damage, derailment risks, maintenance costs, etc. A good management of the wheel–rail interface requires a thorough understanding of the inherent mechanisms and the influence of different parameters. A holistic approach is essential to avoid suboptimization of the system by improving it from one aspect but creating a more costly problem from another aspect.[41] One should also be aware of the fact that some phenomena are so complex that they are very hard to

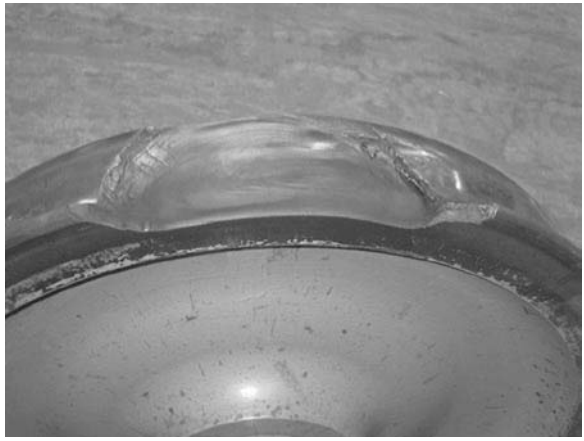
predict and sometimes cannot be explained or quantified in definite manner. A few examples are given here to illustrate the diversity and complexity of the problems.

1.6.1 Matching bogie type to track curvature

Three interesting examples on optimization of wheel–rail profiles have been reported by Kalousek.[54] The first example concerns an inter-city train with rigid bogies which developed high wheel conicity resulting in RCF damage on tangent track. The second example is a heavy haul line with many curves and trains with three-piece bogies where the wheels experienced excessive flange wear and RCF due to insufficient steering of the bogies. In the third example, a mass transit system with bogie hunting resulting in concave worn tread profiles and rapid rail corrugation is studied. All three cases, although very different, were of such a nature that they could be treated by profile optimization measures.

1.6.2 RCF in high-speed

Some years ago, the Swedish high-speed tilting trains X2 (with speeds up to 200 km/h) had a relatively large occurrence of subsurface cracks due to RCF.[23] A number of wheels exhibited serious deep shelling followed by breakaway of large pieces of the wheel rim, see Fig. 1.8. An interesting fact



1.8 Wheel from Swedish high-speed tilting train X2 with a fracture originating from deep shelling which occurred at full speed on the line between Stockholm and Gothenburg in February 2000. A large piece of the tread and flange has broken away. (Photo from Tuzik, 2006).

was that only trailing wheels suffered from RCF and not the powered wheels. It is believed that the problems were caused by excessive vertical contact forces in combination with non-metallic inclusions (size of order 1 mm) in the wheel material. The high-frequency vertical contact forces in the range 100–1000 Hz have been shown to be important in this context. Corrugation on tangent track developed by tractive forces from powered wheels may have contributed. The occurrence of cracks has now been drastically reduced, most likely due to new rail grinding practices.

1.6.3 LL brake blocks

The ERS (Euro Rolling Silently) project of the EU Fifth Framework Programme was run from 2003–2006 with the aim of developing a so-called 'LL' type brake block for tread braked freight wagons.[55] The blocks should be such that, without modifications to the wagons, they could replace the existing cast iron blocks of grade P10 (i.e. a retrofit solution was sought). The new blocks should produce lower rolling noise than existing blocks (by causing only small irregularities on the tread of the wheels) and fulfil a number of requirements on life-cycle costs, dimensions, block/wheel friction, wheel/rail adhesion, temperature distribution, metal pick-up, etc., and they should be recyclable. An unexpected problem encountered during the project was that new composition brake blocks introduced a layer on the wheel tread that reduced the capability of the wheelset to short-circuit the two rails, which is used for signalling purposes. Field tests were set up to find the susceptibility of the blocks when it comes to disturbing the short-circuiting. The testing proved very costly and time-consuming. For this reason a laboratory testing method is now being developed that should be useful for testing and approving new brake blocks.[56]

1.7 Future trends

Demand for railway transportation is increasing. The focus on global climate change during the last years means additional prospects for railways through tram, metro, suburban and high-speed passenger transportation and for freight transportation since these are superior means of transportation from an ecological point of view. However, to attract passengers, customers and investors, train operations need to be more attractive and competitive. To this end, ongoing re-regulation and standardization for interoperability will hopefully mean that different obstacles will be removed and that the railways will be opened up for many new commercially viable transport solutions, both for passenger and freight, nationally and internationally. Further new research findings and technologies are being, and will be, implemented to improve the efficiency, reliability, safety and economy of railway transportation.

Six areas that are deemed (by the authors) to be of major importance, both in research and implementation, during the next 5–15 years are listed here and briefly commented upon.

- **Legislation and standardization** will be a major area in the near future and will have a significant impact on the future development of railways. Interoperability will be the most important part and will affect the wheel–rail interface with regard to materials, geometry, braking systems, maintenance rules, safety requirements, etc. Successively new research results and technological development will be adopted and give rise to new possibilities. Noise emissions will be further restricted calling for more sophisticated methods to reduce noise including, in particular, better brake blocks for freight wagons.
- **Data collection and condition monitoring systems** of good quality already exist to some extent, but will find a major breakthrough during the next decade. The collected data will be used for safety inspections, planning of maintenance, systematic follow-up and evaluation of general performance and management decisions, and also for research. Since vehicle and track may have separate managements, co-ordination will have to be developed. New systems will also be used for determining track fees for train operators.
- **Active control** of the wheel–rail interface and its surrounding systems will be more common. Results of ongoing research on active components and systems will find new applications in improving vehicle dynamics and braking performance and reducing RCF, wear and noise levels, etc. New active components will include magnetostrictive materials for energy-efficient and fast response controlled by distributed computers. It is deemed also that freight wagons will, within the foreseeable future, include electrical systems for monitoring and control.
- **New materials** will continue to be an area of research. Some further refinements of existing materials will be implemented and some new classes, including two-layer materials, will find implementations in curves, switches and crossings, etc. An area that should be the subject for future research is the relevance of the now commonly used material parameters (ultimate stress, hardness, fracture toughness, etc.) for meeting the duty conditions in the wheel–rail interface. Friction modifiers and their use will be further developed and will find more applications.
- **Integrated computer models** for vehicle dynamics and wheel/rail damage and corrugation will be further developed and validated and also integrated into user-friendly systems. Research in areas such as dynamics, contact mechanics, plasticity, fracture mechanics, advanced material modelling, corrugation and tribology will continue and will increase understanding of complicated phenomena and unexpected behaviour. The models will

be used both for finding robust designs and for developing maintenance strategies.

- **Particle emissions** from railways and their impact on human health will be a subject for intensive research during the next decade. Future legislation will include requirements demanding that solutions for reducing emissions be found. The consequences for the wheel–rail interface may depend on its relative contribution and are at this stage hard to predict. However, methods to reduce wear, RCF and noise will also reduce emissions and maintenance costs.

1.8 Sources of further information and advice

The some 1200 references of the present handbook provide a large number of sources of information. In the following some important sources are listed and briefly described.

1.8.1 Journals and textbooks/handbooks

No scientific journal is dedicated solely to the wheel–rail interface. Among the railway-related international scientific journals, *IMechE Part F: Journal of Rail and Rapid Transit* and *Vehicle System Dynamics* contain many papers on the wheel–rail interface. Also journals representing specific scientific fields contain several articles on the wheel–rail interface. Ten such journals are *Engineering Fracture Mechanics*, *Fatigue & Fracture of Engineering Materials & Structures*, *IMechE Part C: Journal of Mechanical Engineering Science*, *International Journal for Numerical Methods in Engineering*, *International Journal of Solids and Structures*, *Journal of Applied Mechanics*, *Journal of the Mechanics and Physics of Solids*, *Journal of Sound and Vibration*, *Journal of Tribology* and *Wear*. Useful articles can also be found in *European Railway Journal*, *International Railway Journal*, *Railway Gazette International* and *Railway Track and Structures*. There are many examples of textbooks/handbooks with important contents on the wheel–rail interface. [7,10,13,20–22, 26,44,57–60]

1.8.2 Conferences

The conference *Contact Mechanics and Wear of Rail/Wheel Systems* is dedicated to the wheel–rail interface. The first conference was held in 1982 and the eighth conference will be held in September 2009 in Florence, Italy (www.cm2009.org). The conference is nowadays held triennially and is dominated by universities, although industry and railway participation has increased. The conference attracts an important part of the research community in the field. Since the fourth conference in 1991, most of the conference

papers have been subject to peer-review and have been published in special issues of the journal *Wear*,[61–63] a system which elevates the attraction, impact and value of the conference.

The *IAVSD Symposia* (International Association for Vehicle System Dynamics, www.iavsd.org) are held biennially and include both road and rail dynamics. The 21st Symposium will be held in August 2009 in Stockholm, Sweden. A number of the papers are related to the wheel–rail interface and the published proceedings are peer-reviewed. Also the *IWRN* (International Workshop on Railway Noise) has relevance for the wheel–rail interface and the papers result in peer-reviewed publications.

Important information on wheel–rail interface issues is provided in proceedings and at coming conferences of *Railway Bogies and Running Gears* (www.railveh.bme.hu), *IHHA* (International Heavy Haul Association, www.ihha.net), *IWC* (International Wheelset Congress, www.iwo16.com), *ASME Joint Rail Conference* (www.asmeconferences.org), *RAD Schiene* (Internationale Schienenfahrzeugtagung Dresden, www.rad-schiene.de), *RCM* (Railway Condition Monitoring, www.theit.org), *STECH* (International Symposium on Speed-up, Safety and Service Technology for Railway and Maglev Systems, www.jsme.or.jp), *UIC High Speed* (Union Internationale des Chemins de Fer, see www.uic.org) and *WCRR* (World Conference of Railway Research, www.wcrr2008.org).

1.8.3 Internet and journal databases

The Internet is a useful source of information. A fairly new phenomenon is on-line publications, see for example Interface – *The Journal of Wheel/Rail Interaction* (www.interfacejournal.com). An internet search in March 2008 using *wheel–rail interface* gave 13 000 hits. For "*rolling contact fatigue*" (which means that the exact phrase *rolling contact fatigue*, not just those three words, are found) there were 26 000 hits, while *rolling contact fatigue* rendered 204 000 hits. A search on the article database *Scopus* (access rights are required), which is a multidisciplinary database in science, engineering, medicine and social science that covers nearly 14 000 scientific journals starting in 1966, gave hits as follows: *wheel/rail interface* 230, *wheel–rail interface* 170, "*rolling contact fatigue*" 1250 and *rolling contact fatigue* 2030 article references.

1.8.4 Organizations, industries, consultancies, research groups and networks

The railway organizations AAR (Association of American Railroads, www.aar.org) and UIC (Union Internationale des Chemins de Fer, www.uic.org) carry out research projects and standardization work and provide

documentation. Also the railway administrations, train operators, national institutions, and major railway industries and consultancies carry out substantial research work and contribute to scientific conferences and journal publications. There are many important research organizations and university groups involved in research on the wheel–rail interface, most of them represented by the authors of the chapters in this book or through the references. Many of them collaborate with parallel groups or with rail organizations and industries, thereby creating many informal networks. The network EURNEX (European Rail Research Network of Excellence, www.eurnex.eu) was financed by EU during 2005–2007. From 2008 it is continuing as EURNEX Association which is being financed by its members.

1.9 Acknowledgements

The Centre of Excellence CHARMEC (CHAlmers Railway MEChanics, www.chalmers.scharmec) at the Department of Applied Mechanics, Chalmers University of Technology in Gothenburg, Sweden, has supported the writing of the present chapter. The assistance from Professors Bengt Åkesson, Jens Nielsen and Birger Karlsson and from Dr Anders Ekberg is acknowledged.

1.10 References

1. Karlsson L-O (1998), Private communication, October 2008.
2. Persson B (2000), *Från ånghäst till "elåsna" – Lok och vagnar på Malmbanan* (*Locomotives and wagons on the Iron Ore Line*), Malmbanans Vänner, Luleå, Sweden (in Swedish).
3. Lundén R, Nordmark T and Paulsson B (2001), Enhancing iron ore transportation in Sweden, *Proceedings 7th International Heavy Haul Conference*, Brisbane, Australia, Qld, 11–14 June, 99–106.
4. Lundén R (1998), LKAB invests in 30 tonne axleloads, *Railway Gazette International*, **154**(9), 585–7 (+ 1 additional page in reprint).
5. Wise S (1987), Railway wheelsets – a critical review, *Proceedings of the Institution of Mechanical Engineers*, **291**(D4), 257–71.
6. Koerfer E (1985), 150 Jahre Schienenerzeugung – Überblick über die technische Entwicklung, *Stahl und Eisen*, **105**(17), 907–12.
7. Profillidis V A (2000), *Railway Engineering* (2nd edn), Ashgate, Aldershot, UK.
8. Timoshenko S P (1953), *History of Strength of Materials*, McGraw Hill, New York, USA.
9. Lundberg G and Palmgren A (1947), Dynamic capacity of rolling bearings, *Acta Polytechnica, Mechanical Engineering Series*, Royal Swedish Academy of Engineering Sciences **1**(3), 50 pp.
10. Iwnicki S (ed.) (2006), *Handbook of Railway Vehicle Dynamics*, CRC, Boca Raton, FL, USA.
11. Knothe K (2008), 'History of wheel/rail contact mechanics: from Redtenbacher to Kalker', *Vehicle System Dynamics*, **46**(1), 9–26.

12. Dowson D (1998), *History of Tribology* (2nd edn), Professional Engineering, London, UK.
13. Suh N P (1986), *Tribophysics*, Prentice-Hall, Englewood Cliffs, NJ, USA.
14. Jernkontoret (2004), *Iron and Steel Production – Development in Sweden 1850 to 2000*, Stockholm, Sweden (in Swedish).
15. Kiessling R and Lange N (1978), *Non-Metallic Inclusions in Steel*, The Metals Society, Book No 194, London, UK.
16. Ekberg A, Kabo E, Nielsen J C O and Ringsberg J W (2003), 'Researchers on the track of wheel–rail interaction', *Railway Gazette International*, **159**(6), 397–99.
17. Shevtsov I Y (2008), *Wheel/Rail Interface Optimization*, Doctoral Dissertation, TU Delft, Delft, The Netherlands.
18. IHHA (2001), *Guidelines to Best Practices for Heavy Haul Railway Operations: Wheel and Rail Interface Issues*, International Heavy Haul Association, Virginia Beach, VA, USA.
19. Cannon D F, Edel K-O, Grassie S L and Sawley K (2003), Rail defects: an overview, *Fatigue & Fracture of Engineering Materials & Structures*, **26**(10), 865–86.
20. Wickens A H (2003), *Fundamentals in Rail Vehicle Dynamics*, Swets & Zeitlinger, Lisse, The Netherlands.
21. Knothe K and Stichel S (2003), *Schienenfahrzeugdynamik*, Springer, Berlin, Germany (in German).
22. Shabana A A, Zaazaa K E and Sugiyama H (2008), *Railroad Vehicle Dynamics – A Computational Approach*, CRC, Boca Raton, FL, USA.
23. Gullers P, Andersson L and Lundén R (2008), High-frequency vertical wheel–rail contact forces – field measurements and influence of track irregularities', *Wear*, **265**(9–10), 1472–78.
24. Nielsen J C O, Ekberg A and Lundén R (2005), Influence of short-pitch wheel/rail corrugation on rolling contact fatigue, *IMechE, Part F: Journal of Rail and Rapid Transit*, **219**(3), 177–87.
25. Steenbergen M J M M (2008), *Wheel–Rail Interaction at Short Wave Irregularities*, Doctoral Dissertation, TU Delft, Delft, the Netherlands.
26. Andersson E, Berg M and Stichel S (2007), *Rail Vehicle Dynamics*, KTH Railway Group, Stockholm, Sweden.
27. Vernersson T A (2006), *Tread Braking of Railway Wheels – Noise-Related Tread Roughness and Dimensioning Wheel Temperatures: Field Tests, Rig Measurements and Numerical Simulations*, Doctoral Dissertation, Chalmers Applied Mechanics, Gothenburg, Sweden.
28. Ekberg A and Kabo E (2005), Fatigue of railway wheels and rails under rolling contact and thermal loading – an overview, *Wear*, **258**(7–8), 1288–300.
29. Deuce R (2007), *Wheel Tread Damage – An Elementary Guide*, Report 100115000, Bombardier Transportation, Netphen, Germany.
30. Nielsen J C O, Lundén R, Johansson A and Vernersson T (2003), Train–track interaction and mechanisms of irregular wear on wheel and rail surfaces, *Vehicle System Dynamics*, **40**(1–3), 3–54.
31. Sandström J (2008), *Analysis of Rail Breaks and Insulated Joints Deterioration*, Licentiate Thesis, Chalmers Applied Mechanics, Gothenburg, Sweden.
32. Larsson D (2004), *A Study of the Track Degradation Process Related to Changes in Railway Traffic*, Licentiate Thesis, Luleå University of Technology, Luleå, Sweden.
33. Esslinger V, Kieselbach R, Koller R and Weisse B (2004), The railway accident of Eschede – technical background, *Engineering Failure Analysis*, **11**(4), 515–35.

34. Smith R A (2003), The wheel–rail interface – some recent accidents, *Fatigue & Fracture of Engineering Materials & Structures*, **26**(10), 901–7.
35. Lundén R (guest ed.) (2003), Wheel/rail safety: special focus on the rail. Special Issue, *Fatigue & Fracture of Engineering Materials & Structures*, **26**(10), 861–1031.
36. Sjövall P (2007), *Identification and Synthesis of Components for Vibration Transfer Path Analysis*, Doctoral Dissertation, Chalmers Applied Mechanics, Gothenburg, Sweden.
37. Karlström A (2006), *On the Modelling of Train Induced Ground Vibrations with Analytical Methods*, Doctoral Dissertation, Chalmers Applied Mechanics, Gothenburg, Sweden.
38. Lane H (2007), *Computational Railway Dynamics – Integrated Track-Train-Subgrade Modeling and Simulations*, Doctoral Dissertation, Chalmers Applied Mechanics, Gothenburg, Sweden.
39. Espling U (2007), *Maintenance Strategy for a Railway Infrastructure in a Regulated Environment*, Doctoral Dissertation, Luleå University of Technology, Luleå, Sweden.
40. Patra A P (2007), *RAMS and LCC in Rail Track Maintenance*, Licentiate Thesis, Luleå University of Technology, Luleå, Sweden.
41. Fröhling R D (2007), Wheel/rail interface management in heavy haul railway operations – applying science and technology, *Vehicle System Dynamics*, **45**(7), 649–77.
42. Kapoor A, Schmid F and Fletcher D (2002), Managing the critical wheel/rail interface, *Railway Gazette International*, **158**(1), 25–8.
43. Tuzik R E (2006), Improving the wheel/rail interface, *Railway Age*, **207**(7), 21–2.
44. Johnson K L (1985), *Contact Mechanics*, Cambridge University Press, Cambridge, UK.
45. Johansson G (2006), *On the Modeling of Large Ratcheting Strains and Anisotropy in Pearlitic Steel*, Doctoral Dissertation, Chalmers Applied Mechanics, Gothenburg, Sweden.
46. Kassa E (2007), *Dynamic Train–Turnout Interaction – Mathematical Modelling, Numerical Simulation and Field Testing*, Doctoral Dissertation, Chalmers Applied Mechanics, Gothenburg, Sweden.
47. Ekberg A and Marais J (2000), Effects of imperfections on fatigue initiation in railway wheels, *IMechE, Part F: Journal of Rail and Rapid Transit*, **214**(1), 45–54.
48. Hiensch M, Larsson P-O, Nilsson O, Levy D, Kapoor A, Franklin F, Nielsen J C O, Ringsberg J W and Josefson B L (2005), Two-material rail development: field test results regarding rolling contact fatigue and squeal noise behaviour, *Wear*, **258**(7–8), 964–72.
49. Seaton A, Cherrie J, Dennekamp M, Donaldson K, Hurley J F and Tran C L (2005), The London Underground: dust and hazards to health, *Occupational and Environmental Medicine*, **62**(6), 355–62.
50. Karlsson H L, Nilsson L and Möller L (2005), Subway particles are more genotoxic than street particles and induce oxidative stress in cultured human lung cells, *Chemical Research in Toxicology*, **18**(1), 19–23.
51. Shevtsov I Y, Markine V L and Esveld C (2005), Optimal design of wheel profile for railway vehicles, *Wear*, **258**, 1022–30.
52. Shevtsov I Y, Markine V L, Li Z and Esveld C (2008), Design of railway wheel profile taking into account rolling contact fatigue and wear, *Wear*, **265**(9–10), 1273–82.

53. Persson I and Iwnicki S D (2004), Optimization of railway profiles using a genetic algoritm, *Vehicle System Dynamics*, **41**, (Supplement) 517–26.
54. Kalousek J (2003), Wheel/rail damage and its relation to track curvature, *Proceedings 6th International Conference on Contact Mechanics and Wear of Rail/Wheel Systems*, Gothenburg, Sweden, 10–13 June, 1–6.
55. Lundén R and Åkesson B (2006), *Chalmers Railway Mechanics – A NUTEK/VINNOVA Competence Centre. Triennial Report 1 July 2003 – 30 June 2006/Review 1 July 1995 – 30 June 2003/Plans 1 July 2006 – 30 June 2009*, Chalmers University of Technology, Gothenburg, Sweden.
56. Bauer H, Dörsch S, Freudenberg H, Geissler H J, Rode C, Staudt V and Viet J J (2007), Development of a laboratory test method evaluating the effect of composite brake blocks in the transition resistance of track circuits, *Proceedings 15th International Wheelset Congress*, Prague, Czech Republic, 23–27 September, on CD.
57. Popp K and Schiehlen W (eds) (2003), *System Dynamics and Long-Term Behaviour of Railway Vehicles, Track and Subgrade*, Springer, Berlin, Germany.
58. Kalker J J (1990), *Three-Dimensional Elastic Bodies in Rolling Contact*, Kluwer, Dordrecht, the Netherlands.
59. Lichtberger B (2005), *Track Compendium – Formation, Permanent Way, Maintenance, Economics*, Eurailpress, Hamburg, Germany.
60. Esveld C (2001), *Modern Railway Track* (2nd edn), MRT Productions, Zaltbommel, the Netherlands.
61. Kimura Y, Beynon J H and Suda Y (guest eds) (2002), Special Issue: Contact mechanics and wear of rail/wheel systems, CM2000, July 2000, *Wear*, **253**(1–2), 1–318.
62. Ekberg A, Ringsberg J W and Lundén R (guest eds) (2005), Special Issue: Contact mechanics and wear of rail/wheel systems, CM2003, June 2003, *Wear*, **258**(7–8), 953–1335.
63. Grassie S L (guest ed.) (2008), Special Issue: Contact mechanics and wear of rail/wheel systems, CM2006, September 2006, *Wear*, **265**(9–10), 1149–555.

Copyrighted Materials
Copyright © 2009 Woodhead Publishing Retrieved from www.knovel.com

2

Basic tribology of the wheel–rail contact

R. LEWIS, University of Sheffield, UK; U. OLOFSSON, Royal Institute of Technology (KTH), Sweden

Abstract: The aim of this chapter is to introduce some of the basics of wheel–rail tribology before the topics are dealt with in greater detail in the subsequent chapters. Wear, contact mechanics, fatigue and adhesion are covered. Great emphasis is placed on the need to consider these subjects together to optimise the management of the wheel–rail contact as they all interact with each other. Frequent references are made to chapters where further information on the issues being discussed can be found.

Key words: wheel–rail tribology, wear, contact mechanics, fatigue, adhesion.

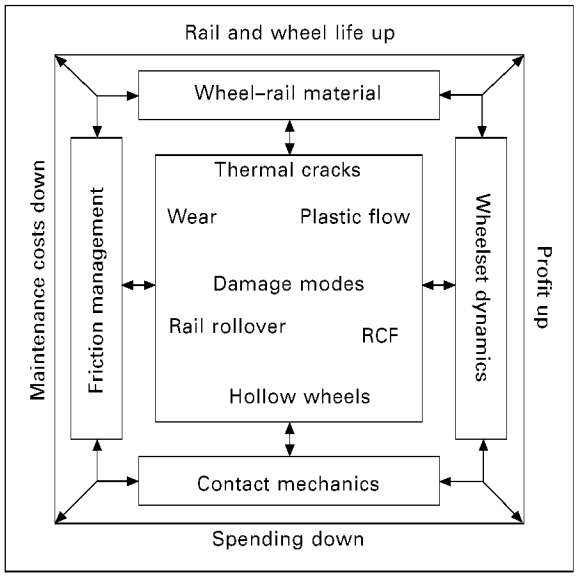
2.1 Introduction

As will be seen throughout this book the wheel–rail contact is extremely complex. A broad interdisciplinary approach is needed to understand, model and optimise it. The contact is absolutely critical to the safe and efficient operation of a railway network.

The area does not really have many analogies with other engineering component contacts, which makes it difficult to transfer knowledge from other areas. The closest perhaps are rolling ball bearing or gears, but these are generally closed systems with fairly well defined contact conditions that do not vary a great deal during the life of the bearing or gear. Much of the complexity of the wheel–rail contact is brought about by the open nature of the system and hence the constantly varying environmental conditions. Along a length of line the position of the contact and its size and the resulting contact stresses also vary constantly and will be different, not just for each railway vehicle, but for each wheel as each, although starting with the same profile, will have worn by different amounts.

As with virtually all engineering systems, costs are paramount, so the wheel–rail contact has to be operating at an optimum, but for the least cost. This is illustrated well in Fig. 2.1, where all the issues are shown.

The critical issues are all interlinked in some way. For example, wear and rolling contact fatigue (RCF); if crack growth is truncated, wear failures will be more likely to be the problem; if wear is reduced, however, cracks may be able to grow to the point where an RCF failure occurs. Lubricants/friction modifiers can be used to control wear, but may have an influence on RCF crack growth. Shakedown limits and also rolling contact fatigue



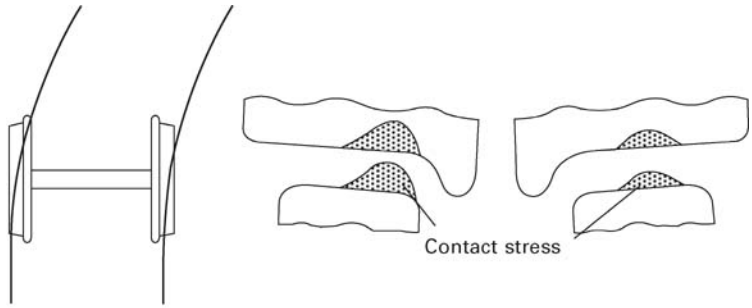
2.1 Systems approach to wheel–rail interface management and research. (Adapted from Kalousek and Magel, (1997)

crack growth are dictated by contact pressure and friction in the contact. The friction, however, can be reduced or increased by friction modifiers and so on. In this chapter, each of these areas will discussed in relatively basic terms to give an overview of the more detailed information included in later chapters. Signposting is given to relevant chapters where further information is given on the various issues outlined.

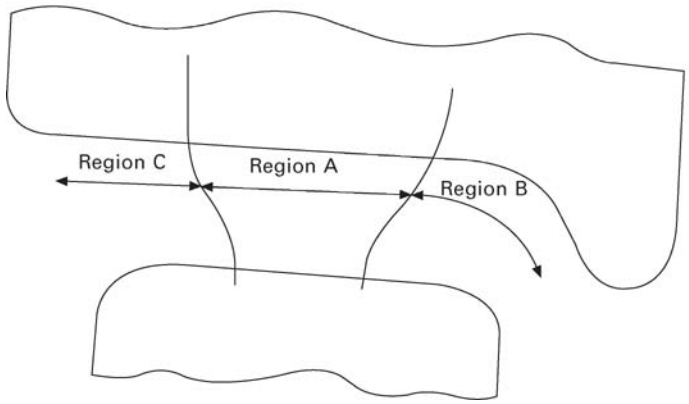
While not explicitly discussed in this chapter, wheel and rail materials have a very important role to play in optimising in wheel–rail contact management. The choice of which rail material is used, for example, can dictate whether wear or fatigue processes dominate (see Section 2.4.2 below). The development of rail materials is described in depth in Chapter 5.

2.2 Contact mechanics

The position of the wheel–rail contact, which is typically 1 cm^2 in size, varies continuously as a train progresses down a section of track. The exact position will depend on the wheel and rail profiles and the degree of curvature of the track and whether the wheel is the leading or trailing wheelset on a bogie, as well as other factors determined by the bogie design. In straight track, however, it is likely that the wheel tread and rail head will be in contact, with wheel flange and rail gauge corner contact occurring in curved track. Figure 2.2 shows how the contact position and stress varies for the two wheels on a leading wheelset entering a right-hand curve.



2.2 Leading wheelset entering a right-hand curve.



2.3 Wheel–rail contact zones. (Adapted from Tournay, 2001)

Three possible regions of wheel–rail contact have been defined (as shown in Fig. 2.3) (Tournay, 2001):

1. **Region A – Wheel tread–rail head**. The wheel–rail contact is made most often in this region and usually occurs as the railway vehicle is running on straight track or very high radius curves. This region yields the lowest contact stresses and lateral forces.
2. **Region B – Wheel flange–rail gauge corner**. The contact in this region is much smaller than that in region A and is often much more severe. Typically, contact stresses and wear rates are much higher. If high wear and material flow occurs, two-point contacts may evolve, where tread and flange contact is apparent.
3. **Region C – Contact between field sides of wheel and rail**. Contact is least likely to occur here and, if it does, high contact stresses are induced, which will lead to undesirable wear features causing incorrect steering of the wheelset.

As a result of the fact the contact position is not spread evenly over the entire wheel or rail profile, the shape of the profiles will change as time progresses, due to wear and material flow (processes largely controlled by the loads and creepage within the contact). In order to be able to predict how profiles may evolve, a good understanding is therefore required of the contact stress. Contact stress is also primarily responsible for driving rolling contact fatigue.

The simplest solution for determining wheel–rail contact geometry and stress is Hertz analysis (see for example Johnson, (1985)), where the wheel and rail can be equated to two cylinders in contact perpendicular to each other. The maximum contact pressure, p, is given by:

$$p = \sqrt[3]{\frac{3PE^2}{2\pi^3 R^2 (1 - \nu^2)^2}} \qquad [2.1]$$

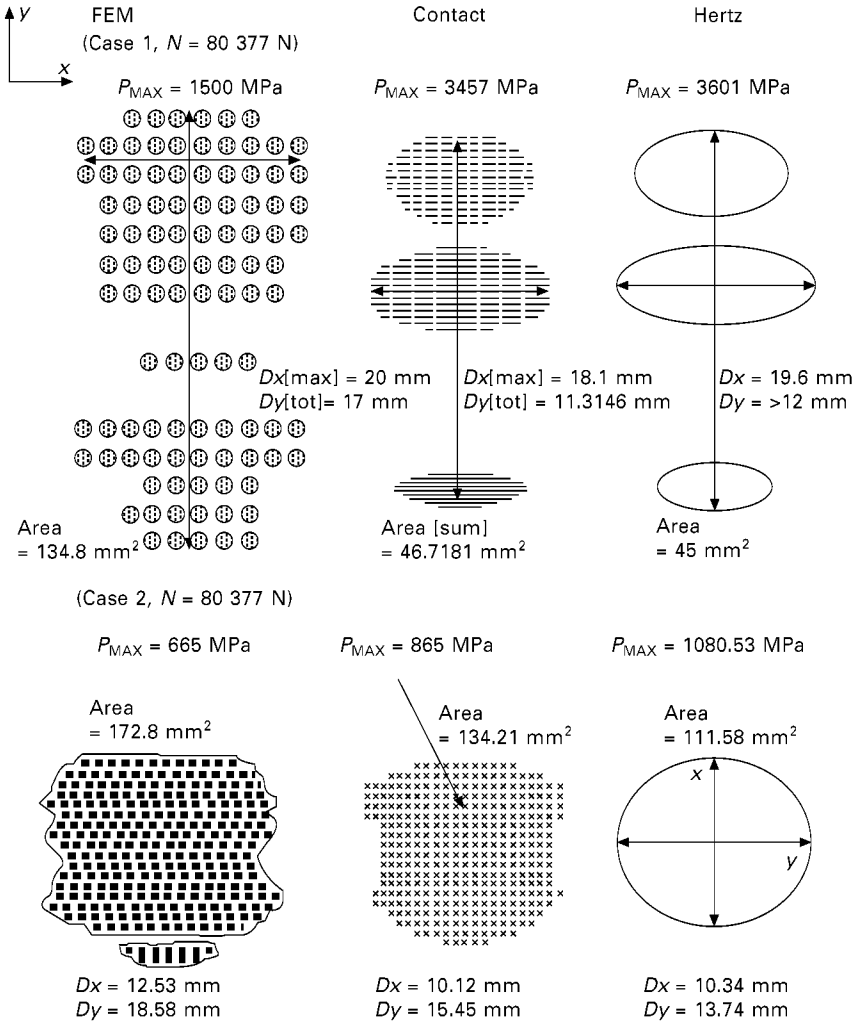
where P is the normal load, E and ν are the Young's modulus and Poisson' ratio, respectively (assumed to be same for wheel and rail materials in this case), and R is the equivalent radius given by:

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} \qquad [2.2]$$

where R_1 and R_2 are the contact radii of the wheel and rail. This approach, however, due to the assumptions made, such as smooth contacting surfaces, a linear elastic material response, a frictionless contact and that the contact dimensions must be small compared to the radii of curvature of the contacting bodies, is limited in accuracy. In the flange contact particularly, the contact radius can be as small as 10 mm, which means this assumption can be invalid. For a commuter train service the Hertzian maximum contact pressure can range from 600 MPa at wheel tread–rail head contact to 2700 MPa at the wheel flange/rail gauge corner contact (Telliskivi and Olofsson, 2001).

Numerical solvers have been developed for calculating area and stress for contacts approximated to Hertzian ellipses, such as FASTSIM (Kalker, 1982), and non-Hertzian contacts, such as CONTACT (Kalker, 1990). CONTACT, however, requires a large amount of computing resource and again is limited by the half-space assumption. Finite element modelling carried out by Telliskivi and Olofsson (2001), including the plastic deformation and actual wheel and rail profiles, showed good correlation in terms of contact area and stress with Hertz and CONTACT for the rail head contact, but considerably different stress and area for the rail gauge corner flange contact (as shown in Fig. 2.4, where Case 1 represents the gauge corner contact and Case 2 the head contact), due to the limiting half-space assumption in Hertz and CONTACT analysis.

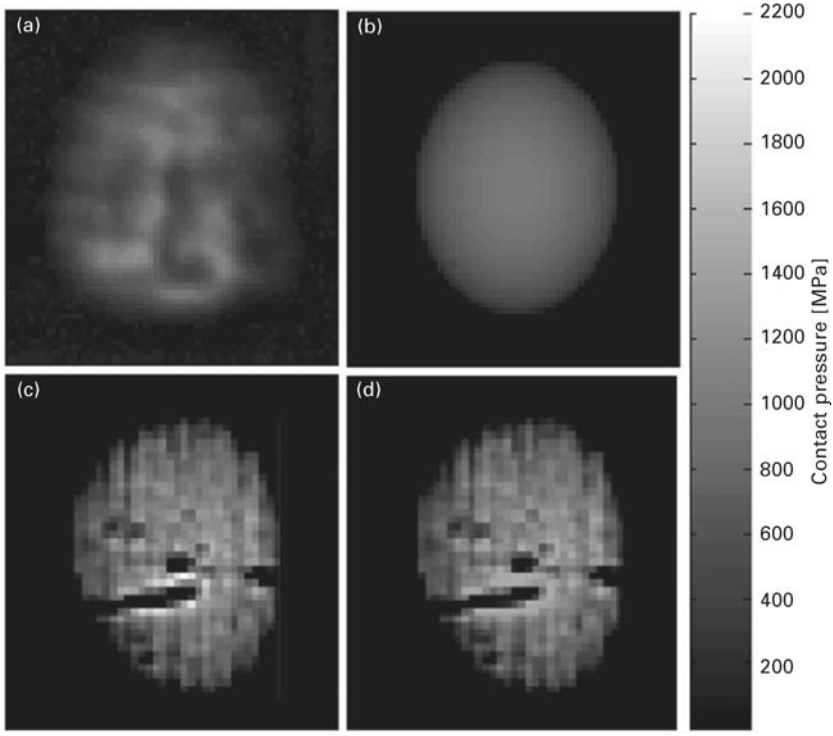
Recently, an innovative ultrasonic technique has been used to study the



2.4 Comparison between FE, CONTACT and Hertz Analysis. (Telliskivi and Olofsson, 2001)

wheel–rail contact (Marshall *et al.*, 2006). Figure 2.5 shows contact pressures derived from ultrasonic scans compared with numerical calculations using actual roughness profiles and Hertz analysis.

There is good global geometric correlation between the ultrasonic results and the numerical model. The degree and fragmentation of the ultrasonic and numerical contacts are qualitatively similar. However, on a local level the ultrasonic results and numerical solutions differ. This is probably due to difficulty in aligning the surfaces to the same orientation in both the



2.5 Contact pressure maps for a load of 80 kN: (a) ultrasonic measurement; (b) Hertzian; (c) elastic model; (d) elastic–plastic model. (Marshall *et al.*, 2006)

experiment and model. The elastic case assumes no localised yielding at the contact; this results in predicted contact pressures in excess of yield for the contacting surfaces. The pressures determined for the elastic case are far in excess of those measured ultrasonically. However, the elastic–plastic case shows similar peak pressures to the experiment. The experimental and elastic–plastic numerical model peak pressures are in excess of the Hertz solution; this is due to the reduced contact conformity attributable to roughening. Much more detail will be given on contact mechanics and dynamic simulations of the wheel–rail contact in Chapter 3, where the merits of the different analysis methods, such as those illustrated in Fig. 2.4 will be discussed.

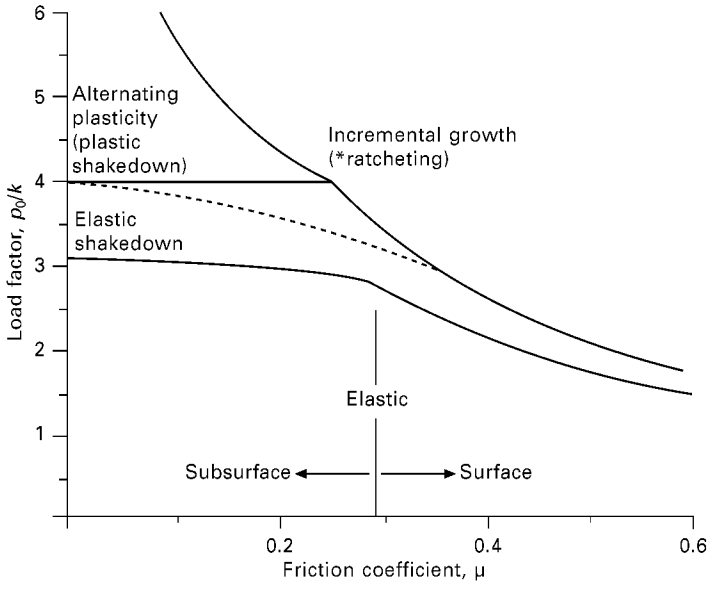
When tractive force is applied at the surface the shear stress increases and the position of the maximum value moves closer to the surface. Because of the rolling/sliding behaviour of a wheel on a rail, a cyclic buildup of plastic deformation occurs beneath the material surfaces. It is this behaviour that leads to rolling contact fatigue and wear occurring.

Figure 2.6 shows a shakedown map, which illustrates the relationship between friction in the wheel–rail contact and the load-carrying capacity of the contact (one of the interactions mentioned in Section (2.1). It shows the limits of material behaviour in terms of non-dimensional contact pressure, p_0/k as a function of friction coefficient, μ (= T/N), where p_0 is the normal contact pressure, k is the shear yield strength, T is the tractive force and N is the normal load.

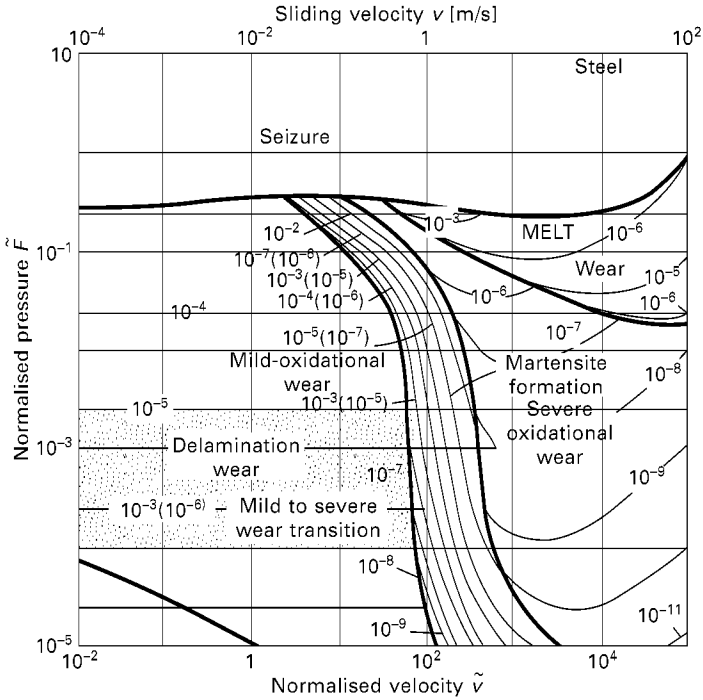
At relatively low friction coefficients, cumulative plastic flow occurs subsurface. For friction coefficients above about 0.3, plastic flow is greatest on the surface. The worst position in terms of damage to the material is in the ratcheting region, where strain is accumulated until the ductility of the material is exceeded and it is lost as wear debris or a surface crack is initiated.

2.3 Wear

Wear is the loss or displacement of material from a contacting surface. A material can wear by a number of different mechanisms or modes. How a material wears depends on the nature of the material and the other elements of the *tribo-system* which include environmental conditions and whether any contaminants are present, for example, wear debris and, in the case of the wheel–rail contact, friction modifiers, friction enhancers like sand, leaves, etc. This is illustrated by the wear map in Fig. 2.7, which shows the relationship



2.6 Shakedown map (see Chapter 7 for greater detail about this plot).



2.7 Steel-on-steel wear sliding wear map. (Lim and Ashby, 1987)

between wear behaviour and operating parameters. This was developed to explain the range of wear behaviour in dry siding tests of steels. In this case the wear behaviour is related to pressure and speed conditions. Regions are shown where particular wear mechanisms dominate and the contours are lines of constant wear rate. This is an extremely useful tool for engineers in designing machine element contacts. A similar map for wear of rail materials is shown later in the chapter.

2.3.1 Wear situations

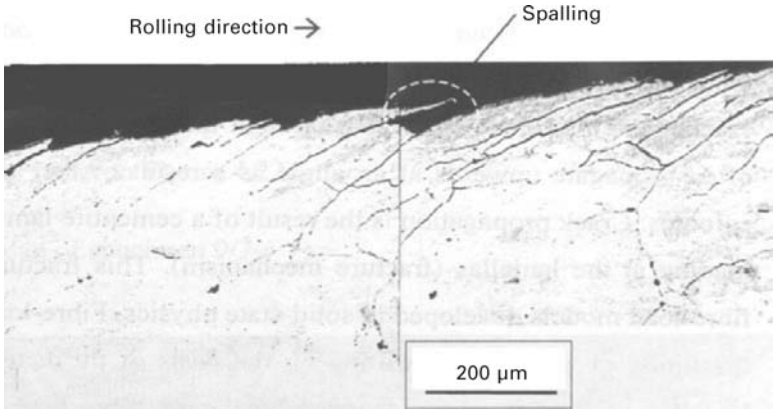
Before looking at actual wear mechanisms it is necessary to first characterise possible wear situations. These are related to the nature of the motion during contact and the number of contact cycles. In the case of the wheel–rail contact there are two of interest, sliding and rolling. While nominally a rolling situation the wheel–rail contact is split into regions of slip and stick (see Chapters 3 and 17, for more details), so some micro-sliding is occurring.

In terms of potential to cause wear, sliding motion (motion tangential to the surface) is more severe than motion perpendicular to the surface, such as occurs with impact or rolling. A number of wear mechanisms can occur as

a result of sliding including *oxidative wear* in mild contact conditions (low load and sliding velocity – typical of rail head–wheel tread contacts) and, in more severe conditions, *adhesive or galling wear* (possible in curves, where the rail gauge corner and wheel flange contact). If particles are present in the contact *abrasive* wear may also occur (e.g. it could occur when friction enhancers, such as sand or Sandite, are used). Very severe sliding conditions can lead to *seizure* and high *heat generation* in the contact, which may cause a thermal breakdown of the surface material (this has been observed in very tight curves).

With rolling motion, the principal mechanisms that dominate wear behaviour are fatigue driven. Generically, these are referred to as surface fatigue and are forms of repeated-cycle deformation mechanisms. Surface fatigue mechanisms involve the formation and propagation of cracks, which ultimately lead to the loss of particles from the surface, a process known as ratcheting (ratcheting is discussed further in Chapter 9 in the context of rail wear and rolling contact fatigue). These cracks tend to form below the surface and propagate to the surface. However, in cases where significant traction is involved, cracks form at the surface. Traction is generated in a rolling contact, when sliding occurs as well as rolling. Sliding is caused by slip as a result of the two components moving at different velocities. The wheel/rail contact is a classic example of a rolling/sliding action.

Figure 2.8 shows a section of a disc made from rail steel that has been rolled against a wheel steel disc in a twin-disc test to simulate the rolling/sliding conditions in an actual wheel–rail contact. The repeated loading cycles have led to the deformation of the surface material and the development of surface cracks, which are leading to material removal. It is very severe



2.8 Cracking and material removal resulting from a rolling/sliding wear situation.

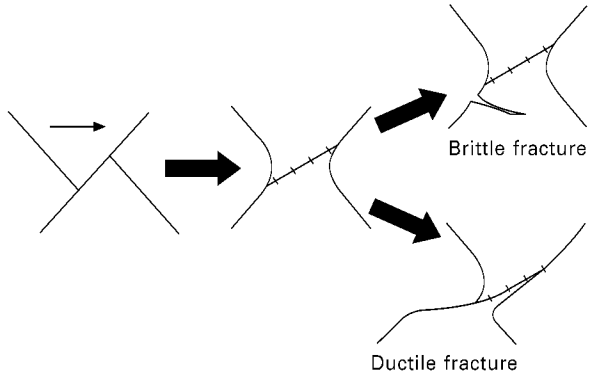
cases, where the cracks turn down as a result of bending moments in a rail, which can cause rail breakages. Surface wear features apparent with surface fatigue are spalls or pits on the material surface.

2.3.2 Wear mechanisms

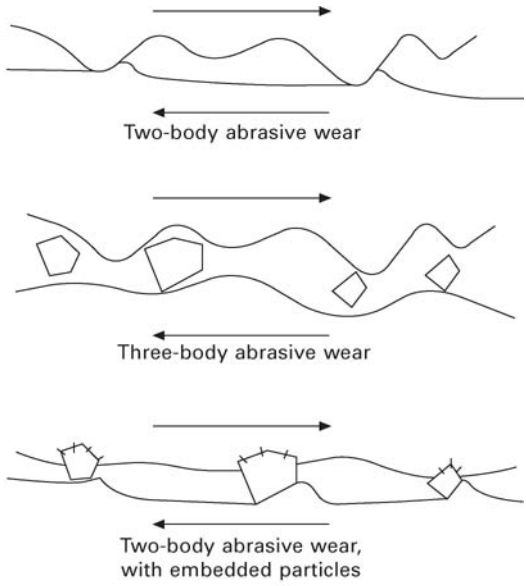
The basic concept for adhesive mechanisms is that actual contact between surfaces occurs at discrete points within the apparent area of contact. At these spots, called junctions, bonding occurs between surface asperities (see Fig. 2.9). When the surfaces move relative to each other, these junctions are broken and new ones formed. Usually the tip is plucked off the softer asperity leaving them adhering to the harder surface. This can be via a ductile or brittle fracture. Subsequently, they become loose and give rise to wear debris. Severe damage can sometimes result in the tearing away of macroscopic chunks of material, and this situation is known as *galling*. If adhesive wear results from the breakdown of lubrication in a contact then the term *scuffing* is used to describe the onset of wear.

Abrasive wear is damage to a component surface which arises because of the motion relative to that surface of either harder asperities (*two-body abrasive wear*) or because of hard particles trapped between the surfaces (*three-body abrasive wear*) (see Fig. 2.10). Such particles may be introduced between the two softer surfaces as a contaminant from the outside environment, or they may have been formed *in situ* by oxidation or by some other chemical or mechanical process.

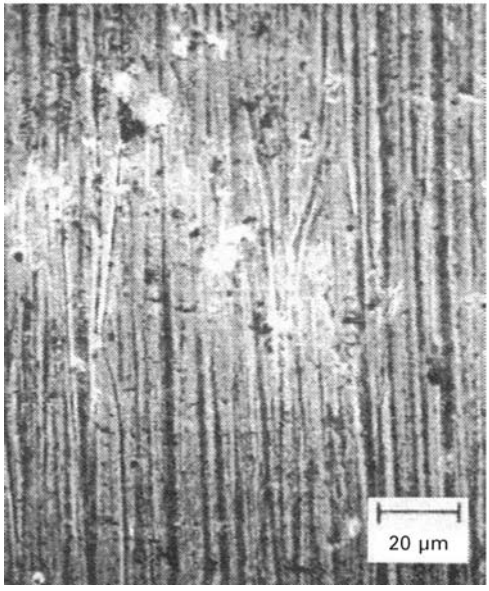
Two-body abrasive wear gives a characteristic surface topography consisting of long parallel grooves running in the sliding direction, as seen in Fig. 2.11. The volume and size of the grooves varies considerably from light scratching at one extreme to severe gouging at the other. Industrial surveys have shown that abrasive wear accounts for up to about 50 % of wear problems.



2.9 Adhesive wear mechanisms.



2.10 Abrasive wear mechanisms.



2.11 Abrasive scratching. (Swanson and Klann, 1981)

The rate of damage is relatively insensitive to hardness of the particles in a three-body situation, as long as they are at least 20 % harder than the surface itself. The most commonly occurring contaminant in industrial machinery is that from quartz or silica (these minerals make up about 60 %

of the Earth's crust) (silica sand is either used on its own as a friction enhancer in the wheel–rail contact or in Sandite). These are likely to have hardness in excess of 8 GPa and consequently do damage to even hardened steels (typically of hardness 7–8 GPa). The hardness of further particulates is shown in Table 2.1.

The oxidative wear process involves the formation of oxides on the surface of the material. It is clearly related to the ability of the wearing material to undergo oxidation and the availability of oxygen. Whether it occurs depends on the temperatures generated in the contact and the relative humidity. Wear rates are lower than those seen with mechanical wear processes. The material removal process is illustrated in Fig. 2.12, as well as the surface of a wheel disc from a twin disc test run under relatively mild contact conditions, where it can be seen that platelets of oxidised material have broken away from the surface.

Thermal wear processes are those directly associated with the increase in temperature caused by frictional heating in the contact. The principal type of wear process in this category is when a material melts or softens to such an extent that it can be displaced like a viscous fluid. Other mechanisms, such as adhesive wear, are also accelerated by a reduction in hardness. Other types are linked with thermal stresses that can cause thermal fatigue and cracking, which lead to loss of material.

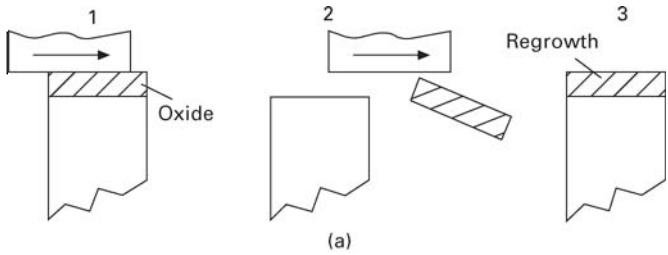
2.3.3 Wear rates and transitions

Wear is often classified as being *mild* or *severe*. This is not based on any particular numerical value of wear rate, but on the general observation that for any pair of materials, increasing the severity of the loading (e.g. by increasing either the normal load, sliding speed or bulk temperature) leads at some stage to a comparatively sudden jump in the wear rate. The differences in the two regimes are shown in Table 2.2.

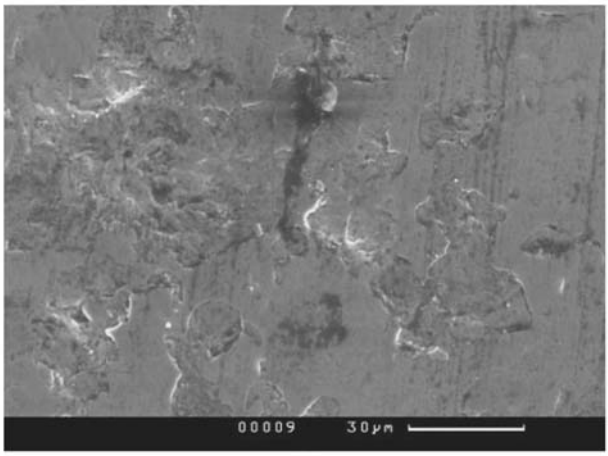
Normalised wear rates of 10^{-4}–10^{-3} are typical for mild wear with wear rates for severe wear reaching levels in the range of 10^{-3}–10^{-2}. The mechanisms

Table 2.1 Particle hardness values

Material	Hardness [GPa]
Diamond	60–100
Born carbide, B4 C	27–37
Silicon carbide	21–26
Alumina	18–29
Quartz	7.5–12
Garnet	6–10
Magnetite, $Fe_3\ O_4$	3.7–6
Soda-lime glass	about 5



(a)



(b)

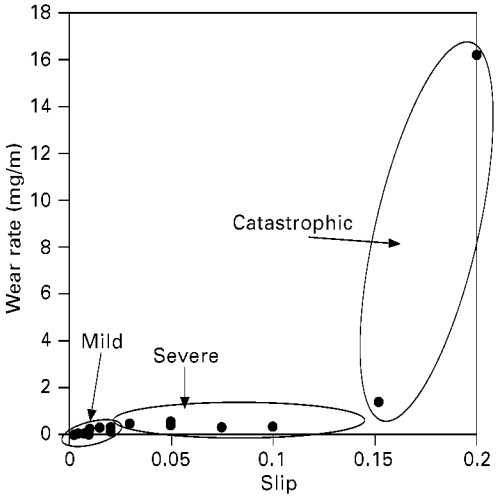
2.12 Oxidative wear: (a) typical surface morphology; (b) wear process.

Table 2.2 Mild and severe wear regimes

Mild wear	Severe wear
Results in extremely smooth surfaces – often smoother than the original	Results in rough, deeply torn surfaces – much rougher than the original
Debris extremely small, typically only 100 nm diameter	Large metallic wear debris, typically up to 0.01 mm diameter

most associated with severe wear are adhesive or thermal mechanisms. Increasing temperatures in the contact and the resulting thermal softening can lead to a further transition into a *catastrophic* wear regime.

The three wear regimes outlined above have been seen during rolling/sliding laboratory tests on wheel and rail materials, as shown in Fig. 2.13 for R8T wheel material. Wear mechanisms for wheel and rail materials are covered in greater detail in Chapters 6, 9 and 10.



2.13 Wear rates of R8T wheel material using a twin-disc rolling/sliding test. (Lewis and Dwyer-Joyce, 2004)

2.3.4 Wear modelling and mapping

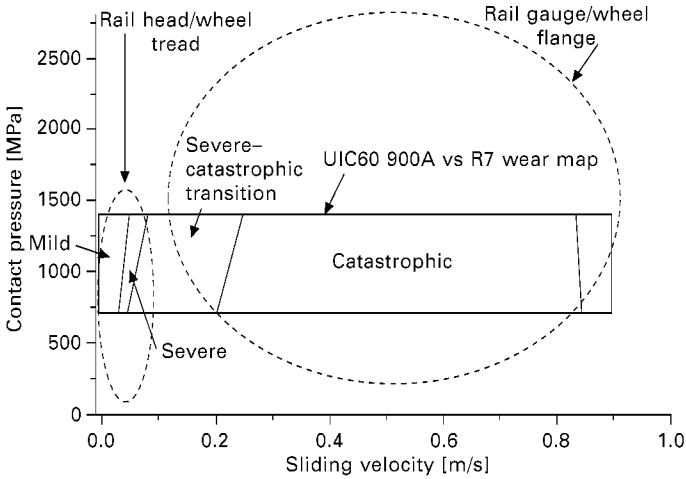
A common starting point for looking at modelling of wear situations is the Archard wear equation, which asserts that wear volume, V, is directly proportional to the load, P, on the contact and the sliding distance, l, but inversely proportional to the surface hardness, H, of the wearing material:

$$V = \frac{kPl}{H} \qquad [2.3]$$

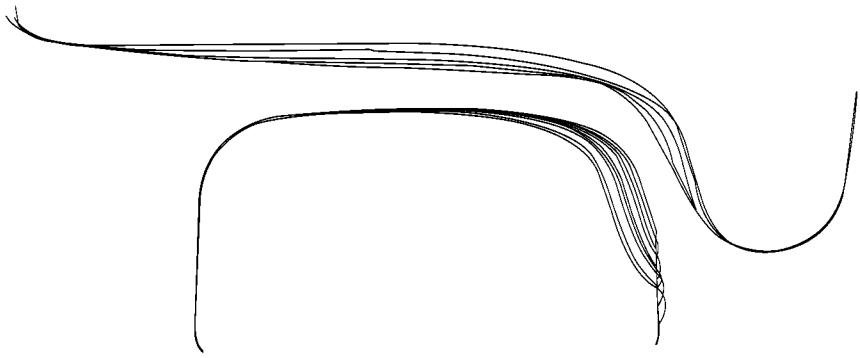
The constant, k, is known as the wear coefficient and varies for different material pairs. It is usually determined by carrying out sliding wear tests using the materials of interest.

Another method for modelling wear that has been widely applied in the wheel/rail context is the $T\gamma$ approach, where T is the tractive force given by the normal force multiplied by the friction in the contact and γ is the slip in the contact. This has been used in rolling contact fatigue modelling as well as wheel and rail wear modelling. Both of these approaches have been used successfully in combination with multi-body dynamics simulations to predict wear of wheel profiles (see Chapter 6) and to study rail corrugation (see Chapter 11).

A good way of displaying wear data is a map, as first illustrated in Fig. 2.9. The map in Fig. 2.14 is for a wheel–rail contact (Lewis and Olofsson, 2004). The wear data were built up using a mixture of twin disc and pin-on-disc testing methods using R7 wheel material and UIC 60 900A rail material. The map has been laid over some predicted wheel–rail contact conditions.



2.14 UIC60 900A rail steel wear map (Lewis and Olofsson, 2004). Plotted over wheel–rail contact conditions derived from GENSYS simulations (Jendel, 2000).

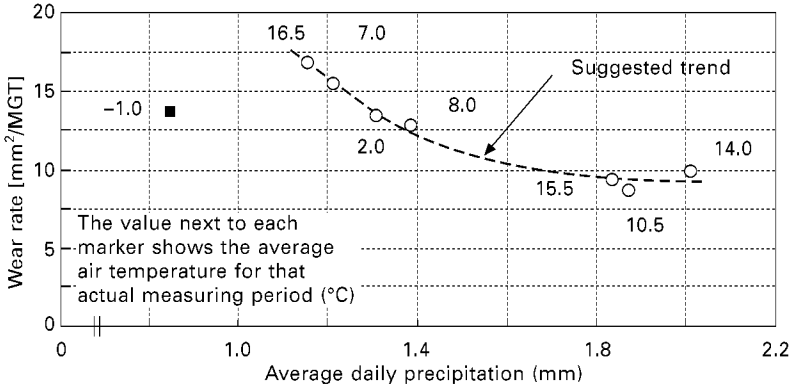


2.15 Form change of wheel and rail from a Stockholm test case. (Nilsson, 2003)

As can be seen the wheel tread–rail head contact falls in the mild to severe wear regime and the wheel flange–rail gauge corner contact is in the severe to catastrophic regimes. This matches what is seen in the field.

An example of the change of wheel and rail profiles over a two-year period in Stockholm local traffic can be seen in Fig. 2.15. Since the wheel–rail contact is an open system, wear will be influenced by factors such as humidity. 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 Fig. 2.16.

Further information on wear predictions using more complex simulations that incorporate rolling contact fatigue is detailed in Chapters 9 and 10.



2.16 Rail wear rate versus average daily precipitation. (From Nilsson, 2003) (MGT = mega gross tonne traffic)

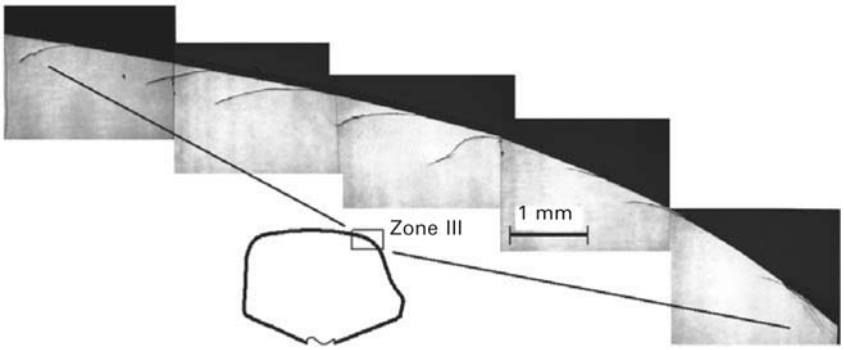
2.4 Fatigue

Fatigue failures for rails and wheels can be categorised into *surface* and *subsurface* failures, which result from crack initiation and growth. Surface fatigue in rails can lead to *head checking* or *squat* formation and subsurface fatigue can result in *shelling* or the formation of *tache ovale*. The initiation and growth of cracks in rail material is described in great detail in Chapters 9 and 10. More information on the features that result from the crack growth (listed above) can be found in Chapters 22–28. Wheel fatigue causes and features are covered in Chapter 7.

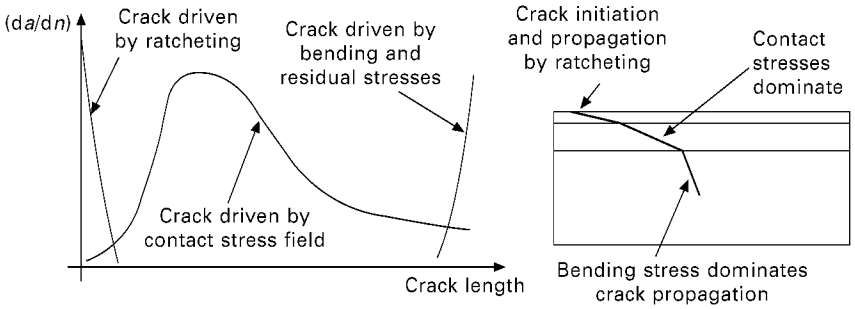
2.4.1 Rail and wheel fatigue

Head checks (shallow cracks at the rail surface) probably present the biggest problem in rails and normally occur on the gauge corner on curves (see Fig. 2.17). They result from accumulation of plastic strain increments (ratcheting), which eventually exhaust the ductility of the surface material, at which point cracks can initiate. The critical contact conditions for this to occur are high load and friction (see ratcheting region on the shakedown map shown in Fig. 2.6).

There are a number of phases present as shown in Fig. 2.18 (Kapoor *et al.*, 2003), which plots surface crack growth rates (da/dn) against crack length and reveals that after initiation crack growth is driven by ratchetting in the plastically deformed layer. As the crack becomes longer and deeper, crack growth is driven by the stress field due to the repeated contact loading. Finally the crack turns downwards and growth is driven by bending stresses in the rail. If the crack reaches a critical crack length at this stage fast fracture can occur resulting in a rail break.



2.17 Head check cracks at gauge corner. (From Olofsson and Nilsson, 2002)

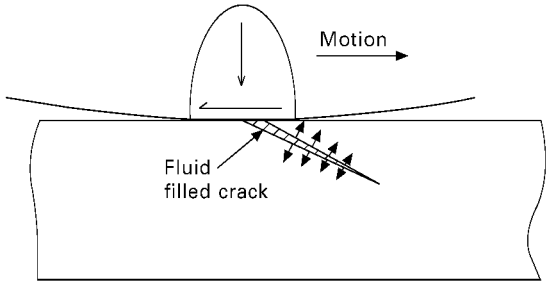


2.18 Phases of crack growth in rail. (Kapoor *et al.*, 2003)

Contamination can have a significant effect on fatigue crack propagation. Water and lubricants trapped in a crack particularly can increase the speed of propagation (Bower, 1988; Bogdanski *et al.*, 1996) (as shown in Fig. 2.19). This is because when they are trapped in the rail cracks, as wheels pass over they cause pressurisation, which increases the crack growth rate. Further information on the effects of oil, water and indentations, as well as friction modifiers, on fatigue processes is given in Chapter 14.

Squats occur on straight track on the surface of the rail head. They appear as darkened areas on the rail. They consist of two cracks, one on the direction of travel and the other in the opposite direction, which is much longer. They can initiate as a result of ratcheting and fluid pressurisation and also from *white etching layers* (WELs) which result from modification of the microstructure of the rail surface material from pearlite to martensite (Pyzalla *et al.*, 2001). They are extremely brittle which means that crack initiation is therefore more likely to occur. WELs normally occur as a result of high temperatures, which may result from wheel skids, for example.

Shelling in rails occurs at the rail gauge corner in curves and is a subsurface initiated defect (Grassie and Kalousek, 1997). Elliptical shell-like cracks



2.19 Crack opening driven by fluid pressurisation.

2.20 Wheel surface fatigue damage.

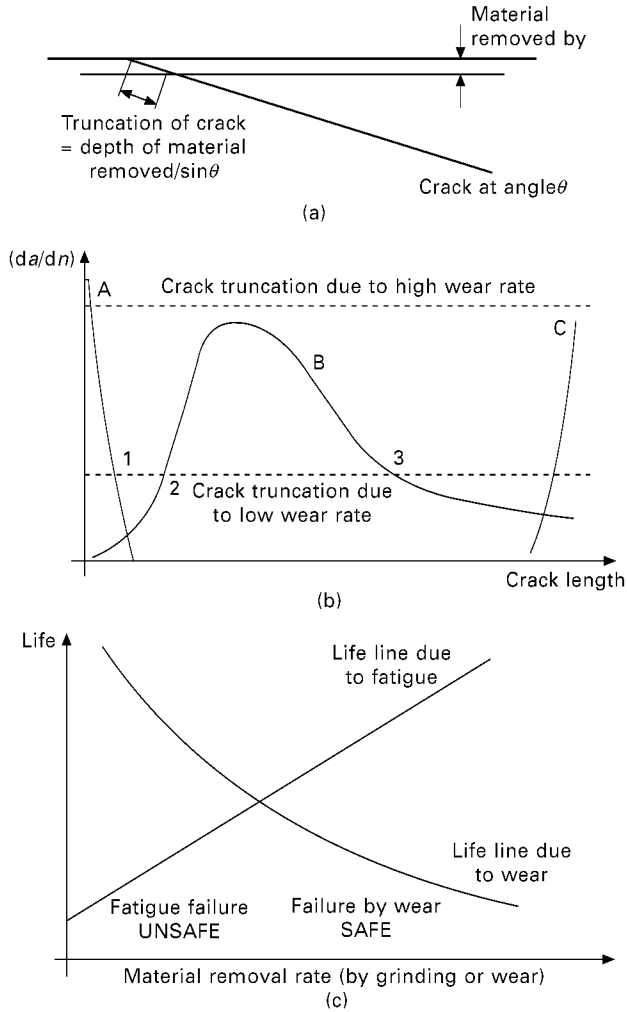
propagate parallel to the rail surface and in many cases cause material to spall away. *Tache ovale* are defects which develop about 10–15 mm below the surface of the rail head from cavities caused by the presence of hydrogen (Grassie and Kalousek, 1997).

In surface-initiated fatigue of railway wheels, fatigue cracks, as with rails, result from plastic flow of the surface material. This will cause crack initiation due to ratcheting and/or low-cycle fatigue of the surface material. Once initiated, the cracks typically grow into the wheel material to a maximum depth of 5 mm. Final fracture occurs as the cracks branch towards the wheel tread. The typical appearance of surface-initiated fatigue failure is shown in Fig. 2.20. Surface-initiated cracks are normally not a safety issue. However, they are the most common type of fatigue damage in wheels. They are costly in requiring reprofiling of the wheel and causing unplanned maintenance.

In the case of subsurface fatigue, cracks will initiate several millimetres below the surface. They continue to grow under the surface and final fracture will normally occur due to branching towards the surface. Such a failure will lead to a large piece of the wheel surface breaking away.

2.4.2 Interaction of wear and fatigue

Wear rates have a strong influence on crack propagation. As shown in Fig. 2.21a, a crack is truncated as more material is removed by wear processes.



2.21 Interaction of wear and fatigue: (a) crack truncation by wear; (b) crack growth rate va crack length; (c) rail life vs material removal rate.

If the crack truncation rate is laid over the crack growth rates the net crack growth rate can be determined. In Fig. 2.21b, two different wear rates are considered. If the wear rate is high, most cracks will probably be worn away before progressing beyond Stage A, where cracks are driven by ratcheting. If the wear rate is low, however, cracks may progress to Point 1, where growth rate equals truncation rate. As these curves will vary considerably with contact conditions, it is possible that the crack may be carried from Point 1 to Point 2 and then continue propagating. The crack reaches a length at which the growth rate declines and it stabilises at Point 3. If the wear rate drops below the intersection of Stages B and C the crack can move to Stage C, which may lead to a dangerous conclusion.

Figure 2.21c shows the effect of material removal rate on rail life. This may be reduced by using harder rail or by lubrication. Indeed Frederick (1993) discusses whether hard rail or soft rails should be used in curves and 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 work carried out in Stockholm (Olofsson and Nilsson, 2002), 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.

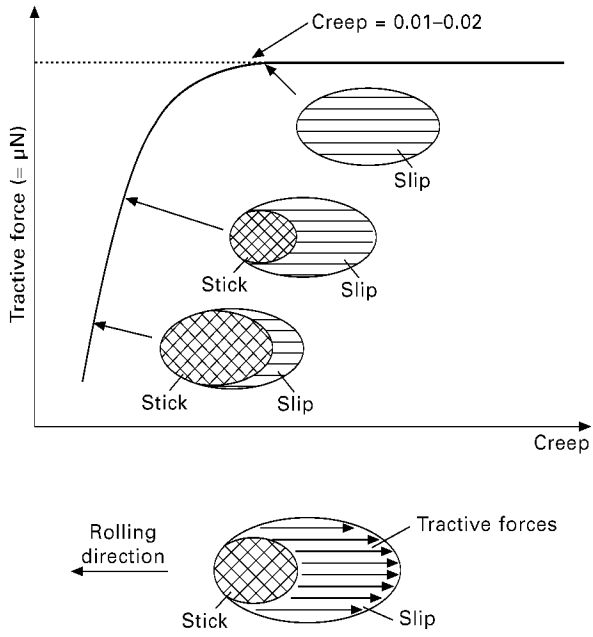
2.5 Adhesion

Maintaining good adhesion between the wheel and rail is imperative for the safe, efficient and reliable operation of a railway network. In train braking maintaining good adhesion is clearly a safety issue; in traction it is a performance issue, as wheel slippage due to low adhesion can lead to delays. Good control of the adhesion levels can also bring about efficiency savings in terms of reduced fuel consumption. Adhesion is covered in detail in Chapter 17.

Going back to basics, the friction force can be defined as the resistance encountered by one body moving over another. This definition covers both sliding and rolling bodies. Pure rolling nearly always involves some sliding, although it may be very small. For sliding bodies the friction force, and thereby the coefficient of friction (friction force divided by normal force), depends on three different mechanisms: deformation of asperities, adhesion of the sliding surfaces and ploughing caused by hard particles/asperities (Suh and Sin, 1981). For most metal pairs, the maximum value of the coefficient of friction ranges from 0.3–1.0 (Czichos, 1992). Friction due to rolling of non-lubricated surfaces over each other is considerably less than dry sliding friction of the same surfaces (Harris, 1991). For the steel wheel–steel rail contact, the rolling coefficient of friction is of the order of 1×10^{-4}.

As shown in Fig. 2.22, 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, μ (ratio of tractive force to normal load, N). Normally wheel–rail traction reaches a maximum at creep levels of 0.01–0.02.

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 friction (friction modifier or lubricant) or by a naturally occurring substance acting to decrease friction (water or leaves, etc.). A great deal of research was carried out on adhesion loss in the UK during the 1970s using both laboratory and field tests (Beagley and Pritchard, 1975; Beagley *et al.*, 1975a,b; Collins and Pritchard, 1972; Broster *et al.*, 1974). 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



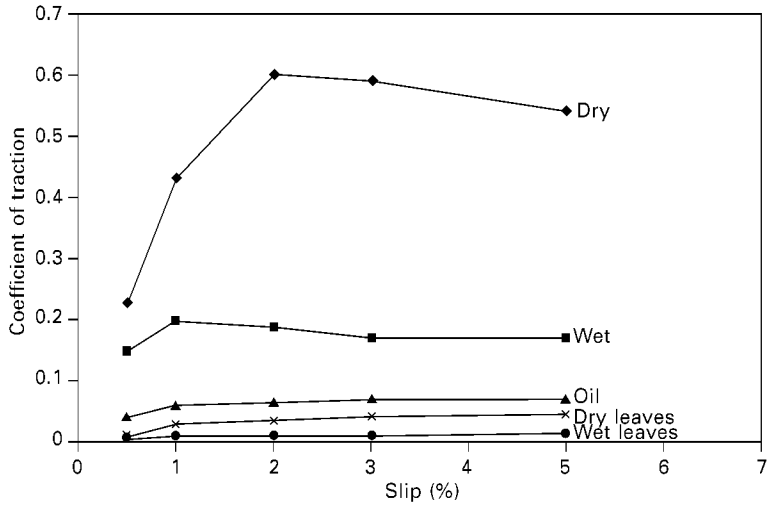
2.22 Relationship between traction and creep in the wheel – rail contact.

wide variety of materials (Demizu *et al.*, 1990). The effect of some of these on traction is shown in Fig. 2.23 (Gallardo-Hernandez and Lewis, 2008). This data was generated during twin disc testing. As can be seen, leaves make a good lubricant. More information on the effects of these contaminants on friction, wear and rolling contact fatigue is given in Chapter 14.

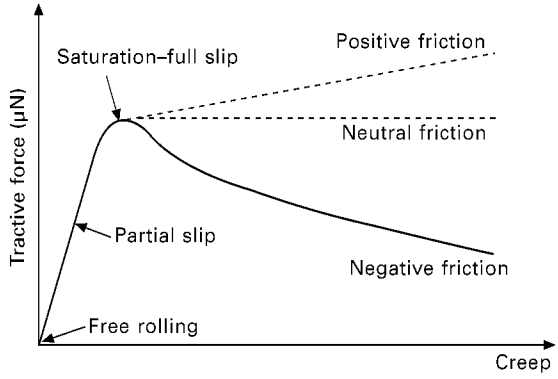
The contaminants mentioned above either occur naturally or are deposited on the track accidentally (i.e. oil). Friction modifiers are products applied to the wheel–rail contact deliberately to generate required coefficients of friction. These may act to increase or decrease friction depending on the situation. A decrease in friction may be needed where wear rates are high, for example in low-radius curves, and an increase may be desired where adhesion loss is prevalent; this would be used in traction and braking. Friction modifiers are divided into three categories (Kalousec and Magel, 1999):

1. **Low coefficient friction modifiers** (lubricants) are used to give friction coefficients less than 0.2 at the wheel flange–gauge corner interface.
2. **High friction modifiers** with intermediate friction coefficients of 0.2–0.4 are used in wheel tread–rail top applications to reduce energy requirements and also reduce noise, corrugation formation and vibrations.
3. **Very high friction modifiers** (friction enhancers such as sand or Sandite) are used to increase adhesion for both traction and braking.

Friction modifiers are classified according to their influence after full slip conditions have been reached in the wheel–rail contact, as shown in Fig. 2.24 (Eadie *et al.*, 2002). If friction increases after the saturation point the



2.23 Creep curves generated during twin disc testing with contaminants applied.



2.24 Behaviour of friction modifiers. (Eadie *et al.*, 2002)

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.

2.6 References

Beagley, T.M., Pritchard, C., 1975, Wheel/rail adhesion – the overriding influence of water, *Wear*, **35**, 299–313.

Beagley, T.M., McEwen, I.J., Pritchard, C., 1975a, Wheel/rail adhesion – the influence of railhead debris, *Wear*, **33**, 141–52.

Beagley, T.M., McEwen, I.J., Pritchard, C., 1975b, Wheel/rail adhesion – boundary lubrication by oily fluids, *Wear*, **33**, 77–88.

Bogdański, S., Olzak, M., Stupnicki, J., 1996, Influence of liquid integration on propagation of rail rolling contact fatigue cracks, in Zobory I, (ed.), *Proceedings 2nd Mini Conference on Contact Mechanics and Wear of Wheel/Rail Systems*, Budapest University of Technology, Budapest, Hungary 134–43.

Bower, A.F., 1988, The influence of crack face friction and trapped fluid on surface initiated rolling contact fatigue cracks, *Transactions of the ASME, Journal of Tribology*, **110**, 704–11.

Broster, M., Pritchard, C., Smith, D.A., 1974, Wheel-rail adhesion: it's relation to rail contamination on British railways, *Wear*, **29**, 309–21.

Collins, A.H., Pritchard, C., 1972, Recent research on adhesion, *Railway Engineering Journal*, **1**(1), 19–29.

Czichos, H., 1992, Presentation of friction and wear data, in Blau, P. J. (ed.), *Friction, Lubrication and Wear Technology*, ASM Handbook, vol. 18, ASM International, Materials Park, OH, USA, 489–92.

Demizu, K., Wadabayashi, R., Ishigaki, H., 1990, Dry friction of oxide ceramics against metals: the effect of humidity, *Tribology Transactions*, **33**, 505–10.

Eadie, D.T., Kalousec, J., Chiddick, K.C., 2002, The role of high positive friction (HPF) modifier in the control of short pitch corrugation and related phenomena, *Wear*, **253**, 185–92.

Frederick, C.O., 1993, Future rail requirements, in Kalker, J.J., Cannon, D.F. and Orringer, O. (eds), *Proceedings of Rail Quality and Maintenance for Modern Railway Operation*, Kluwer Dordrecht, the Netherlands, 3–14.

Gallardo-Hernandez, E.A., Lewis, R., 2008, Twin disc assessment of wheel/rail adhesion, *Wear*, **265**, 1309–16.

Grassie, S., Kalousek, J., 1997, Rolling contact fatigue of rails: characteristics, causes and treatments, *Proceedings 6th International Heavy Haul Conference*, Capetown, South Africa, 7–11 April, 381–404.

Harris, T.A., 1991, *Rolling Bearing Analysis*, Wiley, New York, USA.

Johnson K.L., 1985, *Contact Mechanics*, Cambridge University Press, Cambridge, UK.

Jendel, T., 2000, *Prediction of Wheel Profile Wear – Methodology and Verification*, Licentiate Thesis, TRITA-FKT 2000:9, KTH, Stockholm, Sweden.

Kalker, J.J., 1982, Fast algorithm for the simplified theory of rolling contact, *Vehicle System Dynamics*, **11**, 1–13.

Kalker J.J., 1990, *Three-Dimensional Elastic Bodies in Rolling Contact*, Kluwer, Dordrecht, the Netherlands.

Kalousek, J., Magel, E., 1997, Optimising the wheel/rail system, *Railway and Track Structure*, January.

Kalousek, J., Magel, E., 1999, Modifying and managing friction, *Railway and Track Structures*, **5**, 5–6.

Kapoor, A., Fletcher, D.I., Franklin, F.J., 2003, The role of wear in enhancing rail life, in Dowson, D., Priest, M., Dalmaz, G. and Lubrecht, A.A. (eds) (2003), Tribological research and design for engineering systems, *Proceedings 29th Leeds-Lyon Conference on Tribology, Leeds 2002*, Elsevier, Amsterdam, the Netherlands, 331–40.

Lewis, R., Dwyer-Joyce, R.S., 2004, Wear mechanisms and transitions in railway wheel steels, *IMechE Part J; Journal of Engineering Tribology*, **218**, 467–78.

Lewis, R., Olofsson, U., 2004, Mapping rail wear regimes and transitions, *Wear*, **257** (7–9), 721–9.

Lim, S.C., Ashby, M.F., 1987, Wear mechanism maps, *Acta Metallica*, **35**, 1–24.

Marshall, M.B., Lewis, R., Dwyer-Joyce, R.S., Olofsson, U., Björklund, S. 2006, Experimental characterisation of wheel-rail contact patch evolution, *ASME Journal of Tribology*, **128**(3), 493–504.

Nilsson R., 2003, *Wheel/rail wear and surface cracks*, Licentiate Thesis, TRITA-MMK 2003; 03, KTH, Stockholm, Sweden.

Olofsson, U., Nilsson, R., 2002, Surface cracks and wear of rail: a full scale test and laboratory study, *IMechE Part F: Journal of Rail and Rapid Transit*, **216**, 249–64.

Pyzalla, A., Wang, L., Wild, E., Wroblewski, T., 2001, Changes in microstructure, texture and residual stresses on the surface of a rail resulting from friction and wear, *Wear*, **251**, 901–7.

Suh, N.P., Sin H.C., 1981, The genesis of friction, *Wear*, **69**, 91–114.

Swanson P., Klann, R., 1981, Abrasive wear studies using the wet sand and dry sand rubber wheel tests, in Rhee, S.K., Ruff, A.W. and Ludema, K.C. (eds), *Wear of Materials 1981*, ASME, New York, USA, 379–89.

Telliskivi T., Olofsson U., 2001, Contact mechanics analysis of measured wheel-rail profiles using the finite element method, *IMechE Part F: Journal of Rail and Rapid Transit*, **215**, 65–72.

Tournay, H., 2001, Supporting technologies vehicle track interaction, in *Guidelines to Best Practice for Heavy Haul Railway Operations: Wheel and Rail Interface Issues*, International Heavy Haul Association, Virginia Beach, VA, USA, 2–1–2.73.

Copyrighted Materials
Copyright © 2009 Woodhead Publishing Retrieved from www.knovel.com

3

Wheel–rail contact mechanics

S. IWNICKI, Manchester Metropolitan University, UK;
S. BJÖRKLUND and R. ENBLOM, Royal Institute of Technology, (KTH), Sweden

Abstract: The scope of this chapter is the modelling of the force balance at the wheel–rail interface in railway vehicles. The focus is on simulation of vehicle dynamics and the requirements imposed on numerical efficiency. The introductory section addresses basic concepts related to profile geometry and contact area and is followed by a section on general models for normally and tangentially loaded contacts. Analytical and numerical methods are treated. The section on wheel–rail-specific solutions concentrates on numerical algorithms for dynamic application. Comparisons between elliptic and non-elliptic models are given. The chapter concludes with a section on vehicle dynamics ranging from simulation of contact conditions to a review of available simulation packages.

Key words: wheel–rail contact, railway vehicle dynamics, creep force, Hertz contact.

3.1 Introduction

When bringing two bodies into contact the applied forces will be transmitted over the contact area, which in most cases is very small compared with the size of the bodies. The main goal when analysing such contacts is to calculate the magnitude of stresses and deformations, both at the contact interface and in the interior of the bodies. The size and shape of the contact area may also be of interest.

The subject of contact mechanics is complex, and any solution to a contact problem is restricted to a special case with some fundamental simplifications. Czichos (1978) pointed out that the subject consists of a great variety of situations, depending on the following:

- the number of bodies involved in the contact;
- the geometry of the bodies (a two- or three-dimensional problem, smooth or rough surfaces);
- the material properties (constant or varying material parameters, layers and coatings);
- the deformation mode (purely elastic, elastic–plastic or purely plastic deformation);
- the contact forces (normal and tangential forces);

- the type of relative motion (static, sliding, rolling, etc.);
- the velocity of the relative motion.

This chapter's special focus is on contact mechanics in the field of railway applications. The influence of geometry provides an introduction to the topic and general contact modelling is introduced in Section 3.2, where both analytical and numerical solutions to some important contact problems are presented. Wheel–rail contacts in particular are treated in Section 3.3 and computer tools commonly used in these problems are covered in Section 3.4, while some future trends are discussed in the concluding section, Section 3.5.

3.1.1 Wheel and rail geometry

The nature of the contact patch and the forces between a railway wheel and rail are strongly influenced by the geometry of the wheel and the rail. In particular, the geometry of a vertical cross-section of the rail and a radial cross-section of the wheel are important.

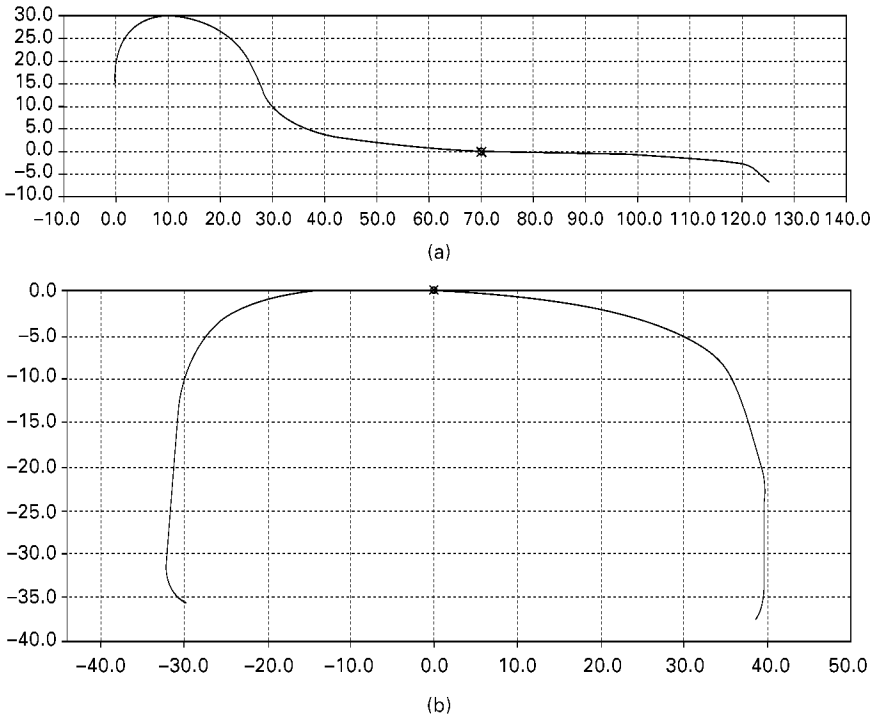
The earliest railway wheels were cylindrical and ran on flanged rails. They were usually fitted to an axle so that both wheels could rotate independently. Fitting the flanges to the wheels instead of the rails must have made a considerable saving of material and probably allowed better guidance of the vehicle although it had the disadvantage of preventing the vehicles from running on the road. Adding a small amount of conicity to the wheels would have enhanced this guidance, and the modern wheelset was formed when the two wheels were joined to the axle and fixed to the vehicle body through bearings in axleboxes.

Wear at the wheel will tend to change the wheel tread from an initial conical profile to a more complex concave shape. Many railway organisations have designed 'worn' profiles, which are intended to maintain the same geometry as the wheel wears (examples are the UK P8 and the UIC S1002 profiles). Most railway organisations incline the rails towards each other by a small angle, and this usually matches the conicity of the wheel so that the normal force with the wheelset in the central position is directed along the web of the rail. In the UK this angle is 1 in 20 but 1 in 30 (for example in Sweden) and 1 in 40 (many countries including Germany) are also common.

The starting point for an analysis of wheel–rail contact is the identification of the size and shape of the contact patch. To do this for new profiles, drawings may be available but, after running for a little while, profiles will deviate significantly from the design case and it is essential to have accurate geometrical information of the worn profiles. Precise measurements can be made using mechanical or laser devices, and Fig. 3.1 shows the 'MiniProf' measuring device which has a small roller connected to two arms with encoders from which the precise locus of the roller and thus the wheel or

(a) (b)

3.1 MiniProf measuring device: (a) rail profile; (b) wheel profile.



3.2 Measured profiles: (a) wheel; (b) rail.

rail geometry can be determined. Typical measured wheel and rail profiles are shown in Fig. 3.2.

3.1.2 The contact patch

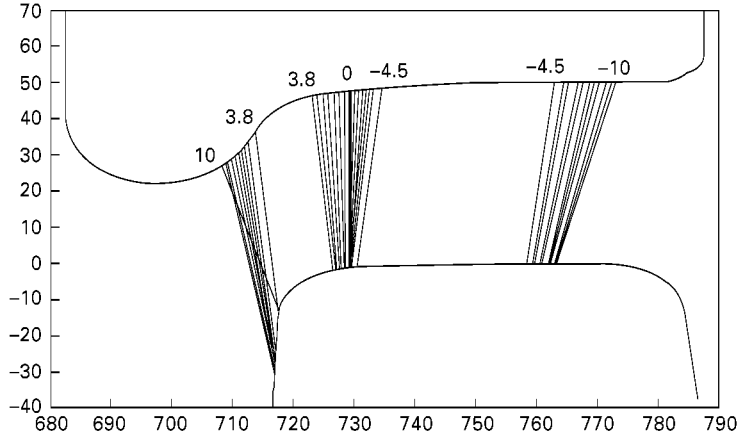
At the point or points where the wheel touches the rail a contact patch develops. The size and shape of this contact patch can be calculated from

the normal force, the local geometry of the wheel and the rail and the material properties. As the wheel and the rail are both bodies of revolution it is possible to describe this geometry by using the radii of curvature in the direction of rolling and for the cross-sectional geometry.

The classical theory of contact was developed by Hertz (1882) when he was a 24 year old research assistant at the University of Berlin. He proved that the contact area between two non-conformal bodies of revolution would be elliptical and he established a method for calculating the semi-axes 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.

The initial point of contact can be obtained considering the bodies to be rigid, or their flexibility can be allowed to influence the position of the contact. Iteration may be required as the roll angle of the wheelset both influences and is a function of the contact position. Many software packages have a contact pre-processor which is run whenever the contact details are required or can be used to set up a table of data from which the properties, such as contact patch location and size and contact angle, can be interpolated. Results of a pre-processor are shown in Fig. 3.3, the red line shows the points in contact with the wheelset in the central position and the blue lines show the contact as the wheel moves towards the flange or the field side of the rail.

Once the initial contact point location has been determined Hertz theory can be used to establish elliptical contact patches around the contact point. The normal load on the contact point is also required and the calculation of this may also be iterative to allow the correct load distribution between the contact points to be found. If the contact is on the tread of the wheel



3.3 Plot of contact position on rail and wheel.

and the head of the rail 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, for example on the flange, or the contact is very conformal the contact patch may be quite non-elliptical and the Hertz method would not produce good results. Knothe and Le-The Hung (1984) set out a numerical method for calculating the tangential stresses for non-elliptical contact.

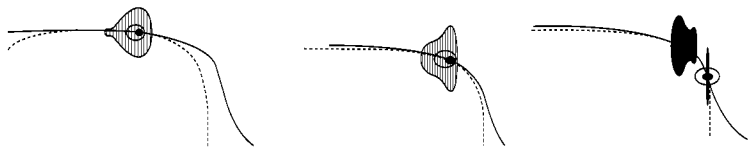
An alternative is to use Multi-Hertzian methods which split the contact patch into strips with Hertz contact being calculated for each strip. Some iteration may be required to establish the correct normal load distribution across the whole contact patch if the contact is being treated as a constraint. Pascal and Sauvage (1991) developed a method using an equivalent ellipse which first calculates the multi-Hertzian contact and then replaces this with a single ellipse which gives equivalent forces. In the methods developed by Kik and Piotrowski (1996) an approximate one-step method is used, and some results for an S1002 wheel and a UIC60 rail are shown in Fig. 3.4. In the 'semi-Hertzian' methods developed by Ayasse and Chollet (2005) the contact is treated as Hertzian for the longitudinal curvatures (along the rail) and non-Hertzian for the lateral curvatures (across the rail).

3.2 General contact modelling

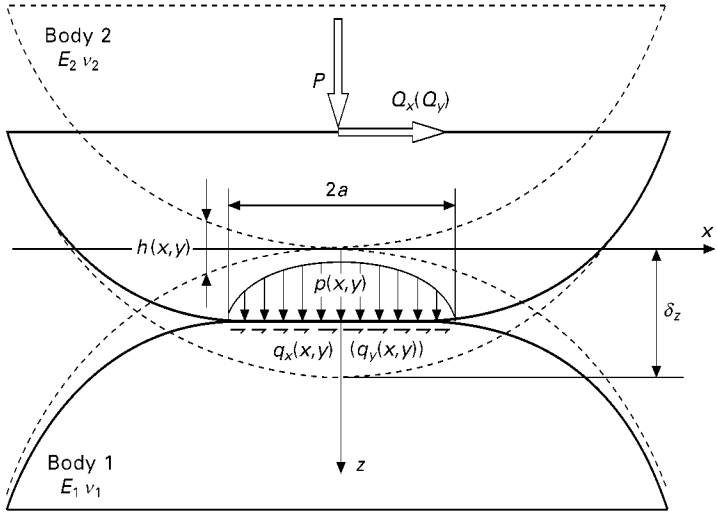
3.2.1 Introduction

The main goal of contact modelling is to determine the magnitude of stresses and deformations that appear when two bodies are brought into contact. At the point of contact there will be a small area where the globally applied forces are transmitted as a normal pressure, as well as via tangential shear stresses in two directions. The notation and treatment in this section will follow the work by Johnson (1985).

Referring to Fig. 3.5, we have a situation that can be described as follows. Initially, two bodies with known topographies and material parameters (E; Young's modulus and υ; Poisson's ratio are brought into contact without transmitting any forces between each other (dashed lines in Fig. 3.5). The surfaces touch each other at the origin of a coordinate system where the z axis



3.4 Non-elliptical contact patches for profile combination S1002/UIC60. (From Kik and Piotrowski, 1996)



3.5 Schematic contact between two bodies.

is in the same direction as the normal to Body 2, and the x–y plane is equal to the surfaces' tangent plane at the contact point. The combined topography is described as the gap, $h(x, y)$, between the undeformed surfaces.

If a normal load, P, is applied, the bodies will deform and a force equilibrium will be maintained by a pressure distribution $p(x, y)$ acting over an area whose maximum size is $2a$. The normal approach of distant points far away from the contact is denoted δ_z, which can also be regarded as the distance the surfaces would have overlapped if they had merged together without any interaction. If, in addition, the tangential loads Q_x and Q_y are applied, the bodies will deform tangentially the distances δ_x and δ_y (not shown in Fig. 3.5), and a tangential force equilibrium will be maintained by shear stresses $q_x(x, y)$ and $q_y(x, y)$ at the contact interface.

The contact problem can be formulated as follows: *Find the distributions of tractions $p(x, y)$, $q_x(x, y)$ and $q_y(x, y)$ for either known deformations (δ_x, δ_y, δ_z) or known loads (P, Q_x, Q_y), or combinations of these.* Finding the traction distributions obviously also includes determining the size and shape of the contact area. Subsequently, one might also want to calculate the stress state in the bulk material, i.e. σ_x, σ_y, σ_z, τ_{xy}, τ_{xz}, and τ_{yz} can be evaluated at any point (x, y, z). The basic assumptions in contact mechanics are as follows:

1. The materials are elastic, isotropic and homogeneous, and the strains are small.
2. The bodies deform like infinite half-spaces. This requires that the contact area is much smaller than the general dimensions of the bodies. As a consequence of this, the theories can mainly be applied to non-conforming

contacts. A contact is said to be *non-conforming* if the bodies initially meet at a point (or a line), and the surfaces do not fit into each other at zero load.

3. The surface slopes are so small that all pressures can be considered to act in the z direction defined in Fig. 3.5.

The boundary conditions for a normally loaded contact are shown in Fig. 3.6. The normal deformations at the surfaces are denoted $\bar{u}_{zk}(x, y)$, with index k for Body 1 and 2, respectively. The relation between the applied normal approach, the gap between the surfaces and the deformations becomes:

$$\bar{u}_{z1}(x, y) + \bar{u}_{z2}(x, y) + h(x, y) = \begin{cases} \delta_z & \text{in the contact area} \\ > \delta_z & \text{outside the contact area} \end{cases} \qquad [3.1]$$

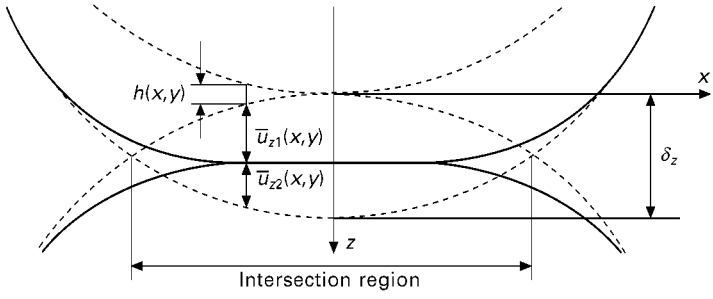
The deformations at the surfaces can be derived by starting from the deformation under a point load which was found by Cerutti (1882) and Boussinesq (1885):

$$\bar{u}_z(r) = \frac{1 - \nu^2}{\pi E} \frac{P_i}{r} \qquad [3.2]$$

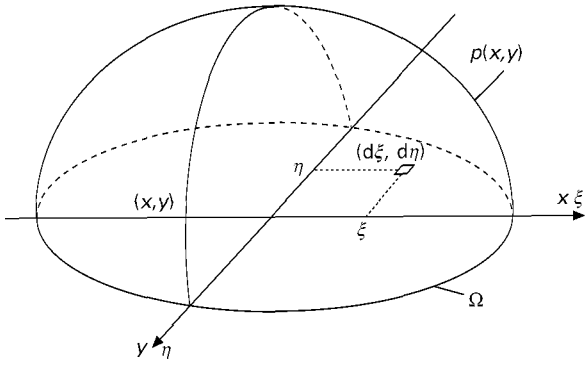
where P_i is the magnitude of the point load and r the distance from the load. The deformation at a point (x, y) due to a pressure distribution $p(\xi, \eta)$ can now be evaluated by using Eq. (3.2) and integrating over the contact region, see Fig. 3.7:

$$\bar{u}_{zk}(x, y) = \frac{1 - \nu_k^2}{\pi E_k} \iint_{\Omega} \frac{p(\xi, \eta) \cdot d\xi \cdot d\eta}{\sqrt{(x - \xi)^2 + (y - \eta)^2}} \qquad [3.3]$$

Using Eq. (3.3) together with Eq. (3.1), we get:



3.6 Boundary conditions.



3.7 Pressure distribution $p(x, y)$ acting over the contact region Ω.

$$\frac{1}{\pi E^*} \iint_{\Omega} \frac{p(\xi, \eta) \cdot d\xi \cdot d\eta}{\sqrt{(x-\xi)^2 + (y-\eta)^2}}$$

$$= \begin{cases} \delta_z - h(x, y) & \text{in the contact area} \\ > \delta_z - h(x, y) & \text{outside the contact area} \end{cases} \qquad [3.4]$$

where E^* is the equivalent modulus of elasticity defined as:

$$\frac{1}{E^*} = \frac{1 - \nu_1^2}{E_1} + \frac{1 - \nu_2^2}{E_2} \qquad [3.5]$$

A pressure distribution that satisfies Eq. (3.4) at all points (x, y) is a solution to the normal contact problem if the influence from tangential shear stresses in the contact can be disregarded. It has been shown that the effect of the coupling between pressures and tangential shear stresses is small; and indeed, it is zero if the bodies have the same material parameters.

When the pressure distribution is known, the total normal load can obviously be found by integration:

$$P = \iint_{\Omega} p(\xi, \eta)\, d\xi\, d\eta \qquad [3.6]$$

A similar procedure to the one above can be used to derive equations relating the shear stresses in the contact to applied tangential displacements.

3.2.2 Normally loaded contacts

Hertzian theory

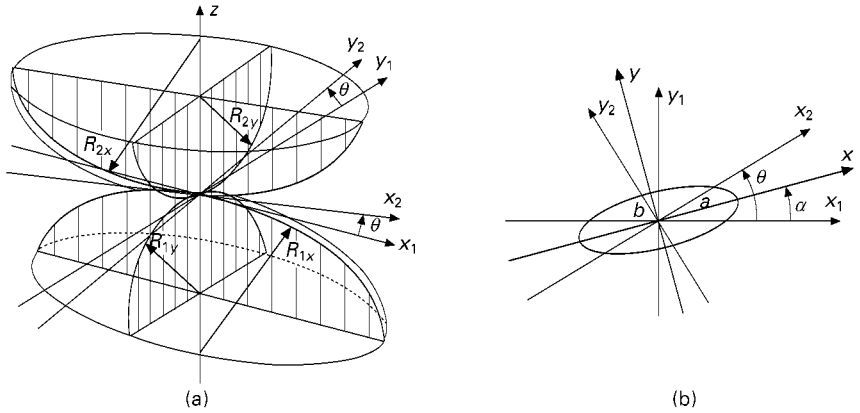
The most widely used equations in contact mechanics were derived by Hertz (1882). They apply to normally loaded and frictionless contacts between solids

with smooth surfaces which, close to the contact region, can be described by second-order polynomials. A general contact of the Hertzian type is shown in Fig. 3.8. Two bodies initially touch at a single point, which defines the origins of two rectangular co-ordinate systems with a common z axis. The geometry of each body can be described by the principal radii, which are defined as the largest and smallest radius of curvature at the contact point. The x and y directions are chosen in the directions of the principal radii for each body. The geometries are thus defined by four principal radii, R_{1x}, R_{1y}, R_{2x} and R_{2y}, and the angle θ between the x_1 and the x_2 axes (see Fig. 3.8a). In Fig. 3.8a, the bodies have convex curvatures in all directions. When the angle $\theta = 0$, the Hertzian theory also applies to cases where one or two radii are concave. The concave radii are then specified with a negative sign.

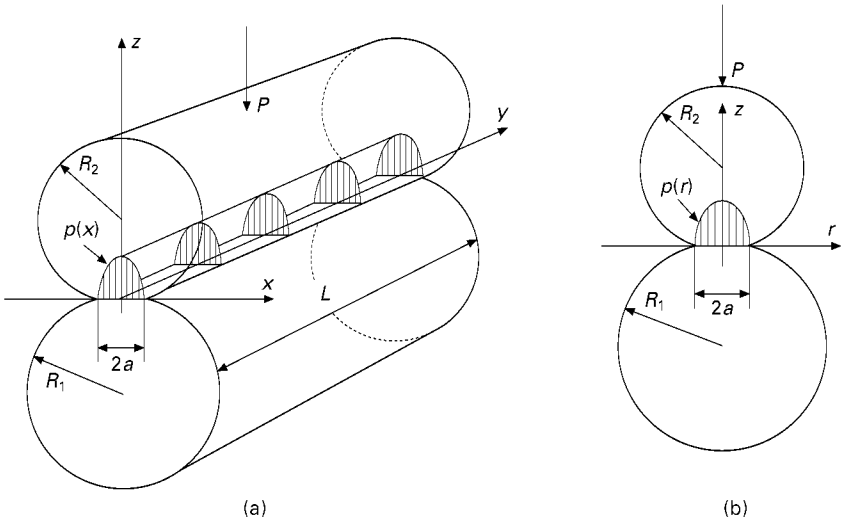
Figure 3.8 shows a general Hertzian contact, but in most cases two special cases are of interest: (i) *line contacts* (contacts between cylinders of length L with parallel axes of revolution, as presented in Fig. 3.9a), and (ii) *circular point contacts* (contacts between spheres, as illustrated in Fig. 3.9b). The Hertzian contact parameters are calculated using the equations in Table 3.1. No explicit equations are available for a general Hertzian contact. The contact shape will be an ellipse with one major axis lying between the x_1 and x_2 axes (see Fig. 3.8b). The angle α between the ellipse and the x_1 axis can be calculated from:

$$\tan 2\alpha = \frac{\sin 2\theta}{\dfrac{1/R_{1x} - 1/R_{1y}}{1/R_{2x} - 1/R_{2y}} + \cos 2\theta} \qquad [3.7]$$

which is used to calculate equivalent radii in the x and y directions:



3.8 General Hertzian contact between two solids: (a) contacting bodies; (b) elliptical contact area.



3.9 Hertzian contacts: (a) line contact; (b) circular contact.

$$\frac{1}{R_x} = \frac{\cos^2 \alpha}{R_{1x}} + \frac{\sin^2 \alpha}{R_{1y}} + \frac{\cos^2 (\theta - \alpha)}{R_{2x}} + \frac{\sin^2 (\theta - \alpha)}{R_{2y}} \qquad [3.8]$$

and

$$\frac{1}{R_y} = \frac{\sin^2 \alpha}{R_{1x}} + \frac{\cos^2 \alpha}{R_{1y}} + \frac{\sin^2 (\theta - \alpha)}{R_{2x}} + \frac{\cos^2 (\theta - \alpha)}{R_{2y}} \qquad [3.9]$$

where the x and y axes coincide with the major axes of the elliptical contact area (see Fig. 3.8b).

Equivalent radii for the contact is:

$$R = (R_x R_y)^{1/2} \qquad [3.10]$$

which is used together with the parameter ξ:

$$\xi = \begin{cases} \left(\dfrac{R_x}{R_y}\right)^{1/2} & \text{if } R_x \geq R_y \\ \left(\dfrac{R_x}{R_y}\right)^{1/2} & \text{if } R_y > R_x \end{cases} \qquad [3.11]$$

as input to the elliptical contact equations in Table 3.1 and Fig. 3.10.

The size of the elliptical contact is calculated from the parameter c (Table 3.1) and the function F_3 in Fig. 3.10:

Table 3.1 Summary of the equations by Hertz. The functions F_1 and F_2 are plotted in Fig. 3.10

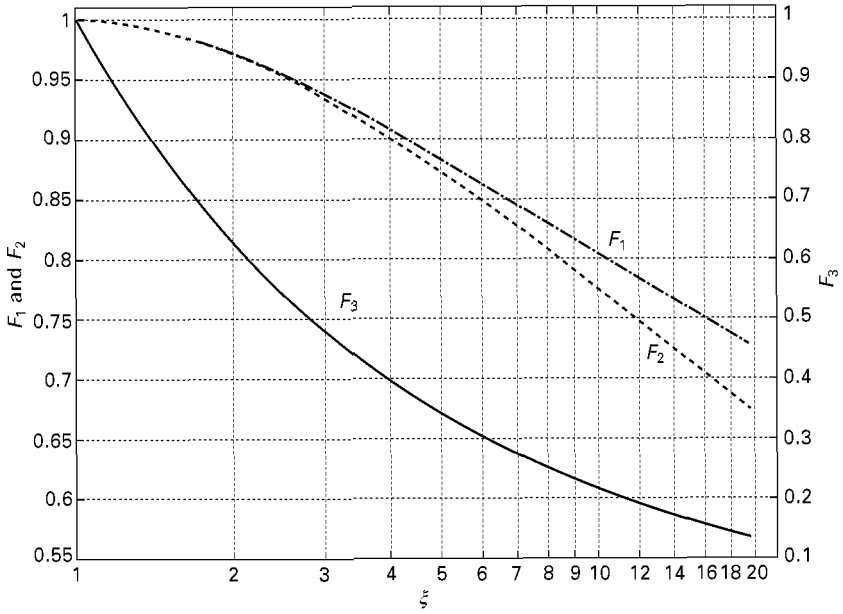
	Line contacts	Circular contacts	Elliptical contacts
Equivalent modulus of elasticity		$\frac{1}{E^*} = \frac{1-\nu_1^2}{E_1} + \frac{1-\nu_2^2}{E_2}$	
Equivalent radius	$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2}$	$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2}$	See Eqs (3.7)–(3.10)
Contact size	$a = \left(\frac{4PR}{\pi LE^*}\right)^{\frac{1}{2}}$	$a = \left(\frac{3PR}{4E^*}\right)^{\frac{1}{3}}$	$c = \sqrt{ab} = \left(\frac{3PR}{4E^*}\right)^{\frac{1}{3}} F_1(\xi)$
Maximum pressure	$p_0 = \left(\frac{PE^*}{\pi LR}\right)^{\frac{1}{2}}$	$p_0 = \left(\frac{6PE^{*2}}{\pi^3 R^2}\right)^{\frac{1}{3}}$	$p_0 = \left(\frac{6PE^{*2}}{\pi^3 R^2}\right)^{\frac{1}{3}} \frac{1}{F_1^2(\xi)}$
Deformation	Not applicable	$\delta_z = \left(\frac{9P^2}{16RE^{*2}}\right)^{\frac{1}{3}}$	$\delta_z = \left(\frac{9P^2}{16RE^{*2}}\right)^{\frac{1}{3}} F_2(\xi)$
Pressure distribution	$p(x) = p_0\sqrt{1-(x/a)^2}$	$p(r) = p_0\sqrt{1-(r/a)^2}$	$p(x, y) = p_0\sqrt{1-(x/a)^2-(y/b)^2}$
Max shear stress, $\tau_{1,\max}$	$\tau_{1,\max} = 0.30\, p_0$	$\tau_{1,\max} = 0.31\, p_0$	See Fig. 3.11
Location of, $\tau_{1,\max}$, z	$z = 0.78a$	$z = 0.48a$	See Fig. 3.11

$$a = \begin{cases} c/F_3(\xi) \text{ if } R_x \geq R_y \\ c/F_3(\xi) \text{ if } R_y > R_x \end{cases} \qquad b = \begin{cases} cF_3(\xi) \text{ if } R_x \geq R_y \\ c/F_3(\xi) \text{ if } R_y > R_x \end{cases} \qquad [3.12]$$

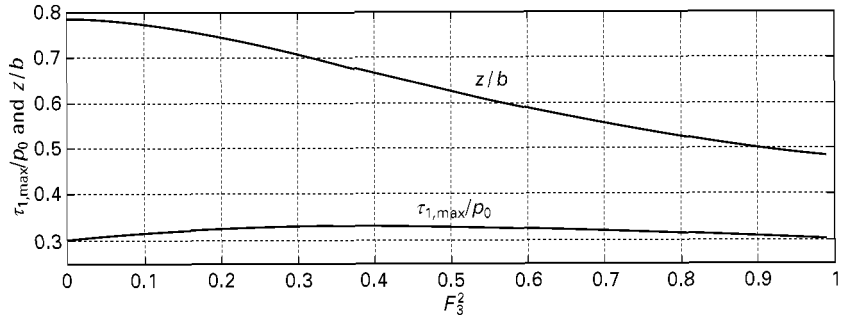
The pressure distribution at the contact interface gives rise to stresses within the two solids. The most interesting stress is the maximum shear stress, $\tau_{1,\max}$, since plastic yielding is expected to initiate when $\tau_{1,\max}$ exceeds the plastic limit. The maximum shear stress and its location can be found in Table 3.1 for line and circular contacts, and in Fig. 3.11 for elliptical contacts.

Numerical methods

Besides the Hertzian theory, there are only a few analytical solutions available in contact mechanics. A large number of numerical methods for solving general contact problems have been described. Liu *et al.* (1999) present a survey of



3.10 Correction factors for elliptical contacts.



3.11 Magnitude and location of the maximum shear stress as functions of the correction factor F_3, from Fig. 3.10.

numerical methods with a special focus on contacts between rough surfaces. Most methods work by replacing the continuous pressure distribution in Eq. (3.4) with a discrete set of pressure elements. In this section, a common numerical method for normally loaded contacts is presented.

The contact area is divided into n rectangular cells, each subjected to a uniform unknown pressure, p. A discrete version of Eq. (3.4) for a cell located at (x_i, y_i) with gap $h(x_i, y_i)$ then becomes:

$$\sum_{j=1}^{n} C_{ij} p_j = \delta_z - h(x_i, y_i) \qquad [3.13]$$

where C_{ij} is the influence coefficient that relates the deformation at cell i to a unit pressure in cell j. Equation (3.13) provides a relation between the deformation and the pressures for each cell. Writing this relationship for all cells gives n equations for the n unknown pressures. This equation system is conveniently written in matrix form as:

$$\mathbf{Cp} = (\delta_z - \mathbf{h}) \qquad [3.14]$$

where $\mathbf{C}$ is the symmetric matrix of influence coefficients, $\mathbf{p}$ and $\mathbf{h}$ are vectors respectively containing the unknown pressures and the gap before deformation. δ_z is a scalar, describing the normal approach between the surfaces. The influence coefficients for a uniform pressure on a rectangular cell with size $2a \times 2b$ were derived by Love (1929):

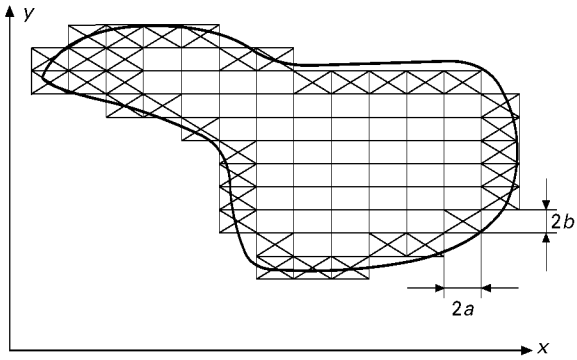
$$\begin{aligned}
\pi E^* C_{ij} &= (x + a) \ln \left[\frac{(y + b) + ((y + b)^2 + (x + a)^2)^{1/2}}{(y - b) + ((y - b)^2 + (x + a)^2)^{1/2}} \right] \\
&+ (y + b) \ln \left[\frac{(x + a) + ((y + b)^2 + (x + a)^2)^{1/2}}{(x - a) + ((y + b)^2 + (x - a)^2)^{1/2}} \right] \\
&+ (x - a) \ln \left[\frac{(y - b) + ((y - b)^2 + (x - a)^2)^{1/2}}{(y + b) + ((y + b)^2 + (x - a)^2)^{1/2}} \right] \\
&+ (y - b) \ln \left[\frac{(x - a) + ((y - b)^2 + (x - a)^2)^{1/2}}{(x + a) + ((y - b)^2 + (x + a)^2)^{1/2}} \right]
\end{aligned} \qquad [3.15]$$

where $x = |x_i - x_j|$ and $y = |y_i - y_j|$.

As the shape of the contact region is not known in advance, it is necessary to start with an assumed region shape that is large enough to enclose the true contact region. An appropriate choice is the intersection curve between the two bodies, i.e. the contour for which $h(x, y) = \delta_z$ (see Figs 3.6 and 3.12).

In order to keep contact over the whole assumed contact area, the solution to Eq. (3.14) will give negative pressures for cells outside the true contact area. These cells are excluded and the system is solved again until all pressures are positive. The remaining cells then represent the true contact region to the accuracy of the mesh size.

In contacts between rough surfaces, it is often found that the pressure in some cells is excessively large, implying that the deformation is plastic rather than elastic. An approximate method to account for this is to limit the allowable pressure by a yield pressure p_y. Thus, Eq. (3.14) is first solved according to the procedure outlined above, removing all cells having negative pressures. The resulting pressures are inspected and all pressures exceeding p_y are set to equal the yield pressure and then removed from the subsequent



3.12 Intersection region (curved solid line) divided into pressure cells. Non-contacting cells are marked with an X.

iterations. However, the cells with plastic behaviour do still contribute to the deformation at the elastic cells, which has to be taken account of in the right-hand side of Eq. (3.14). The equation system is repeatedly solved until all pressures are positive and less than or equal to p_y. Marshall *et al.* (2006) used this method to calculate the contact area and pressure distribution in wheel–rail contacts.

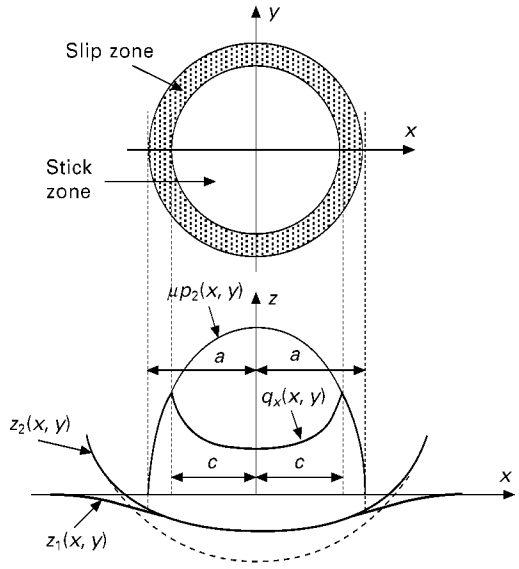
The method is simple and is only an idealisation of the true plastic behaviour in rough contacts. It is, however, commonly accepted that when the plastic deformations are small and limited to a minor part of the contact area, they have only small effects on the overall contact behaviour. Thereby even a purely elastic model can be used to calculate global contact parameters, like size and shape of the contact area, with acceptable accuracy. Several more elaborate methods to account for plasticity in contact pressure calculations have been described, see for example Johnson (1985), Persson (2001) and Busher (2002).

3.2.3 Tangentially loaded contacts

When a normally loaded contact is subjected to an increasing tangential force, the contacting bodies will eventually start to slide against each other. Before this gross sliding occurs, the sliding will start locally at some contact spots while the rest of the contact will stick. This phenomenon, called *micro-slip* or *micro-displacement*, is important in many applications involving friction, such as screw joints, couplings, shrink fits, rolling bearings and gears.

Analytical solutions for line and circular contacts

Slip will start where the pressure is low, i.e. at the edges of the contact. This means that the stick region will be located at the centre of the contact. The



3.13 Tangentially and normally loaded contact between a sphere and a plane.

stick region has a rectangular form for a line contact and a circular form for a circular contact. Consequently, the slip region surrounds the stick region and has an annular shape for circular contacts (see Fig. 3.13).

Equations for calculating the sizes and shapes of the stick and slip regions and the tangential traction distribution have been derived by Cattaneo (1938) and independently by Mindlin (1949). For a contact that is subjected to a normal load P, and subsequently to a tangential load, Q_x, the size of the stick region is given by:

$$c = a\left(1 - \frac{Q_x}{\mu P}\right)^{1/2} \qquad [3.16]$$

for a line contact, and

$$c = a\left(1 - \frac{Q_x}{\mu P}\right)^{1/3} \qquad [3.17]$$

for a circular contact. μ is the coefficient of friction and a is the contact size given by the Hertzian theory. The tangential traction distribution in the slip region is given by $\mu p(x, y)$, while in the stick region it becomes:

$$q_x(x) = \frac{\mu p_0}{a}\left[(a^2 - x^2)^{\frac{1}{2}} - (c^2 - x^2)^{\frac{1}{2}}\right] \qquad [3.18]$$

for a line contact, and

$$q_x(x, y) = \frac{\mu p_0}{a}\left[(a^2 - r^2)^{\frac{1}{2}} - (c^2 - r^2)^{\frac{1}{2}}\right] \qquad [3.19]$$

for a circular contact, where $r = (x^2 + y^2)^{1/2}$. For circular contacts, it is also possible to find a relation between applied tangential load and tangential deformation:

$$\delta_x = \frac{3\mu P}{16\, aG^*}\left(1 - \left(1 - \frac{Q_x}{\mu P}\right)^{\frac{2}{3}}\right) \qquad [3.20]$$

where G^* is a combined shear modulus:

$$\frac{1}{G^*} = \frac{2 - \nu_1}{G_1} + \frac{2 - \nu_2}{G_2} \qquad [3.21]$$

The situation in tangentially loaded elliptical contacts is very similar to the circular contacts described above. The stick region becomes an ellipse with the same eccentricity as the contact ellipse. Expressions for the elliptical case have been derived by Deresiewicz (1957).

Numerical methods for tangentially loaded contacts

Numerical methods can also be used to study tangentially loaded contacts. Such methods have been described by Kalker (1990) and Jaeger (1992), for example. In this section, a method by Björklund and Andersson (1994) is presented, which in many ways is comparable with the method for normally loaded contacts described in Section 3.3.2.

In addition to the unknown pressures and the applied normal displacement, the tangential problem also includes unknown tangential tractions in two directions, $q_x(x, y)$ and $q_y(x, y)$, and applied tangential displacements, δ_x and δ_y. The tractions are again solved by an equation system, in this case with three equations for each cell:

$$\begin{bmatrix} \mathbf{C}_{xx} & \mathbf{C}_{xy} & \mathbf{C}_{xz} \\ \mathbf{C}_{xy} & \mathbf{C}_{yy} & \mathbf{C}_{yz} \\ \mathbf{C}_{xz} & \mathbf{C}_{yz} & \mathbf{C}_{zz} \end{bmatrix} \begin{bmatrix} \mathbf{q}_x \\ \mathbf{q}_y \\ \mathbf{p}_z \end{bmatrix} = \begin{bmatrix} \delta_x - \mathbf{s}_x \\ \delta_y - \mathbf{s}_y \\ \delta_z - \mathbf{h} \end{bmatrix} \qquad [3.22]$$

There are three influence matrices for each traction direction. $\mathbf{s}_x$ and $\mathbf{s}_y$ represent the unknown slip distances for each cell. The first step in the solution of Eq. (3.22) is the same procedure as that for solving Eq. (3.14), i.e. the true

contact region and the pressures are calculated on the assumption that the induced normal displacements from the tangential tractions are negligible. When the true contact region has been found, the regions of stick and slip can be achieved by an iterative procedure, similar to that for finding the true contact regions. Equation (3.22) is solved by assuming that all cells stick ($\mathbf{s}_x = \mathbf{s}_y = 0$), i.e. for the case of an infinite friction coefficient. Cells for which the resulting tangential traction violates Coulomb's law of friction:

$$\sqrt{q_{xi}^2 + q_{yi}^2} > \mu p_i \qquad [3.23]$$

belong to a slipping region and their tangential tractions are known. Equation (3.22) can now be reduced and rewritten in consideration of the known tangential tractions and solved again. This procedure is repeated until the solution contains only the sticking cells. When all tractions are known, the sliding distances can be solved from the original Eq. (3.22).

3.3 Wheel–rail contact analysis

3.3.1 Introduction

In general terms, wheel–rail contact mechanics is a wide field of study that ranges from overall vehicle dynamics and assessment of global contact forces, to micro-mechanics of the surface materials in contact. This section, however, confines itself to a few areas of importance for the rapid calculation of contact stress distributions.

In railway applications, the wheel–rail contact is of the rolling-sliding kind, and the stress problem involves both normal and tangential loads. Generally, the corresponding normal and shear stress distributions are not independent. However, it is possible to explore some fundamental properties of the wheel–rail contact, allowing decoupling of the solutions for the normal and tangential stress distributions. Traditionally, Hertz' (1882) theory of elliptic contact has been applied to the normal problem, implying the following assumptions:

1. ideally smooth and frictionless surfaces (a reasonable simplification for the normal problem in combination with assumption 2);
2. identical material stiffness properties of both contacting bodies (necessary for decoupling of the tangential problem);
3. linear elastic material (applies to steel structures up to the yield limit – under certain loading conditions not valid in wheel–rail contact, but is usually a reasonable simplification, justified by the comparably small volume of plastic material);
4. constant curvature of the bodies close to the contact area (usually a reasonable assumption, but may be violated close to profile radius transitions);

5. the extension of the contact area being small in comparison to body dimensions and radii (necessary for the underlying half-space assumption. – may be violated at the gauge corner contact).

Nevertheless, the Hertzian approach to the normal contact dominates in vehicle dynamics simulations due to its simple closed-form evaluation, which is important for the numerical efficiency in non-linear transient dynamics.

A more general basis for contact stress calculation is provided by influence functions on an elastic half-space, formulated with integral equations by Cerruti (1882) and Boussinesq (1885). The application of this method does not assume elliptic contact or elastic identity between the contacting bodies. Numerical integration is, however, required (see Section 3.3.2).

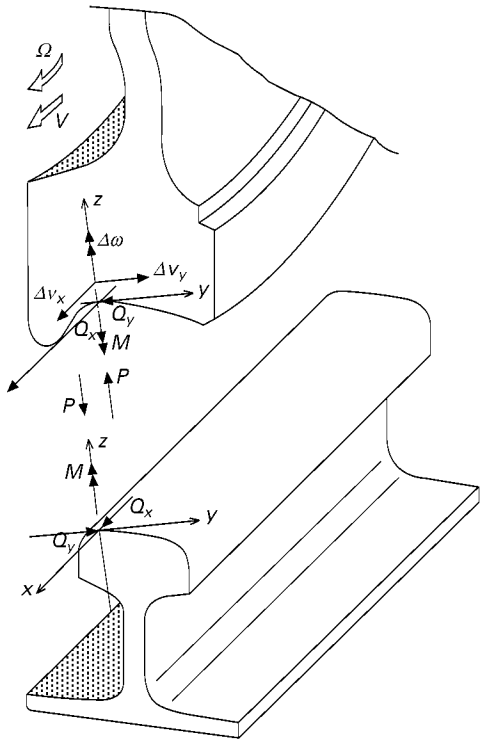
A simplified but convenient method is provided by the concept of elastic foundation, as proposed by Winkler (1867). Here, the two-dimensional contact area is discretised by a grid of mutually independent springs. However, the spring stiffness needs to be calibrated by some suitable theory.

During recent years much attention has been paid to the improvement of the solution of the general wheel–rail contact problem. Modern numerical methods, like finite element or boundary element, offer the possibility of detailed modelling. Thus, it is possible to describe complicated geometric shapes, even down to the scale of surface roughness, as well as non-linear material properties and different contact and friction conditions. In this case, it does not make sense to separate the normal and tangential contact. However, such models tend to become demanding in terms of computer power, and are thus not suitable for use online in transient vehicle simulations. This has given rise to a second trend searching for simplified numerically efficient solutions still retaining sufficient accuracy.

Fundamental wheel–rail contact quantities

The contact quantities are expressed in a contact patch related co-ordinate frame (x, y, z), where x denotes the longitudinal (rolling) direction, z the normal direction positive rail-outwards, and y the lateral direction (Fig. 3.14). The notation used in the figure is as follows:

R = wheel rolling radius
V = wheel travelling speed
Ω = wheel rotational speed
P = normal force
Q_x = longitudinal creep force
Q_y = lateral creep force
M = spin moment
Δv_i = relative in-plane translation velocity between contacting bodies, $i \in \{x, y\}$
$\Delta\omega$ = relative rotation velocity between contacting bodies



3.14 Wheel–rail contact quantities.

Futhermore:

a = longitudinal contact ellipse semi-axis
b = lateral contact ellipse semi-axis
t = time
υ_x = longitudinal creep ratio
υ_y = lateral creep ratio
φ = spin ratio
u_x = longitudinal displacement
u_y = lateral displacement
s_x = longitudinal normalised slip
s_y = lateral normalised slip
p = contact pressure
q_x = tangential stress component in longitudinal direction
q_y = tangential stress component in lateral direction
μ = coefficient of friction
α = traction coefficient

The creep and spin ratios are defined as the relative velocities normalised with respect to a reference speed, usually taken to be the travelling speed or, alternatively, the average of travelling speed and nominal rolling speed:

$$\upsilon_i = \frac{2\Delta v_i}{V + \Omega R}, \quad i \in (x, y)$$

$$\varphi = \frac{2\Delta\omega}{V + \Omega R} \qquad [3.24]$$

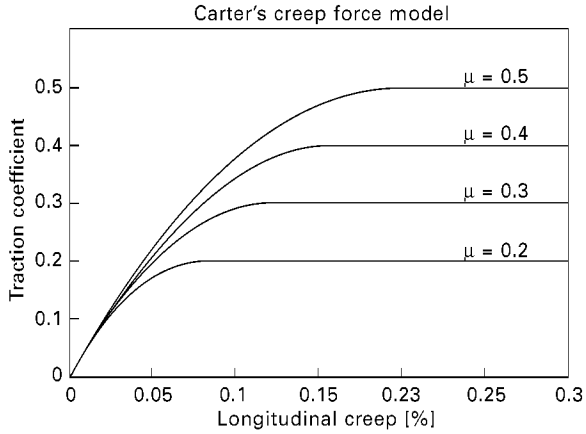
The traction coefficient is defined as the ratio of tangential to normal forces:

$$\alpha = \frac{\sqrt{Q_x^2 + Q_y^2}}{P} \qquad [3.25]$$

Creep forces

Creep forces act in the plane of the contact patch and are related to the friction between wheel and rail. For a creep force to develop, a certain amount of slip (creep) is required. In general the tangential forces thus depend on the normal load, friction conditions and relative motion between the contacting surfaces.

An early contribution to the theoretical development was Carter's (1926) creep force law (Eq. 3.26). Carter paid attention to the behaviour of a locomotive wheel transmitting large traction or braking forces. To capture the longitudinal action only, it was sufficient to approximate the system by a cylinder rolling on an infinite half-space. Carter showed that the tangential force is bounded by zero at zero slip and the maximum force given by Coulomb's friction law (Fig. 3.15):



3.15 Traction coefficient according to Carter for some levels of friction.

$$\frac{Q_x}{\mu P} = \begin{cases} -k\upsilon_x + \frac{1}{4}k^2\upsilon_x|\upsilon_x| & \text{if } k|\upsilon_x| \le 2 \\ -\operatorname{sgn}(\upsilon_x) & \text{if } k|\upsilon_x| > 2 \end{cases} \qquad [3.26]$$

$$k = \frac{4R}{\mu a}$$

where k is called *Carter's creep coefficient.*

In general, the normalised slip speed may be expressed in terms of rigid body creep and surface strain:

$$\left.\begin{aligned} s_x(x, y, t) &= \upsilon_x(t) - \varphi(t) \cdot y - \frac{\partial u_x(x, y, t)}{\partial x} + \frac{1}{V} \cdot \frac{\partial u_x(x, y, \mathrm{t})}{\partial t} \\ s_y(x, y, t) &= \upsilon_x(t) + \varphi(t) \cdot x - \frac{\partial u_y(x, y, t)}{\partial x} + \frac{1}{V} \cdot \frac{\partial u_y(x, y, \mathrm{t})}{\partial t} \end{aligned}\right\} \qquad [3.27]$$

Furthermore, Coulomb's friction law is applied:

$$\left.\begin{aligned} &|\mathbf{q}|(x, y, t) \le \mu \cdot p(x, y, t) \\ &\mathbf{q} = [q_x(x, y, t), q_y(x, y, t) \end{aligned}\right\} \qquad [3.28]$$

If the typical wavelength of the motion quantities is sufficiently large in relation to the size of the contact patch, the condition may be considered as steady-state and the time-dependent term disregarded. The calculation methods used in practice and described in the following all assume steady-state conditions. Should this condition not be fulfilled, the more complicated non-steady-state contact mechanics apply (Kalker, 1971). This topic is, however, outside the scope of this chapter.

3.3.2 Comparing numerical approaches

In this section, an overview of different approximate solutions to the wheel–rail contact problem is given. Firstly, the contributions by Kalker are addressed, since these methods have been dominating for decades. Secondly, a list of other approaches is given; and thirdly, some numerical comparisons between selected methods are presented.

Kalker

For small values of creep, the quadratic term in Carter's formula Eq. (3.26) may be disregarded. In the 1960s, Kalker (1967) developed a linear theory

based on this approach. Both longitudinal and lateral creep rates were considered, and the contact patch was assumed to be elliptic according to Hertz' theory:

$$\begin{Bmatrix} F_x \\ F_y \\ M \end{Bmatrix} = Gc^2 \begin{bmatrix} C_{11} & 0 & 0 \\ 0 & C_{22} & cC_{23} \\ 0 & cC_{23} & cC_{33} \end{bmatrix} \begin{Bmatrix} \upsilon_x \\ \upsilon_y \\ \varphi \end{Bmatrix} \qquad [3.29]$$

$$c^2 = ab$$

where C_{ij} are called *Kalker's creep coefficients*, $i, j \in \{1, 2, 3\}$.

Kalker tabulated the creep coefficients C_{ij} being functions of Poisson's ratio and the ratio of the contact ellipse semi-axes. Kalker's linear theory is extensively used in practice in rail vehicle dynamics applications, both directly in linear simulations, where the creep is sufficiently small, and as a basis for non-linear extensions.

Further developments resulted in a generalised theory for two arbitrary bodies in contact. The bodies are still described as elastic half-spaces, but the contact geometry is not restricted to the elliptic case. Kalker also provided a numerical implementation using boundary element discretisisation (Kalker, 1990). The contacting surfaces are, *a priori*, partitioned in rectangular elements in which the traction is constant. The displacement and stress fields are calculated by superposition of the contribution from each element, using the influence function approach. This code, known as CONTACT, is often used as a reference, but it requires too much computer time for online use in vehicle dynamics simulations.

The third important contribution to this field is a simplified theory based on the concept of a thin elastic layer, rigidly supported. In this approach, the equations are decoupled by the application of a Winkler bed. The corresponding calculation algorithm is referred to as FASTSIM (Kalker, 1990), coined from its capacity of being suitable for fast computer simulation.

In addition to providing creep forces, the FASTSIM algorithm has become a standard tool for calculating shear stresses. The simplified theory assumes a constant stress gradient until the traction bound is reached, whereafter the resulting shear stress assumes the value of that limit. The traction bound is defined as the contact pressure multiplied by the coefficient of friction. Due to the two-dimensional calculation grid, the influence of the spin creepage may easily be taken into consideration. The tangential stress distribution is calculated by numerical integration over the contact patch area. In the longitudinal (x) integration loop the tangential stresses are determined as follows:

$$\left.\begin{aligned} q_x^{(k)}(x,y) &= q_x^{(k-1)}(x,y) - \left(\frac{v_x}{L_x} - \frac{\varphi}{L_z}y\right)\mathrm{d}x \\ q_y^{(k)}(x,y) &= q_y^{(k-1)}(x,y) - \left(\frac{v_y}{L_y} + \frac{\varphi}{L_z}x\right)\mathrm{d}x \\ \boldsymbol{q}^{(k)}(x,y) &= [q_x^{(k)}(x,y), q_y^{(k)}(x,y)] \end{aligned}\right\} \quad [3.30]$$

where L_j are the tangential flexibility parameters, $j \in \{z, y, z\}$ and k is the numerical integration step index. The tangential flexibility parameters are defined with respect to Kalker's linear theory. The traction bound $f(x, y)$ is calculated from the contact pressure distribution and the coefficient of friction:

$$f(x,y) = \frac{2\mu P}{\pi ab}\left(1 - \left(\frac{x}{a}\right)^2 - \left(\frac{y}{b}\right)^2\right) \quad [3.31]$$

If the local stress magnitude exceeds the traction bound, the actual point is in the slip zone and the stress is reduced accordingly:

$$\left.\begin{aligned} &\text{If } r^{(k)}(x,y) = \frac{|\boldsymbol{q}^{(k)}(x,y)|}{f(x,y)} > 1 \\ &\text{then} \\ &\boldsymbol{q}_s^{(k)}(x,y) = \frac{\boldsymbol{q}^{(k)}(x,y)}{r^{(k)}(x,y)} \end{aligned}\right\} \quad [3.32]$$

where $\boldsymbol{q}_s$ is the tangential stress in the slip zone.

Overview of further approximate methods

In an early contribution, Knothe and Le-The (1984) use an iterative method to determine the contact area by means of the virtual approach of an equivalent body. The contact patch is divided into strips. To reduce the number of unknowns while still satisfying the contact condition, an elliptic stress distribution along each strip is assumed, and the advantage of neighbouring strip lengths not being independent is recognised.

Kik and Piotrowski (1996) propose a non-iterative approach to estimate the contact patch, defined by the interpenetration area at a body approach, taken as the elastic deformation calculated using Boussinesq's influence function. The contact stress distribution is assumed to be elliptic in the direction of rolling. A similar method is used by Linder and Brauchli (1996),

the difference being the calculation of interpenetration depth – here taken as a given fraction of the Hertzian body approach. Both methods use an adapted FASTSIM implementation for calculating tangential stresses.

A somewhat different application of the interpenetration approach is presented by Ayasse and Chollet (2005) and Quost *et al.* (2006). Here the contact patch is again divided into strips and a virtual ellipse is assigned to each strip. The contact patch ellipticity as well as body approach is compensated due to the differing geometry relations compared with those defined by the Hertzian theory. In addition, the orientation of the strips takes the local contact angle into consideration. The resulting contact patch is not flat, therefore, but follows the lateral curvature of the profile combination. The contact stress distribution is assumed to be parabolic, with reference to Kalker's simplified theory. For the tangential problem, an adapted version of FASTSIM is used.

Pascal and Sauvage (1993) investigate the possibility of describing two-point contact with either two separate ellipses or one equivalent. Both have been compared with Kalker's results. The double ellipse results agree reasonably well with CONTACT, while the single ellipse results deviate significantly.

Harder (1999) presents a normal contact solution for varying railhead radius. An equivalent ellipse is defined from a bicurvature geometry. A similar approach is reported by Alonso and Giménez (2005). Here, the non-constant curvature problem is solved by using a Winkler foundation, non-linearly calibrated with respect to Hertz by using the parameters *ellipticity* and *contact area*.

Comparison of three principally different approaches

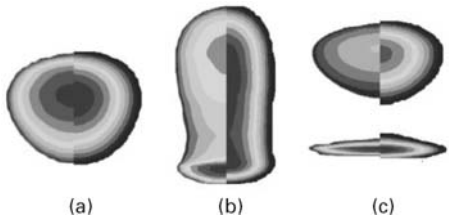
To highlight the impact of the underlying assumptions, contact area and contact pressure distributions are compared for three principally different methods (Enblom and Berg, 2008):

1. The CONTACT code devised by Kalker relies on a rigorous theory valid for an elastic half-space, and is widely accepted as a reference.
2. The approximate method proposed by Ayasse and Chollet relies on the interpenetration approach, assigning a virtual ellipse to each strip. This method, labelled STRIPES, has been used here to indicate the influence of the contact patch curvature.
3. Reference to an elliptic approach is given by the multibody simulation code GENSYS (www.gensys.se). Here, the Hertzian theory is applied with an equivalent contact ellipse, taking into account a reasonable elastic deformation of the contacting bodies.

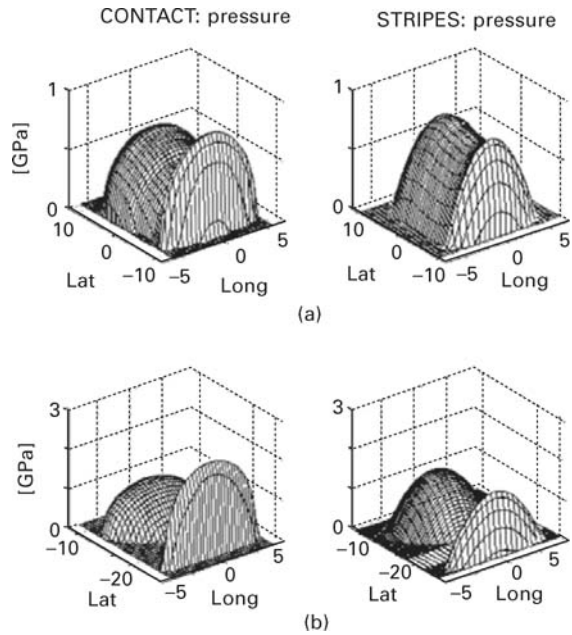
All methods assume elastic material behaviour.

The first example shows an S1002 wheel profile resting on a UIC60

rail with a 1:40 inclination. A comparison of the contact patch shapes as calculated by CONTACT and STRIPES is given below for different lateral wheel positions (Fig. 3.16). With respect to contact patch shape, there is a high degree of agreement between STRIPES and CONTACT, whereas the normal stress distributions differ to some extent. Sample distributions are given in Fig. 3.17 for the neutral and gauge corner cases. Reasonably good agreement between STRIPES and CONTACT is obtained. The parabolic stress distribution assumed by STRIPES shows a higher maximum stress at



3.16 Contact area comparison with wheelset lateral positioning. Left: CONTACT; Right: STRIPES. S1002 wheel on UIC60 rail inclined at 1:40: (a) 3 mm offset towards the field side; (b) zero lateral offset; (c) 5 mm offset towards the gauge corner.



3.17 Contact pressure comparison: (a) zero lateral offset; (b) 5 mm offset towards gauge corner. S1002 wheel on UIC60 rail inclined at 1:40.

the tread contact. Despite this, the relative rank notably switches towards the gauge corner, probably because of the increasing curvature not taken into consideration in the half-space approximation used by CONTACT.

As a second example, a similar comparison was performed on four simulated cases at different curve radii (Table 3.2, Fig. 3.18). The contact conditions were taken from simulations with a commuter vehicle. Results for the leading outer wheel during quasistatic curve negotiation are shown. The vehicle has a comparably soft wheelset guidance.

It should be noted that, for the R600 case, the contact position is close to the railhead radius transition from 80 mm to 13 mm. The large area predicted by STRIPES for the R300 case includes a low stress connection between the two major contact regions, indicating almost conformal contact.

For the tangent case, there is a high degree of agreement among the contact area sizes for all three cases. However, due to the parabolic distribution, STRIPES predicts higher maximum stress for the same normal force. For the R600 and R400 cases, the Hertzian approach and STRIPES both indicate

Table 3.2 Curving cases – profile pairing: wheel S1002 on rail UIC60 inclined at 1:30

Case	Radius [m]	Load [kN]	Speed [km/h]	Displ. [mm][a]
Tangent	–	78	120	0
R600	600	99	120	5.9
R400	400	98	93	6.3
R300	300	101	80	6.5

(a) Basic data

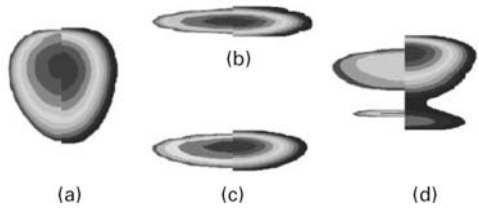
Case	Contact area [mm^2]		
	Contact	Stripes	Elliptic
Tangent	130	125	135
R600	51	70	74
R400	59	84	72
R300	77	155	77[b]

(b) Contact area

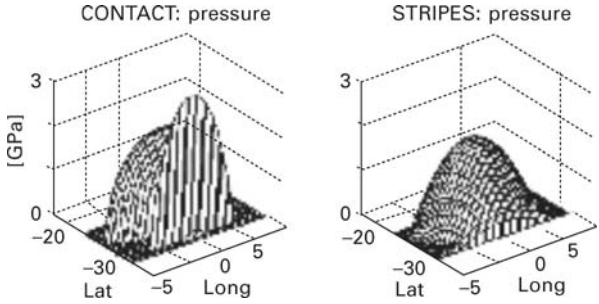
Case	Contact pressure [MPa]		
	Contact	Stripes	Elliptic
Tangent	929	1229	867
R600	3242	2995	1994
R400	2694	2417	2038
R300	3114	1782	2132

(c) Maximum contact pressure

[a] Lateral displacement towards gauge corner.
[b] Two-point contact.



3.18 Contact area comparison during curving. Left: CONTACT; Right: STRIPES. S1002 wheel on UIC60 rail inclined at 1:30: (a) tangent, 120 km/h; (b) R = 600 m, wheelset displacement 5.9 mm; (c) R = 400 m, wheelset displacement 6.3 mm. (d) R = 300 m, wheelset displacement 6.5 mm.



3.19 Contact pressure comparison in a 300 m curve. S1002 wheel on UIC60 rail inclined at 1:30.

larger contact areas than CONTACT does. The non-constant rail curvature in this region, together with the different influence on the geometry by the elastic deformation, may have an influence. Again, STRIPES shows higher maximum stresses than the elliptic case. For the final R300 case, the comparison turns out differently, probably due to the smaller lateral radii and the reduced applicability of the half-space assumption.

The contact area shapes are in strong agreement as long as the curvatures are shallow enough. For the R300 case, the influence of the curvature becomes evident – as is also illustrated by the stress distribution (Fig. 3.19).

3.4 Computer simulation tools for railway vehicle dynamics

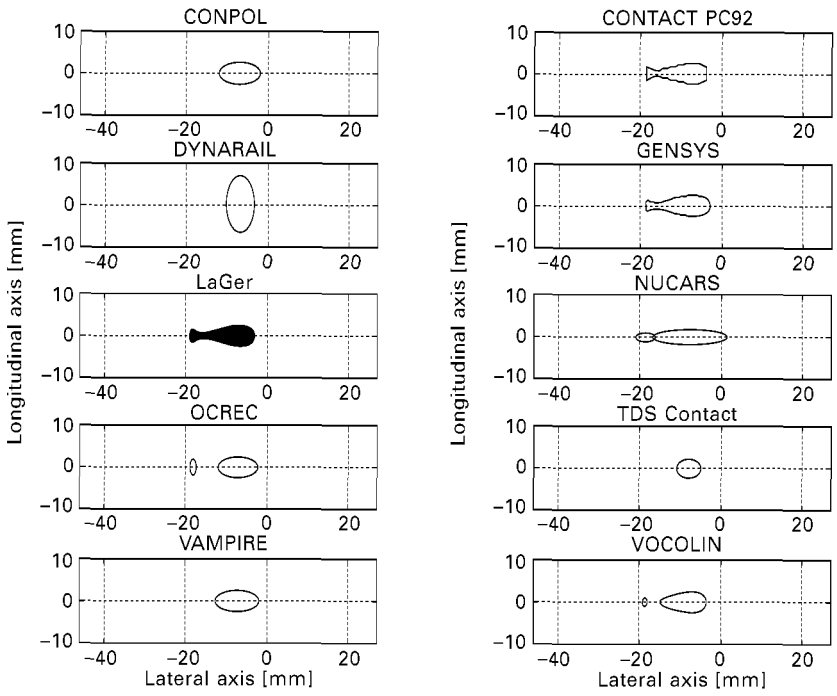
3.4.1 Simulation of contact conditions and creep forces

A number of computer algorithms for the calculation of creep forces have been produced, most notably by Kalker (1990). The program CONTACT is based on Kalker's 'complete theory' which includes non-Hertzian contact but which is relatively slow and not practical for use at every time step in a

numerical integration. Interpolation routines are available such as USETAB, which interpolate between values of creep force pre-calculated by CONTACT. FASTSIM is based on Kalker's 'simplified theory', which assumes an elliptical contact patch with a flexible layer between the two rigid bodies.

When surface roughness effects are significant the creepage-creepforce relationship is affected and Bucher *et al.* (2002) have proposed methods for dealing with this situation. Knothe and Gross-Thebing (1986), 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 (1996).

A comparison of the different software tools used to calculate wheel–rail contact has been carried out by Shackleton and Iwnicki (2008). This shows that all the codes give a similar prediction for the location of the contact patch but that there are some significant differences in the predicted contact patch size and shape and in the creep forces. Figure 3.20 shows the predicted contact patch from 10 simulation tools for new wheel and rail profiles with the wheelset in the central position on the track and a 20 kN wheel load.



3.20 The contact patch size and shape predicted by different contact simulation tools. (From Shackleton and Iwnicki, 2008.) Wheel profile: New S1002; rail profile; New UIC 60.

3.4.2 The effect on vehicle dynamics

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 de Possel *et al.* Matsudaira (1960). All of the prize winners had used a linear analysis of the problem, but de Pater (1961) formulated the hunting behaviour as a non-linear problem. Van Bommel (1964) later published non-linear 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 non-linearities to be included more easily. Wickens (1965) 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 UK.

Müller (1966) carried out one of the first simulations with analogue computers of a bogie vehicle running into a curve. Müller also recognized the importance of the inclusion of non-linear wheel profiles and included tabulated geometric data which was measured for a combination of worn wheels and rails. This was taken further by Cooperider and Law (1975) who produced an algorithm for combining measured wheel and rail profiles to produce the non-linear 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. Time stepping integration could be carried out if the systems were non-linear, 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.

Multi-body 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 multi-body simulation packages and the methods that they used was carried out by Schielen (1990).

True (1994) applied the theory of non-linear dynamics to the behaviour of a railway vehicle and showed that failure to consider the non-linearity of the wheel–rail contact can lead to an inaccurate estimate of the critical speed of the vehicle. Schupp (2004) described a method using numerical bifurcation analysis to simulate the non-linear behaviour of railway vehicles. These methods have resulted in the software PATH which has been used together with SIMPACK

3.4.3 Current computer packages

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 originally written by research institutes have been developed into powerful, validated and user-friendly packages. Examples are: ADAMS/Rail, GENSYS, NUCARS, SIMPACK and VAMPIRE. A comparison of these five packages is presented in the Manchester Benchmarks by Iwnicki (1999).

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.

Inputs to the model are usually made at each wheelset. Typical inputs are cross level, gauge and vertical and lateral alignment of the track. 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 recording vehicle. In the UK the high-speed track recording coach (HSTRC) runs over the whole network collecting track data at regular intervals. Additional forces may be specified such as wind loading or powered actuators (e.g. in tilting mechanisms). Depending on the purpose of the simulation a wide range of outputs, for example displacements, accelerations, forces at any point, can be extracted.

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.

One of the early complete packages, MEDYNA (Mehrkörper-Dynamik) (Wallrapp and Führer, 1990) was developed at the German Aerospace Research organisation DLR together with MAN and the Technical University of Berlin. MEDYNA was based on a multi-body system 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 as well as rail vehicles, it allowed non-linear 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 are developed including second-order terms to allow stress stiffening effects to be taken into account.

MSC ADAMS is amongst the most widely used multi-body dynamics software and was originally developed by Mechanical Dynamics Inc. (MDI) for the automotive industry in Michigan, USA. It was first commercially released in 1980 in the USA, and in 1995 the railway module, known as ADAMS/Rail, was developed through a partnership with NS Materieel of Netherlands Railways (now NedTrain) and later by licensing the wheel–rail contact elements from MEDYNA. This package is currently marketed by VI-Grade as VI-Rail.

In the USA, 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 (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 multi-body 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 in 1991 allowed simulation of the wheel–rail contact with a multi-Hertzian approach.

In the UK, 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 now been brought together into one coherent package VAMPIRE (Evans, 1999) which is now supported by Delta Rail.

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 non-linear, 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 (Andersson, 1977). 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 post-programs, became in 1993 the railway vehicle analysis tool called GENSYS.

3.5 Future trends

A major trend in railway engineering is the strong momentum towards improved performance in terms of higher speed and higher load. Since the only load-carrying interface to the ground is still the wheel–rail contact, the stresses on the contact patch are increasing in several respects. Despite significant advances in material technology, a thorough analysis of the consequences of the interface forces is required.

To be able to adequately predict different kinds of possible wheel and rail damage mechanisms, improved contact models are also required. Such mechanisms may be different kinds of rolling contact fatigue, thermal damage, material collapse due to local overloads and material loss with geometry alteration due to wear. All these phenomena are sensitive to local load variations; consequently, their prediction is reliant on the accuracy of available models. Thus, the requirements for contact modelling will be the ability to resolve the stress state to a sufficient degree, including spatial distribution, high-frequency loads and rapid thermal transients. These demands are considerably more severe than in traditional vehicle dynamics simulations, where the main purpose is to calculate low-frequency contact forces. The feedback needed is criteria relevant to the optimum design of vehicle and track with respect to wheel–rail interface performance.

3.6 Sources of further information and advice

- Johnson K L (1987), *Contact mechanics*, Cambridge, UK, Cambridge University Press. A comprehensive textbook on general contact mechanics covering the topics: Motion and forces at a point of contact; Line loading of an elastic half-space; Point loading of an elastic half-space; Normal contact of elastic solids – Hertz theory; Non-Hertzian normal contact of elastic bodies; Normal contact of inelastic solids; Tangential loading and sliding contact; Rolling contact of elastic bodies; Rolling contact of inelastic bodies; Calendering and lubrication; Dynamic effects and impact; Thermoelastic contact; Rough surfaces.
- Jacobson B and Kalker J J (eds) (2000), *Rolling Contact Phenomena*, Springer, Vienna, Austria, New York, USA. A set of in-depth lecture

notes covering the following topics: Rolling contact phenomena – linear elasticity; Finite element methods for rolling contact; Plastic deformation in rolling contact; Non-steady-state rolling contact and corrugations; Modelling of tyre force and moment generation; Rolling noise; Lubrication.

- Knothe K, Wille R and Zastrau B W (2001), Advanced contact mechanics – road and rail, *Vehicle System Dynamics*, **35**(4–5), 361–407. A review paper on rolling contact modelling.
- Gross-Thebing A (1993), Lineare Modellierung des instationären Rollkontaktes von Rad und Schiene, *VDI Fortschritt-Berichte*, Reihe 12, 199, VDI-Verlag, Düsseldorf, Germany. A PhD thesis on the linearised analysis of non-steady-state rolling contact.

3.7 References

Alonso A and Giménez J G (2005), A new method for the solution of the normal contact problem in the dynamic simulation of railway vehicles, *Vehicle System Dynamics*, **43**(2), 149–60.

Andersson E (1977), Simulation von Spurkräften und Laufeigenschaften, *ZEV-Glasers Annalen*, **101**(8/9), 339–47.

Ayasse J B and Chollet H (2005), Determination of the wheel-rail contact patch in semi-Hertzian conditions, *Vehicle System Dynamics*, **43**(3), 161–72.

Björklund S and Andersson S (1994), A numerical method for elastic contacts subjected to normal and tangential loading, *Wear*, **179**, 117–22.

van Bommel P (1964), *Application de la théorie des vibrations nonlineaires sur le probleme du mouvement de lacet d'un vehicule de chemin de fer*, Doctoral Dissertation, Technische Hogeschool Delft, Delft the Netherlands.

Boussinesq J (1885), *Application des potentials à l'etude de l'equilibre et du movement des solides elastiques*, Paris, France, Gautier–Villars.

Bucher F, Knothe K and Theiler A (2002), Normal and tangential contact problem of surfaces with measured roughness, *Wear*, **253**, 204–18.

Busher F (2002), *The contact between micro-rough rails and wheels*, Doctoral Thesis, University of Technology, Berlin, Germany.

Carter F W (1926), On the action of a locomotive driving wheel, *Proceedings of the Royal Society London*, **A112**, 151–7.

Cattaneo C (1938), Sul contattodi due corpi elastici: distribuzione locale degli sforzi, *Rendiconti dell'Accademia Nazionale dei Lincei*, **27**, 342–8, 434–6, 474–8.

Cerruti V (1882), *Accademia dei lincei*, Roma, Mem fis mat.

Cooperrider N K, Hedrick J K, Law E H, Kadala P S and Tuten J M (1975), *Analytical and Experimental Determination of Nonlinear Wheel/rail Geometric Cconstraints*, Report FRA-O&RD 76-244, US Deptartment of Transportation, Washington DC, USA.

Czichos H (1978), *Tribology*, Elsevier, Amsterdam, the Netherlands

Deresiewicz H (1957), Oblique contact of nonspherical elastic bodies, *Journal of Applied Mechanics*, **24**, 623.

Enblom R and Berg M (2008), Impact of non-elliptic contact modelling in wheel wear simulation, *Wear*, **265**, 1532–41.

Evans J R (1999), Rail vehicle dynamic simulations using VAMPIRE, *Vehicle System Dynamics* **31** (Supplement), 119–40, Swets & Zeitlinger, Lisse, the Netherlands.

Harder R F (1999), Creep force – Creepage and frictional work behaviour in non-Hertzian counter formal rail/wheel contacts, *Proceedings International Heavy Haul Association Specialist Technical Session on Wheel/Rail Interface*, Moscow, Russia, 14–17 June, 207–14.

Hertz H (1882), Über die Berührung fester, elastischer Körper, *J für die reine und angewandte Mathematik*, **92**, 156–71.

Iwnicki S D (1999), *The Manchester Benchmarks for Rail Vehicle Simulation*, CRC, Boca Raton, FL, USA.

Jaeger J (1992), *Elastic Impact with Friction*, Thesis, Delft University of Technology, Delft, the Netherlands.

Johnson K L (1985), *Contact Mechanics*, Cambridge, UK, Cambridge University Press.

Kalker J J (1967), *On the Rolling Contact of Two Elastic Bodies in the Presence of Dry Friction*, Dissertation, Delft University of Technology, Delft, the Netherlands.

Kalker J J (1971), A minimum principle for the law of dry friction with application to elastic cylinders in rolling contact; Part 1: Fundamentals, application to steady-state rolling; Part 2: Application to non-steady rolling cylinders, *Journal of Applied Mechanics*, **38**, 875–80.

Kalker J J (1990), *Three-dimensional Elastic Bodies in Rolling Contact*, Dordrecht, the Netherlands, Kluwer Academic.

Kik W and Piotrowski J (1996), A fast, approximate method to calculate normal load at contact between wheel and rail and creep forces during rolling, in Zobory I (ed.), *Proceedings 2nd Mini Conference on Contact Mechanics and Wear of Rail/Wheel Systems*, Budapest, University of Technology Budapest, Hungary, 52–61.

Knothe K and Gross-Thebing A (1986), Derivation of frequency dependant creep coefficients based on an elastic half-space model, *Vehicle System Dynamics*, **15**(3), 133–53.

Knothe K and Le-The H (1984), A contribution to the calculation of the contact stress distribution between two elastic bodies of revolution with non-elliptical contact area, *Computers and Structures*, **18**(6), 1025–33.

Linder C and Brauchli H (1996), Prediction of wheel wear, in Zobory I. (ed.), *Proceedings 2nd Mini Conference on Contact Mechanics and Wear of Rail/Wheel Systems*, Budapest, University of Technology Budapest, Hungary, 215–23.

Liu G, Wang Q and Lin C (1999), A survey of current models for simulating the contact between rough surfaces, *Tribology Transactions*, **42**, 581–91.

Love A E H (1929), Stress produced in a semi-infinite solid by pressure on part of the boundary, *Philosophical Transactions of the Royal Society*, **A228**, 377.

Marshall M B, Lewis R, Dwyer-Joyce R S, Olofsson U and Björklund S (2006), Experimental characterization of wheel-rail contact patch evolution, *Journal of Tribology*, 493–505.

Mindlin R D (1949), Compliance of elastic bodies in contact, *Trans ASME, Journal of Applied Mechanics*, **16**, 259–68.

Pascal J-P and Sauvage G (1991), New method for reducing the multicontact wheel/rail problem to one equivalent rigid contact patch, in Sauvage G (ed.), *The Dynamics of Vehicles on Roads and Tracks, Proceedings 12th IAVSD Symposium*, Swets and Zeitlinger, Lisse, the Netherlands, 475–89.

Pascal J-P and Sauvage G (1993), The available methods to calculate the wheel/rail

forces in non-Hertzian contact patches and rail damaging, *Vehicle System Dynamics*, **22**(3/4), 263–75.

de Pater A D (1961), The approximate determination of the hunting movement of a railway vehicle by aid of the method of Krylov and Bogoljubov, *Applied Scientific Research*, **10**, 205–28.

Persson B N J (2001), Elastoplastic contact between randomly rough surfaces, *Physical Review. Letters*, **87**, 1161.

de Possel R, Beautefoy J and Matsudaira T (1960), *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, the Netherlands.

Quost X, Sebes M, Eddhahak A, Ayasse J B, Chollet H, Gautier P E and Thouverez F (2006), Assessment of a semi-Hertzian method for determination of wheel-rail contact patch, *Vehicle System Dynamics*, **44**(10), 789–814.

Schielen W (1990), *Multibody Systems Handbook*, Springer-Verlag, Berlin.

Schupp G (2004), Computational Bifurcation analysis of mechanical systems with applications to railway vehicles, *Vehicle System Dynamics*, **41**(Supplement), 458–67.

Shackleton P and Iwnicki S D (2008), A comparison of wheel-rail contact codes for railway vehicle simulation: an introduction to the Manchester Contact Benchmark and initial results, *Vehicle System Dynamics*, **46**, 129–49.

True H (1994), Does a Critical Speed for Railroad Vehicles exist? RTD-Vol. 7, *Proceedings 1994 ASME/IEEE Joint Railroad Conference*, Chicago, IL, USA, 22–24 March, American Society of Mechanical Engineers, New York, USA, 125–31.

Wallrapp O Führer C (1990), MEDYNA – an interactive analysis program for geometrically linear and flexible multibody systems, in W. Schielen, ed., *Multibody Systems Handbook*, Springer-Verlag, Berlin, Germany, 203–23.

Wickens A H (1965), The dynamic stability of railway vehicle wheelsets and bogies having profiled wheels, *International Journal of Solids and Structures*, **1**, 319–41.

Winkler E (1867), *Die Lehre von der Elasticität und Festigheit mit besonderer Rücksicht auf ihre Anwendungen in der Technik, 1. Teil*, Prague, Czech Republic, H Dominicus.

Copyrighted Materials
Copyright © 2009 Woodhead Publishing Retrieved from www.knovel.com

4

Friction and wear simulation of the wheel–rail interface

S. ANDERSSON, Royal Institute of Technology (KTH), Sweden

Abstract: The conditions in wheel rail contacts can be both mild and severe. Predicting the friction and the wear in such contacts is generally thought to be rather difficult and uncertain. This paper, however, addresses these tasks and will outline some possibilities for predicting friction and wear in rolling and sliding contacts as in wheel–rail contacts. In a rolling and sliding contact, the two interacting surfaces characteristically move at different speeds in a tangential direction. The Tribology Group at KTH Machine Design has worked on simulating friction and wear in rolling and sliding contacts for more than 20 years. The modelling principles the group has successfully used are based on (i) the single-point observation method and (ii) treating friction and wear as initial-value processes. Simple examples will be presented, demonstrating how these principles can be used.

Key words: friction, wear, simulation, models.

4.1 Introduction

Both rolling and sliding occur in wheel–rail contacts. The wheel tread is mainly in contact with the rail head on straight tracks, but in the curves the flange may be in contact with the gauge corners of the wheel. The contact area in rolling and sliding contacts is often divided into stick and slip regions, depending on the tangential deformations in the contacts relative to the ideal displacements in the contacts. In curves the flanging often gives nearly complete sliding between the interacting contact surfaces, but for straight tracks the contact between will be mainly rolling.

Friction is the restrictive force or the tractive force that hinders the relative tangential motion of two interacting surfaces. In this case, the friction between the wheels and the rails should be as high as possible, although with some restrictions regarding stability and wear. The forces are dependent on the material and the properties of the interacting surfaces, the materials between the surfaces and the contact intensities. The form and the topography of contact surfaces are important and have a significant influence on the behaviour of the contacts. The topography can often be looked at as being random. When the contact surfaces interact the friction and wear may be random as well. The surface tops, i.e. the asperities, of the interacting surfaces will often

interact randomly. The asperities will come into and generate and leave local contacts randomly and the result will be a strongly varying friction force. That force is often strongly mechanically filtered in service, but in friction measuring machines you may see some variations, which afterwards are often further filtered numerically. Friction models as well as wear models seldom consider the randomness of the friction force, but rather represent a mean value. The value of the friction force, in the same manner as wear, is dependent on whether the interacting surfaces are sliding or both sliding and rolling. In order to understand and build up models for the friction and wear of contact surfaces, we have found that a mental model, which we have named the single point observation method or principle (see below) has made it easier to understand the processes by following what happens at a particular point of one of the interacting contact surfaces during operation.

In a contact between a railway wheel and a rail, we would like to have as much friction as possible, since the propulsion of the train is based on the frictional contacts between the wheels and the rails. However, we have to restrict the general goal in order to get a reliable and robust solution to the interaction of the contact between railway wheels and rails. That the friction should be as high as possible is still a general goal, but it has to be considered together with some secondary goals, i.e. that friction should not vary too much and that wear should be as little as possible.

Wear can be defined as the removal of material from solid surfaces by mechanical action. Wear can appear in many ways, depending on the material of the interacting contact surfaces, the operating environment and the running conditions. In engineering terms, wear is often classified as either mild or severe. *Mild* wear is what engineers strive for and can be obtained by creating contact surfaces of appropriate form and topography. Choosing adequate materials and sometimes lubrication is also often necessary in order to obtain mild wear conditions. The problem is that in order to achieve mild wear the contacts must often be lubricated in some way. Lubrication will reduce wear but also give low friction, which must sometimes be accepted. Mild wear results in smooth surfaces, and that type of wear dominates at the rail head. *Severe* wear may occur at times, producing rough or scored surfaces which will often generate a surface rougher than the original one. Severe wear can be acceptable although rather extensive, but it can also be catastrophic which is always unacceptable. Severe wear may be found at the rail edges in curves.

The potential to predicting friction and wear is often thought to be limited. Even so, many friction and wear models are found in the literature.[1,2] These models are often simple ones describing a single friction and wear mechanism from a fundamental point of view or empirical relationships fitted to particular test results. Most of them represent a mean value, i.e. the random characteristic of both friction and wear is seldom considered. In this

chapter we will present some of the most commonly used models or models that could be used in simulations of friction and wear in railway wheel and rail contacts.

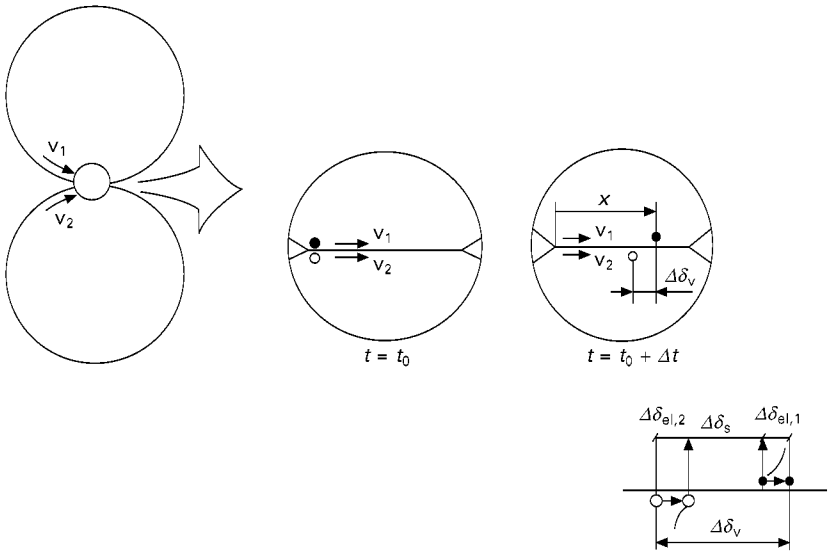
4.2 Single-point observation method

The single-point observation method has been found to be very useful in improving understanding and modelling of many friction and wear processes. The method was developed and successfully used in many works at KTH Machine Design in Stockholm. It was initiated as a concept in simulation work of mild wear in boundary-lubricated spur gears.[3] The theoretical application of the method was based on formulas for the sliding distances in gears developed long before in the author's PhD thesis.[4] The author then found that the distances over which a point on a gear flank slides against an opposite flank during one mesh vary depending on the position on the flank, the gear ratio, the size of the gears and the loads applied on the gear tooth flanks. The author also found that if the sliding distance must be considered, gear contacts cannot be replaced with rolling and sliding rollers in contact. Results obtained from some gear tests indicate that the amount of wear on the gear flanks seems to be in line with the simulated wear which has been based on, amongst other factors, determination of sliding distances. That observation and many years of pin-on-disc tests have inspired the author and others to simulate friction and wear in rolling and sliding contacts of different types. The findings of these simulations have been verified by experimental results. All these works have also improved our understanding of what is going on in contacts.

The single-point observation method can be illustrated by a rolling and sliding contact according to Fig. 4.1. The contact surfaces move with peripheral speeds of v_1 and v_2 with $v_1 > v_2$. We observe a point on surface 1, P_1, that has just entered the contact and follow that point through the contact. We also note a point on surface 2, P_2, which is opposite the first observed point P_1 on surface 1 when it enters the contact. As P_1 moves through the contact the interacting opposite surface will not move as fast as surface 1, since $v_1 > v_2$. A virtual distance $\Delta\delta_v = x \cdot (v_1 - v_2)/v_1$ in the tangential direction will occur between P_1, the observed point, and a point P_2, opposite the first one when entering the contact. That distance is first compensated by tangential elastic deformations of the contact surfaces $\Delta\delta_{el,1} + \Delta\delta_{el,2}$, but, when that is not possible any longer, the observed point will slide against the opposite surface a distance $\Delta\delta_s$ equal to:

$$\Delta\delta_s = \Delta\delta_v - \Delta\delta_{el,1} - \Delta\delta_{el,2} \qquad [4.1]$$

The frictional shear stress in the contact depends on the process which means that it will first be dependent mainly on the elastic deformations and,



4.1 Illustration of the single-point observation method and the elastic deformation and sliding in a rolling and sliding contact.

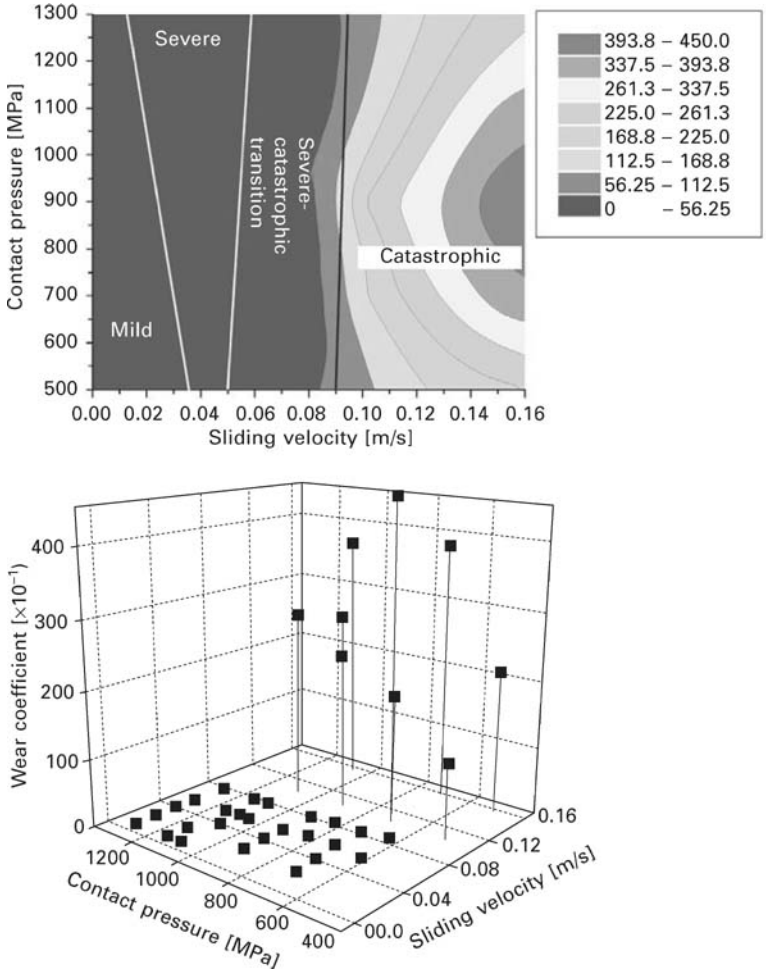
at higher torques or higher slip, mainly on the sliding between the surfaces. However, since these phenomena are always active in rolling and sliding contacts, it is interesting to analyse the effect the stick zone represented by the elastic deformation has on the friction force in the contact. In many applied cases, however, the elastic deformation effect on friction can be neglected.

4.3 Wear maps and transition diagrams

As mentioned in the introduction, friction and wear can be of different types. We would like to know what type of friction and wear can be expected in a particular contact and when and why the transitions between different types occur. Some interesting research on that subject has been carried out and is on going. We will therefore briefly present some results from work done on transitions between different friction and wear modes. The diagrams to look for are often called wear maps or transition diagrams. The most frequently referenced paper about wear maps is that by Lim and Ashby,[5] who classified different wear mechanisms and corresponding wear models for dry sliding contacts. They studied the results of a large number of dry pin-on-disc experiments and developed a wear map, based on the parameters: $\tilde{Q} = V/As$, $p = \dfrac{F_N}{AH}$ and $v = \dfrac{vr_0}{a_0}$, where V is the wear volume, A is the apparent contact area, F_N is the normal load, H is the hardness of the softer material

in the contact, v is the sliding velocity, r_0 is the radius of the pin, and a_0 is the thermal diffusivity of the material. The friction in the contacts is directly related to the different wear mechanisms and can thereby be formulated for the different mechanisms. That is not, however, done directly.

For contacts between railway wheels and tracks, Lewis and Olofsson,[6] have carried out a similar investigation. The goal of their investigation was 'to produce tools in the form of maps of rail material wear data for identifying and displaying wear regimes and transitions'. They collected wear data from laboratory as well as field tests, but found that data are often lacking for rail gauge and wheel flange contacts. Independent of that, they collected available data and structured the data in different ways. Figure 4.2 shows



4.2 Wear coefficient map according to Lewis and Olofsson.[6]

an example of a wear coefficient map developed by Lewis and Olofsson.[6]

The wear coefficient they used was determined using the common wear model:

$$\frac{V}{s} = K \cdot \frac{N}{H} \qquad [4.2]$$

where s is sliding distance, N is the normal contact force and K is a dimensionless wear coefficient. This model used to be named Archard's Wear Law, but could just as well be named Holm's Wear Law. An ongoing work on transitions in wheel–rail contacts is being carried out by Sundh.[7] His goal is to construct wear maps that include the contact between rail gauge and wheel flange. A further goal is to study how transfers from mild to severe wear are dependent on the type of lubricant, surface coating and topography. He studies both dry and lubricated contacts.

For lubricated contacts, the degree of surface separation by a lubricant very strongly influences both the friction and the wear. The degree of separation in a contact is often divided into boundary lubrication, mixed lubrication and full-film lubrication.

Boundary lubrication refers to lubrication in which the load is supported by the interacting surface asperities and the lubrication effect is mainly determined by the boundary properties of the lubricant between the interacting asperities. In *mixed* lubrication, the lubricant film itself supports some of the load in the contact, although the boundary properties of the lubricant are still important. In this case, the hydrodynamic and elastohydrodynamic effects are also important. Mixed lubrication is therefore sometimes referred to as partial lubrication or partial elastohydrodynamic lubrication (EHL). In *full-film* lubrication, the interacting contact surfaces are fully separated by a fluid film. In the literature, full-film lubrication is sometimes referred to as EHL, since the film-formation mechanism of high-performance contacts and local asperity contacts is probably elastohydrodynamic.

In contacts between railway wheels and rails the lubrication, if any, is mainly boundary, where the effect of the lubricant is mainly related to the boundary properties of the lubricant. Lubricants used are all from oils, greases and pastes. The lubricating effect from the environment in the form of crushed leaves is an old and well known problem in the autumn. The crushed leaves seem to react with the rail material, generating a very slippery low-friction coating on the surfaces.

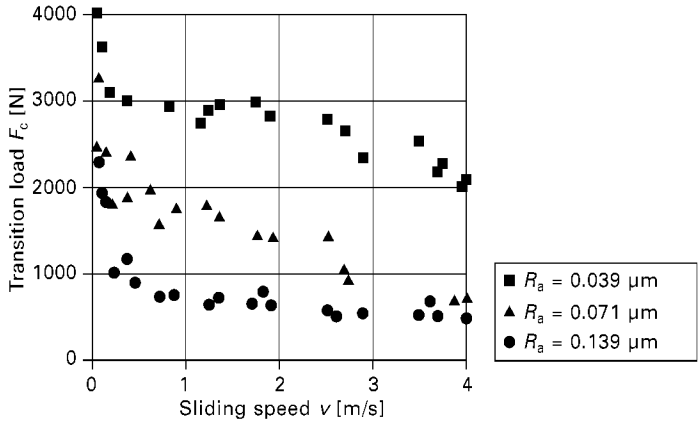
As mentioned in the introduction, the transition from the desired mild situation to a severe situation should be avoided. Research has been conducted to determine when and under what conditions transitions from one kind of friction and wear to another may occur in lubricated contacts as well. In an initial such study, a so-called IRG transition diagram was developed.[8] In that diagram, one can identify the different lubrication regimes: a mixed

or partial EHL regime, a boundary lubrication regime and a failure regime. The last regime is sometimes called the scuffed or unlubricated regime and is a severe condition. The other regimes are mild. The transition from a desired mild regime to a severe regime has also been studied by Andersson and Salas-Russo (see Fig. 4.3).[9] The criteria they used for the transition is the track appearance. When a significant part of the track is scored, seized or strongly plasticized, severe conditions are obtained. Andersson and Salas-Russo found that for bearing steels, the surface topography has a stronger influence on the mild to severe transition level than does the viscosity of the lubricant.

4.4 Friction models

Friction is the force between interacting surfaces that hinders the relative tangential motion of the surfaces. The deformations of the contact surfaces may play some role in contacts, but can normally be neglected. The general trend toward increased use of product models and simulations during both the development and use of products has created a need for powerful and relevant product models. Since friction has a very strong influence on the performance and behaviour of mechanical systems, good representation of friction phenomena is important in all simulations and analyses.

The Machine Elements Groups at the Royal Institute of Technology in Stockholm and at the Luleå Technical University in Luleå, along with a number of Swedish companies, have pursued a research program named INTERFACE. The goal of the program was to develop relevant friction and wear models



4.3 The influence of surface roughness R_a on the transition load of a lubricated sliding steel contact. Ball (d = 10 mm) and disc material: SAE52100, $H_{v,ball}$ = 8000–8500 MPa, $H_{v,disc}$ = 5800–6300 MPa, $R_{a,ball}$ = 0.008 μm. Lubricant: ISO VG 46 mineral oil.[23]

for simulations in industry of different types of mechanical devices. The program was based on previous work by Sellgren,[10,11] who developed general principles for modelling systems. His approach was modular, and laid down strict guidelines for behavioural models of machine elements, modules and interfaces. Sellgren defined an interface as an attachment relation between two mating faces.[11] That definition was elaborated on by Andersson and Sellgren in terms of an interaction relation between two functional surfaces.[12] A functional surface is a carrier of a function.

Friction in mechanical contacts is influenced by many parameters, including the geometry of the contact surfaces, their properties, the running conditions and any lubricants used. This chapter presents different friction models for transient sliding contacts running under different conditions.

4.4.1 Common friction models for pure sliding and oscillating contacts

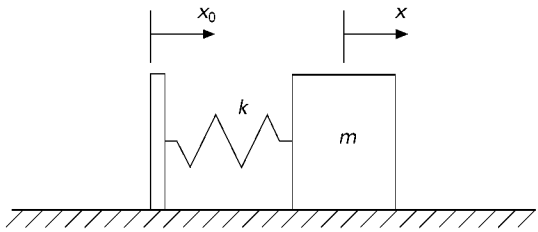
The behaviour of the different friction models presented below was studied using the 1 DOF system shown in Fig. 4.4, where the motion of the left wall, x_0, is independent. It can thus also have an oscillating motion with different frequencies. The equation of motion for the system can be formulated as:

$$m \cdot \ddot{x} = k \cdot (x_0 - x) - F_f \qquad [4.3]$$

where m is the mass of the moving body, x is the position of the body, x_0 is the independent position of the wall, k is the spring constant and F_f is the friction force, equal to F_C if we assume Coulomb friction, F_V if we assume viscous friction, etc. The friction force on the body acts in the opposite direction to the sliding velocity of the upper body.

Coulomb friction model

The most commonly used friction model is the Coulomb friction model, which can be formulated as:



4.4 A 1 DOF dynamic system with a moving wall, a spring, and a body sliding on another fixed body. The friction in the contact between the bodies can be represented by different models.

$$F_C = \begin{cases} \mu N \cdot \text{sign}(v) \text{ if } |v| > 0 \\ \min(|F_{app}|, \mu N) \cdot \text{sign}(F_{app}) \text{ if } v = 0 \end{cases} \quad [4.4]$$

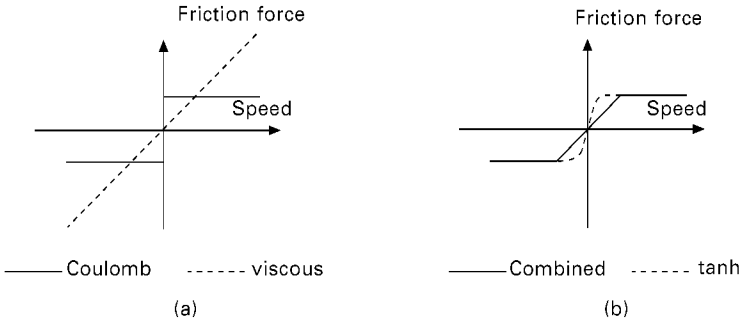
where F_C is the Coulomb friction force, $v = \dot{x}$ is the sliding speed, μ is the coefficient of friction and N is the normal contact force. F_{app} is the applied force on the body. The Coulomb friction model is often simplified as:

$$F_C = \mu N \cdot \text{sign}(v) \quad [4.5]$$

although this can cause problems in simulations due to the properties of the sign function. The Coulomb friction model is illustrated in Fig. 4.5a.

Coulomb friction is often referred to as dry friction, but the model is used for dry contacts as well as boundary and mixed lubricated contacts. Although it is known that a Coulomb friction model does not always a good representation of the friction behaviour in a contact, such models are often used to describe the friction in mechanical contacts. It must be remembered that the representation of friction data for different contacts by coefficients of friction does not necessarily mean that the contacts exhibit Coulomb friction behaviour; rather, these coefficients represent a certain value for a particular running condition. In practice, however, these coefficients are often used as if the friction behaviour is in accordance with Coulomb, which sometimes results in deviations and poor behaviour prediction. All in all, it is surprising that the use of the coefficient of friction is so common despite the fact that the use of a Coulomb friction model in analyses and simulations is often rather troublesome.

There are advantages to using a modularized model architecture of a system in a product realization process with natural question-driven simulations. Such an architecture makes it possible to assemble different component and interface models into a system model and to directly simulate the behaviour



4.5 (a) Coulomb and viscous friction forces as function of sliding speed. (b) Combined Coulomb and viscous friction and combined Coulomb and tanh friction as function of sliding speed.

of the system, without extra condition checks on the state of the system or interruptions in the model. However, in order to construct such a model, some extra considerations have to be taken into account in order to find useful friction models.

4.4.2 Viscous friction model

Since the equation of motion for dynamic systems is strongly non-linear with a Coulomb friction model, a viscous friction model is often used instead. Such a model is considerably easier to simulate, but the representation of the friction is often poor. A viscous friction model can be formulated as:

$$F_V = k_V v \tag{4.6}$$

where F_v is the viscous friction force, v is the sliding speed and k_v is a viscous coefficient. In this case the equation of motion for the system in Fig. 4.6b is a linear differential equation, which can be solved both numerically and analytically. Although the simulation is easy to perform, the validity of the viscous model is doubtful. However, in some cases, such as full-film contacts, the viscous model may offer the best representation of the behaviour. In other cases, the viscous model is a poor representation but, by tuning the viscous coefficient, the model can represent such things as damping rather well under particular running conditions (see Fig. 4.6a). A step response with the 1 DOF system shown in Fig. 4.4 with viscous friction is shown in Fig. 4.6a. The damping of the oscillations varies with the viscous friction coefficient.

4.4.3 Combined Coulomb and viscous friction model

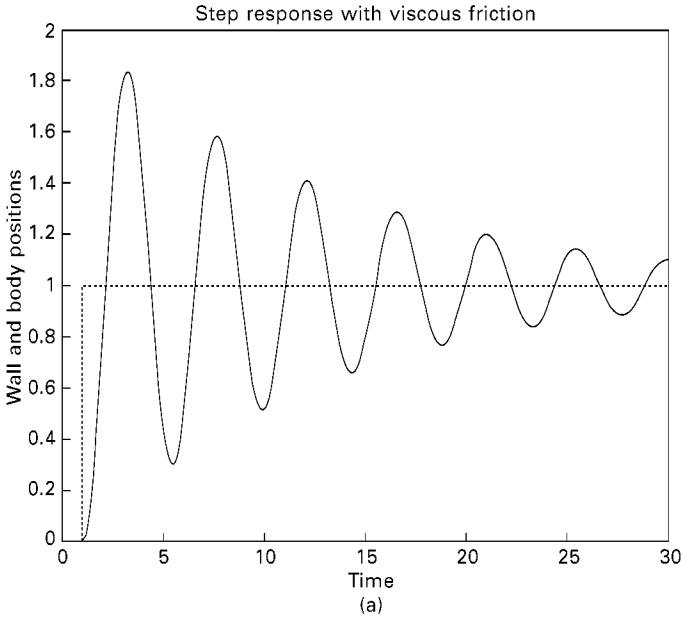
Since the Coulomb friction model is problematic as regards both the analysis and simulation of a system's behaviour, a combination of the viscous friction model and the Coulomb friction model could be advantageous. Such a model will have the following form:

$$F_{CV} = \min(k_{sat}|v|, \mu N) \cdot \text{sign}(v) \tag{4.7}$$

where k_{sat} is a coefficient that determines the speed of the transition from – to + (see Fig. 4.5b).

The combined Coulomb and viscous friction model can easily be modelled in MATLAB/Simulink by using the saturation block. The 'sat' function can be defined as follows:

$$\text{sat}\,(k_{sat} v) = \begin{cases} \min\,(k_{sat} v, 1) \text{ if } v \geq 0 \\ \max\,(k_{sat} v, -1) \text{ if } v < 0 \end{cases} \tag{4.8}$$



4.6 Results from a step response simulation of a 1 DOF system as shown in Fig. 4.4, with (a) a viscous friction model, (b) a combined Coulomb and viscous friction model.

The combined Coulomb and viscous friction model can then be formulated as:

$$F_{CV} = \mu N \cdot \text{sat}(k_{sat} v) \quad [4.9]$$

Using the combined Coulomb and viscous friction model eliminates the difficulty in determining the friction force at zero sliding speed both at startup and at direction change, although we still have a strongly non-linear system model to solve. However, simulation tests do not indicate any problem with the non-linearity and the step response shown in Fig. 4.6b behaved as expected. The combined Coulomb and viscous friction model is thus a plausible candidate to represent friction in sliding and oscillating contacts. Another possibility, also used in some systems models, is to replace the sign function in the simplified Coulomb friction model shown in Eq. (4.3) with a tanh function, and to let the new model be valid for all v:

$$F_{C\,\tanh} = \mu N \cdot \tanh(k_{\tanh} v) \quad [4.10]$$

where $k_{\tanh}$ is a coefficient that determines how fast the tanh function changes from near –1 to near +1. This model behaves similarly to the other combined model, but is probably more numerically stable (Fig. 4.5b).

The proposed combined Coulomb and viscous model and the combined Coulomb and tanh friction model are both functions of the sliding speed. This means that the final position of an oscillating motion of the body will be zero when all applied forces are zero. That is not what is expected in a real case of Coulomb friction. Thus the proposed combined friction models are convenient for simulations provided we are not interested in a very accurate value for the final position. If the final position is important, we have to consider the small displacements (i.e. the micro-displacements) that are functions of the tangential deformations or displacements between the surfaces (see below).

Stribeck friction model

Most sliding contacts are lubricated or contaminated. The friction force will then vary with the sliding speed depending on the extent to which the interacting contact surfaces are running under boundary, mixed or full-film conditions. Even dry contacts show some behaviour similar to that in lubricated contacts in that they have a higher static friction than dynamic or sliding friction. In lubricated sliding contacts, the friction decreases with increased sliding speed until a mixed or full-film situation is obtained, after which the friction in the contact can either be constant, increase or decrease somewhat with increased sliding speed due to viscous and thermal effects. This behaviour was described some 100 years ago by Stribeck,[13] whose name is often associated with sliding friction in lubricated contacts

running under boundary, mixed and full-film conditions. A model can be formulated as:

$$F_S = (\mu N + (F_{S0} - \mu N) \cdot e^{-(|v|/v_S)^i}) \cdot \text{sign}(v) + k_v v \quad [4.11]$$

where F_S is the Stribeck friction force, v is the sliding speed, μ is the coefficient of friction and N is the normal load in the contact. F_{S0} is the maximum static friction force, v_s is a sliding speed coefficient, k_v is a viscous coefficient and i is an exponent.

The Stribeck friction model can provide very good representation of the friction between sliding surfaces. It covers everything from Coulomb friction to viscous friction, depending on the choice of parameter values. However, the Stribeck friction model presents the same problem as the Coulomb friction model when it comes to changing sliding direction when the friction value is dependent on the applied force and the contact surfaces are sticking to each other. As a result, we tried to formulate a combined Stribeck and viscous model and a combined Stribeck and tanh friction model, both accommodating the transition to pure sliding in either direction. We replaced the sign function with a $\text{sat}(k_{sat}v)$-function and a $\tanh(k_{tanh}v)$-function, respectively:

$$F_{SV} = (\mu N + (F_{S0} - \mu N) \cdot e^{-(|v|/v_S)^i}) \cdot \text{sat}(k_{sat} v) + k_v v \quad [4.12]$$

and

$$F_{S\tanh} = (\mu N + (F_{S0} - \mu N) \cdot e^{-(|v|/v_S)^i}) \cdot \tanh(k_{tanh} v) + k_v v \quad [4.13]$$

These combined Stribeck models facilitate simulations in the same way as the combined Coulomb models, but also share their disadvantages (see above).

4.4.4 Friction at small displacements

The main disadvantage of the Coulomb and Stribeck friction models is the undetermined friction force at zero-sliding speed. Although the behaviour at these points can be modelled fairly well, the representation often complicates simulations and sometimes necessitates extra condition checks of the system states or interruptions of the simulations. Various tricks can be used to overcome the problems. These modifications improve the ability to simulate the behaviour of systems, but do not represent small displacements and position at stop very well. Small displacements and position at low speeds are important in many high-precision and control applications, and so we also have to look for a friction model that can handle these conditions.

The friction behaviour of contact surfaces under pure or oscillating sliding conditions with very small displacements has been extensively studied at

the Machine Element Group at KTH for many years. All the results show that the transition from positive to negative motion or the start of a motion is always associated with small relative tangential displacements between the interacting surfaces before full friction or gross slip is obtained. The displacement before gross slip occurs is often 1–10 μm. The explanation of this behaviour is that before gross slip occurs, there is local sliding in the contact. This phenomenon used to be named micro-slip or micro-displacement.

Micro-slip models based on basic surface models represent the micro-slip phenomenon rather well, but are not always easy to use in dynamic simulations.[14–16] Drawing on existing models, Sellgren and Olofsson developed a constitutive model for micro-slip in finite element (FE) analysis.[17] The derived model is, however, rather complex to use in other simulation tools, because the displacements have to be set equal to zero at turning points. This problem can be overcome by formulating the micro-slip model as a differential equation according to the Dankowicz model (see below). Another way to accommodate micro-slip in contacts is by combining Coulomb friction with elastic deformation, which is often used in FE codes.

4.4.5 Dankowicz friction model

Dankowicz represented friction behaviour by a first-order differential equation and a help variable z.[18] The Dankowicz friction model has the following form:

$$z = x \cdot \left(1 - \frac{z}{\delta} \cdot \text{sign}\,(x)\right)$$
$$F_D = \frac{F_{max}}{\delta} \cdot z \qquad [4.14]$$

where F_D is the Dankowicz friction force, z is the help variable, δ represents the micro-slip displacement before gross slip and F_{max} is the maximum friction force, which here can be considered as the sliding friction force.

In order to find out to what extent the Dankowicz model is numerically appropriate in different situations, an analysis was made using the 1 DOF system shown in Fig. 4.4. To reduce the number of simulations, the equation of motion and the friction model were made non-dimensional by introducing the non-dimensional variables

$$X = \frac{x}{\delta}, X_0 = \frac{x_0}{\delta}, Z = \frac{z}{\delta}, T = t \cdot \omega_e = t \cdot \sqrt{k/m}$$

which yield the non-dimensional differential equations for the 1 DOF system in Fig. 4.4, with the Dankowicz friction model:

$$\begin{cases} X = (X_0 - X) - CZ \\ Z = X\,(1 - Z \cdot \text{sign}\,(X)) \end{cases} \quad [4.15]$$

where

$$C = \frac{F_{\max}}{\delta \cdot k}$$

The non-dimensional equation of motion was solved numerically using MATLAB/Simulink with the solver ODE23. Two different motions of the wall were analyzed: an oscillating motion with continuously increasing frequency and an oscillating motion with a constant frequency. Each motion was first simulated with an amplitude of the same size as δ and then with an amplitude ten times δ. For each combination of motion type and amplitude, three cases were simulated: $C = 1$, $C = 10$, and $C = 100$. The simulations were carried out without any trouble, leading to the conclusion that the Dankowicz model seems to work very well under different conditions.

Dahl friction model

The Dahl model is frequently used in control engineering.[19] Both Dahl and Dankowicz based their friction models on the fact that the friction force is a function of displacement only, and thus the time rate of friction can be expressed as:

$$\frac{dF_{\text{Dahl}}(x)}{dt} = \frac{dF_{\text{Dahl}}(x)}{dx} \cdot \frac{dx}{dt} \quad [4.16]$$

where $F_{\text{Dahl}}(x)$ is a friction force. The general Dahl friction model has the form:

$$\frac{dF_{\text{Dahl}}}{dx} = \sigma_0 \left| 1 - \frac{F_{\text{Dahl}}}{\mu N} \text{sign}\,(\dot{x}) \right|^i \cdot \text{sign}\left(1 - \frac{F_{\text{Dahl}}}{\mu N} \text{sign}\,(\dot{x}) \right) \quad [4.17]$$

where F_{Dahl} is the Dahl friction force and σ_0 is a coefficient. In the literature, the Dahl model is often simplified with the exponent $i = 1$:

$$\frac{dF_{\text{Dahl}}}{dt} = \sigma_0 \left(1 - \frac{F_{\text{Dahl}}}{\mu N} \text{sign}\,(\dot{x}) \right) \cdot \dot{x} \quad [4.18]$$

Canudas de Wit et al. *friction model*

The Canudas de Wit *et al.* friction model[20,21] has many similarities with the Dahl and Dankowicz models. They, too, based their model on the fact that

the friction force is dependent on deformation between the surfaces and formulated the model as follows, in its simplest form:

$$\dot{z} = \dot{x} - \frac{|\dot{x}|}{g(x)} \cdot z \qquad [4.19a]$$

$$g(v) = \frac{1}{\sigma_0}(\mu N + (F_{S0} - \mu N)e^{-(|v|/v_S)^2}) \qquad [4.19b]$$

$$F_{CW} = \sigma_0 z + \sigma_1 \frac{dz}{dt} + k_v v \qquad [4.19c]$$

where F_{CW} is the friction force and σ_0 and σ_1 are coefficients.

Combined Coulomb, Stribeck, viscous and Dankowicz friction model for small transient motions

Based on the findings above, we propose that in many cases a combined friction model for sliding motions should be used, based on the following equations:

$$\left.\begin{aligned} F_{SD} &= \left(1 + \left(\frac{F_{S0}}{\mu N} - 1\right) \cdot e^{-(|v|/v_S)^i}\right) \cdot \mu N \cdot \frac{z}{\delta} + k_v v \\ z &= v \cdot \left(1 - \frac{z}{\delta} \cdot \operatorname{sign}(v)\right) \end{aligned}\right\} \qquad [4.20]$$

This model incorporates micro-slip, Coulomb friction, Stribeck effects and viscous friction. The δ parameter can be determined from micro-displacement tests. The other parameters (i.e. μN, F_{S0}, v_s, i and k_v) can be determined from dynamic friction tests. Normally we then make oscillating tests and identify the parameters from the test results.

4.4.6 Friction as a random process

Friction between sliding contact surfaces is a result of random interactions between rubbing asperities. The rubbing asperities cause friction by shearing surface materials, lubricants or surface coatings. Yet while the friction force in nearly all friction tests is highly random in nature, with significant variations in both amplitude and frequency, most friction models do not take these variations into account and, instead, represent the friction forces by a smooth mean value. Furthermore, test results seldom compensate for the dynamic behaviour of the test equipment, although most friction measurements are subjected to mechanical filtering by the test machine dynamics.

The random nature of friction can be represented in the friction model by adding noise to the smooth mean friction force. White noise is not recommended since it contains frequencies that are infinitely high. A stochastic signal with controllable frequencies and amplitude with a standard deviation can be created using the method described in.[22] The friction model can then be written as:

$$F_{\text{random}} = F_{\text{smooth}} + P(A, f) \qquad [4.21]$$

where F_{random} is the random friction force, F_{smooth} is the friction force determined by any smooth friction model, $P(A, f)$ is a random function representing the friction force noise, A is the amplitude standard deviation and f is a typical frequency. Since the random friction force variations are often produced by asperity interactions during sliding, the noise frequency may be a function of the sliding speed v.

4.4.7 Concluding remarks regarding friction models

There are advantages to generating behavioural models for mechanical systems using a modularized model structure.[10–12] The structure is built up of component, subsystem and interface models. By connecting these models, a system model is constructed that can be used for simulations and analyses. In this chapter, the term interface refers to mechanical contacts interacting under transient and oscillating sliding motions.

Friction has a strong influence on both the behaviour and performance of a system. This chapter has presented different friction models and discussed their relevance and convenience as models in a modularized system model structure and as part of a system model in simulations. The different friction models presented were as follows:

- commonly used friction models such as the Coulomb friction model, viscous friction model and Stribeck friction model;
- combined friction models such as the Coulomb and viscous friction model, Coulomb and tanh friction model, Stribeck and viscous friction model and Stribeck and tanh friction model;
- friction models for small displacements such as the Dankowicz model and models often used in control engineering such as the Dahl model and Canudas de Wit model;
- a friction model that takes into account the random nature of friction.

4.5 Wear simulation

Predicting the amount of wear is generally thought to be rather difficult and uncertain. This chapter addresses this difficult task, outlining some

possibilities for predicting wear in rolling and sliding contacts, i.e. common wheel–rail contacts. Wheel–rail contacts are often dry, i.e. not lubricated, but they are, however, often contaminated with different materials. If the rolling and sliding contacts are running under boundary or mixed conditions, the wear of the contact surfaces is often relatively low. If the surfaces are contaminated with particles, however, wear may be rather extensive. Different environmental contaminations may both reduce and increase friction and wear, but they always have a strong influence on both friction and wear.

In a rolling and sliding contact, the two interacting surfaces characteristically move at different speeds in a tangential direction. The Tribology Group at KTH Machine Design has worked on simulating friction and wear in rolling and sliding contacts since 1987. The modelling principles the Group has successfully used are based on (i) the single-point observation method (see above) and (ii) treating wear as an initial-value process. Two simple examples will be presented below, demonstrating how these principles can be used.

Wear can be defined as the removal of material from solid surfaces by mechanical action. The amount of material removed is often quite small for mainly rolling contacts but will increase with the degree of sliding. Wear can appear in many ways, depending on the materials of which the interacting contact surfaces are made, the operating environment and the running conditions. In engineering terms, wear is often classified as either mild or severe. *Mild* wear is what engineers strive for by creating contact surfaces of appropriate form and topography. Choosing adequate materials and surface treatments is also necessary in order to obtain mild conditions. Often, however, the surface lubrication is the most important factor in ensuring that mild wear conditions are obtained. Sometimes, *severe* wear may occur, producing rough or scored surfaces.

Wear can also be classified in terms of the fundamental wear mechanisms involved, the wear mechanisms described in the literature being adhesive wear, abrasive wear, corrosive wear and surface fatigue wear.

- **Adhesive wear** is a type of wear that occurs due to adhesive interactions between rubbing surfaces. Such wear is also referred to as scuffing, scoring, seizure and galling due to the appearance of the worn surfaces. Adhesive wear is often associated with severe wear, but is probably also a mechanism involved in mild wear.
- **Abrasive wear** occurs when a hard surface or hard particles plough a series of grooves in a softer surface. The wear particles generated by adhesive or corrosive mechanisms are often hard and will act as abrasive particles, wearing the contact surfaces as they move through the contact.

- **Corrosive wear** occurs when the contact surfaces chemically react with the environment and form reaction layers on their surfaces, layers that will be worn off by the mechanical action of the interacting contact surfaces. The mild wear of metals is often thought to be of the corrosive type. Another corrosive type of wear is fretting, which is due to small oscillating motions in contacts. Corrosive wear generates small sometimes flake-like wear particles, which may be hard and abrasive.
- **Surface fatigue wear**, which can be found in rolling contacts, appears as pits or flakes on the contact surfaces; in such wear, the surfaces become fatigued due to repeated high contact stresses.

4.5.1 Classic wear models

Surfaces in rolling and sliding contact may wear if they rub against each other and are not completely separated by a clean oil film; they may also wear if the oil film separating them contains abrasive particles. The amount of wear is dependent on the properties of the surfaces, surface topography and lubrication and running conditions. The wear models that have been formulated often describe sliding contacts. The best-known such wear model is:

$$\frac{V}{s} = K \cdot \frac{F_N}{H} \qquad [4.22]$$

where V is the wear volume, s is the sliding distance, K is the dimensionless wear coefficient, H is the hardness of the softer contact surface and F_N is the normal load. This model is often referred to as Archard's Wear Law,[23] even though the basic form of the model was first described by Holm.[24] The wear coefficient, K, however, is interpreted differently by Holm and by Archard.

By dividing both sides of Eq. (4.1) by the apparent contact area, A, and by replacing K/H with a dimensional wear coefficient, k, we get the following, often used wear model:

$$\frac{h}{s} = k \cdot p \qquad [4.23]$$

where h is the wear depth and p is the contact pressure.

Some scientists have tried to analyse the validity of the wear model according to Eqs. (4.22) and (4.23), and one result of this is the wear map mentioned in Section 4.3. In the wear map of Lim and Ashby[5] two wear mechanisms, namely delamination wear and mild oxidational wear, were mentioned. Both these mechanisms can be considered mild wear mechanisms, in engineering terms, and both produce thin, plate-like wear debris. The intention of delamination wear theory, as developed by Suh,[25] is to explain flake debris generation. He based his theory on the fact that there is a high

density of dislocations beneath the contact surfaces. Under sliding interactions between the contact surfaces, these dislocations form cracks that propagate parallel to the surfaces. The total wear volume was assumed to equal the sum of the wear volume of each contact surface; the basic wear model, developed by Suh,[25] is as follows:

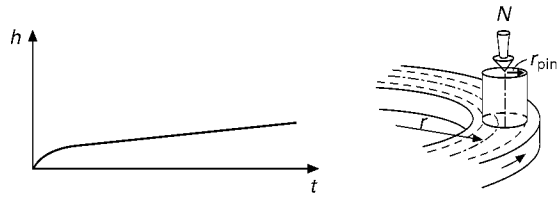
$$V = N_1 \cdot (s/s_{01}) \cdot A_1 \cdot h_1 + N_2 \cdot (s/s_{02}) \cdot A_2 \cdot h_2 \qquad [4.24]$$

where V is the wear volume, N_i is the number of wear sheets from surface i, A_i is the average area of each sheet, h_i is the thickness of the delaminated sheet, s_{0i} is the necessary sliding distance to generate sheets and s is the actual sliding distance. It is noticeable that the wear volume from each contact surface is building up the total wear volume, which was not clearly formulated before. Suh stated that a certain sliding distance is needed before a wear particle is formed. However, the sliding distance is set equal for both surfaces which indicates that he has not been aware of the single-point observation method.

Another interesting sliding wear mechanism is the oxidative wear mechanism proposed by Quinn.[26] Quinn stated that the interacting contact surfaces oxidize. The oxide layer will gradually grow until the thickness of the oxide film reaches a critical value, when it will separate from the surface as wear debris. Even in this case a certain sliding distance is needed before wear debris will be formed. Depending on whether the oxide growth is linear or parabolic, the wear is direct proportional or proportional to the power of the sliding distance. Experimental observations indicate that under steady-state mild conditions, the wear is nearly directly proportional to the sliding distance.

4.5.2 Sliding wear in a rolling and sliding contact

Wear in rolling and sliding contacts can be of different types, as mentioned above. If a surface is subject to high repeated dynamic loading, surface fatigue may occur and pits of different forms may form on the surface. In this chapter, however, we will not deal with surface fatigue; instead, we will focus our attention on sliding wear. To illustrate the wear process, a typical wear curve obtained in a pin-on-disc testing machine using a flat-ended cylindrical pin rubbing against a disc under any condition is shown in Fig. 4.7. A typical wear process always starts with a short running-in period in which the highest asperities and the contact surfaces in general are probably plastically deformed and initially worn; this is followed by a steady-state period in which the wear depth is directly proportional to the sliding distance. The initial running-in period is rather brief but not very well understood. The general appearance of a wear curve seems to apply to dry as well as boundary and mixed lubricated contacts; it also applies to contacts



4.7 A schematic wear curve from a pin-on-disc test with a flat-ended cylindrical pin.

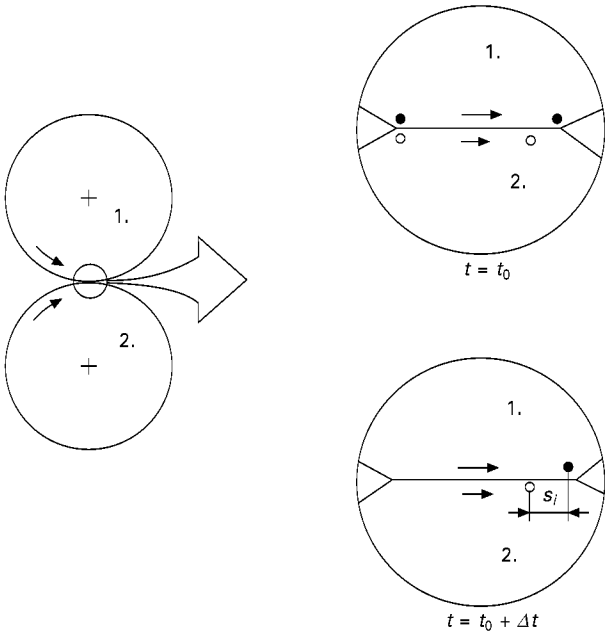
lubricated with lubricants contaminated with abrasive particles. Aside from ease of testing, the pin-on-disc configuration is a popular testing geometry because most of the wear is on the pin. The distance a point on the pin's contact surface slides against the disc is much longer than the corresponding distance a contact point on the disc slides against the pin during a single revolution of the disc.

Simple pin-on-disc test results indicate that sliding distance is an important parameter determining sliding wear. For rolling and sliding contacts, the sliding part of the surface interactions, although not obvious, is therefore of interest. Some people maintain that the effect of sliding is negligible in most rolling and sliding contacts. Various investigations have demonstrated, however, that the distances the contacts slide against the opposite interacting surfaces during a mesh are sufficient to form wear debris in most rolling and sliding contacts. For this reason, we will show how much a point on a contact surface slides against an opposite contact surface during a mesh.

Consider two discs that are pressed together and run at different peripheral velocities (see Fig. 4.8). This is a typical situation in tractive rolling contacts. The absolute value of the sliding distance, s_i, is for $i = 1$ a point on the contact surface of body 1 and for $i = 2$ a point on the contact surface of body 2. The sliding distance, s_i, during one mesh at a point on one of the contact surfaces sliding against the opposite interacting surface is equal to:

$$s_i = 2a \cdot \left| \frac{v_1 - v_2}{v_i} \right| \qquad [4.25]$$

where a is the half width of the contact, v_1 is the peripheral velocity of surface 1 and v_2 is the peripheral velocity of surface 2. The sliding distances in rolling and sliding contact according to Eq. (4.25) apply to rollers. For contacts between other bodies, such as gears and railway wheels and rails, determining the sliding distances may be more complicated. The principle, however, is the same, namely, to study how great a distance a point on a contact surface slides against the opposite surface during a single mesh. In the examples shown, the elastic deformations of the contact surfaces in the tangential direction are ignored; those displacements would reduce the



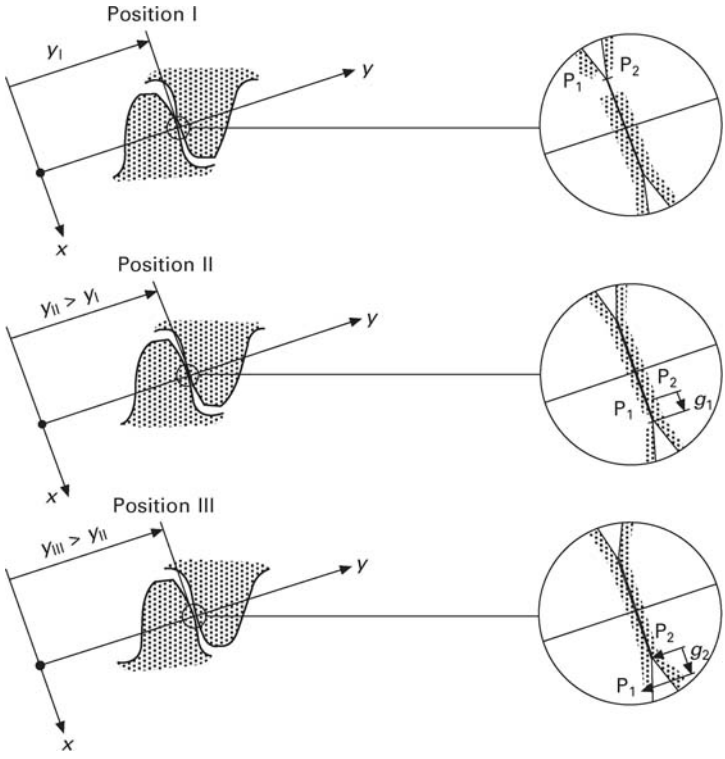
4.8 The basic principle for determining of the sliding distance in a rolling and sliding contact between two discs.

sliding distance a little, but micro-displacements normally have very little effect on the contact conditions in wheel–rail contacts.

4.5.3 The application of the single-point observation method in wear simulation

The single-point observation method was initially found to be very useful during our work on simulating friction and wear of boundary-lubricated spur gears[3] as mentioned above. The distances a point on a gear flank slides against an opposite flank during one mesh vary depending on the position on the flank, the gear ratio, the size of the gears and the loads applied on the gear tooth flanks. The principle for determining these sliding distances is shown in Fig. 4.9. The sliding distance is there referred to as g, although s is used to refer to sliding distance elsewhere in this chapter.

Test results obtained indicate that the amount of wear on the gear flanks seems to be in line with the sliding distances recorded. That observation and many years of pin-on-disc tests have inspired the author and others to try to simulate wear in rolling and sliding contacts. Our first such effort was a simulation of the mild wear of tooth flanks working under boundary-lubricated conditions.[3] The wear simulation was then based on the following wear model:



4.9 The distance, g_1, point P_1 on the pinion flank and the distance, g_2, point P_2 on the gear flank slide during one mesh; position I corresponds to the moment in time when P_1 and P_2 come into contact with each other, while positions II and III correspond to the moments in time when P_1 and P_2 disengage, respectively.[24]

$$\frac{h}{s} = k \cdot p \qquad [4.26]$$

The simulation was simplified by assuming that the wear coefficient was constant throughout the process. The initial running-in period was not considered. The contact pressure between the flanks was assumed to be constant, i.e. the mean contact pressure was determined and used. By means of these simplifications and the sliding distances determined according to derived equations, it was possible to simulate the wear depth at a particular point on a gear flank (the wear simulation was run in a simple spreadsheet). The wear distribution and estimated wear coefficient were found to be in reasonably good agreement with the experimental observations from tests conducted. Our awareness of the risk that the principle used and the simplifications made might only be relevant to the studied case, however,

motivated us to continue our research into simulations of wear in rolling and sliding contacts. Further studies were therefore successfully conducted to find out how generally applicable the principle and the simplifications are.

4.5.4 Wear as an initial-value process

A wear process is seldom a steady-state process, even if steady-state conditions are desirable and often comprise a dominant part of the wear process. Normally, the running-in wear is greater than the ensuing wear. The forms of the contact surfaces are often such that the wear depth will vary with time. Furthermore, it was found that in slow-running thrust roller bearings, mild wear of the contact surfaces caused geometrical changes that initiated other wear processes. Olofsson[27] found that mild wear of the contact surfaces of spherical thrust roller bearings caused an increase in the contact pressure at the pure rolling points.[27] Consequently, the increased contact pressure initiated surface fatigue wear at the pure rolling points much earlier than expected.

As a direct result of that finding and because wear simulations often contain many simplifications, we looked at wear simulations from a mathematical–numerical point of view. We found that simulations of wear processes can advantageously be regarded as initial-value problems.[28] We know the initial conditions and properties of the contacts fairly well, so if we can formulate how the surfaces change, it should also be possible to predict the states of the surfaces at any time during operation. The wear of contact surfaces can thus be treated as an initial-value problem. The wear rate may then be formulated according to the following model:

$$\frac{\mathrm{d}h}{\mathrm{d}t} = f \begin{array}{l}(\text{material, topography, lubricant, load,} \\ \text{velocity, temperature, } \ldots)\end{array} \qquad [4.27]$$

where h is the wear depth at a particular point on an interacting surface and t is time. This formulation is in agreement with that of the dynamic behaviour of mechanical systems and can easily be numerically integrated. A model often used in many wear simulations is

$$\frac{\mathrm{d}h}{\mathrm{d}t} = k \cdot p \cdot v_{\mathrm{s}} \qquad [4.28]$$

where v_{s} is the sliding velocity.

The wear model in Eq. (4.28) may be regarded as a generalization of Archard's wear law, i.e. Eqs (4.22) and (4.23) above.

Equation (4.28) is often reformulated thus:

$$\frac{\mathrm{d}h}{\mathrm{d}s} = k \cdot p \qquad [4.29]$$

since

$$ds = v_s \cdot dt \qquad [4.30]$$

4.5.5 Numerical integration of a wear model in a rolling and sliding contact

A commonly used wear model is Eq. (4.28). Numerically integrating a wear model entails making geometry and time discrete. The simplest numerical integration method is the Euler method. The wear depth at a chosen point on a gear flank or a roller, for example, is then determined by:

$$h_{i,\text{new}} = h_{i,\text{old}} + k_i \cdot p_i \cdot |v_1 - v_2| \cdot \Delta t \qquad [4.31]$$

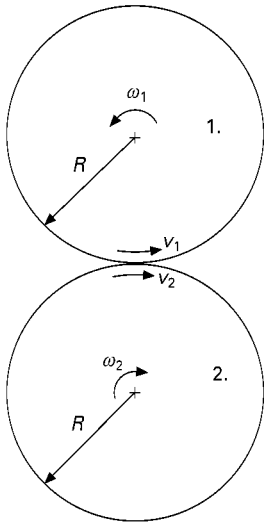
where $h_{i,\text{new}}$ is the obtained wear depth on surface i, $h_{i,\text{old}}$ is the wear depth on i in the simulation loop before the actual loop, k_i is the dimensional wear coefficient multiplied with the number of meshes or revolutions before geometry is changed, p_i is the local pressure at i when the actual time step starts and Δt is the time step. The Euler integration method is the simplest numerical integration method. Other numerical integration methods can of course be used in the same way, as different schemes are used in behavioural simulations of technical systems. After a simulation, one must always check the accuracy of the simulation. Common tests for doing so are the k and Δt checks although, if the values chosen for these are too great, the results may not be correct. A common way to handle this is to see whether the same results are obtained using half the values of k and Δt. One of the most difficult and time-consuming parts of a simulation is determining the pressure at a particular point in each simulation loop, pressure at any point being dependent on the pressure at all other points in the contact.

4.5.6 Determining the wear of interacting rollers

We consider two cylindrical rollers both of radius R (see Fig. 4.10). The rollers are pressed together with force F_N and rotated at angular velocities ω_1 and ω_2, respectively. The peripheral velocities of the contact surfaces are $v_1 = \omega_1 \cdot R$ and $v_2 = \omega_2 \cdot R$. The wear of the contact surfaces is assumed to be properly described by the following wear model:

$$\frac{dh_i}{dt} = k_i \cdot p \cdot v_{s,i} \qquad [4.32]$$

where $i = 1$ for roller 1 and $i = 2$ for roller 2; h_i is the wear depth at a point on surface i when it rubs against the opposite contact surface, k_i is the wear coefficient for a point on surface i when it rubs against the opposite contact surface, p is the local contact pressure and $v_{s,i}$ is the sliding velocity at a



4.10 Interacting cylindrical rollers.

point on surface i sliding against the opposite interacting surface. The sliding velocity, $v_{s,i}$ for points on both contact surfaces equals:

$$v_{s,i} = |v_1 - v_2| \qquad [4.33]$$

We assume that the rollers are subject to a constant load and that the angular velocities are constant. The wear model will then have the following form after integration:

$$\int_0^{h_i} \mathrm{d}h_i = k_i \cdot |v_1 - v_2| \int_0^{\Delta t} p\,\mathrm{d}t \qquad [4.34]$$

with

$$h_i = k_i \cdot |v_1 - v_2| \cdot \int_0^{\Delta t} p\,\mathrm{d}t \qquad [4.35]$$

If we study complete meshes, the contact pressure, p, can be replaced by the mean contact pressure, p_m. The wear depth is small compared with the radius of the rollers; p_m can thus be determined once and used for all simulated revolutions. Equation (4.35) can thus be reformulated according to the following:

$$h_{i,\mathrm{new}} - h_{i,\mathrm{old}} = k_i \cdot p_m \cdot |v_1 - v_2| \cdot \Delta t \qquad [4.36]$$

If Δt is brief, so that only one point on each of the contact surfaces has passed the contact once, then the wear of each surface per mesh will be:

$$h_{1/\text{mesh}} = k_1 \cdot p_m \cdot \frac{|v_1 - v_2|}{v_1} \cdot 2a \qquad [4.37a]$$

$$h_{2/\text{mesh}} = k_2 \cdot p_m \cdot \frac{|v_1 - v_2|}{v_2} \cdot 2a \qquad [4.37b]$$

respectively. Surface 1, however, is moving faster than surface 2. In the long run, points on surface 1 will be in contact more often than points on surface 2. Consequently, the wear of the two surfaces will only differ in relation to the wear coefficients. This can be demonstrated by the following relationships: Assume that the mechanism has been running for a fairly long time and that roller 1 has rotated n_1 revolutions. Roller 2 has then rotated $n_2 = n_1 \cdot (\omega_2/\omega_1)$ revolutions. The wear of the rollers will then be as follows:

$$h_{1/\text{longtime}} = k_1 \cdot p_m \cdot \frac{|v_1 - v_2|}{v_1} \cdot 2a \cdot n_1 \qquad [4.38a]$$

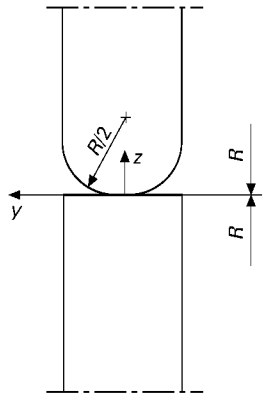
and

$$h_{2/\text{longtime}} = k_2 \cdot p_m \cdot \frac{|v_1 - v_2|}{v_2} \cdot 2a \cdot n_2 = k_2 \cdot p_m \cdot \frac{|v_1 - v_2|}{v_2}$$

$$2a \cdot \frac{\omega_2}{\omega_1} \cdot n_1 = k_2 \cdot p_m \cdot \frac{|v_1 - v_2|}{v_1} \cdot 2a \cdot n_1 \qquad [4.38b]$$

since $v_1 = \omega_1 \cdot R$ and $v_2 = \omega_2 \cdot R$.

From an experimental point of view, it is advantageous to change the form of roller 1 so that the contact surface will have a radius of $R/2$ perpendicular to the direction of motion of the contact surface (see Fig. 4.11). The contact will then be a point contact instead of a line contact as in the previous example.



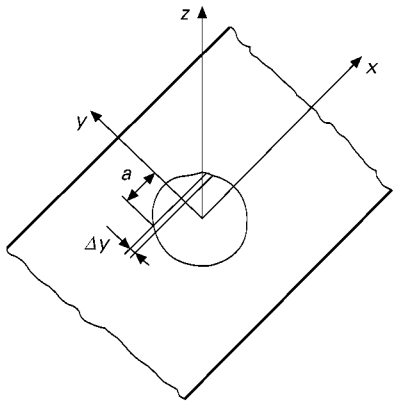
4.11 Interacting modified roller and cylindrical roller.

This change will improve the experimental set-up, but unfortunately make the wear simulation more difficult. The assumption that the wear coefficients for points on each surface are constant throughout the whole process, however, is relevant even in this case. The sliding velocity can also be assumed to be constant and the contact pressure variation at a point on a surface during a mesh can, as in the previous example, be assumed to be replaced by a mean pressure. An important difference, however, is that the mean pressure does not remain constant throughout the whole wear process, since the wear of the contact surfaces will change the pressure distribution in the contact. We assume that the same wear model as in the previous example is valid in this case as well, and that the developed equation, after considering the simplifications, will be as follows:

$$h_{i,\text{new}} - h_{i,\text{old}} = k_i \cdot p_m \cdot |v_1 \cdot v_2| \cdot (t_{\text{new}} - t_{\text{old}}) \qquad [4.39]$$

When the contact surfaces wear, the forms of the surfaces will change and thereby also the pressure distribution in the contact. This means that we cannot assume that the pressure is constant, so we cannot, as in the previous example, run a large number of revolutions in one simulation loop. The main problems in this case, therefore, are determining the local pressure at an observed point in every simulation loop and deciding on the duration of each loop before a new local pressure determination must be made.

In this case, we do not have a standard Hertzian contact case, so we make the contact width b equal to the half axle of a sphere of radius $R/2$ against a plane. A Winkler surface model of rod stiffness K_N has been used to simulate the wear process of a modified roller interacting with a cylindrical roller. The contact surfaces are then divided into a number of sliced surfaces of width $\Delta y = b/10$ perpendicular to the sliding direction (see Fig. 4.12). We assume that we can simplify the wear simulation by determining the wear



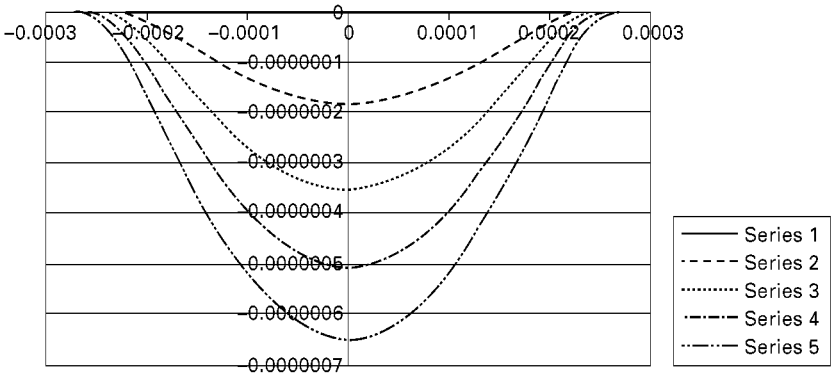
4.12 Contact point and co-ordinate system–sliced contact.

for each slice in the same way as above. The penetration, d, of the modified upper roller against the lower cylindrical roller is determined so that the sum of the load of each slice support equals the applied force, F_N. The local wear is now determined and the geometry of the contact surfaces is modified. Thereafter, a new penetration, d, is determined, and so on. Figure 4.13 presents some simulation results for a modified upper roller interacting with a cylindrical roller. The rollers will wear during running, The wear of the discs will increase in both depth and width with time (see Fig. 4.13).

4.5.7 Determination of the pressure distribution

Determining the contact pressure at a particular point is often the trickiest and most time-consuming part of a wear simulation. The deformation at one point is dependent on the deformation at all other points around the observed point. This implies a rather complex process for accurately calculating the pressure distribution. Today, three different ways are commonly used to determine the contact pressure.

Finite element (FE) calculation is one method that is becoming increasingly popular to use as computer power increases and FE programs are improved. The main drawback of the FE method is that determining the pressure distribution often entails considering a great many small elements on the surfaces. That is often difficult to do, since the combination with the body models often leads to a huge number of elements and a very long calculation time. The FE method will probably be used more in the future for interface-related problems than it is today.



4.13 Simulated wear of an ellipsoidal roller 1 interacting with a cylindrical roller 2. $R_{1x} = 25.10^{-3}$ m, $R_{1y} = 125.10^{-3}$ m, $R_{2x} = 25.10^{-3}$ m, $F_N = 100$ N, $v_1 = 1.25$ m/s, $v_1 - v_2 = 6E\text{-}2$ m/s. Series 1 is at $t = 0$, Series 2 after n revolutions, Series 3 after $2n$ revolutions, Series 4 after $3n$ revolutions and series 5 after $4n$ revolutions. The figure shows the form of changes of the contact surface of the rollers during running.

To determine the micro-topography in the contact zone, boundary element (BE) methods are commonly used. The BE programs used are often based on the same assumptions as Hertz used when he derived his equations. This means that most BE programs cannot be used in all applications. The BE method ends up in a numerical process which is solved in different ways in order to obtain a reasonably accurate result as fast as possible.[29–31] Some smart combinations of BE and FE methods will probably be used in future.

A common way to simplify the determination of local pressure is to use a Winkler surface model. The surfaces are then replaced by a set of elastic bars; the shear between the bars is neglected and the contact pressure at a point is then dependent only on the deformation at that point according to:

$$p = K_N \cdot u_z \qquad [4.40]$$

where u_z is the deformation of the elastic rod. The spring constant, K_N, can be determined by:

$$K_N = C_w \cdot \frac{E'}{b} \qquad [4.41]$$

where $C_w \approx 1$, E^* is the combined elastic modulus of the contact surfaces and b is approximately the width of the elastic half axle according to Hertz. An extention of the the Winkler model is the Brush model, which also considers the tangential deformations. The Brush model is often used for friction simulations of complex contacts.

4.5.8 Concluding remarks regarding wear simulations

The wear of rolling and sliding contacts can be simulated. The basic principles we recommend are (i) the single-point observation method and (ii) treating the wear process as an initial-value problem. Using these principles, nearly any practical case can be simulated if you have a relevant wear model that can imitate the behaviour in a particular case. The most common wear model is the so-called Archard's generalized wear model:

$$\frac{dh}{dt} = k \cdot p \cdot v_s \qquad [4.42]$$

How well that model describes the wear process has yet to be investigated in detail in some cases. In many cases, however, the simulated wear distributions agree fairly well with experimental observations. Wear simulations are quite often performed stepwise, with repeated determinations of pressure, sliding velocities, etc. Determining the pressure distribution in the contact is often considered the most difficult and time-consuming task. Most simulations are done numerically, and choosing the appropriate surface element size and time step is critical. Too long a time step may produce incorrect results or

an unstable simulation, while too short a time step, on the other hand, may result in excessively time-consuming calculations.

4.6 References

1. Andersson S., Söderberg A. and Björklund S. (2007) Friction models for sliding dry, boundary and mixed lubricated contacts, *Tribology International*, **40**, 580–87.
2. Meng H.-C. (1994) *Wear Modeling: Evaluation and Categorisation of Wear*, Dissertation, University of Michigan, Ann Arbor, MI, USA.
3. Andersson S. and Eriksson B. (1990) Prediction of the sliding wear of spur gears, *Nordtrib'90*, Hirtshals, Denmark, 10–13 June.
4. Andersson S. (1975) *Partial EHD Theory and Initial Wear of Gears*, Doctoral Thesis. KTH, Stockholm, Sweden.
5. Lim S.C. and Ashby M.F. (1987) Wear mechanism maps, *Acta Metallica*, **35**(1), 1–24.
6. Lewis R. and Olofsson U. (2004) Mapping rail wear regimes and transitions, *Wear*, **257**, 721–9.
7. Sundh J. (2007) *An Experimental Study on Wear Transitions in the Wheel-Rail Contact*, Licentiate Thesis, TRITA-MMK 2007:02, RTH, Stockholm, Sweden.
8. Begelinger A. and deGee A.W.J. (1981) Failure of thin film lubrication. *ASME Journal of Lubrication Technology*, **103**, 203–8.
9. Andersson S. and Salas-Russo E. (1994) The influence of surface roughness and oil viscosity on the transition in mixed lubricated sliding contacts, *Wear*, **174**, 71–9.
10. Sellgren U. (1999) *Simulation Driven Design: Motives, Means and Opportunities*, Doctoral thesis, KTH, TRITA-MMK 1999:26.
11. Sellgren U. (2003) Architecting models of technical systems for non-routine simulations, *Proceedings International Conference on Engineering Design, ICED 03*, Stockholm, Sweden, 19–21 August.
12. Andersson S. and Sellgren U. (2004) Representation and use of functional surfaces, *7th Workshop on Product Structuring–Product Platform Development*, Chalmers University of Technology, Gothenburg, Sweden 24-25 March.
13. Stribeck R. (1902) Die Wesentlichen Eigenshaften der Gleit- und Rollenlager, *Z* Verin. Dtsch. Ing., **45**(36) 1341–8, 1432–8, 1463–70.
14. Hagman L.A. (1997) *Measurement and Modeling of Microslip for Engineering Surfaces in Contact*, Doctoral Thesis, KTH, stockholm, Sweden TRITA-MMK 1997:11.
15. Olofsson U. and Hagman L. (1997) A model for micro-slip between flat surfaces based on deformation of elliptical elastic bodies, *Tribology International*, **30**(8), 599–603.
16. Björklund S. (1997) A random model for micro-slip between nominally flat surfaces, *Transactions ASME, Journal of Tribology*, **119**, 726–32.
17. Sellgren U. and Olofsson U. (1999) Application of a constitutive model for micro-slip in finite element analysis, *Computational Methods in Applied Mechanical Engineering*, **170**, 65–77.
18. Dankowicz H. (1999) Modelling of dynamic friction phenomena, *ZAMM*, **79**, 399–409.
19. Dahl P.R. (1977) Measurement of solid friction parameters of ball bearings, *6th Annual Symposium on Incremental Motion, Control System and Devices*, University of Illinois, Champaign – Urbana, IL, 24–27 May, 49–60.

20. Olsson H. (1996) *Control Systems with Friction*, Doctoral Thesis, Lund Institute of Technology, Lund, Sweden.
21. Canudas de Wit C., Olsson H, Åström K.J. and Lischinsky P. (1995) A new model for control of systems with friction, *IEE Transactions on Automatic Control*, **40**(3), 419–25.
22. Patir N. and Cheng H. S. (1978) An average flow model for determining effects of three-dimensional roughness on partial hydrodynamic lubrication, *ASME Journal of Lubrication Technology*, **100**, 12–17.
23. Archard J.F. (1980) Wear theory and mechanisms, in Peterson M.B., Winer W.O. (eds), *Wear Control Handbook*, ASME, New York, USA, 161–78.
24. Holm R. (1946) *Electric Contacts*, Almqvist & Wiksells Boktryckeri AB, Uppsala Sweden.
25. Suh N.P. (1973) The Delamination Theory of Wear, *Wear*, **25**, 111–24.
26. Quinn T.F.J. (1962) Role of oxidation in the mild wear of steel, *British Journal of Applied Physics*, **13**, 33–7.
27. Olofsson U. (1997) Characterisation of wear in boundary lubricated spherical roller thrust bearings, *Wear*, **208**, 194–203.
28. Strang G. (1986) *Introduction to Applied Mathematics*, Wellesley-Cambridge Press, Wellesley, MA, USA.
29. Björklund S. and Andersson S. (1994) A numerical method for real elastic contacts subjected to normal and tangential loading, *Wear*, **197**, 117–22.
30. Kalker J.J. (1990) Three-dimentional elastic bodies in rolling contacts, Kluwer Academic, Dordrecht, The Netherlands.
31. Telleskivi T. (2003) *Wheel Rail Interaction Analysis*, Doctoral Thesis, TRITA-MMK 2003:21, KTH, Stockholm, Sweden.

Copyrighted Materials
Copyright © 2009 Woodhead Publishing Retrieved from www.knovel.com

5
Rail materials

J. E. GARNHAM and C. L. DAVIS,
University of Birmingham, UK

Abstract: Railway rail metallurgical developments are briefly reviewed historically to the present day. The development of pearlitic rail microstructure during rail production is examined, higher strength grades are described and the work-hardening of such structures examined. Welding of these rails is briefly described together with their joining to rail switches and crossings made from austenitic steel. Also described are the changes in rail microstructure due to cyclic, compressive, rolling-sliding contact which lead to rolling contact fatigue and wear. New developments in rail microstructures are examined such as bainitic rail, high-carbon, hyper-eutectoid pearlitic rail and coated rail.

Key words: rail steel metallurgy, history, production, rolling contact fatigue and wear, welding.

5.1 Introduction

5.1.1 Historical review (after Perez-Unzueta, 1992)

The first known guides for wheeled vehicles date from the ancient Greeks (grooves into aligned stone blocks) and around the same time they developed bearings for wheel movements. Guidance systems with raised wooden rails developed during the Renaissance in Bohemia and in the UK, wagon-ways using wooden rails are reported to exist at the beginning of the 17th century. In the 18th century the use of cast iron for rails was developed, initially as plates on wooden rails (early surface engineering!) and also at this time, rails began to be shaped to facilitate vehicle steering. In 1789, W. Jessop developed cast (presumably, edge lipped) wheels to run on cast rails, with the latter having the beginnings of the rounded railhead on a flange configuration seen in modern rails.

With the development of steam locomotives from 1803, particularly George Stephenson's first passenger railway in 1825, these early pioneers found the existing cast iron rails were inadequate as they easily fractured therefore Stephenson pioneered the use of shaped wrought iron rails with a seven-pass rolling process to form them, developed by Birkenshaw in 1820. At this time, one area where cast iron rail was found superior to wrought iron rail was railhead wear resistance. This focused attention on (economically) improving wrought iron rail quality with regard to steel processing. In 1856, Henry Bessemer developed his steel-making process and this was further

improved by R. Mushet to give ingots which could be forged and rolled. In 1856 he produced the first rails from such steels for use around Derby Station, greatly improving rail life. Within the next 20 years, rail transverse profiles were rapidly developed with the eventual adaptation of the general shape used worldwide today.

5.1.2 Subsequent rail development

From this history, the railway system has been developed with the rail as a beam support for the vehicles' weight on spaced sleepers, with the wheel and rail profiled to facilitate steerage around bends and to give minimal frictional resistance to rolling movement (*cf* a bearing race) whilst still remaining mainly unlubricated, so as to facilitate adequate driving and braking traction. The railhead material is subject to compressive rolling and sliding forces, with a very small area of contact with the wheel, thus generating wear and eventual rolling contact fatigue (RCF) on both contacting surfaces. The complete rail is subject to bending moments between sleepers. Also, certain parts of the rail system, such as the 'nose' at switches and crossings, are additionally subject to impact forces.

For a given maximum contact stress and vector of maximum creepage, the resistance of the railhead material to wear and fatigue is determined by its chemical composition, steel manufacturing process and associated hot forming process(es). In some countries in the past, some rails have been separately heat-treated, but this is not common practice. Most rail for several decades, on nearly all global networks, has been manufactured from low-alloy, carbon–manganese steel with medium to near-eutectoid levels of carbon; a primary reason being the low cost of these steels. The normal structure is pearlite (a lamella structure of carbide and ferrite) with a varying degree of pro-eutectoid (PE) ferrite at the prior austenite (PA) grain boundaries; with higher carbon grades the percentage of PE ferrite will be very low and, in some recent grades, non-existent. Recently, on a small scale, other steel structures have been trialled on railway networks, such as low carbon, higher alloy, carbide-free bainitic rails, with some trials still ongoing. These will be described later.

Some past and recent rail steel specifications are shown in Table 5.1. There can be some small specification variance dependent upon the steel making process. Together with the subsequent rolling practice, this could affect the degree of homogeneity in the rail microstructure and the size and distribution of brittle and ductile inclusions. In Japan, Europe and many other countries today, most rail steel blooms are produced by continuous-casting. This gives a high degree of homogeneity to the product, although the size and distribution of non-metallic inclusions will be dependent upon the cross-sectional (transverse) area of the continuously cast bloom and the

Table 5.1a UK rail steel – selected examples of pearlitic rail steel compositions (some historic) (wt%)

Steel	C	Si	Mn	S	P	Cr	Mo	Ni	Cu	Al	N	O	V	Ti	Comment
BS11-equivalent (1922)	0.55 0.65	0.10 0.30	0.80 max	0.050 max	0.040 max										Sawley, 1989
BS11 (1985) Normal Grade	0.45 0.60	0.05 0.35	0.95 1.25	0.040 max	0.040 max	–	–	–	–	–	–	–	–	–	BSI, 1985
BS11 (1985) Wear Resisting Grade A	0.65 0.80	0.10 0.50	0.80 1.30	0.040 max	0.040 max	–	–	–	–	–	–	–	–	–	BSI, 1985
BS11 (1985) Wear Resisting Grade B	0.55 0.75	0.10 0.50	1.30 1.70	0.040 max	0.040 max	–	–	–	–	–	–	–	–	–	BSI, 1985
'1% Cr' wear-resistant grade S1100 nominal	0.75	0.70	1.00	0.040 max	0.030 max	1.00	–	–	–	–	–	–	–	–	Gladman, 1992; Hodgson, 1985
1% Cr-V wear-resistant grade S1200 nominal	0.75	0.80	1.00	0.040 max	0.030 max.	1.00	–	–	–	–	–	–	0.10	–	Gladman, 1992
EN 13674-1 (2003), grades 200 to 350 HT						Ref. Corus (www.corusrail.com/en) and Railtrack Line Specification RT/CE/S/061 (1996)*									
'200'	0.40 0.60	0.15 0.58	0.70 1.20	0.008 0.035	max. 0.035	max 0.15	–			max 0.004			max 0.03		H = 3 ppm max 200–400 HB
'220'	0.50 0.60	0.20 0.60	1.00 1.25	0.008 0.025	max 0.025	max 0.15	max 0.02	max 0.1	max 0.15	max 0.004	max *0.008	max *20 ppm	max 0.03	max *0.025	H = 3 ppm max 220–260 HB
'260'	0.60 0.82	0.15 0.58	0.65 1.25	0.008 0.025	max 0.025	max 0.15	max 0.02	max 0.1	max 0.15	max 0.004	max *0.01	max *20 ppm	max 0.03	max *0.025	H = 2.5 ppm max 260–300 HB
'260Mn'	0.55 0.75	0.15 0.60	1.30 1.70	0.008 0.025	max 0.025	max 0.15				max 0.004			max 0.03		H = 2.5 ppm max 260–300 HB
'320 Cr'	0.60 0.80	0.50 1.10	0.80 1.20	0.008 0.025	max 0.020	0.80 1.20				max 0.004			max 0.18		H = 2.5 ppm max, 320–360 HB

Table 5.1a Cont'd

Steel	C	Si	Mn	S	P	Cr	Mo	Ni	Cu	Al	N	O	V	Ti	Comment
'350HT'	0.70	0.15	0.70	0.008	max	max				max			max		H = 2.5 ppm max,
	0.82	0.58	1.20	0.025	0.020	0.10				0.004			0.03		350–390 HB
Corus head-	0.72	0.10	1.00	max.	max	0.14				max.					H = 2.5 ppm max,
hardened SHH	0.82	0.50	1.25	0.025	0.025	0.30				0.004					360–388 HB
Corus head-	0.72	0.40	0.70	max.	max	0.40				max.					H = 2.0 ppm max,
hardened MHH	0.82	0.80	1.00	0.025	0.020	0.60				0.004					350–410 HB

Note: Not all minor elements specified are shown.

Table 5.1b Non-UK rail steel – examples of pearlitic compositions (wt%)

Steel	C	Si	Mn	S	P	Cr	Mo	Ni	Cu	Al	N	O	V	Ti	Comment
Europe UIC 860-0	0.40	0.05	0.80	0.050	0.050	–	–	–	–	–	–	–	–	–	UIC, 1986
(1986) 700	0.60	0.35	1.25	max	max										
900A	0.60	0.10	0.80	0.040	0.040	–	–	–	–	–	–	–	–	–	UIC, 1986
	0.80	0.50	1.30	max	max										
900B	0.55	0.10	1.30	0.040	0.040	–	–	–	–	–	–	–	–	–	UIC, 1986
	0.75	0.50	1.70	max	max										
1100	0.60	0.30	0.80	0.030	0.030	0.80	–	–	–	–	–	–	–	–	UIC, 1986
	0.82	0.90	1.30	max	max	1.30									
North America	0.67	0.10	0.70	0.035	0.037	0.20	–	–	–	–	–	–	–	–	Gladman, 1992
AREA 90/114 (1983)	0.80	0.50	1.00	max	max	max									
AREA improved (1983)	0.74	0.25	0.90	0.030	0.030	0.20									Gladman, 1992
	0.82	0.50	1.25	max	max	max									
Japan															
JIS E1101 (1990);	0.55	0.10	0.60	0.050	0.045	–	–	–	–	–	–	–	–	–	
37N	0.70	0.35	0.95	max	max										
JIS E1101	0.63	0.15	0.70	0.025	0.030	–	–	–	–	–	–	–	–	–	
(40N, 50N, 60N)	0.75	0.30	1.10	max	max										
Nippon high carbon steels, comparison trials ↓ ref (Uchino *et al.*, 1998)															
0.8% C heat-treated	0.80	0.55	1.07	0.010	0.020	0.17	–	–	–	–	–	–	–	–	
0.9% C hyper-eutectoid	0.90	0.48	0.60	0.010	0.013	0.19									
0.9% C hyper-eutectoid with boron	0.90	0.50	0.60	0.011	0.013	0.18	–	–	–	–	–	–	–	–	+ 0.0029 Boron
1.0% C hyper-eutectoid	1.00	0.50	0.74	–	–	0.21	–	–	–	–	–	–	–	–	Iwano *et al.*, 2006

Note: Not all minor elements specified are shown.

Table 5.1c UK austenitic grades for rail crossings

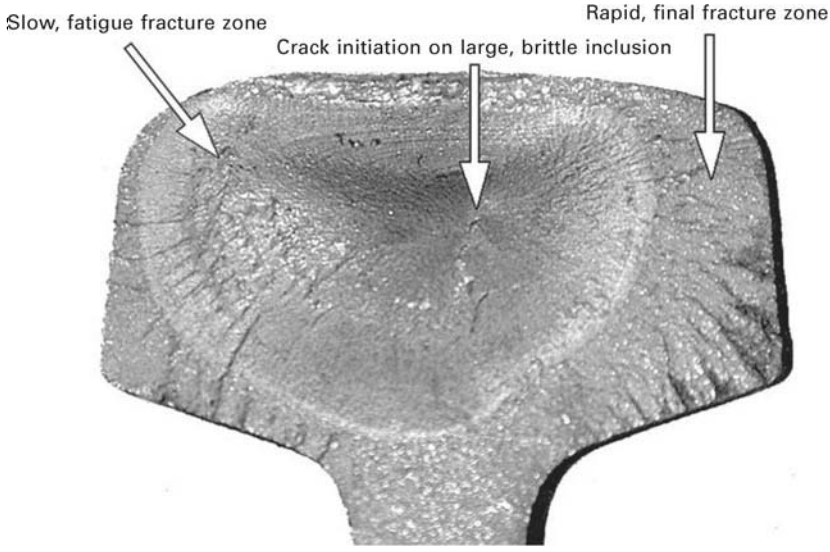
Steel	C	Si	Mn	S P	Cr	Mo	Ni	Cu	Al	N	O	V	Ti	Comment
Hadfield	0.90 1.20	0.40 max	11.0 14.0	Both 0.050 max. but target as low as possible.	–	–	–	–	–	–	–	–	–	Sawley, 1989; Hodgson, 1985
Low-carbon grade	0.70 0.90	0.40 max.	12.0 16.0		–	–	–	–	–	–	–	–	–	Sawley, 1989, Hodgson, 1985

subsequent rolling reduction required to give the cross-sectional area of the specified rail profile (Boulanger, 2003).

5.2 Pearlitic rails

5.2.1 Rail production

The basic method of rail steel bloom production in France and the UK can be found in Boulanger (2003). The stages are: (i) liquid steel production by blast furnace and then basic oxygen converter process or by electric arc remelting; (ii) secondary steel-making where the melt is further refined by vacuum degassing to bring non-metallic inclusions levels down to specified levels and where alloy additions are made; (iii) continuously casting (typically with a 330 by 254 mm cross-section) and cutting into blooms. Blooms can then be slow cooled from 600 °C in sealed boxes to minimise hydrogen entrapment levels (Llewellyn, 1992). Such methods produce close uniformity of chemical composition throughout and across each rail length with associated uniformity of mechanical properties (British Steel Corporation, 1985). With the introduction of the rail steel manufacturing routes outlined above, the size, volume and dispersion of deleterious oxide and silica-based, brittle inclusions have been greatly reduced; consequently, sudden rail fracture due to subsurface initiated fatigue failure in the body of a rail (tache ovale type failures – Fig. 5.1) rarely occur nowadays. With recent specifications



5.1 Typical 'tache ovale' type fatigue fracture found in old rail. (Adapted from an image, courtesy of H.C. Eden)

(Table 5.1), phosphorus and sulphur levels have been reduced, hence also a reduction in the volume of ductile MnS-based inclusions, and additionally nitrogen and oxygen levels are now more tightly controlled and specified (Table 5.1).

After steel manufacture, blooms are transported to rail-rolling facilities, either on-site or elsewhere, for rail rolling. In the recent Corus facility, rail is manufactured by reheating the blooms and rolling through seven rolling stands via shaped rolls (www.corusrail.com) followed by air-cooling and straightening. Methods of straightening are optimised via bending reversals (Llewellyn, 1992). Initial runs of the product have 'qualifying tests' for fracture toughness; fatigue crack growth rate, fatigue life, residual stress, variation of centreline running surface hardness (of heat-treated rails), tensile strength and elongation, segregation. Production runs have 'acceptance tests' for chemical composition, microstructure, decarburisation, oxide content and distribution, sulphide content and distribution, hardness, mechanical properties, dimension tolerances (straightness, surface flatness and twist), any internal quality and surface breaking defects (Boulanger, 2003) via metallurgical, mechanical, ultrasonic, eddy current and surface laser inspection techniques.

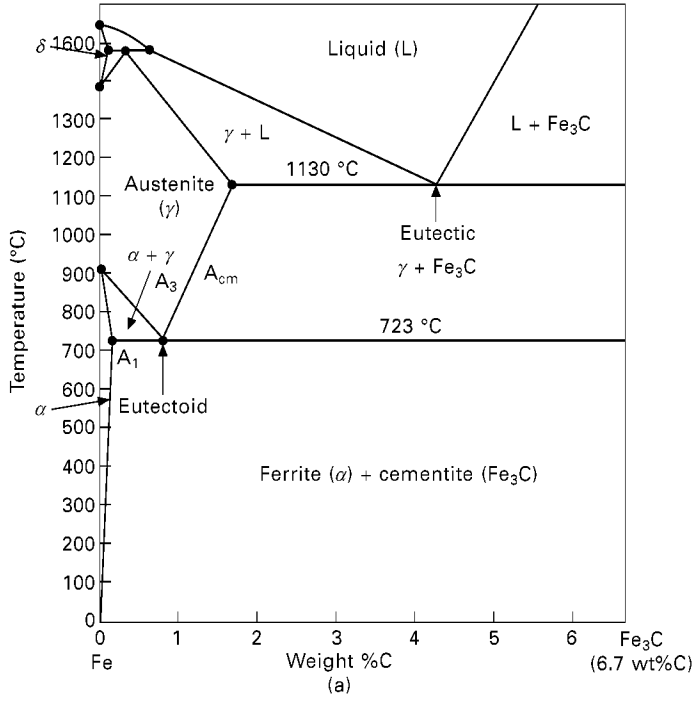
5.2.2 Pearlitic plain carbon and low-alloy structures developed from air-cooling

The structure of steel can vary greatly dependent upon alloy composition and time–temperature history (hot forming, cold forming and further heat-treatment, if required). A guide to some of this is given by the iron–carbon phase diagram (Fig. 5.2). Most railway rail steels have a carbon content between 0.5 wt% and that of the eutectoid composition (~0.8 wt%). From Fig. 5.2 it can be seen that during cooling or reheating, between around 1200 °C down to the A_3 temperature, the steel is in the austenite phase. This has a faced-centred cubic atomic structure with high carbon solubility. On cooling below the eutectoid (A_1) transformation temperature, except at highly accelerated cooling rates, the atomic structure changes to body-centred cubic which has lower carbon solubility and consequently carbon comes out of solution in various ways, dependent upon cooling rate. For plain and low-alloy carbon steels, normal air-cooling results in carbon coming out of solution by a diffusional process to separate in the form of iron–carbide (Fe_3C)-based lamellae and eutectoid ferrite lamellae, a structure termed 'pearlite' due to its 'pearl-like' visual appearance (Fig. 5.3).

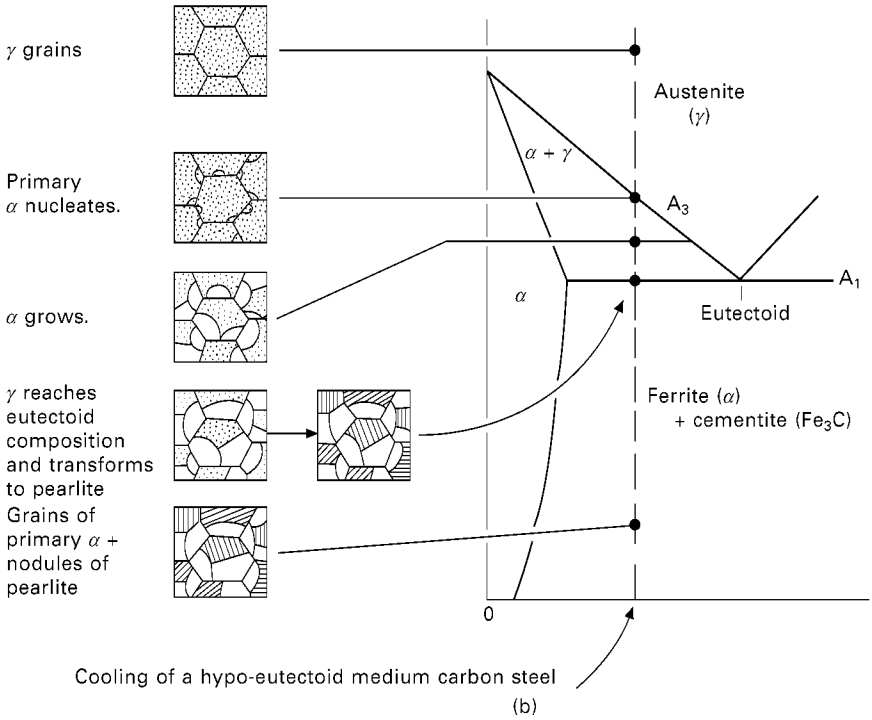
As shown in Fig. 3.2b, at carbon contents less than eutectoid composition, some PE ferrite will transform from the austenite first, the lower the carbon content, the higher the amount of PE ferrite that forms. Most pearlitic rail steels are medium (~ 0.5 wt%) to high/eutectoid (~ 0.8 wt%) carbon steels with small alloying additions of manganese and silicon (Table 5.1). The small

amount of PE ferrite in most of these steels forms at first on the austenite grain boundary junctions and then along these boundaries, thus defining the PA grain size after transformation on cooling. For high-carbon 'hyper-eutectoid' steels cooled in a similar fashion, with carbon contents in excess of the eutectoid content, brittle cementite nucleates first along the austenite grain boundaries, rather than ferrite.

With standard rails, reheat temperatures, rolling temperatures and cross-section reductions at each rolling stage are controlled to give the resultant specified grain size. With head-hardened rails, the railhead can be control cooled straight from a rolling stand or rails can be reheated and normalised, and then control cooled. Normalising involves reheating to the lower temperature region of the austenite phase to minimise austenite grain-growth and then controlled cooling. Restriction of grain-growth in the austenite phase can also be achieved by minor additions of alloying elements as these form a stable dispersion of oxide/nitride-based particles which restrict grain boundary movements (Honeycombe and Bhadeshia, 1995). In some alloys



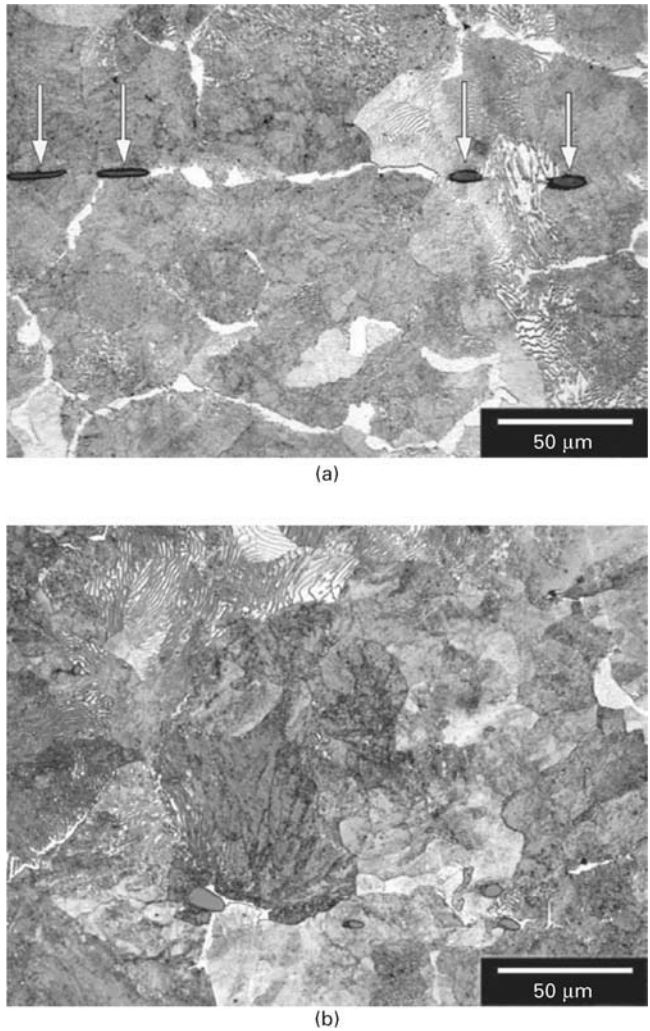
5.2 Iron–carbon phase diagram (adapted from Ashby and Jones, 1998). (a) Complete diagram covering steel and iron. (b) Detail of diagram plus schematic sketch showing the development of microstructure on slow cooling of a hypo-eutectoid, medium carbon steel.



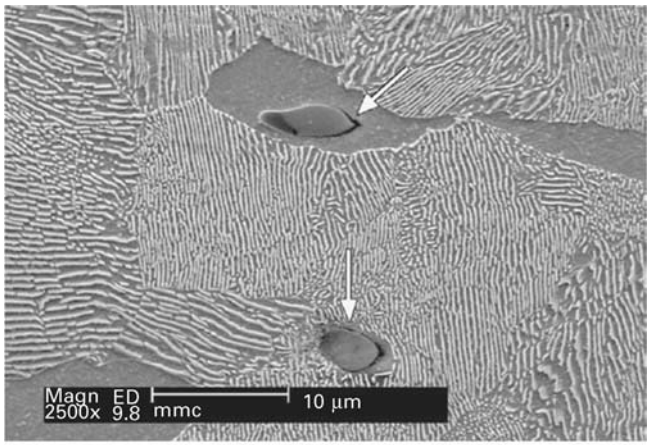
5.2 Cont'd

(including one rail grade), small additions of vanadium are also added for precipitation hardening (Table 5.1). For a given austenite grain size, the final structure is then dependent upon the cooling rate (Fig. 5.4). For low-alloy, carbon steels, very rapid quenching results in all the carbon being 'frozen' in the structure, which accommodates this by a displacive (non-diffusional) change to a body-centred tetragonal atomic lattice with a lath-type rather than a granular structure; namely martensite. Slightly slower cooling also gives displacive transformations in the form of bainite, but with time for some carbon to come out of solution, as carbide particles within the laths at a fast rate (lower bainite), between the laths at a slower rate (upper bainite) or as a mix of the two at intermediate rates. For all these displacive transformation products, further low-temperature heat-treatment brings further carbon out of solution. All these transformations are comprehensively described in Bhadeshia (2001).

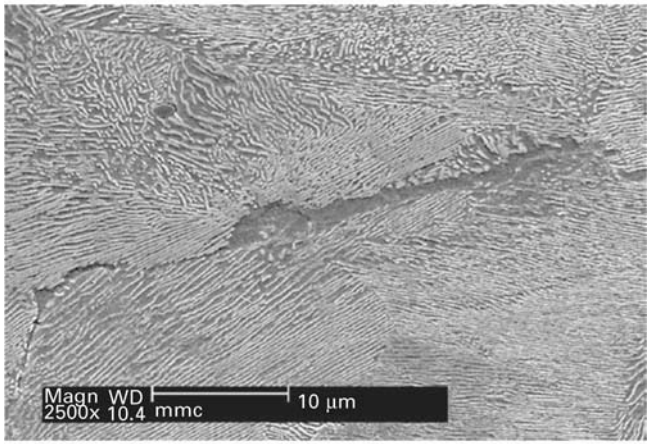
Slower than this, at air-cooling rates, pearlite forms by a diffusional transformation, however, for a given austenite grain size, the refinement of the lamellae spacing is determined by the cooling rate. Within the prior austenite grains, pearlite initiates and grows from grain boundaries and inclusions into



5.3 Railhead microstructures showing PE ferrite and pearlite. (a) 220 grade rail (0.55 wt% C). Optical image (*therefore ferrite light and carbide dark*) showing PA grain boundaries defined by PE ferrite around colonies of pearlite lamellae (individual lamella are not visible at this magnification). MnS-based inclusions can also be seen (arrowed). (b) 260 grade rail (0.81 wt% C). Optical image showing only a very small volume of PE ferrite, thus PA grain boundaries are only partially defined. (c) BS11 grade rail (0.46 wt% C). Electron microscope image (*therefore ferrite now dark and carbide light*) showing pearlitic lamellae colonies between PE ferrite at PA grain boundaries. MnS-based inclusions can also be seen (arrowed). (d) 260 grade rail (0.81 wt% C). Electron microscope image showing pearlitic lamellae colonies, plus some PE ferrite at a PA grain boundary.



(c)

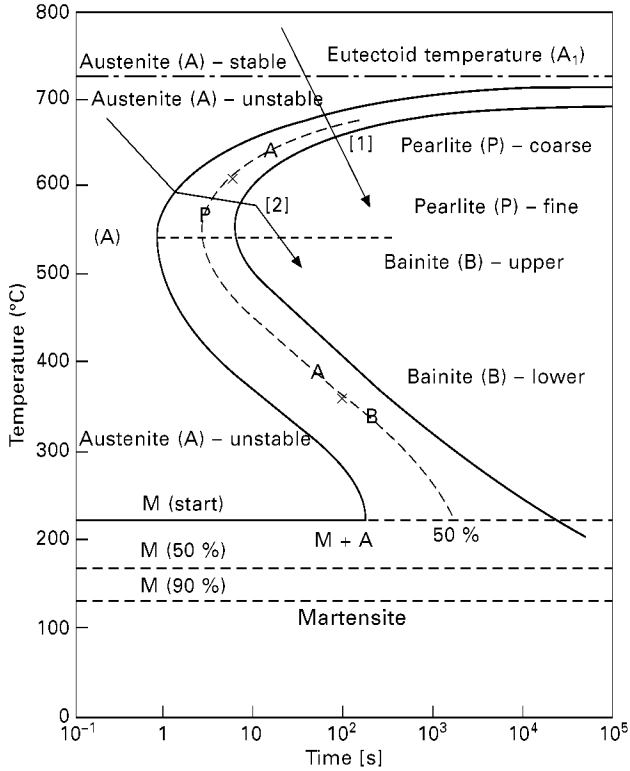


(d)

5.3 Cont'd

colonies, each with lamellae at a certain orientation (Fig. 5.2b). More rapid air-cooling progressively delays the transformation to temperatures lower than the eutectoid temperature of both PE ferrite and pearlite with the result that less PE ferrite forms than indicated by the phase diagram (for hypo-eutectoid steels) and more refined pearlite forms, however, although the rapid cooling refines the pearlitic structure (smaller inter-lamella spacing), it is also 'diluted', i.e. has less carbon content (as Fe_3C) than indicated by the phase diagram.

A large austenite grain size similarly suppresses transformation, but strengthening aspects of pearlite refinement are more than offset by the detrimental effects of the large PA grain size; at slow cooling rates, coarse ferritic transformation products may result; at faster cooling rates



5.4 Schematic representation of an isothermal '*time–temperature–transformation* (TTT)' diagram for a eutectoid composition, plain carbon steel [adapted from Callister (1994), Pickering (1992), Gladman (1992), Krauss (1980), Honeycombe and Bhadeshia (1995)]. Most rail steels have slightly lower carbon content and some alloying, but this shows the general pattern of isothermal cooling. With continuous cooling, the transformation curve is shifted right to slightly longer times. Arrow [1] represents the normal air-cooling of hot rolled and normalised rail; arrow [2] represents the accelerated and controlled cooling required for head-hardening the rail.

Widmanstätten ferrite may form and possibly bainite (Gladman, 1992). The inter-lamellar spacing of the pearlite has be shown to be inversely proportional to the degree of under-cooling below the eutectoid transformation temperature (Honeycombe and Bhadeshia, 1995) and, as can be seen developed in Eqs (5.1)–(5.3), this spacing has a significant effect on strength. This factor is used in the production of head-hardened steels, where the railhead is subject to a separate, tightly controlled cooling regime (Fig. 5.4) so as to refine the pearlitic structure in the railhead to give enhanced wear and rolling contact fatigue resistance.

The strength of pearlitic microstructures is expressed by modifications of

the Hall-Petch 'yield stress–grain size' relationship; for ferrite yield strength $\{\sigma_y(f)\}$ vs ferrite grain size (d_f) in mild steel:

$$\sigma_y(f) = \sigma_0 + k_y d_f^{-1/2} \quad [5.1]$$

where σ_0 and k_y are constants. Pickering (1992) and Gladman (1992) developed further quantitative relationships for ferrite–pearlite steels, for example, for ultimate tensile strength (UTS):

$$\text{UTS (MPa)} = 294 + 27.7(\text{wt\% Mn}) + 83.2\ (\text{wt\% Si}) + 3.85\ (\%\ \text{pearlite}) + 7.7d_f^{-1/2} \quad [5.2]$$

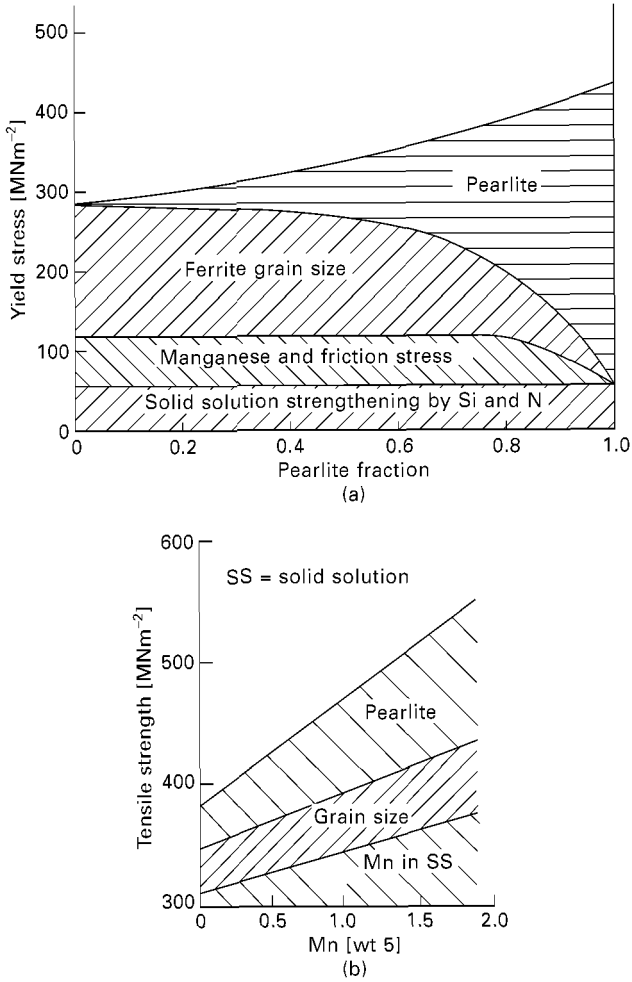
Full discussion can be found in Pickering (1992), Gladman (1992) and Llewellyn (1992). For fully pearlitic steels, Krauss (1980) showed that for the pearlite yield stress $\{\sigma_y(p)\}$, the pearlite colony size (p) has a very minor effect compared to lamellar spacing (s) and PA grain size (d_a):

$$\sigma_y(p) = 52.3 + 2.18(s^{-1/2}) - 0.40(p^{-1/2}) - 2.88(d_a^{-1/2}) \quad [5.3]$$

The influences of some of the factors described above on strength, along with the effect of solid solution strengthening (from small additions of inexpensive alloying elements) are shown schematically in Fig. 5.5. The strongest influence is carbon content (i.e. pearlite volume fraction). Manganese additions result in the eutectoid composition being at lower carbon contents, thereby increasing the proportion of pearlite in the matrix.

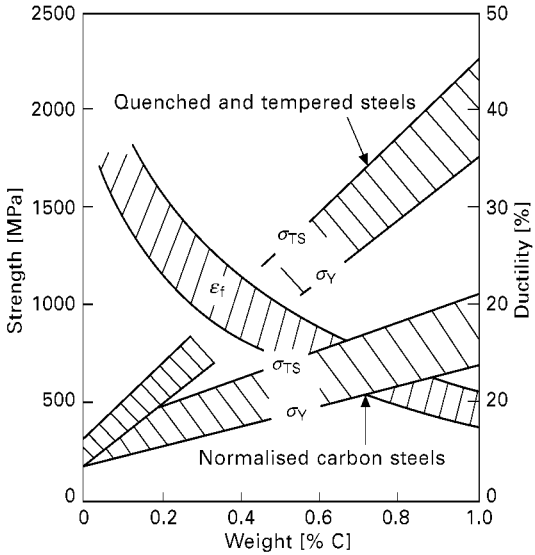
Figure 5.6 also shows the effect of carbon content on tensile strength and true strain (at fracture) and, for comparison, the relationship of carbon content of the steel in the quenched and tempered condition is also shown. The comparative mechanical properties shown indicate that this latter condition would be preferred for rail and in a few cases it has been used, but generally this is not the case and pearlitic structures are preferred, both for reasons of cost and for the comparative wear and fatigue resistance properties of pearlite microstructures when modified by compressive rolling and sliding stresses. These factors will be discussed later.

The change in ductility, in the form of true strain at fracture, with pearlite fraction (i.e. carbon content) is also shown in Fig. 5.6. Pickering's work indicated that inter-lamellar spacing had little effect on 'maximum uniform strain' as the effects of such spacing on flow stress and work-hardening rate self-cancelled, however, it did effect 'true strain at fracture' as larger spacing, specifically thicker cementite lamella (thickness = 0.14 spacing for an equilibrium eutectoid composition), can fracture at smaller strains and more readily generate cavities. There is therefore a limiting factor to strength improvement with refinement of inter-lamella spacing and the balance with cementite lamella thickness (the pearlite dilution factor); for a eutectoid composition, a lamella spacing is optimised at ~ 0.32 μm for the least damaging impact on the ductile–brittle transition temperature. This



5.5 Influence of carbon and manganese contents, plus other factors, on strength in C-Mn steels. (a) Effect of pearlite fraction on yield stress of carbon manganese steels (adapted from Pickering, 1992). (b) Effect of manganese content on tensile strength of C–Mn steels. (Adapted from Honeycombe and Bhadeshia, 1995) (Steel in normalised condition, grain size referred to is ferrite, 'N' refers to free nitrogen)

optimum spacing increases as pearlite dilution increases, as does toughness. For further qualification and detail, see Gladman (1992). To summarise, the impact toughness of pearlitic steels is related to the pearlite fraction, with an increase in the amount of pearlite resulting in a decrease in toughness, both in terms of the ductile–brittle transition temperature and the fracture toughness (K_{1C} value). This has implications for the length of fatigue crack



5.6 Influence of carbon content on tensile strength and ductility of normalised C–Mn steel with the tensile strength relationship in the 'quenched and tempered' condition super-imposed. The ductility relationship in the quenched and tempered condition is broadly similar. (Adapted from Ashby and Jones, 1998)

that can develop in the railhead until the critical size is reached, for the loading experienced, and the rail fails in a sudden and catastrophic manner.

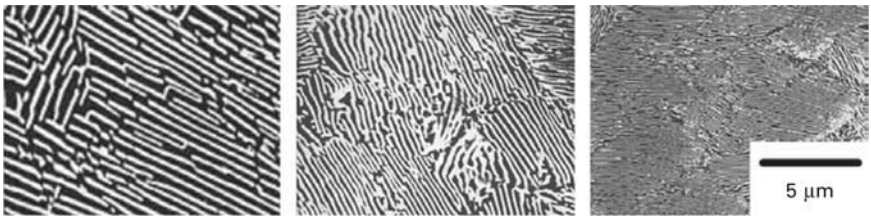
For tests in tension of fully pearlitic steels, Kavishe and Baker (1986) found that with decreasing inter-lamellar spacing, tensile strength, proof stress and cleavage fracture stress all increased, with proof stress related to the mean free distance in the ferrite lamellae by a Hall-Petch type equation, however, plane strain fracture toughness (in tension) first decreased but then increased for very fine spacing. They found no clear relationship between prior austenite grain size and these factors, but as grain size increased tensile ductility decreased as the latter represents the strain required to develop a nodule sized micro-crack which initiated unstable cleavage fracture. Pearlite nodule size is related to prior austenite grain size.

The bending fatigue strength of pearlitic rail increases at a modest linear rate with increases in tensile strength, for example a tensile strength increase from 700 to 1100 MPa gives a bending fatigue strength increase of around 300–400 MPa (for 2×10^6 cycles) (Gladman, 1992); however, the fatigue strength is reduced where crack initiation is facilitated, for example by corrosion, surface defects/micro-flakes, large near- and subsurface, deleterious inclusions and by any installed tensile stresses in the rail.

5.2.3 Higher strength pearlitic grades

Whereas the rail shape has to have enough ductility and toughness to be resistant to bending fatigue in its role as an 'I-beam' between sleepers, the wheel contact zone on the railhead requires maximum resistance to wear and rolling contact fatigue. The traditional solution to this was to increase the alloy content in the form of ~ 1wt% Cr which increased and refined the pearlite content; hardness of these rails was ~ 350 HB compared with ~ 250 HB of standard rail (Hodgson, 1985; Sawley, 1989). Strength could be further enhanced by a small vanadium addition and careful control of post-rolling thermal treatment (Hodgson, 1985). A disadvantage of these grades, apart from higher cost, is poor weldability and the fact that such rails located in high wear zones required bolted joints (Sawley, 1989), although modified welding techniques have subsequently been developed.

In the 1980s an alternative, more cost-effective way of giving a low-alloy rail-head more wear resistance (i.e. refining the lamella pearlite) was developed, 'head-hardening'. Rolled rail is either reheated (normalised) or taken direct from the rail-rolling stand and then cooled via an array of highly controlled, water mist sprays such that the railhead contact zone has a harder, more refined pearlitic structure whilst minimising adverse affects on straightness. This generates hardness of between 300 and 400 HB for a depth of 30 cm plus across the railhead. A precise knowledge of the chemical composition of a particular batch and its exact input temperature into the spray zone are required (Sawley, 1989), and additional techniques are required for welding and straightening such rails. In the past, a method of achieving this off-line has been to induction-heat the railhead and control cool with water mist sprays (Hodgson, 1985; Boelen and Jago, 1990). The refinement of the pearlitic structure, compared with normal and higher grade as-rolled rail, is shown in Fig. 5.7.



5.7 Pearlite lamellae structure for (left to right) 220 grade, 260 grade and 350HT (head-hardened) grade rail steel (images from Institute of Rail Welding job knowledge paper: IORW_Paper_15 March2003.doc). (*Electron microscope images, therefore carbide light and ferrite dark*)

5.2.4 Work-hardening of pearlitic structures

Pickering (1992) has shown that the work-hardening rate of ferrite–pearlite steel increases linearly with an increasing proportion of pearlite in the matrix (the 'pearlite fraction'). Decreasing inter-lamellar spacing similarly increases the work-hardening rate and also the flow stress and fracture stress. These determinations were made with *tensile* stress applied. With regard to hypo-eutectoid ferrite–pearlite steels, Umemoto *et al.* (2000, 2001) have shown that the work-hardening exponent for both ferrite and pearlite, with static loading, decreases with increasing strain and that they are near-equivalent in value (~ 0.5). However, this work did not examine the cyclic compressive stress and directional strain cycles that a railhead is subject to. Research by Eden *et al.* (2005), Garnham and Davis (2006) and Garnham *et al.* (2007) has indicated that with compressive rolling-sliding loads, strain partitioning between the PE ferrite and pearlite occurs at the railhead. This is further discussed in Section 5.7.

Compressive stress–strain characteristics and cyclic deformation tests on three near-eutectoid rail steels were carried out by Boelen and Jago (1990) by using PTFE (polytetrafluoroethylene) interface films at the compression die–sample interface. They compared a standard (0.8 wt% C, 1.0 wt % Mn) heavy-haul grade with an equivalent head-hardening (HH) structure and a vanadium micro-alloyed (VMA), head-hardened structure, where the PA grain size was more refined. Respective strength coefficients were (MPa) 1393, 1582 and 1655 and work-hardening coefficients were 0.101, 0.060 and 0.066. Cyclic deformation tests were carried out at constant stress amplitude, and the consequent deformation rates decreased with increasing number of cycles; a characteristic of pearlitic structures initially explored by Bower and Johnson (1991). At lower stress cycles there was 'plastic shakedown'; respective 'shakedown limits' were (MPa) 1093, 1350 and 1450. 'Dark bands', transversing several pearlite colonies, were observed in the standard grade structure and, to a lesser extent, in the HH structure; none were seen in the VMA structure. Transmission electron microscopy (TEM) examination revealed that these were shear bands of displaced cementite lamellae and heavily dislocated ferrite. These were first seen at low strains (0.05) with the standard grade and at higher strains (0.10) with HH grade. At a strain of 0.21, the VMA grade was free of such bands and deformation (i.e. dislocation density) appeared homogeneous. More recently, Canadinc *et al.* (2008) have tested samples taken from two high-carbon, pearlitic rails and 'J6' bainitic rail (carbide-free) in tension, compression and torsion in order to further develop and verify their crystal plasticity model for rail steel deformations and deformed microstructures. The model predicts slip behaviour in body-centred cubic, strained lattices between restrictions (pearlite lamellar carbide or bainite inter-lath austenite) for the three modes of stress and strain.

To summarise, at present pearlitic steels are the mostly widely used for rail because they are cost-effective in providing acceptable wear and fatigue performance of the railhead, due to the progressive work-hardening and microstructural alteration generated by compressive, rolling-sliding contact, plus providing adequate tensile mechanical properties for a structural beam between sleepers. Present market requirements for increased reliability and longer life, combined with higher axle loads and more traffic, has driven research into finding more refined pearlitic structures as well as alternative microstructures.

5.3 Austenitic rails for switches and crossings

In addition to the compressive rolling-sliding loads, parts of the railhead situated within switches and crossings are subject to severe, dynamic impact loads for which standard rail grades are not adequate due to their relatively low impact fracture toughness. Although some higher grade pearlitic steels are used, the preferred material is 'Hadfield' manganese steel, cast in monobloc sections, due to their high toughness and work-hardening rate, from 180–210 HB up to over 400 HB (Hodgson, 1985). These types of steel contain high levels of manganese (Table 5.1), which acts as an austenitic stabiliser, thus allowing the face-centred cubic austenitic structure to exist at room temperature. This structure has a significantly higher toughness than the pearlitic steels.

These austenitic steels have a tendency to form brittle carbide particles on cooling, which need to be minimised if the desired high toughness is to be achieved. This can give the steel very poor weldability as the cooling conditions after welding can result in the formation of these undesirable carbides. These steels also have a 60 % higher thermal expansion coefficient than pearlitic grades, which also causes problems on welding. During the 1980s, a lower carbon version was developed to lessen the risk of carbide formation (Table 5.1). These problems also stimulated research into alternative steel microstructures, for example bainitic structures (see Section 5.6). In the past the switches and crossings sections were bolted into track, however, present welding practice is to use a stainless steel 'inter-lay' between the austenitic steel crossing and the pearlitic rail, thus two separate welds are required (Zhang *et al.*, 2007).

5.4 Welding rail

Welding of rails is only carried out when absolutely necessary, for example to join rail sections together and to replace/repair damaged rail, since weld breaks can account for about 25 % of total number of rail breaks (IoRW, 2006). Different welding methods (alumino-thermic welding, flash-butt

welding and manual metal arc welding) are regularly used in the industry depending on the circumstances.

Alumino-thermic welding is generally used to repair rail breaks and is based on the generation of superheated liquid metal, produced by an exothermic reaction between iron oxide and aluminium powders. The powders are added to a mould in the proportions required by the rail type (composition and properties) and ignited using a fuse. The resulting weld requires trimming/grinding to achieve the required railhead profile. The microstructure generated in the weld is pearlitic (due to the slow cooling arising from the insulating effect of the mould), thereby giving matching properties with the parent rail. As the alumino-thermic weld is a casting, typical casting type defects can arise, such as hot tears and porosity (Fig. 5.8). Other defects can arise from incorrect powder proportions, incorrect pouring, poor weld trimming, etc.

Flash-butt welding is used for re-railing and renewals. The two rail sections are brought together and resistance heating is used to soften and melt the rail edges (arcs form across the interface leading to flashing). A physical force is required to push the two rail sections together, and molten metal is extruded at the join, which subsequently needs to be removed. Computer control allows for greater reproducibility in the weld properties, compared to an alumino-thermic weld, with greater productivity being possible (60+

5.8 Rail failure due to excessive porosity in an alumino-thermic weld. (Image courtesy of Dr Martin Harvey, 2007)

welds compared to 4+ alumino-thermic welds in the same time). The pearlitic weld microstructure generally has well matched properties to the parent rail, and fewer defects compared to alumino-thermic welds, although achieving the correct set-up (e.g. alignment, rail preparation) is essential; for example, poor surface preparation can result in porosity and lack of fusion.

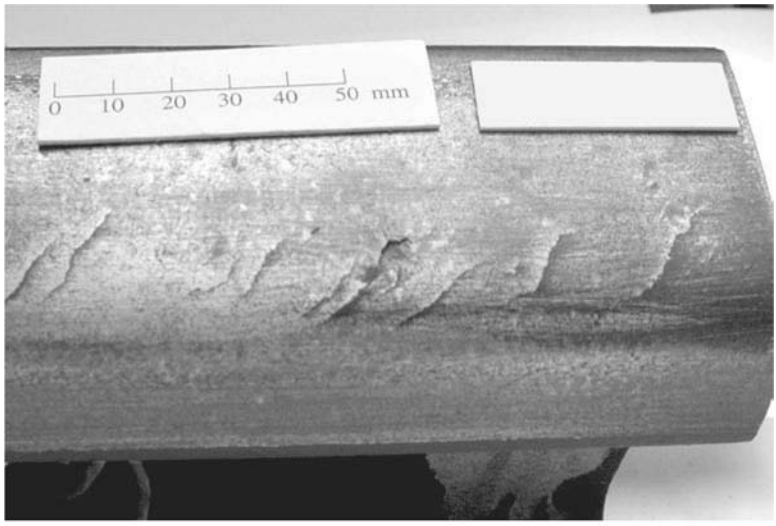
Manual metal arc (MMA) welding is used for rail repair and involves an arc being struck between the rail and a consumable electrode that deposits molten metal onto the rail. The weld is protected from the atmosphere by a slag and evolved gas layer. MMA welding is generally a multi-pass operation with the slag layer needing to be removed after each pass. The resulting microstructure is typically bainitic or martensitic due to the fast cooling rates experienced. Defects from gas porosity (e.g. oxygen, nitrogen) or entrained slag inclusions can arise.

Welding of high-strength rails (refined pearlitic grades and bainitic grades) using alumino-thermic and flash-butt techniques can result in a softer heat-affected zone adjacent to the weld. This soft zone may have reduced wear and fatigue properties leading to the formation of depressions on the running surface. The depressions will affect ride comfort for passengers and can cause increased dynamic loading on the rail resulting in a greater incidence of RCF. Welding of Hadfields austenitic steel switches and crossings is problematic due to the concerns about forming brittle carbide particles, hence a stainless steel insert can be used (Zhang *et al.*, 2007).

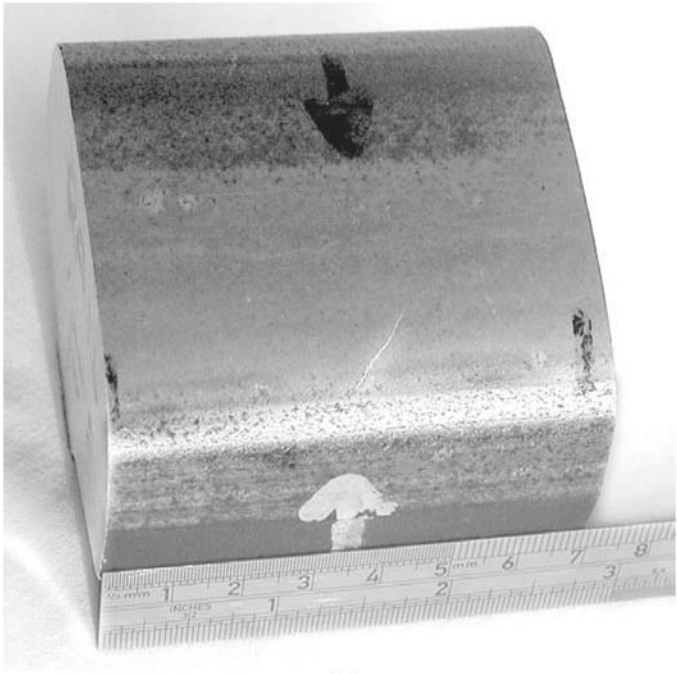
Other rail welding techniques, such as gas pressure welding and enclosed arc welding are being developed. Further information can be obtained from the Institute of Rail Welding on all aspects of rail joining.

5.5 Wear and rolling contact fatigue of pearlitic rail

When pearlitic steel microstructures are subject to compressive, directional loading, the structure is altered and performs mechanically in a manner not fully indicated by its tensile mechanical properties. Apart from rail, this can be seen with pearlitic wire ropes and cables (Gladman, 1992). The material work-hardens and the Fe_3C-based carbide lamellae align, bend, fracture and compact with the applied strain. The degree of work-hardening/microstructural strain varies across the railhead profile and is dependent upon contact conditions. It can be particularly severe on the high rails of curves where there is a large transverse component to the mix of longitudinal, transverse and spin creepage (sometimes called 'slippage'). On these curves, the gauge *side* of the rail is subject to very high creepage and subsequent wear, but not high contact stresses. The mix of high contact stresses and transverse creepage occurs just *above* the gauge corner, and in this region RCF cracks are initiated. In Fig. 5.9 examples of such cracks can be seen for low-carbon

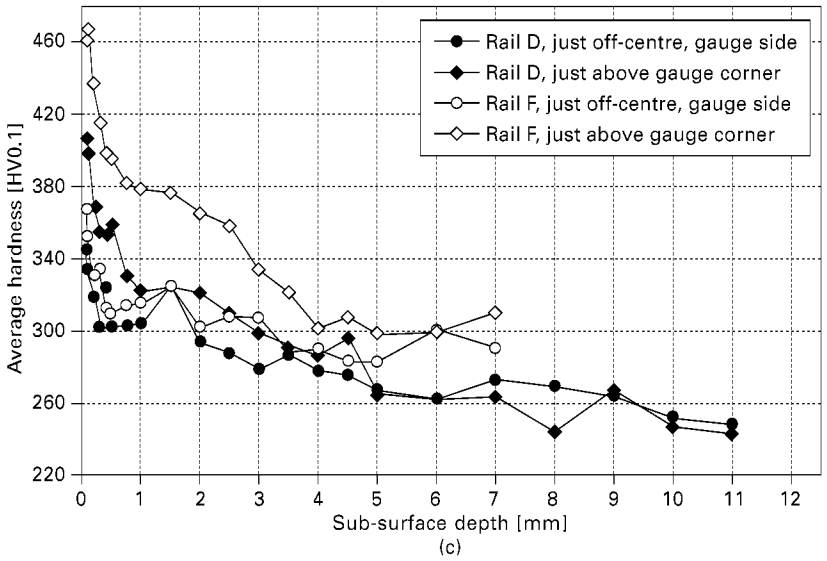


(a)



(b)

5.9 RCF cracks located just above the gauge corner on curved high rails. (a) Multiple cracks in low carbon, BS11 rail (Rail 'D'). (b) Singular crack in high carbon, '260 grade' rail (Rail 'F'). (c) Micro-hardness plots for these rails, at the track centres and at the respective 'RCF' crack zones, just above the gauge corners.

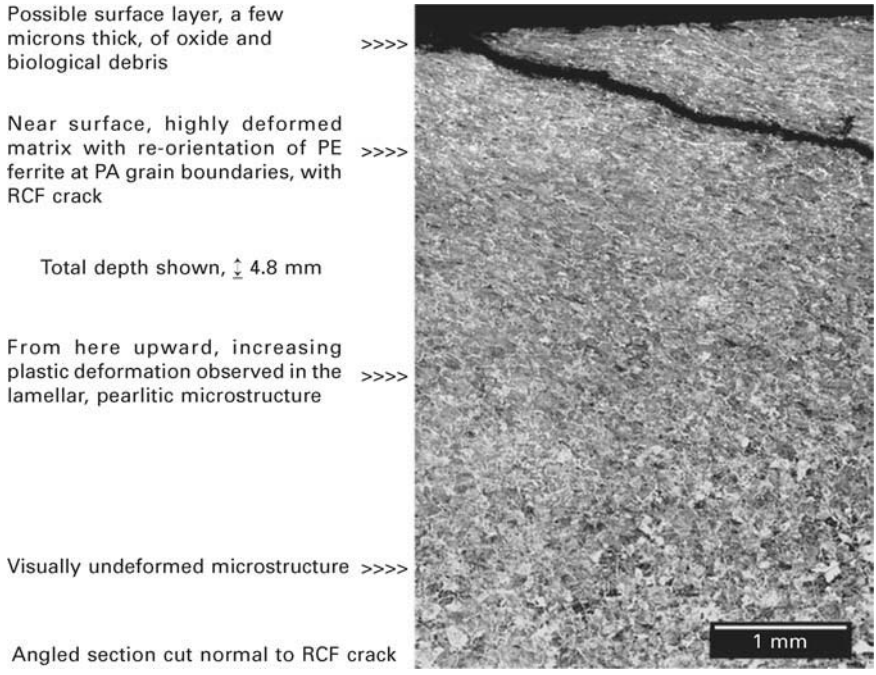


5.9 Cont'd

BS11 rail and a high-carbon 260 grade rail, together with respective micro-hardness plots comparing these 'RCF zones' with respective railhead centres; note the higher degree of hardening at the zones susceptible to RCF cracking. The accumulation of microstructural strain approaching the surface of the BS11 rail, at this 'RCF zone', is shown in Fig. 5.10. Comparing Figs 5.9c and 5.10, it can be seen that the depth of strain-hardening extends far deeper than the visual microstructural strain.

The PE ferrite at PA grain boundaries is strained between the pearlite zones, with the degree of strain dependent upon shape and orientation. Note that most new rail has a soft decarburised surface; however, its depth is not sufficient to be detrimental to rail life prior to its removal by wear (Carroll and Beynon, 2006). In Fig. 5.11 the microstructure of the gauge side of a worn BS11 rail is shown, with both the highly worn, wheel flange contact part and the non-contact part beneath showing the original decarburised surface structure.

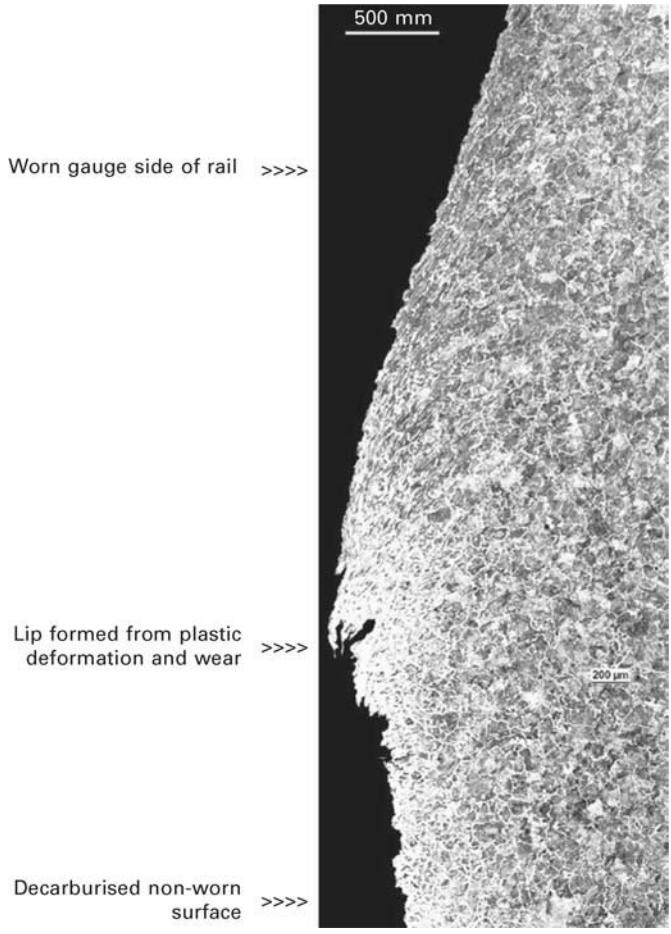
Examination of worn rail and test disc microstructures, together with micro- and nano-hardness surveys, has shown that there is strain partitioning between the PE ferrite and pearlite with more rapid hardening and strain exhaustion of the PE ferrite leading to primary (RCF) crack initiation at the edges of these PE ferrite zones (Eden *et al.*, 2005; Garnham and Davis, 2006; Garnham *et al.*, 2007). Pearlite carbide lamella flex considerably under compressive, directional loading so as to accommodate the strain, initially within 'strain bands' and then with eventual compaction and possible fracture, the latter



5.10 Angled section showing maximum strain distortion of the pearlite matrix of BS11 Rail 'D' at the 'RCF zone' just above the gauge corner (see Fig. 5.9). It can be seen from the micro-hardness plots (see Fig. 5.9) that the strain-hardening depth extends beyond that indicated by matrix strain distortion.

dependent upon lamella thickness and inter-lamella spacing (Langford, 1977; Perez-Unzueta and Beynon, 1993; Wetscher *et al.*, 2004). Mechanical properties become directional due to lamella compaction (Wetscher *et al.*, 2007). Figure 5.12 shows the features discussed above. Distortion of the lower, grade rail pearlite lamellae and RCF crack initiation on strain-aligned and compacted PE ferrite at PA grain boundaries is shown in Figs 5.12a and 5.12b, for BS11 Rail D and a 220 grade rail test disc, respectively. The ferrite lamellae within pearlite are supported by the harder carbide lamellae and deform with them, with dislocations buildup and crystallographic alterations resulting in a substructure developing with a preferred crystallographic planes aligned parallel to the surface (Satoh and Iwafuchi, 2005).

Perez-Unzueta (1992) and Perez-Unzueta and Beynon (1993) examined the compressive pure sliding and rolling-sliding (cooled, dry) wear of four pearlitic rail steels, mainly differing in inter-lamellar spacing. They confirmed a Hall-Petch type relationship for bulk hardness with inter-lamellar spacing (instead of grain size), found wear rate inversely linear to that hardness and spacing considerably reduced approaching the wear surface, due to more

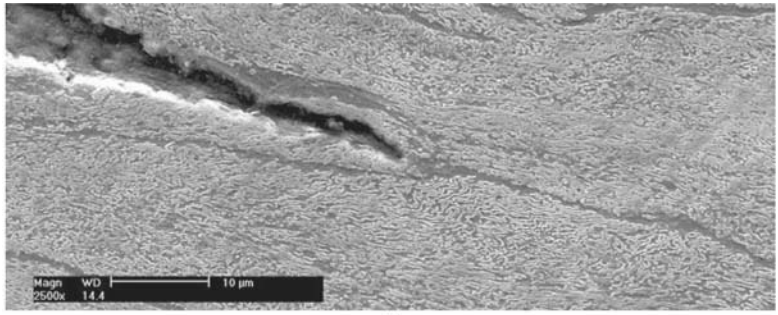


5.11 Gauge side of BS11 Rail 'D' (transverse section) showing decarburisation of non-contact region, worn gauge side in contact region and lip formed by plastic creepage and wear.

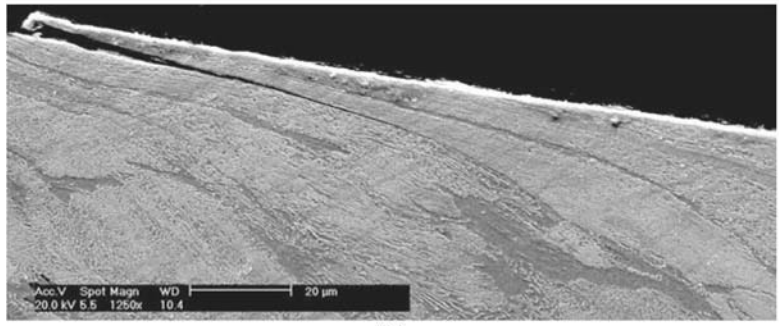
severe deformation of the ferrite lamellae. Perez-Unzueta (1992) also explored if and how the bent and/or fractured carbide platelets aligned themselves three-dimensionally at and near the contact surface with hardness tests in the surface plane and the section plane. He suggests that the effect of re-alignment was to present an increased area fraction (over the bulk fraction) of hard carbide at the contact surface, and it was this that made pearlitic steels superior to martensitic and (carbide-containing) bainitic structures for rail steels, as the carbide particles in the latter two microstructures do not adapt in this way with compressive rolling-sliding contact.

At the contact surface, where material has been strained beyond the plastic shakedown limit, angled surface micro-cracks are generated above which

Previous Page
Copyrighted Materials
Copyright © 2009 Woodhead Publishing Retrieved from www.knovel.com



(a)



(b)

5.12 Pearlitic rail lamella structures and the microstructural alterations. (*Electron microscope images, therefore carbide light and ferrite dark*) (a) BS11 Rail 'D', RCF crack tip, 80 µm subsurface initiated at PE ferrite boundary bounded by pearlite lamellae folding, fracture and alignment (angled section, normal to RCF cracks). (b) 220 grade rail test disc; pearlite distortion approaching surface and 'strain bands' within pearlite. PE ferrite boundaries here, approximately normal to strain, are not so distorted. Disc tested at 1500 MPa maximum contact stress and 1 % creepage, water lubricated, to 25 % RCF life).

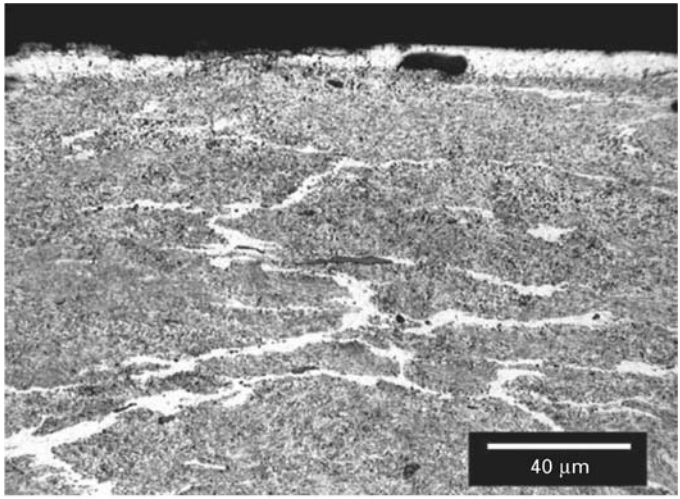
surface flakes form, the tips of which are no longer constrained by attachment to material below. These extrude by a 'ratcheting process' (Kapoor and Franklin, 2000) and wear by abrasion, adhesion, oxidation and very localised low-cycle fatigue (e.g. a section through a flake is shown in Fig. 5.12b). This process is accelerated for dry contact compared with lubricated. Reversing the contact direction reduces the wear rate and RCF crack propagation rate (Tyfour and Beynon, 1994a, b). Mixing dry and lubricated cycles, the situation for most rail, accelerates the wear and fatigue rates as lubricant is trapped within small cracks on the slower moving surface where rolling and sliding are in opposite directions (Tyfour *et al.*, 1996, Fletcher and Beynon, 2000). Trapped lubricant both hydraulically loads the crack tip and reduces friction between the crack faces.

Sliding and rolling-sliding contact surfaces of steels, subject to high and/or extended, usually cyclic, thermomechanical forces, can transform incrementally (where there is low wear), or suddenly (as in rail vehicle wheel locking), to a form hard, brittle form of martensite commonly described as 'white etching constituent' (WEC), although there is still academic debate on the exact nature of such transformations. In other words, iron carbide has been re-absorbed into solution by a continuum of nanothermal events at the asperity level. This transformation has been examined in depth by Lojkowski *et al.* (2001). On rail, apart from wheel slides, such structures are found on tangent track crowns and corrugation peaks (Clayton and Allery, 1982; Baumann *et al.*, 1991), i.e. where there has been low wear. Two examples are shown in Fig. 5.13. Fracture of these zones can initiate cracking and spalling at the surface. Recent rail contact simulation, twin disc tests, where WEC regions have been generated on a disc (Caroll and Beynon, 2007a, b), has shown that crack growth is dependent upon the subsurface deformation of the pearlite, with longer cracks where WEC was present. Where there was no cracking or spalling, wear of the WEC region was much lower than the pearlite surface and wear of the contacting wheel disc was also reduced, however, if the WEC region cracks and spalls, wear rates are increased due to this.

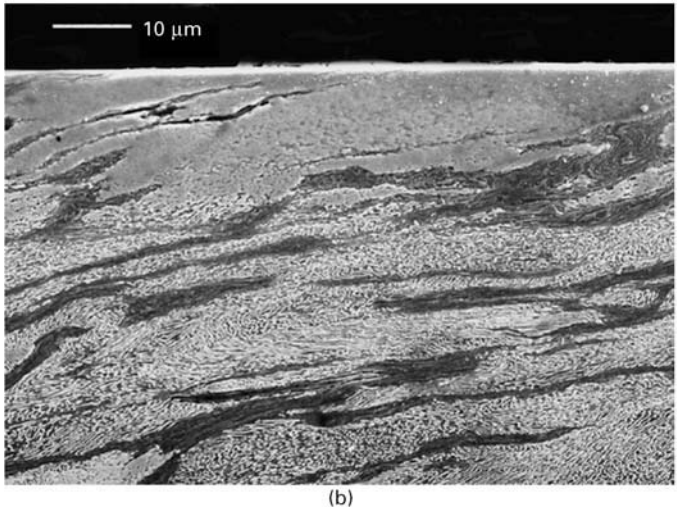
Brittle inclusions (alumina- and silica-based) are detrimental to rail life. If large enough, subsurface fatigue crack initiation can occur deep within the railhead leading to 'tache ovale' type, transverse fracture of the rail (Fegredo *et al.*, 1988; Shur *et al.*, 2005) as shown in Fig. 5.1. Due to quality control in modern rail production, such occurrences are rare nowadays, 'however' on a far finer scale, very fine, micro- or nano-cracks initiate on these small brittle inclusions near-surface. Ductile (MnS-based) inclusions are generally benign; however, if the rod-shaped inclusions stringers, aligned along the rail, are distorted within the microstructure by a high vector of transverse creepage, their shape can be altered to that of a very thin disc, i.e. a 'plane of weakness' is formed; voids and cracks can initiate on such planes (Fig. 5.14). Sulphur content has been shown to be significant with regard to rail life, since increased sulphur content leads to an increase volume fraction of MnS inclusions (Fegredo *et al.*, 1988; Liu *et al.*, 1993).

5.6 Bainitic rail

The bainitic steel structure has been defined thus (Krauss, 1992): 'Bainite is nominally a two phase microstructure which is formed by austenite transformation between the temperature ranges at which pearlite and martensite form.' For plain carbon steels, the area for bainite formation with continuous cooling is limited. Where a bainitic microstructure is required, alloy additions are made to extend the bainite part of the isothermal and continuous cooling



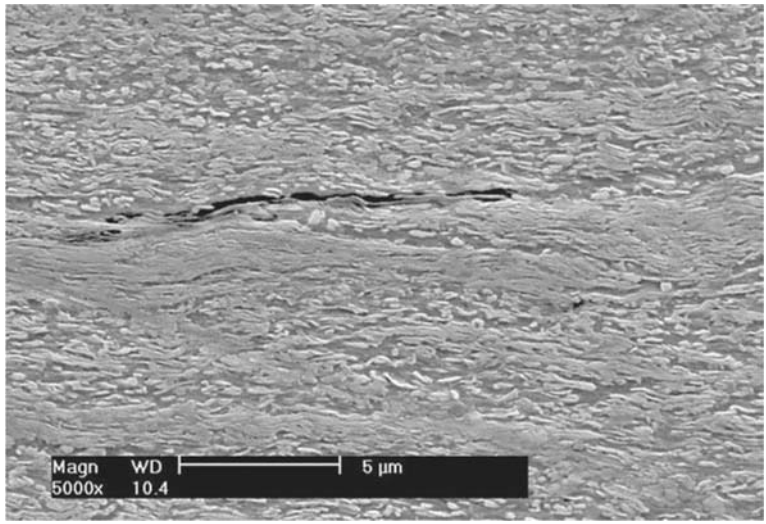
(a)



(b)

5.13 Formation of white etching constituent (WEC) on railhead surface. (a) Intermittent WEC formation at track centre of BS11 Rail 'D' (axial section). (b) WEC formation and pearlitic carbide dissolution at track centre of an old, 0.44 wt%C BS11 rail on tangent track (axial section). *(Electron microscope image, therefore carbide light, ferrite dark, WEC grey)*

curves (the bainite 'nose') over a wider range of cooling rates, as shown by comparing Fig. 5.15 with Fig. 5.4, particularly small additions of molybdenum. The 'nose' can be further extended by very small additions of boron which improves toughness and fatigue resistance (Honeycombe and Bhadeshia,

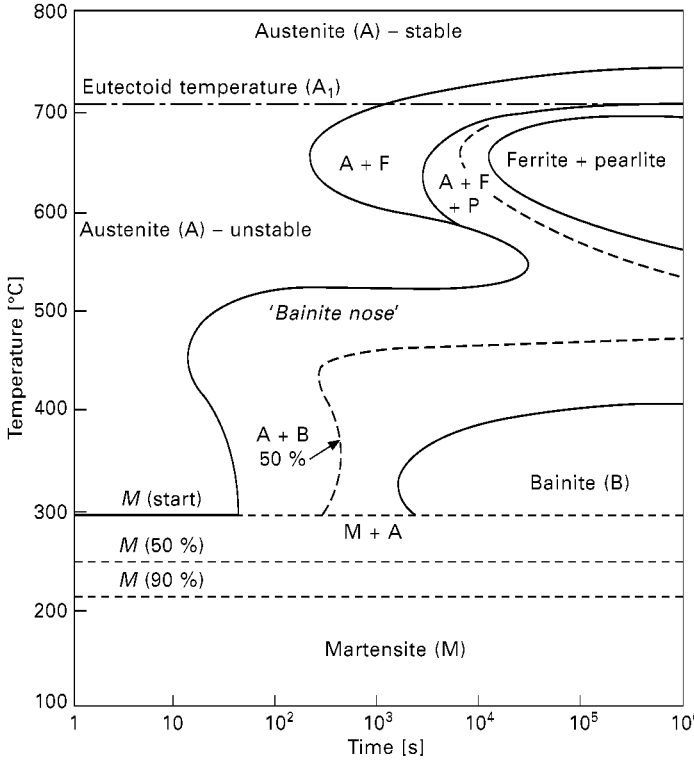


5.14 Very near-surface of 260 grade Rail 'F'; angled section through strain-flattened, MnS based inclusion showing void formation. *(Electron microscope image, therefore carbide light and ferrite dark)*

1995). Additions of oxidising elements, such as aluminium and titanium, are also required to ensure that the boron does not become oxidised during steel manufacture and thus becomes ineffective.

For plain carbon, low-alloy steels, the bainitic lath structure forms within the prior austenite grain boundaries by an initially displacive transformation from austenite to ferrite, but then the cooling rate is such that there is time for some carbon to come out of solution as carbide particles, as aligned particles within the laths at a fast rate (lower bainite) and as semi-lamellae between the laths at a slower rate (upper bainite). Past research showed that such carbide-containing bainitic structures, or tempered martensite structures, were not superior to pearlitic structures for normal rail wear, particularly the higher grades of 1 wt% Cr or head-hardened (Ghonem *et al.*, 1982).

With low-carbon and alloying, other forms of bainite are possible with no carbon coming out of solution. These have superior (tensile) mechanical properties to medium- and high-carbon pearlitic steels. The lower carbon alloys have retained austenite at inter-lath sites and the higher carbon alloys, zones of retained austenite with their centres partially transformed to martensite – sometimes called 'massive bainite'. Some electron micrographs and schematics of these structures are shown in Fig. 5.16 and respective hardness data, engineering stress–strain and ductile–brittle transformation temperature plots can be seen in Fig. 5.17. These plots show the greatly increased toughness and ductility of the carbide-free bainitic steels, and the



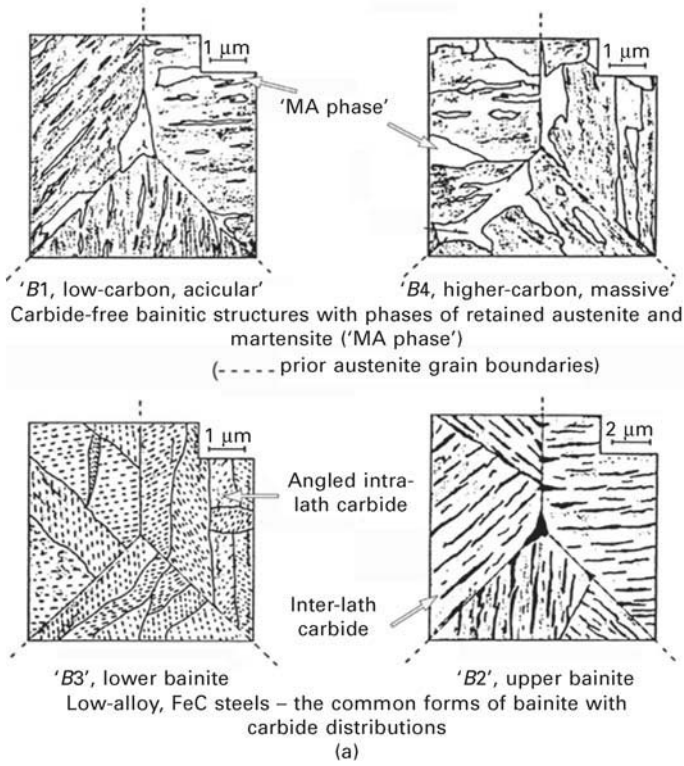
5.15 Schematic representation of a TTT curve for an alloy steel [adapted from Callister (1994) and Gladman (1992)], showing the comparative shift of the ferrite–pearlite transformation to slower times, leaving an extended 'bainite nose' covering a wider range of cooling rates. One of the functions of designing alloy compositions for low carbon, bainitic rails, was to extend this nose over a wide range of (air) cooling rates.

sharp drop in toughness in the bainitic grade containing inter- and intra-lath carbide (plus a little martensite). Some authors grade steel bainitic structures into four types, 'B1' to 'B4', as shown in Fig. 5.16. Experimental bainitic rail steel compositions from several geographic sources are shown in Table 5.2.

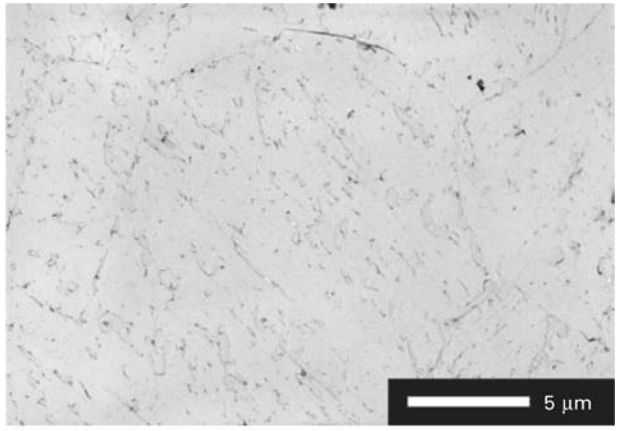
Harbraken and Economopoulos (1967) were early investigators into these forms of bainite in which carbides did not precipitate. In conjunction with British Rail Research (BRR), Callender (Callender, 1983; Callender *et al.*, 1983a,b) used work such as this to investigate the suitability of such low-alloy, molybdenum–boron, bainitic microstructures for rail steels, particularly rail crossings and further British Rail Research (BRR) work examined the suitability of carbide-free bainitic steels for rail and monobloc wheels (Sawley *et al.*, 1988). The attraction was the combination of hardness, toughness

and air-cooling characteristics (i.e. 'weldability', air-cooling to the desired microstructure). This would facilitate the welding of bainitic crossings directly into similarly structured track and the possibility of direct disc breaking on to bainitic wheels without thermal damage.

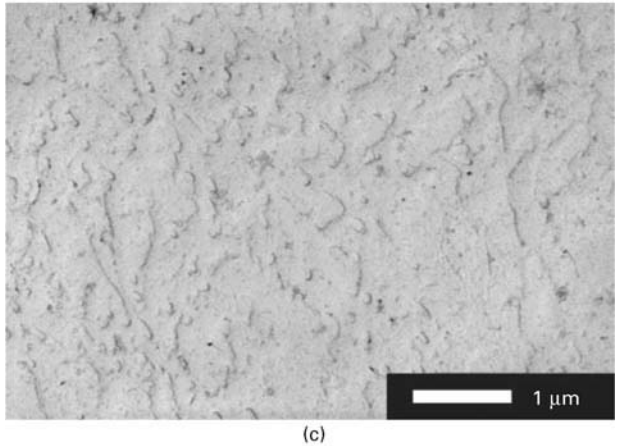
With pure sliding, pin-on-disc testing, wear rate was linearly related to hardness, irrespective of structure (Clayton *et al.*, 1987). Early rolling-sliding experiments on these materials indicated that under normal rolling-sliding conditions, its wear rate was not better than standard pearlitic rail and there was increased wear on the counter-face wheel material (Heller and Schweitzer, 1982; Garnham and Beynon, 1992; Garnham, 1995); however, under very severe rolling-sliding conditions such as high creepage at the gauge face,



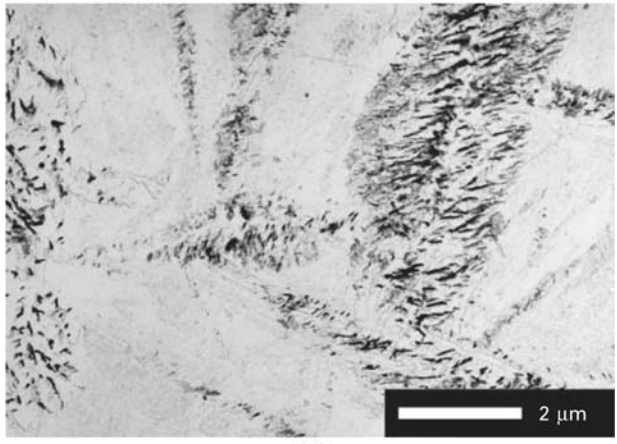
5.16 Experimental bainitic steel microstructures, two carbide-free (after Garnham, 1995). (a) Schematic of four types of bainitic steel microstructure. (b) TEM image of 0.04 wt%C, carbide-free bainite 'B04', showing 'MA phase'. (c) TEM image of 0.20 wt%C, carbide-free bainite 'B20', showing 'MA phase'. (d) TEM image of *minority* mixed phase of lower bainite laths (with inter-lath carbide) and martensite in 0.52 wt%C bainite 'B52'. (e) TEM image of *majority* upper bainite phase (with intra-lath carbide) in 'B52'.



(b)

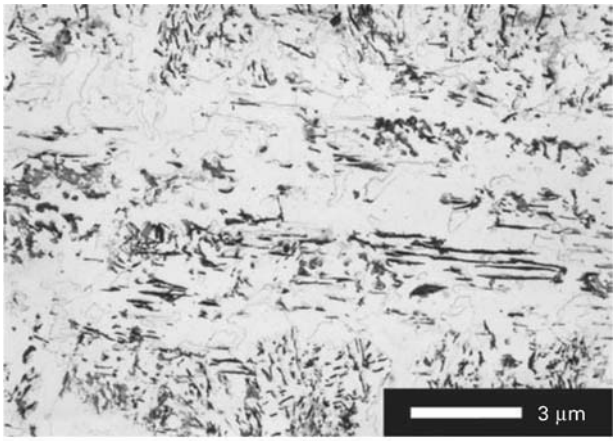


(c)



(d)

5.16 Cont'd

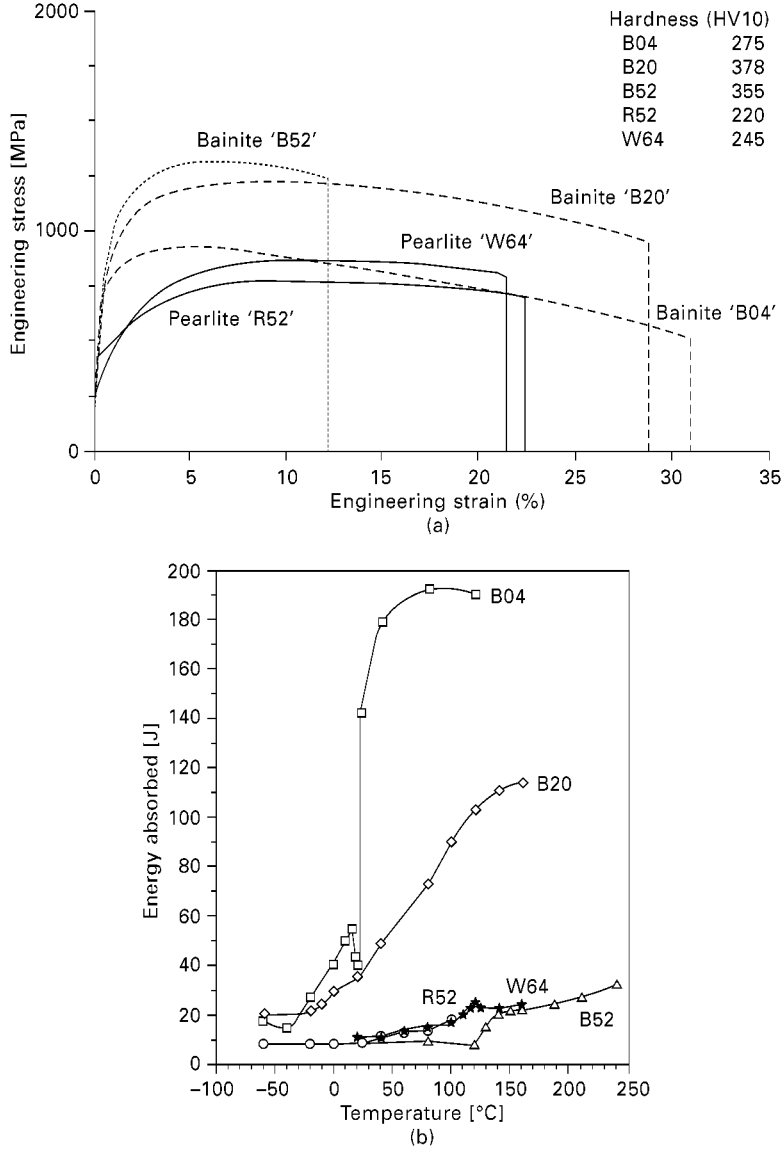


(e)

5.16 Cont'd

the bainitic grades were superior, particularly higher carbon grades which remained carbide-free by control of alloying additions and cooling rates (Devanathan and Clayton, 1991; Clayton and Jin, 1996; Jin and Clayton, 1997; Sawley and Scholl, 1998). This can be explained by alignment of pearlitic carbide under normal rail contact conditions giving high wear and RCF resistance for the material hardness; however, under severe conditions where that alignment is disturbed, then the wear-resistant properties of the carbide, free bainites excel, due to their structure of ferrite with a fine distribution of higher carbon, martensite-austenite (MA) phase (i.e. particles of retained austenite and/or particles of retained austenite surrounding centres transformed to martensite). Further twin-disc tests, water lubricated and at a high creepage of 10 % over a range of contact stresses indicated that, for equivalent mechanical properties, the bainitic structures performed similar to the pearlitic steel (Clayton and Su 1996; Su and Clayton, 1996; Yokoyama *et al.*, 2002) compared standard carbide-containing bainites with equivalent hardness pearlitic structures, both for dry and water lubricated twin disc tests; they found the bainites inferior for wear resistance and superior for RCF resistance. From all this previous work, Bhadeshia (2001, 2002) examined and modelled the bainitic structures in-depth and suggested an optimum composition for rail use (Table 5.2). This work stimulated a patent application from Corus (Table 5.2) covering a range of bainitic compositions.

Wear of the railhead consists of both material loss from the surface and profile loss by creeping plastic deformation of the railhead edges (Fig. 5.11). Another advantage of the harder, carbide-free bainitic structures over normal grade pearlitic rail was that they were more resistant to this form of wear, reflecting findings for track profile distortion with twin disc tests (Allery and Sawley, 1989; Garnham, 1995).



5.17 Comparative properties of experimental bainitic steels (*cf* Fig. 5.16) and two pearlitic steels, 0.52 wt%C rail (R52) and standard 0.64wt%C wheel (W64) (after Garnham, 1995). (a) Engineering stress–strain curves and hardness. (b) Ductile–brittle, transformation plots.

Research was stimulated in several countries (Table 5.2) and bainitic rails have been installed on track. Their use is now routine for certain high-wear locations in parts of Europe and Japan, but they are still in the development stage in the UK (Corus Rail, 2007), North America and elsewhere, partly due

Table 5.2a UK experimental *bainitic* steels for rail and rail crossings

Steel UK	C	Si	Mn	S	P	Cr	Mo	Ni	Cu	Al	N	H	V	Ti	B	Comment
Proposed optimum composition for bainitic rail crossings	< 0.12	0.25	1.0	0.007	< 0.01	1.2–1.8	0.4	3.0	–	0.04	–	–	–	0.04	0.003	Callender, 1983; Callender *et al.*, 1983a,b UTS nom. 1050 Mpa Hardness nom. 320 HV
Novel bainitic steel, Callender *et al.*	0.08	0.3	0.7	0.01 max	0.01 max	2.0	0.5	3.0	2.0* Cl.8*	–	–	–	–	–	0.002	Appl: British Steel Corp. and Edgar Allen Eng, Inventor: Callender W. R., Sawley K. J. Brook R. (*Callender et al., 1983b)*
Garnham and Beynon experimental casts B04 (carbide-free, MA phase)	0.04	0.19	0.80	0.009	0.009	2.76	0.25	1.93	–	0.033	–	–	–	0.028	0.0023	Garnham and Beynon, 1992; Garnham, 1995; Devanathan and Clayton, 1991
B20 (carbide-free, MA phase)	0.20	0.16	0.67	0.009	0.012	2.29	0.27	1.68	0.03	0.010	–	–	–	0.002	0.0023	Garnham and Beynon, 1992; Garnham, 1995.
B52 ($\alpha' + \alpha_{lb} + \alpha_{ub}$)	0.52	0.22	0.37	0.011	0.013	1.70	0.27	1.44	–	0.028	–	–	–	0.028	0.0022	Garnham and Beynon, 1992; Garnham, 1995; Devanathan and Clayton, 1991.
Clayton experimental cast extra to 3 above B10 (carbide free, MA phase)	0.10	0.27	0.59	0.02	0.008	1.71	0.58	4.09	–	–	–	–	–	–	< 0.01	Also used by Devanathan and Clayton, 1991.

Table 5.2a Cont'd

Steel UK	C	Si	Mn	S	P	Cr	Mo	Ni	Cu	Al	N	H	V	Ti	B	Comment
Experimental bainitic rail steel	0.09	0.20	1.00	–	–	–	0.50	–	–	0.03	–	–	–	0.03	0.003	Bhadeshia, 2001
Experimental bainitic rail steel	0.30	0.20	2.00	–	–	1.0	0.50	–	–	0.03	–	–	–	0.03	0.003	Bhadeshia, 2001
Experimental bainitic rail steel	0.04	0.20	0.75	–	–	2.8	0.25	2.0	–	0.03	–	–	–	0.03	0.01	Bhadeshia – 'best available' composition for rail use. *(Similar to Garnham B04; see above)*
Novel bainitic steel, Bhadeshia	0.05 0.50	1.0–3.0 and/ or Al	0.50–2.50	< 0.025	–	0.25–2.5	< 1.0	< 3.0	<3.0	1.0–3.0 and/ or Si	–	–	<0.5	<0.1	<0.005	W < 0.01, Bhadeshia and Jerath 1995, 1996 Appl. British Steel plc; Inventor: Bhadeshia HKDH and Jerath V, *Improvements in and relating to carbide-free bainitic steels*

Table 5.2b Non-UK experimental *bainitic* steels for rail

Steel	C	Si	Mn	S	P	Cr	Mo	Ni	Cu	Al	N	H	V	Ti	B	Comment
Germany																
Krupp Stahl – experimental bainitic rail steels	0.07	0.30	4.50	–	–	–	0.50	–	–	–	–	–	–	–	–	Experimental bainitic rail steel – *Carbide free bainite* Heller and Schweitzer, 1982 (*cf* Bhadeshia, 2001)
Krupp Stahl – experimental bainitic rail steels	0.30	1.00	0.70	–	–	2.7	0.20	–	–	–	–	–	–	–	–	Experimental bainitic rail steel – *Bainite with some carbide ppt'n* (*cf* Bhadeshia, 2001)
Thyssen Stahl – experimental bainitic rail steels	0.40	0.70	1.50	–	–	1.10	0.80	–	–	–	–	–	0.10	–	–	Experimental bainitic rail steel – Used on Narvik-Kiruna line (*cf* Bramfitt, 1998).
USA																
TTCI experimental bainitic rail steel, J6	0.26	1.84	2.00	0.006	0.008	1.94	0.44	0.11	0.17	0.033	–	1.1 ppm	0.01	0.04	0.0026	*cf* Kristan Sawley *et al.*, 2003
Japan																
Early experimental bainitic rail steel	0.33	0.31	1.2	0.015	0.014	1.25	0.20	–	–	0.028	–	–	0.05	–	0.0029	Ichinose *et al.*, 1978
Nippon Steel – experimental bainitic rail steels – selection	0.29	0.35	1.16	–	–	2.21	–	–	–	–	–	–	–	–	–	Nb 0.04 Bramfitt, 1997

Table 5.2b Cont'd

Steel	C	Si	Mn	S	P	Cr	Mo	Ni	Cu	Al	N	H	V	Ti	B	Comment
Nippon Steel – experimental bainitic rail steels – selection	0.34	0.32	0.76	–	–	2.51	–	–	–	–	–	–	–	–	0.0015	
Nippon Steel – experimental bainitic rail steels – selection	0.32	0.29	0.41	–	–	2.81	0.59	–	–	–	–	–	–	–	–	
Nippon Steel – experimental bainitic rail steels – selection	0.25	0.15	0.31	–	–	2.98	–	2.41	–	–	–	–	–	–	–	
Nippon Steel – experimental bainitic rail steels – selection	0.35	1.98	0.74	–	–	2.41	–	–	–	–	–	–	–	0.032	–	
Nippon Steel – experimental bainitic rail steels – selection	0.38	0.51	1.99	–	–	0.51	–	–	0.11	–	–	–	–	–	–	
Nippon Steel – experimental bainitic rail steels – selection	0.15	0.51	1.41	–	–	0.95	0.41	3.89	–	–	–	–	–	–	–	

to patent application issues and concerns over wear rates under less severe loading. One drawback of course is their increased cost over pearlitic grades due to their higher alloy content. Track trials at TTCI in the USA with a bainitic rail initially proved promising; it has superior track life (wear and RCF resistance) compared with head-hardened pearlitic rail of hardness 350 HV (Kristan *et al.*, 2003). This 'J6' bainite has a hardness of 445 HV and a structure of carbide-free, lath ferrite with inter-lath retained austenite; the composition is shown in Table 5.2. Its laboratory wear tests gave contradictory results with slightly higher wear rate under standard rail contact conditions, as other authors have found (Heller and Schweitzer, 1982; Garnham and Beynon, 1992). It was surmised that maybe this, in removing initiating cracks more readily, might contribute to its superior RCF crack resistance in track. More recent TTCI wear tests (i.e. pure sliding ball on disc tests and dry test-track rails), comparing 'J6' bainitic rails with high-carbon, fully pearlitic rails, have indicated that this bainite has inferior wear resistance compared with the premium grade pearlite rails as it is work-hardened less (Hernández *et al.*, 2007).

Track trials in Austria found hard bainitic grades superior to head-hardened pearlitic grades in resistance to RCF cracking on gauge corners (Pointner, 2006). He reported no detrimental effect on wheel wear when using harder rail steel grades. He also commented on the balance between wear and rolling contact fatigue; although with harder rail grades initiating cracks are not worn away so rapidly, their growth rates are far lower than those of normal grade rail and generally their resistance to significant RCF cracking is far higher. With bainitic grades, which have higher wear rates than equivalent hardness pearlitic grades, under normal rail conditions, early RCF cracks are removed and wheel–rail contact becomes conformal earlier, however, these grades have a very high toughness and resistance to crack growth. In 1979/80, bainitic rail from Thyssen Stahl was tried on a severe bend in the iron ore line of Northern Scandinavia (composition is shown in Table 5.2). This bainite had higher hardness– 436 HV, superior weldability and a service life eight times better than that of UIC860 *900A* pearlitic rail of hardness 290 HV (De Boer *et al.*, 1995; Bramfitt, 1998; Kern *et al.*, 1998). Research into bainitic grades has been ongoing in Japan for many years (Ichinose *et al.*, 1978; Bramfitt, 1998). Some compositions are shown in Table 5.2; Bramfitt (1998) gives more detail within a comprehensive assessment of bainitic rail.

To summarise, carbide-free bainitic steels have been developed with some mechanical properties superior to those of premium grade pearlitic steel rails. They appear to have superior RCF performance. Their comparative wear performance, *except* under severe sliding conditions (with rolling-sliding), is disappointing compared to that of softer pearlitic rail; this can be explained by the carbide lamellae redistribution patterns of pearlite under rolling-sliding

contact. The wear performance on track is even worse than indicated (by profile area loss), as the bainitic rail does not plastically deform as much as the pearlitic rail. However, this poor wear performance may enhance RCF resistance by clearing incipient cracks as they form, facilitating the forming of a more conformal contact (thus reducing contact stress) and, if RCF cracks have initiated, by having a very high resistance to crack propagation. Tests and track assessments for severe wear, comparing carbide-free bainitic rails with premium grade, high-carbon, pearlitic rails, have given mixed results. It is additionally worth noting that carbide-free bainitic rails are far more weldable than carbide-containing grades as they air-cool to the desired microstructure.

5.7 Recent rail material developments

5.7.1 High-carbon, hyper-eutectoid rail steel, but with carbide-free, prior austenite grain boundaries

These have been primarily developed by Nippon Steel, and initial developments are described by Uchino *et al.* (1998) with compositions shown in Table 5.1b. Their target was to increase the wear resistance of (heavy-haul) rail by increasing the density and thickness of lamella carbide without a significant increase in hardness compared with standard high-grade rail; i.e. to facilitate the compaction of carbide lamella platelets at and near the rolling contact surface. Their steels were subject to accelerated cooling or salt bath quenched so as to prevent brittle iron carbide forming along PA grain boundaries and to control hardness. With the 0.9 wt% C steel the pearlite nose (*cf* Fig. 5.3) moved to a higher temperature range than that of the 0.8 wt% C steel and with a boron addition, to an even higher range, thus giving stable pearlite transformation over a wide cooling range. The 0.8 wt% C grade had a little PE ferrite at PA grain boundaries; both 0.9 wt% C steels were fully pearlitic.

Ductility (tensile elongation) was compared with regard to PA grain size. Both the 0.8 wt% C and 0.9 wt% C steels showed a linear relationship with similar slopes, the finer the PA grain size, the higher the ductility. The 0.9 wt% C steel had only 80 % of the ductility of the 0.8 wt% C steel over the grain size range; for example at 62.5 μm grain size (G_A5) elongations were 13.0 % and 10.4 %, respectively; therefore, for use as rail the higher carbon steel needs a more refined PA grain size. The experimentation included a hardness survey across a flash-butt rail weld. The weld itself is only 10 HV points below the rail hardness of ~ 390 HV, however, the flanking heat-affected zones drop to 270 HV, although these each only extend for ~ 10 mm axially along the track. This softer zone may cause greater wear and subsequent localised track undulations; however, no data have been published on this.

Dry wear tests were carried out on a twin-disc machine at 680 MPa maximum contact stress and 20 % creepage, with the rail disc running faster (the opposite of track conditions where the rail is the slower surface). For both steels, wear rate was linearly related to hardness and at equivalent hardness, the 0.9 wt% C hyper-eutectoid steel had significantly higher wear resistance compared with the 0.8 wt% C steel.

Progress on the development of these steels can be seen in Iwano *et al.* (2006). Rails are rolled and heat-treated giving a deeper hardening into the head than seen with conventional, in-line, head-hardened rails. An even higher carbon version with 1.0 wt% C (*cf* Table 5.1b) has been developed and tested. By even tighter control of rolling and austenitising, elongations > 10 % were achieved. These have been installed on heavy-haul track in North America and are being assessed on heavy-haul track in Australia and South America. With the twin-disc tests described above, with all steels having an initial hardness of ~ 400 HV, after 700 cycles the surfaces had work-hardened to 670, 705 and 750 HV for the 0.8, 0.9 and 1.0 wt% C steels, respectively, with this hardening reflected in the wear resistance results. No specific test data were given for RCF resistance, although RCF cracks were observed for one material condition during field trials. To summarise, these high-carbon steels can function as rail but, compared with standard pearlitic rail grades, far closer control of rail production, quality and welding is required to avoid the risk of brittle carbide forming at PA grain boundaries as this could lead to sudden rail breaks.

5.7.2 'Two-material rail' – railheads with a coating

Two methods have been considered to produce coatings on the railhead, a rolled in-line process by Corus (www.corusrail.com) and a laser cladding method by a commercial company, Duroc (www.duroc.se) (Franklin *et al.*, 2005; Hiensch *et al.*, 2005). Corus have developed a process where the surface of the railhead is coated in 1 mm plus of stainless steel to prevent corrosion. The coating is applied prior to rail-rolling. This has a function in urban networks where there is a requirement for good electrical contact direct on to the rail, plus anti-squeal and anti-wear properties. No wear or fatigue resistance data are given. So far, only data for the Duroc method have been presented, data from twin-disc tests and track assessments (the heavy-haul, iron ore railway of Northern Scandinavia).

With the Duroc process, the railhead of standard high-grade rail is ground away and then a super-hard, high-alloy, 'tool-steel' coating, with some ceramic additives, is applied by laser and the surface reground to profile (compositions not given). Track trials and twin-disc tests of two types of Duroc coating on 260 grade rail, with respective hardness values of 370 HV and 410 HV, showed that such coatings were highly resistant to wear and rolling contact

fatigue (Franklin *et al.*, 2005; Hiensch *et al.*, 2005). Cladding depth data for rail are not given, but for twin-disc tests the coating was around 0.7 mm initially and then ground to 0.35 mm thick for dry tests and 0.1 mm for interface bonding tests. These tests were wet, dry and a combination of both, all at –1 % creepage, at 1500 and 2500 MPa maximum contact stress. Both had good bonding quality and out-performed 260 grade pearlitic rail in all tests. At the higher contact stress, one of the two types did eventually delaminate, whereas in track trials this type was crack-free whereas small cracks eventually developed on the other type. The companion 900A grade rail was comprehensively cracked. Neither coating lessened squeal noise. The effect on a conformal wheel material was not explored; as could be seen with bainitic steel twin-disc tests where the counter disc was of wheel material, there is a possibility that wheel wear will increase with harder, non-pearlitic rail (Garnham, 1992, 1995) although, from track observations, Pointner (2006) indicates that it may reduce.

5.8 Conclusions

Due to historic and cost reasons, most rail has been manufactured, rolled and air-cooled from plain carbon–manganese steels with structures of pearlite and some PE ferrite at PA grain boundaries. The strain distortion of these structures under cyclic, compressive loading and rolling-sliding contact has resulted in wear and fatigue resistance properties superior to those indicated by their mechanical properties as determined by standard tensile tests.

In order to facilitate longer rail life under harsher conditions, refinement of such structures has been developed, via reduction or elimination of PE ferrite by using higher carbon grades and/or refinement of the pearlitic structure, in particular inter-lamella spacing, by controlled cooling (head-hardening), alloying (e.g. chromium) and refinement of PA grain size. Other steel structures have been explored historically, but only recent developments of bainitic structures, mainly low-carbon, alloyed, carbide-free bainites, have been followed in depth. Although these give no benefits in wear resistance under normal conditions, they are more resistant to RCF and air-cool to structure, thus greatly facilitating weld joining. Other, more expensive, novel solutions have also been recently explored, such as hard-cladding the rail contact surface.

5.9 References

Allery, M. B. P., Sawley, K. J. (1989), *Results of British Rail Research field trials where experimental bainitic rails have been laid in track curves amongst standard BS11 pearlitic rails,* personal communication to J. E. Garnham.

Ashby, M. F., Jones, D. R. H. (1998), *Engineering Materials 2* (2nd edn), Butterworth & Heinemann, Oxford, UK.

Baumann, G., Fecht, H. J., Liebelt, S. (1996), Formation of white-etching layer on rail treads, *Wear*, **191**, 133–40.

Bhadeshia, H. K. D. H. (2001), *Bainite In Steels* (2nd edn), Institute of Materials, London, UK.

Bhadeshia, H. K. D. H. (2002), 'Steels for Rails', in Buschow, K., Cahn, R. W., Flemings, M. C., Ilschner, B., Kramer, E. J., Mahajan, S. and Veysierre, P. (eds), *Encyclopedia of Materials: Science and Technology*, Pergamon Press, Elsevier Science, Amsterdam, the Netherlands, 1–7.

Bhadeshia H. K. D. H. Jerath V. (1995), *Improvements in and relating to carbide-free bainitic steels*, UK patent 9501097.1.

Bhadeshia, H. K. D. H. and Jerath, V. (1996), *Method for producing carbide-free bainitic steels*, European Patent EP0804623.

Bramfitt, B. L. (1998), The metallurgical characteristics of bainite as applied to a rail steel, *Proceedings 4th Internation Symposium on 'Rail Steels' as part of the 39th International Symposium on 'Mechanical Working and Steel Processing'*, Indianapolis, IN, USA, 20–21 October 1997, Iron & Steel Society, Warrender, PA, USA and Association of American Railroads, Washington DC, USA, 989–96.

British Steel Corporation (1985), Rail Products Technical Information Brochure, *Rail Metallurgy*, 28–54.

Boelen, R., Jago, R. A. (1990), Structure-property relationships in eutectoid rail steels, *Seminar on Microstructure Control to Achieve Properties in Modern Steels*, Monash University, Melbourne, VIC, Australia, 30 Jan. – 1 Feb., organised by The Institute of Metals and Materials, Australasia.

Boulanger, D. (2003). Rail metallurgical developments to address the changing needs of railway, *Proceedings 6th World Congress and Railway Research*, Edinburgh, UK, 28–30 September, 1403–18.

Bower, A. F., Johnson, K. L. (1991), Plastic flow and shakedown of the rail surface in repeated wheel-rail contact, *Wear*, **144**, 1–14.

BSI (1985) *BS 11:1985 Specification for railway rails*, British Standards Institution, London, UK.

BSI (2003) *BS EN 13674-1:2003 +A1:2007 Railway applications. Track. Rail. Vignole railway rails 46 kg/m and above*, British Standards Institution, London, UK.

Callender, W. R. (1983), *Bainitic steels for Railway Applications*, PhD Thesis, University of Sheffield, Sheffield, UK.

Callender, W., R., Sawley, K. J., Brook, R. (1983a), *Bainitic steels,* Patent Application GB8320622.

Callender W. R., Sawley K. J. and Brook R. (1983b), *Bainitic steels*, EU patent application EP0136004.

Callister, W. D. (1994), *Materials Science and Engineering: An Introduction* (3rd edn), Wiley, New York, USA.

Canadinc, D., Sehitoglu, H., Maier, H. J., Kurath, P. (2008), On the incorporation of length scales associated with pearlitic and bainitic microstructures into a visco-plastic self-consistent model, *Materials Science and Engineering A*, **A485**, 258–71.

Carroll, R. I., Beynon, J. H. (2006), Decarburisation and rolling contact fatigue of a rail steel, *Wear*, **260**, 523–37.

Carroll, R. I., Beynon, J. H. (2007a), Rolling contact fatigue of white etching layer: Part 1 – crack morphology, *Wear*, **262**, 1253–66.

Carroll, R. I., Beynon, J. H. (2007b), Rolling contact fatigue of white etching layer: Part 2 – Numerical results, *Wear*, **262**, 1267–73.

Clayton, P., Allery, M. B. P., (1982), Metallurgical aspects of surface damage problems in rails, *Canadian Metallurgical Quarterly*, **21**, 31–46.

Clayton, P., Jin, N. (1996), Unlubricated sliding and rolling/sliding wear behaviour of continuously cooled, low/medium carbon bainitic steels, *Wear*, **200**(1–2), 74–82.

Clayton, P., Su, X. (1996), Surface initiated fatigue of pearlitic and bainitic steels under water lubricated rolling/sliding contact, *Wear*, **200**(1–2), 63–73.

Clayton, P., Sawley, K. J., Bolton, P. J., Pell, G. M. (1987), Wear behaviour of bainitic steels, *Wear*, **120**(2), 199–220.

De Boer, H., Datta, S. R., Kaiser, H-J., Lundgren, S. O., Müsgen, B., Schmedders, H., Wick, K. (1995), Naturharte bainitische Scheinen mit hoher Zugfestigheit, *Stahl und Eisen*, **115**(2), 93–8 (p.138 English summary).

Devanathan, R., Clayton, P. (1991), Rolling-sliding wear behaviour of three bainitic steels, *Wear*, **151**(2), 255–67.

Eden, H. C., Garnham J. E., Davis, C. L. (2005), Influential microstructural changes on rolling contact fatigue crack initiation in pearlitic rail steels, *Materials Science and Technology*, **21**(6), 623–9.

Fegredo, D. M., Shehata, M. T., Palmer, A., Kalousek, J. (1988), The effect of sulphide and oxide inclusions on the wear rates of a standard C-Mn and a Cr-Mo alloy rail steel, *Wear*, **126**, 285–306.

Fletcher, D. I., Beynon, J. H. (2000), The effect of intermittent lubrication on the fatigue life of a pearlitic rail steel in rolling-sliding contact, *Proceedings of the IMechE, Part F (Journal Rail and Rapid Transit)*, **214**(F3), 145–58.

Franklin, F. J., Weeda, G-J., Kapoor, A., Hiensch, E. J. M. (2005), Rolling contact fatigue and wear behaviour of the 'Infrastar' two-material rail, *Wear*, **258**, 1048–54.

Garnham, J. E. (1995), *The Wear of Bainitic and Pearlitic Steels*, PhD Thesis, University of Leicester, Leicester, UK.

Garnham J. E., Beynon, J. H. (1992), Dry rolling-sliding wear of bainitic and pearlitic steels, *Wear*, **157**(1), 81–109.

Garnham J. E., Davis, C. L. (2006) The role of deformed microstructure on rolling contact fatigue initiation, *Proceeding 7th International Conference on Contact Mechanics and Wear of Rail/Wheel Systems*, Brisbane, Qld, Australia, 24–27 September, 399–409.

Garnham, J. E., Franklin, F. J., Fletcher, D. I., Davis, C. L., Kapoor, A., (2007), Predicting the life of rail steels, *Proceedings of the IMechE, Part F (Journal of Rail and Rapid Transit)*, **221**, 45–58.

Ghonem, H., Kalousek, J., Stone, D. (1982), On the fracture and wear characteristics of chromium-molybdenum bainitic steel in Comins, N. R. and Clark, J. B. (eds), *Specialty Steels and Hard Materials*, Pergamon Press, Oxford, UK and New York, USA, 259.

Gladman, T. (1992), Medium/high carbon steels for rails, rods, bars and forgings, *Volume 7, Materials Science & Technology: A Comprehensive Treatment* (Cahn, R. W., Haasen, P. and Kramer, E. J. (eds)), VCH-Wilers Weinheim, Germany, 401–32.

Harbraken, L. J., Economopoulos, M. (1967), Bainitic microstructures in low carbon alloy steels and their mechanical properties, *Proceedings International Symposium on Transformation and Hardenability in Steels*, University of Michigan, Ann Arbor, MI, USA, Climax Molybdenum Co., Ann Arbor, MI, USA.

Heller, W., Schweitzer, R. (1982), Hardness, microstructure and wear behaviour of rail steels, Proceedings *2nd International Heavy Haul Railways Conference*, Colorado Springs, CO, USA, 25–29 September, 282–86.

Hernández, F. C. R., Demas, N. G., Davis, D. D., Polycarpou, A. A., Maal, L. (2007), Mechanical properties and wear performance of premium rail steels, *Wear*, **263**, 766–72.

Hiensch, E. M. J., Kapoor, A., Josefson, B. L., Ringsberg, J. W., Nielson, J. C. O., Franklin, F. J. (2005), Two-material rail development: field test results regarding rolling contact fatigue and squeal noise behaviour, *Wear*, **258**, 964–72.

Hodgson, W. H. (1985), *Rail Metallurgy*, personal communication, British Steel Track Products booklet.

Honeycombe R. W. K. and Bhadeshia, H. K. D. H. (1995), *Steels: Microstructure and Properties* (2nd edn), Metallurgy & Materials Science Series, Edward Arnold, London, UK.

Ichinose, H., Takehara, J., Iwasaki, N., Ueda, M. (1978). An investigation on contact fatigue and wear resistance behaviour in rail steels, *Proceedings 1st Heavy Haul Railways Conference*, Perth, WA, Australia, 18–22 September, 1–10.

Institute of Rail Welding (2006), *Good news on failures and rejects*, February, available at http://www.iorw.org/article-feb0601b.html, accessed March 2009.

Iwano, K., Ueda, M., Karimine, K., Yamamoto, T. (2006), Recent development of rails in Nippon Steel, *Proceedings 7th International Conference on Contact Mechanics and Wear of Rail/Wheel Systems*, Brisbane, Qld Australia, 24th–27th September, 287–93.

Jin, N., Clayton, P. (1997), Effect of microstructure on rolling/sliding wear of low carbon bainitic steels, *Wear*, **202**(2), 202–7.

Kapoor, A., Franklin, F. J., (2000), Tribological layers and the wear of ductile materials, *Wear*, **245**, 204–15.

Kavishe, F. P., Baker, T. J. (1986), Effects of prior austenite grain size and pearlite inter-lamellar spacing on the strength and fracture toughness of a eutectoid rail steel, *Materials Science & Technology*, 2, 816–22.

Kern, A., Schmedders, H., Zimmerman, A., The development of bainitic steels for special railway system requirements, *Proceedings 4th International Sympostium on 'Rail Steels' as part of the 39th International Symposium on 'Mechanical Working and Steel Processing'*, 20–21 Oct 1997, Indianapolis, IN, USA, Iron & Steel Society, Warrender, PA, USA and Association of American Railroads, Washington DC, USA, 1015–22.

Krauss, G. (1980), *Principles of Heat-Treatment of Steels*, ASM International, Metals Park, OH, USA.

Krauss, G. (1992), Microstructure and transformations in steels, in Pickering F. B. (ed.), *Materials Science & Technology: A Comprehensive Treatment*, Cahn, R. W., Haasen, P. and Kramer, E. J. (eds), VCH-Wiley, Weinheim, Germany, 1–40.

Kristan, J., Sawley, K. J., Canadine, D., Lee, K. M., Polycarpou A. A., Sehitoglu, H. (2003), Wear and rolling contact fatigue in bainitic steel microstructures, *Proceedings 6th International Conference on Contact Mechanics and Wear of Rail/Wheel Systems* Gothenburg, Sweden, 10–13 June, 413–22.

Langford, G. (1977), Deformation of Pearlite, *Metallurgical Transactions A*, **8A**, 861–75.

Liu, C. D., Bassim, M. N., St. Lawrence, S. (1993), Evaluation of fatigue-crack initiation at inclusions in fully pearltic steels, *Materials Science and Engineering*, **A167**, 107–13.

Llewellyn, D. T. (1992), *Steels: Metallurgy and Applications*, Butterworth-Heinemann, Oxford, UK.

Lojkowski, W., Djahanbakhsh, M. Bürkle, G. Gierlotka, S. Zielinski, W. Fecht, H-J.,

(2001), Nanostructure formation on the surface of railway tracks, *Materials Science and Engineering* **303A**, 197–208.

Perez-Unzueta, A. J. (1992), *Wear Resistance of Pearlitic Rail Steel.*, PhD Thesis, University of Leicester, Leicester UK.

Perez-Unzueta, A. J., Beynon, J. H. (1993), Microstructure and wear resistance of pearlitic rail steels, *Wear*, **162–164**, 173–82.

Pickering, F. B. (1992), Structure–property relationships in steels/high strength low alloy steels, in Pickering, F. B. (ed.), *2, Volume 7, Materials Science & Technology: A Comprehensive Treatment*, Cahn R W, Haasen P and Kramer EJ (eds) VCH Wiley, Weinheim, Germany, 41–44.

Pointner, P. (2006), High strength rail steels – The importance of material properties in contact-mechanics problems, *Proceedings 7th International Conference on Contact Mechanics and Wear of Rail/Wheel Systems*, Brisbane, Qld Australia, 24–27 September, 295–307.

Satoh, Y., Iwafuchi, K. (2005), Crystal orientation analysis of running surface of rail damaged by rolling contact, *Wear*, **258**, 1126–34.

Sawley, K. J., Benyon, J. A., Jones, E. G. (1988), Bainitic steels for railway wheels, *Proceedings 9th International Wheelset Congress*, Montreal, QC, Canada, 12–15 Septermber, 1–12.

Sawley, K. J. (1989), Materials in railway engineering, *Metals and Materials*, April, 210–14.

Sawley, K. J., Scholl, M. (1998), The development of bainitic steels for rails, *Proceedings of the 4th International Symposium on 'Rail Steels' as part of the 39th International Symposium on 'Mechanical Working and Steel Processing'*, 20–21 Oct 1997, Indianapolis, IN, USA, Iron & Steel Society, Warrender, PA, USA and Association of American Railroads, Washington DC, 1007–14.

Shur, E. A., Bychkova, N. Y., Trushevsky, S. M., (2005), Physical metallurgy aspects of rolling contact fatigue of rail steels, *Wear*, **258**, 1165–71.

Su, X. C., Clayton, P. (1996), Surface-initiated rolling contact fatigue of pearlitic and low carbon bainitic steels, *Wear*, **197**(1–2), 137–44.

Tyfour, W. R., Beynon, J. H. (1994a), The effect of rolling direction reversal on fatigue crack morphology and propagation, *Tribology International* **27**, (4), 273–82.

Tyfour, W. R., Beynon, J. H. (1994b), The effect of rolling direction reversal on the wear rate and wear mechanism of pearlitic rail steel, *Tribology International*, **27**(6), 401–12.

Tyfour, W. R., Beynon, J. H., Kapoor, A. (1996), Deterioration of rolling contact fatigue life of pearlitic rail steel due to dry-wet rolling-sliding line contact, *Wear*, **197**, 255–65.

Uchino, K., Ueda, M., Kutaragi, K., Babazono, K. (1998), Development of hyper-eutectoid steel rail for heavy haul railways, *Proceedings 4th International Symposium on 'Rail Steels' as part of the 39th International Symposium on 'Mechanical Working and Steel Processing'*, 20–21 Oct 1997, Indianapolis, IN, USA, Iron & Steel Society, Warrender, PA, USA and Association of American Railroads, Washington DC, USA 1047–56.

UIC (1986), *Technical specification for the supply of rails*, UIC leaflet 860–0, 8th edn (1986-07-01), International Union of Railways, Paris, France.

Umemoto, M., Liu, Z. G., Sugimoto, S., Tsuchiya, K. (2000), Tensile stress-strain analysis of single-structure steels, *Metallurgy and Materials Transactions A*, **31A**, 1785–94.

Umemoto, M., Liu, Z. G., Tsuchiya, K., Sugimoto, S., Bepari, M. M. A. (2001), Relationship between hardness and tensile properties in various single structured steels, *Materials Science and Technology*, **17**, 505–11.

Wetscher, F., Vorhauer, A., Stock, R., Pippan, R. (2004), Structural refinement of low alloyed steels during severe plastic deformation, *Materials Science and Engineering* **387–389A**, 809–16.

Wetscher, F., A., Stock, R., Pippan, R. (2007), Changes in the mechanical properties of a pearlitic steel due to large shear deformation, *Materials Science and Engineering*, **445–446A**, 237–43.

Yokoyama, H., S. Mitao, Yamamoto, S. and Fujikake, M. (2002), Effect of the angle of attack on flaking behaviour in pearlitic and bainitic rail steels, *Wear*, **253**(1), 60–66.

Zhang, F., Lv, B., H, B., Li, Y. (2007), Flash butt welding of high manganese steel crossing and carbon steel rail, *Materials Science and Engineering*, **454–455A**, 288–92.

Copyrighted Materials
Copyright © 2009 Woodhead Publishing Retrieved from www.knovel.com

6
Railway wheel wear

F. BRAGHIN and S. BRUNI, Politecnico di Milano, Italy;
R. LEWIS, University of Sheffield, UK

Abstract: This chapter presents the basic principles governing the process of regular wear in railway wheels that represents one of the major sources of maintenance costs in the management of a fleet of railway vehicles and has direct implications on rolling stock performance as well as on vehicle impact on the track. To be able to predict wheel wear, a multidisciplinary approach is required since tribology, behaviour of metallic materials, solid mechanics and vehicle dynamics have to be accounted for. A review of existing modelling approaches is proposed and a description of possible measures to reduce wheel wear and its effects is presented.

Key words: wear process, wear testing, wear regimes, wear modelling, effect of wear on railway dynamics, means to reduce wear.

6.1 Introduction

Wear of wheel profiles is one of the most critical issues affecting cost and performance of railway transportation systems. Since the early days of the railways, wheel profile design has been driven by addressing problems of strength and wear, as well as by ensuring ride safety and performance (Wickens, 2003). By altering the profile from the shape originally chosen by the designer, wear has a very large effect on vehicle dynamics (see Chapter 15), so that excessive modifications of the original profile have to be recovered by grinding the wheel to its original profile. Clearly, a limited number of re-profiling operations are possible for a wheel, before the height of the rim is reduced to a value which is no longer compatible with the mechanical resistance of the wheel. When this limit is reached, the wheel has to be replaced. Wheel re-profiling and wheel replacement have a very large impact on maintenance costs and therefore a large number of investigations have been carried out to mitigate this problem.

The term 'wheel wear' is generally used to denote any kind of damage occurring on the rolling surface of railway wheels which involves loss of material. This can be caused by the effect of adhesive and/or abrasive wear, rolling contact fatigue (RCF) and (to a minor extent) material relocation due to plastic deformation. Since RCF effects on railway wheels are specifically addressed by Chapter 7, this chapter will focus on the effect of adhesive and abrasive wear.

Two main manifestations of wear typically occur on railway wheels: a

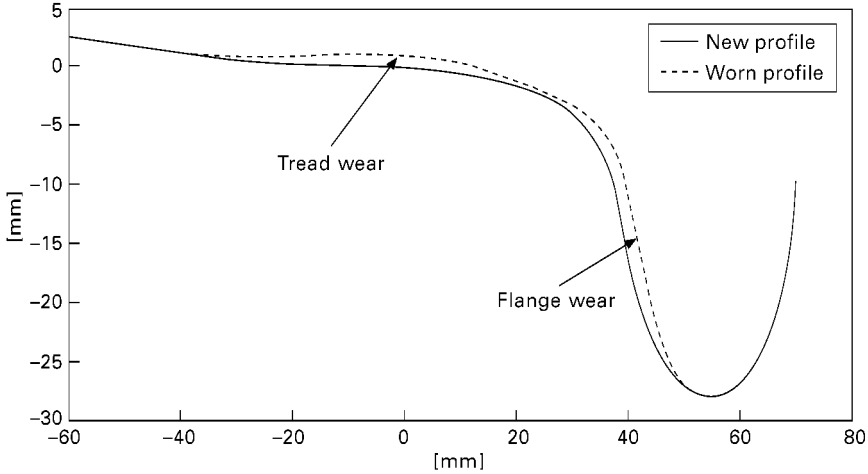
change in the transversal profile, called 'regular wear', and the formation of periodic wear patterns in circumferential direction, called wheel out-of-roundness or 'irregular wear' (Enblom, 2006). Regular wear is produced by slow variations of contact forces and creepages associated with the longitudinal and lateral motion of the wheelset on the track. On the other hand, irregular wear is produced by fast variations of wheel–rail contact conditions, which are produced by train–track interaction and are strongly affected by the vertical vibrations of the system. In this chapter, the presentation will be confined to regular wear, since the effect of wheel out-of-roundness is covered by Chapter 8.

The aim of this chapter is hence to describe the 'regular wear' process occurring on railway wheels. Section 6.2 outlines the main features of this phenomenon, which has a 'natural feedback' nature, introducing the main ingredients for its understanding. Section 6.3 focuses on tribological issues, describing wear regimes and how they relate to wheel–rail contact, as well as wear testing techniques that are used to characterise material behaviour. Section 6.4 addresses wheel–rail contact mechanics in view of its effect on wheel wear: a presentation is made of the most widely used theories and models describing the relationship between the local contact parameters affecting the wear process, the geometry of wheel–rail contact and global contact parameters, namely wheel–rail contact forces and creepages. Section 6.5 reviews the state-of-the-art of wheel wear modelling. Finally, Section 6.6 reviews some possible means to reduce wheel wear.

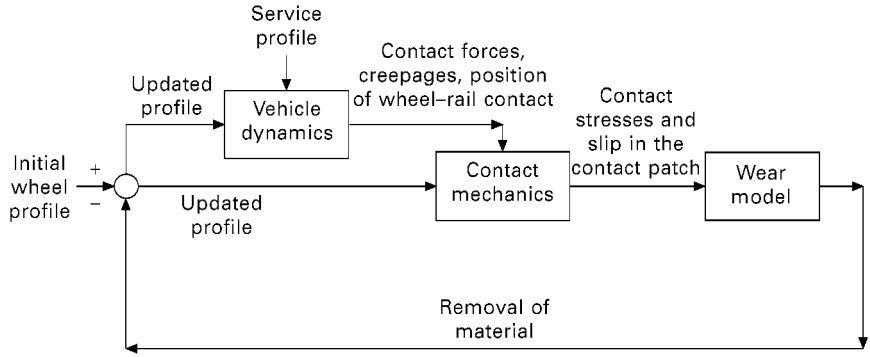
6.2 The wheel wear process

Regular wear of railway wheels is the result of the wheel exchanging frictional forces with the rail. Hence, particularly relevant to the formation of regular wear are vehicle running conditions where large frictional forces occur on the wheel: traction, braking and especially curve negotiation. While in traction and braking, wear is produced by longitudinal tractive/braking efforts and tends to concentrate on the central portion of the wheel tread, the process of wear formation during curve negotiation is more complicated, and (at least in the case of a solid wheelset) involves both longitudinal frictional forces caused by the rolling radius difference between the inner and outer wheels and transversal frictional forces caused by the wheelset travelling along the curve with a non-zero angle of attack (Roberti and Bruni, 2001; Wickens, 2003).

As a consequence of this complex wheel–rail contact condition, two distinct wear regions appear on the wheel profile: one on the flange and the other the thread, see Fig. 6.1. Wear rates in these two regions depend on a variety of factors, such as the vehicle design, the service profile (e.g. track curvature, cant deficiency, etc.), the status of the contacting surfaces



6.1 Wear regions on the wheel profile.



6.2 Schematic representation of the wheel wear process.

(including the possible presence of natural or artificial lubrication) and the mechanical properties of wheel and rail materials. The effect of some of these parameters is discussed in more detail in Section 6.6.

The flowchart presented in Fig. 6.2 provides a schematic representation of the process of wheel wear formation (Braghin *et al.*, 2006; Enblom, 2006): starting from the initial rail profile, and considering a specific service condition, vehicle dynamics gives rise to dynamic variations of wheel loads, creepages and frictional forces, and to a migration of the contact point across the surfaces of the wheel and of the rail. This 'global' information about wheel–rail contact corresponds to 'local' information regarding the shape, location and size of the contact patch, and the distribution of the normal and shear stresses and of the slip in the contact area, the correspondence between 'global' and 'local' contact parameters being established by the laws

of contact mechanics and depending upon the actual shape of the contacting bodies. Finally, the action of normal and tangential stresses in the presence of slip is responsible for material removal, according to a given wear law that takes into account the behaviour with respect to wear of the wheel material, the wear regime established as a function of the contact conditions, the presence of contaminants creating natural or artificial lubrication.

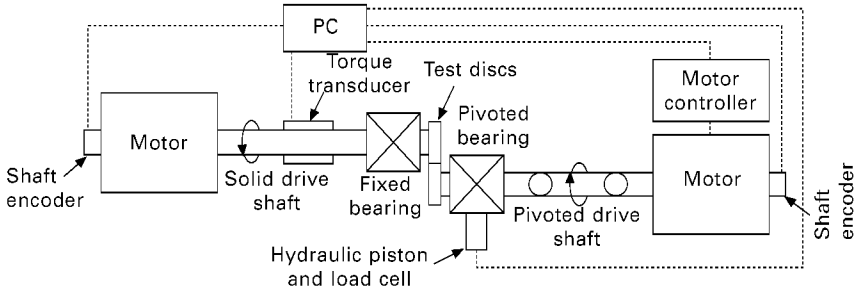
The removal of material results in a modification of the wheel profile, which affects both vehicle dynamics and contact mechanics, so that the wheel wear process takes the form of a natural feedback loop. Unfortunately, each of the three main blocks of the feedback loop (vehicle dynamics, contact mechanics and wear model) establishes a highly non-linear relationship between input and output, which makes it difficult to understand how the process will evolve with cumulated service and how the change of any relevant parameter (like e.g. initial wheel profile, service condition, vehicle design, etc.) will affect the progress of wear.

6.3 Tribological issues in wheel wear

6.3.1 Wear testing

A number of different techniques have been used for studying wear occurring in the wheel–rail contact. The simplest of these are specimen tests carried out on standard tribological test equipment such as pin-on disc machines (Olofsson and Nilsson, 2002). This type of rig will give a contact with 100 % slip, which would represent only the most severe conditions found in very tight curves. In order to vary this slip value and better replicate the wheel–rail contact, twin-disc testing can be used (see for example Bolton and Clayton, 1984; Krause and Poll, 1986; Garnham and Beynon, 1992; Lewis and Dwyer-Joyce, 2004). In this type of test, two discs made from wheel and rail material are loaded together (in parallel) and rolled at varying speeds to give the required slip. An example of such a rig is shown in Fig. 6.3. The test discs are hydraulically loaded together and driven at controlled rotational speed by independent electric motors. Shaft encoders monitor the speeds continuously. A torque transducer is assembled on one of the drive shafts and a load cell is mounted beneath the hydraulic jack. The slip ratio required is achieved by adjustment of the rotational speeds. Further details of the machine can be found in Lewis and Dwyer-Joyce (2004).

A further refinement of this testing method has involved using larger discs machined to give scaled down wheel and rail profiles (see for example Kumar and Rao, 1984). This has the added advantage of better replicating the actual contact shape and slip distribution within it. Full-scale roller test rigs have also been used. Examples of these include rigs where a wheel is repeatedly rolled along a short piece of rail (McEwen and Harvey, 1985)



6.3 Schematic diagram of the SUROS twin disc test machine. (For more details see Lewis and Dwyer-Joyce, 2004)

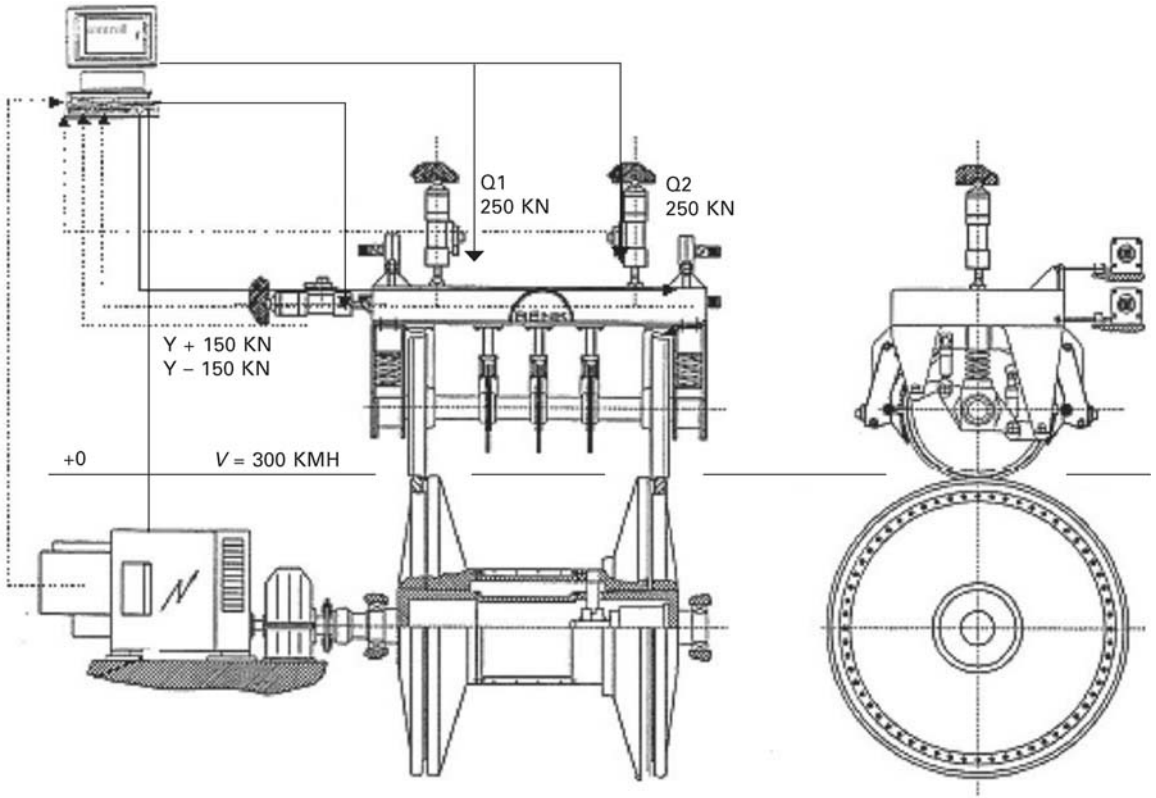
or larger rigs where whole wheelsets are loaded and rolled against rotating loops of rail, an example of which is shown in Fig. 6.4.

This is the BU300 roller rig, owned by Lucchini RS. It is composed of two discs driven by a DC motor and bearing rail profiled steel rings. The wheelset is placed on these discs and is connected, by the primary suspension, to a transverse beam representing the half-bogie. Two hydraulic actuators are placed vertically over the transversal beam, one for each side, and are used to apply different vertical loads on each wheel, thus reproducing the vertical load acting on the wheelset as well as the load transfer from one wheel to the other occurring during curve negotiation. A third hydraulic actuator applies a lateral force on the transversal beam to reproduce the lateral force acting on the wheelset in different service conditions such as curve negotiation. At each side of the transversal beam two electric servomotors are placed longitudinally at two different heights. These actuators are used to control the yaw movement of the transversal beam and therefore the wheelset's angle of attack. More details on the test rig can be found in Bruni *et al.* (2000).

This kind of apparatus, while of limited use for carrying out a large-scale investigation into different materials and wear mechanisms, is essential in validating multi-body models produced to simulate wheel profile evolution (see for example Braghin *et al.*, 2006) as all test conditions are known and can be modelled accurately, something that is impossible with trains running in service.

Field measurements, however, are still useful and have been used to study the causes of wheel and rail wear (Dearden, 1960). Additionally, a large amount of data have been gathered from simulated field experiments carried out on specially built test tracks (Steele, 1982).

It is clear that a hierarchy of tests exists from very simple tests, where test parameters can be easily controlled and measurements are more straightforward to take, but where the contact simulation is less accurate, to highly complex

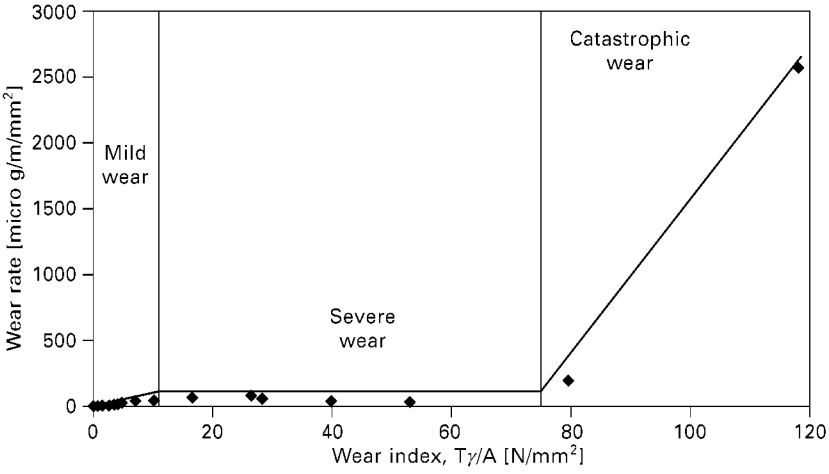


6.4 Schematic diagram of the BU300 full-scale roller rig for mounted wheelsets.

tests where test parameter control is hard to achieve, but where actual contact conditions are being used. It is important though, when trying to rank wear behaviour or isolate wear mechanisms, that control can be achieved during tests and that they are repeatable. For this reason and because it offers the ability to test a wide range of materials at lower cost, the twin-disc approach has been used more than most. This type of approach has also been widely used for studying RCF, adhesion and wheel–rail isolation.

6.3.2 Wear regimes

Twin-disc testing carried out to study the wear behaviour of railway wheel and rail steels has led to the identification of a number of wear regimes (Beagley, 1976; Bolton *et al.*, 1982; Kumar and Rao, 1984; Bolton and Clayton, 1984; Lewis and Dwyer-Joyce, 2004). Early tests demonstrated that two wear regimes existed (Beagley, 1976; Bolton *et al.*, 1982). These were designated *mild* and *severe*. Later work led to the identification of a further regime, designated *catastrophic* wear (Bolton and Clayton, 1984; Lewis *et al.*, 2004). Figure 6.5 shows results of twin-disc testing of R8T wheel steel against UIC60 900A rail steel (data from Lewis and Dwyer-Joyce, 2004). The results are plotted in terms of wear rate (μg material lost/m rolled/mm^2 contact area) against an index based on work done in the contact, $T\gamma/A$, where T is tractive force (normal load divided by friction coefficient), γ is slip and A is contact area. This is a useful way to plot the data as it allows comparison of different test geometries. Wheel tread and railhead wear are thought to fall within the mild regime and wheel flange and rail gauge corner

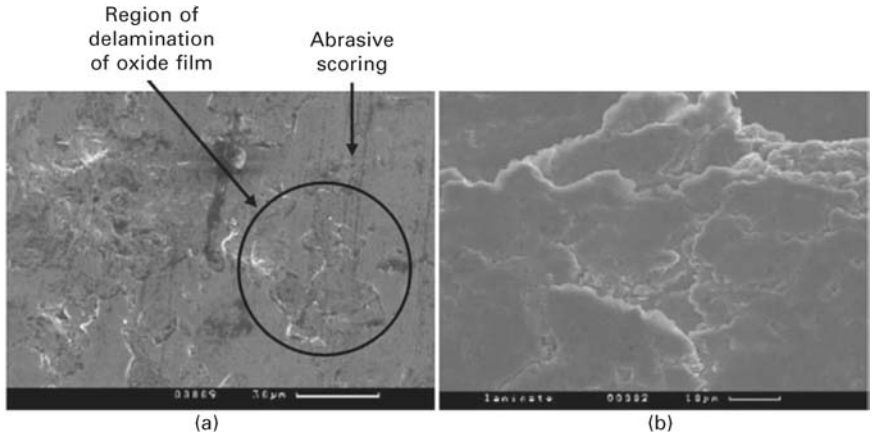


6.5 Wear rates and regimes for R8T wheel steel tested against UIC60 900A rail steel. (Lewis and Dwyer-Joyce, 2004)

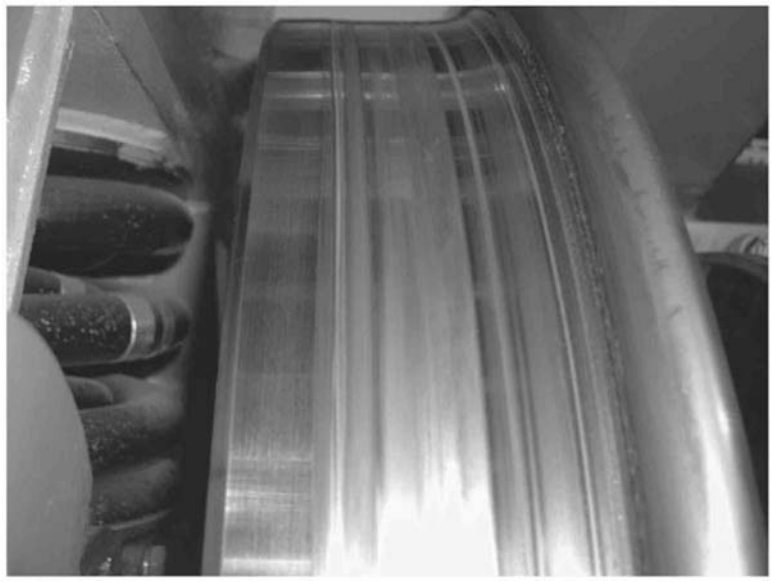
wear in the severe or catastrophic regime. This has been verified with field measurements carried out by Olofsson and Nilsson (2002).

The regimes for wheel materials have been characterised in terms of wear rate and wear mechanism (Lewis and Dwyer-Joyce, 2004). At low $T\gamma/A$, in the mild wear regime, oxidative wear occurs on both wheel and rail discs. The disc surfaces turned a rusty brown colour. Closer examination of the wear surfaces revealed abrasive score marks and evidence of the oxide layer breaking away from the surface (see Fig. 6.6a). This ties in with observations made in the field that on straight track where low slip occurs on the high rail, oxidative wear is prevalent generating magnetite (Fe_3O_4) (Broster *et al.*, 1974) and in full-scale test-rig results, where reddish oxide film appeared for low slip conditions (see Fig. 6.7) (McEwen and Harvey, 1985; Braghin *et al.*, 2006). As $T\gamma/A$ was increased, the wear mechanism of the wheel discs altered. The wheel disc appeared to be wearing by a ratcheting process (deformation followed by crack growth and subsequent material removal). Closer examination of the wheel disc surfaces revealed that this was the case (see Fig. 6.6b).

Figure 6.8a shows a section through a wheel disc, run at low $T\gamma/A$, parallel to the rolling direction. At the surface the oxide layer is just visible. There is a very small amount of deformation just below the wear surface of the disc. At higher levels of $T\gamma/A$, observation of the subsurface morphologies revealed that a larger amount of plastic deformation was occurring below the wheel disc wear surface (see Fig. 6.8b) and crack formation just below the surface was visible which was leading to thin slivers of material breaking away from the surface. The slivers have a similar thickness to the oxide layer and could indicate that a severe oxidative wear mechanism is occurring



6.6 Wheel disc surface run at: (a) $T\gamma/A$ = 0.21; (b) $T\gamma/A$ = 4.1. (Lewis and Dwyer-Joyce, 2004)



6.7 Wheel rolling surface on a full-scale roller rig after 2000 km. (Braghin *et al.*, 2006)

where fracture occurs between the oxide layer and the metal. As $T\gamma/A$ was increased further, far greater cracking was visible at the wear surface and some of these cracks were seen to alter direction from running parallel to the wear surface and turning up to turning down into the material causing larger chunks of material to break away (see Fig. 6.8c).

6.3.3 Wear transitions

While steel wear regimes have been defined well in terms of wear rate, metallographic features and wear debris (see Fig. 6.9), there is not a great understanding of what mechanisms lead to the changes in wear rate that occur. In order to further understand the wear mechanisms, the transitions between these regimes were studied in more detail (Lewis and Dwyer-Joyce, 2004). It was proposed that the first transition is associated with the onset of fully sliding contact conditions and the second is a result of surface temperature effects.

Figure 6.10 shows friction measurements taken during UIC60 900A versus R8T wear tests carried out at 1500 MPa plotted against slip, i.e. a creep curve. As would be expected the friction reaches a threshold. This transition represents the change from partial slip in the disc interface to full slip conditions. Also shown is the Carter creep curve for an assumed limiting

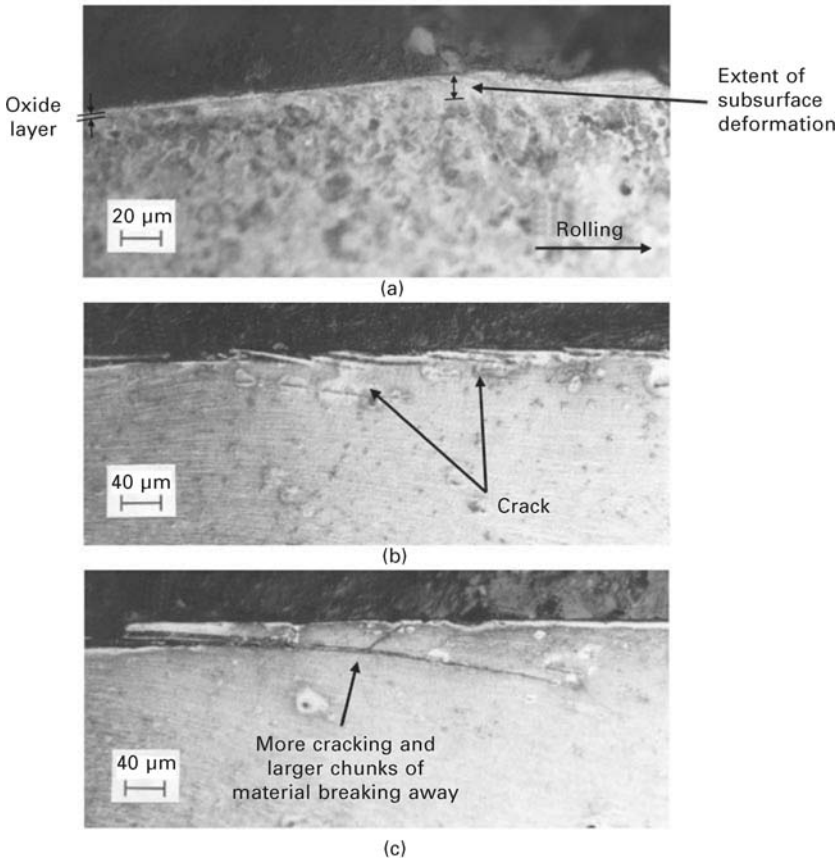
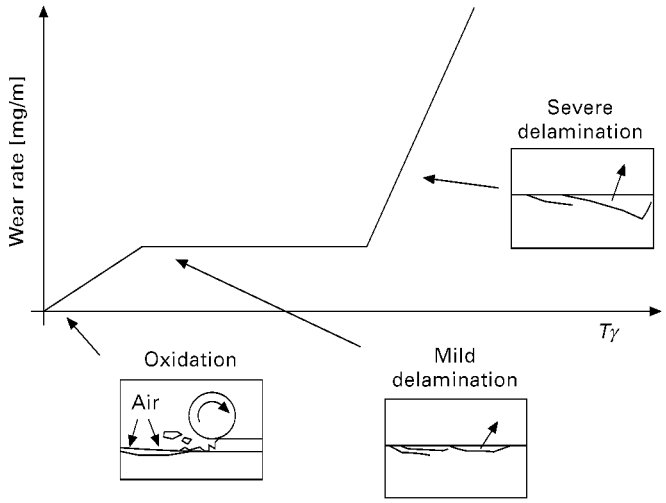


Fig. 6.8 Sections parallel to rolling direction through wheel disc run at: (a) $T\gamma/A$ = 0.21; (b) $T\gamma/A$ = 16.6; (c) $T\gamma/A$ = 28.3. (Lewis and Dwyer-Joyce, 2004)

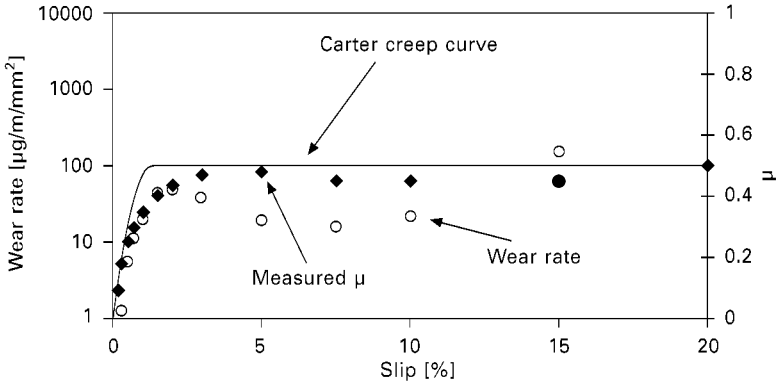
friction of 0.5. This model creep curve is based on smooth elastic cylinders in contact (Carter, 1926).

The wear data (also shown in Fig. 6.10) follow a similar pattern indicating that at the point of transition from partial slip to full slip a wear transition also occurs. After the full slip condition has been reached, increasing the magnitude of slip has no effect. As shown in Fig. 6.11, depth of plastic deformation (from sections of test discs) follows a similar pattern with respect to slip.

At the second wear transition (where the wear data break from the pattern of the friction measurements) another mechanism must be initiated, leading to the much higher wear rates observed. Once the slip has reached about 15 % the wear rate increases rapidly. The theoretically smooth surface stress state, however, remains unaltered since the normal load and traction coefficient



6.9 Schematic diagram of wear features and regimes.

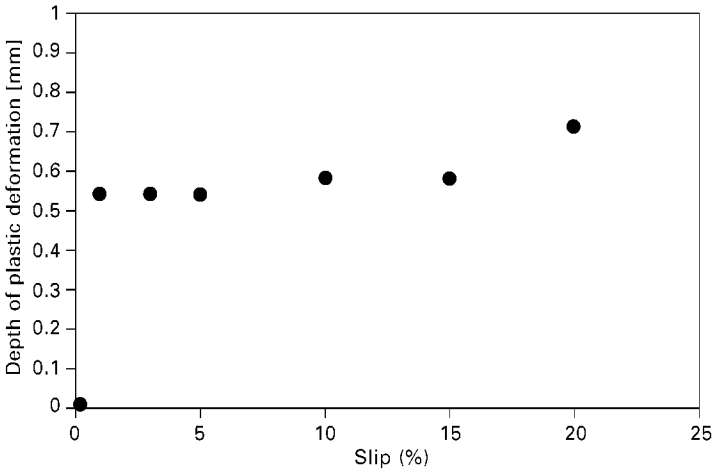


6.10 Friction versus slip in the twin-disc contact (for tests carried out at 1500 MPa). (Lewis and Dwyer-Joyce, 2004)

are unchanged. Clearly the increase in sliding speed has resulted in some mechanism changes. In this section the associated surface temperature rise is considered.

An approach for determining the temperature in a twin-disc experimental contact has been outlined previously (Lewis and Dwyer-Joyce, 2004; Gallardo-Hernandez *et al.*, 2006). The disc surface temperature is considered to be made up of three components:

1. the body temperature, T_b, of the material under the contacting surfaces of the discs, which is the average temperature of the outer surface of the discs;



6.11 Depth of subsurface plastic deformation against slip for R8T wheel material (for twin-disc tests carried out at 1500 MPa). (Lewis *et al.*, 2004)

2. the instantaneous temperature rise above the bulk temperature at the point of contact, called the flash temperature, T_f (Blok, 1963);
3. the ambient temperature of the air surrounding the discs, T_a.

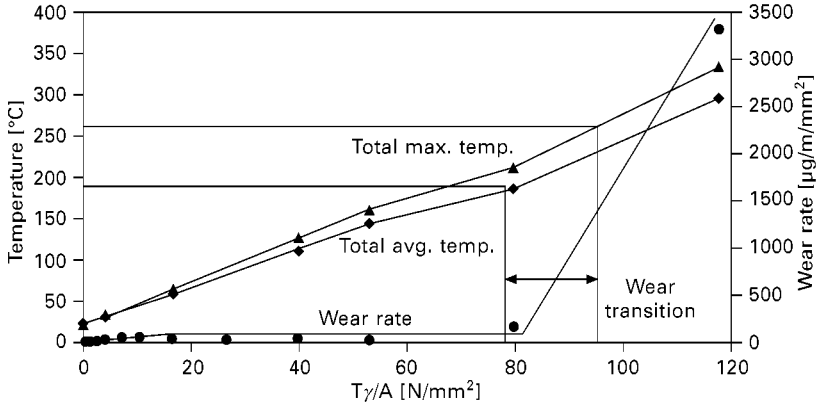
In order to calculate the body (or average surface) temperature of the two discs in contact, T_b, heat generated due to sliding friction in the contact is equated to heat loss due to convection, conduction and radiation giving:

$$\mu F_n u_s = 2(\dot{Q}_1 + \dot{Q}_2 + \dot{Q}_3) \qquad [6.1]$$

where μ is the coefficient of friction, F_n is the normal load applied to the discs, u_s is the sliding speed and $\dot{Q}_1$, $\dot{Q}_2$ and $\dot{Q}_3$ are heat loss due to convection, conduction and radiation, respectively. The dissipated heat is multiplied by two as there are two discs. Calculations were carried out to determine temperatures in the twin-disc contact for R8T versus UIC60 900A rail steel. The results, shown in Fig. 6.12, indicate that the transition from severe to catastrophic wear occurs between 200 and 300 °C. These temperatures correspond with those causing a drop in the yield strength of carbon–manganese steels similar to wheel and rail steels (British Steel Makers Creep Committee, 1973).

6.3.4 Wear mapping

While using the $T\gamma A$ method for plotting wear rate data enables wear transitions to be identified easily and comparisons of different material combinations to be made, it does not help in fully understanding how the



6.12 Twin-disc contact temperatures and wear coefficients for UIC60 900A rail material versus R8T wheel material. (Lewis and Dwyer-Joyce, 2004)

individual contributions of different parameters such as contact pressure and slip affect wear rate. In order to allow a more complete analysis of the effect of individual parameters a mapping method was used for plotting wear data. The approach used was similar to that developed by Lim and Ashby (1987) for mapping sliding wear mechanisms and previously used for studying rail steel wear transitions by Lewis and Olofsson (2004). Wear coefficients were calculated from the wheel steel wear data using Archard's equation (Archard, 1953):

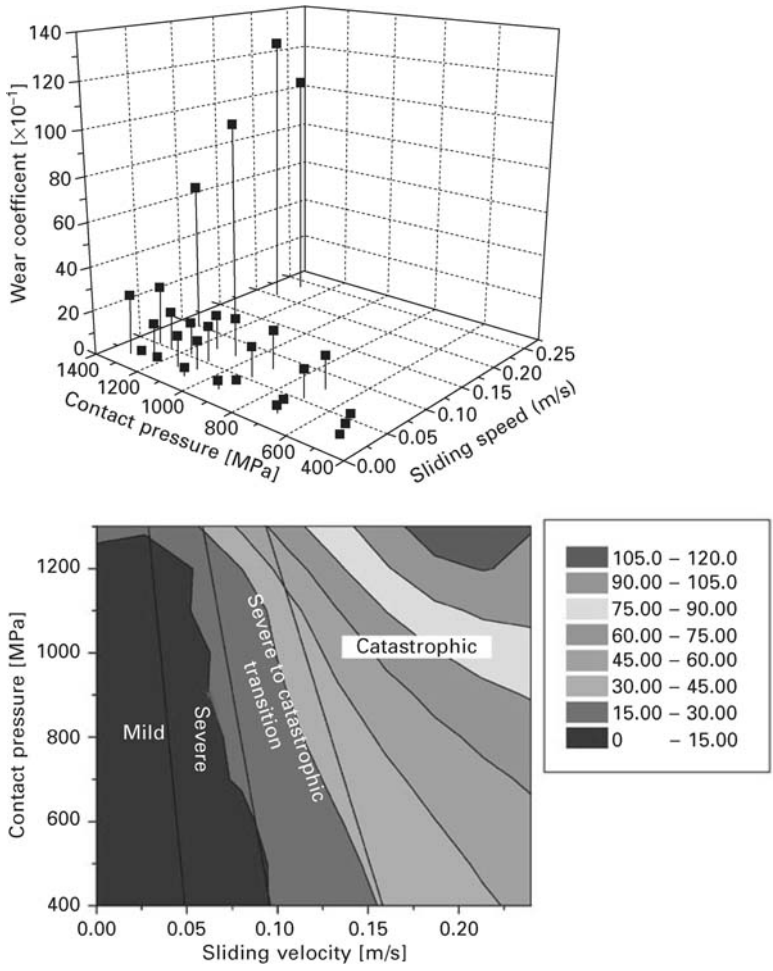
$$K = \frac{Vh}{Pd} \qquad [6.2]$$

where K is the wear coefficient, V is the wear volume, P is the normal load, d is the sliding distance and h is the material hardness.

Wear coefficients were then plotted against contact pressure and sliding speed in the contact. Two types of plots were constructed; contour maps and 3D point graphs. Obviously the accuracy of the contour map is limited by the amount of data available. The accompanying 3D graphs give an indication of where data are lacking on a particular map. Transitions based on those outlined above were marked on the contour plots.

Figure 6.13 illustrates data from twin-disc testing using Class D tyre versus BS11 rail material (details in Bolton and Clayton, 1984). Data were available at high sliding velocities giving a good picture of wear likely under severe wheel–rail contact conditions. Figure 6.14 shows data from twin-disc testing for R8T wheel material versus UIC60 900A rail (for details see Lewis and Dwyer-Joyce, 2004).

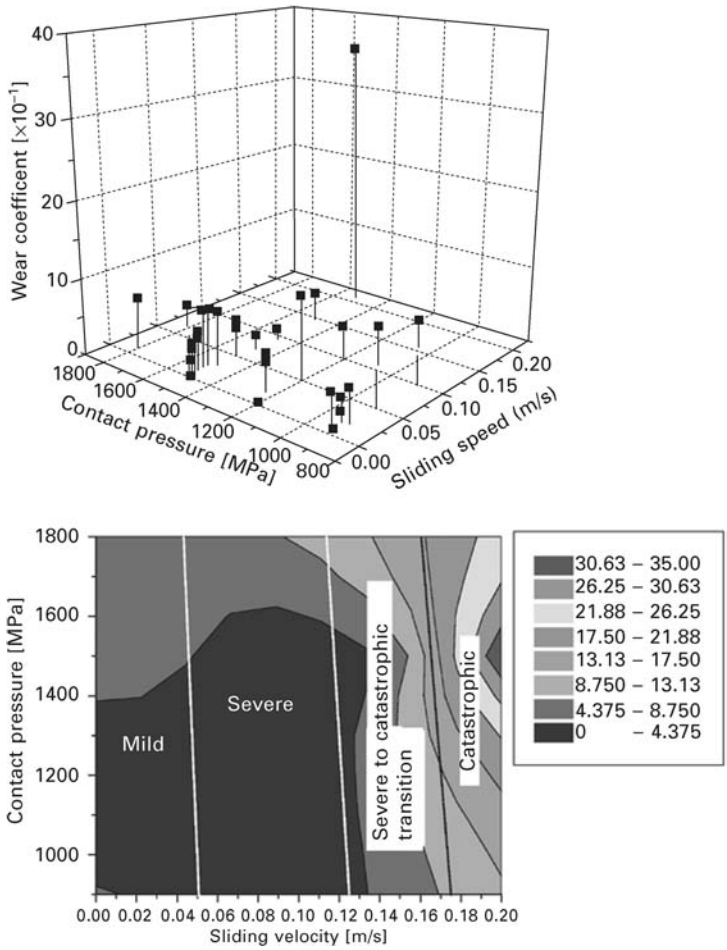
In previous work on wear of wheel steel no attempt has been made to correlate wear data to wheel–rail contact conditions. The wheel–rail contact



6.13 Wear coefficient maps for Class D tyre material versus BS11 rail material. (Lewis *et al.*, 2004) (data from Bolton and Clayton, 1984)

conditions illustrated in Fig. 6.15 resulted from a study using GENSYS train dynamic modelling software (Jendel, 2002). As can be seen a clear difference exists between the railhead–wheel tread and the rail gauge–wheel flange contacts.

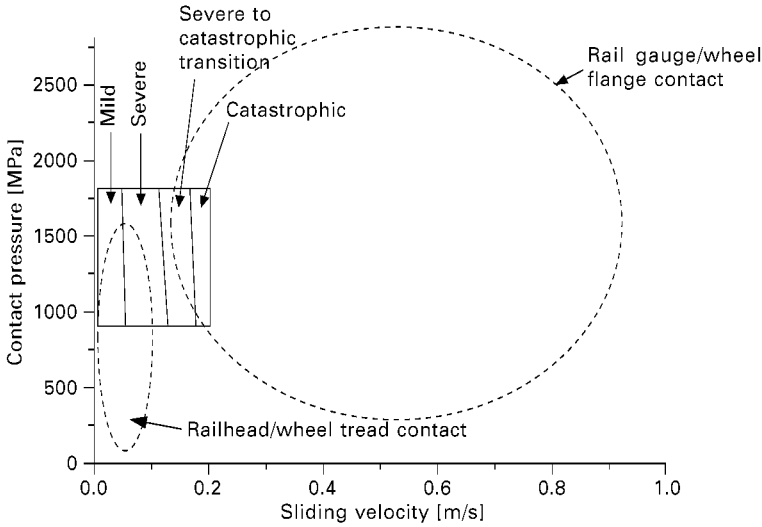
To study how the wear regimes identified above fit in with the wheel–rail contact conditions, the wear map of R8T wheel steel versus UIC60 900A rail steel has been overlaid (from Fig. 6.14). This indicates that the railhead–wheel tread contact will experience mild to severe wear and the rail gauge–wheel flange contact will experience severe to catastrophic wear. This backs up previous suppositions regarding the wear regimes that the railhead–wheel tread and rail gauge–wheel flange contacts fall into. Full-scale test results



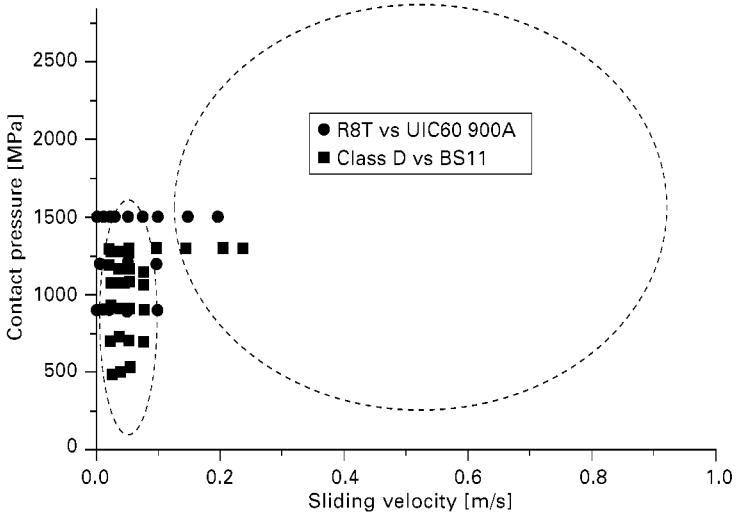
6.14 Wear coefficient maps for R8T wheel material versus UIC60 900A rail steel. (Lewis *et al.*, 2004) (data from Lewis and Dwyer-Joyce, 2004)

support this view (Nilsson, 2000). The wheel flange wear was significantly higher compared with wheel tread wear. It could also be seen that the tread wear rate was higher for powered wheels than for trailing wheels.

Figure 6.16 shows the wear data points collected for the wheel materials in terms of the contact conditions. It can be seen that there is a large amount of wear test data for conditions typical of a railhead–wheel tread contact, but very little for the rail gauge–wheel flange contact. This clearly identifies an area that needs to be addressed in future research, especially as axle loads are increasing and rolling stock is being used on track with low-radius curves as well as the high-radius curves on high-speed lines, which means it is likely



6.15 R8T wheel material wear map (from Fig. 6.10) plotted over wheel–rail contact conditions derived from GENSYS simulations. (Lewis *et al.*, 2004) (Simulation data from Jendel, 2002)



6.16 Available wheel material wear data plotted over typical wheel–rail contact conditions. (Lewis *et al.*, 2004)

that the severity of the wheel–rail contact conditions will rise. Increasing the wear data available will also improve the accuracy and applicability of the wear maps.

6.3.5 Wear modelling

A couple of approaches to modelling the rate of material removal from a wheel material under rolling sliding conditions have been mentioned in previous sections. These and others will be discussed here. The models are those that could be used in the wheel wear modelling process illustrated in Fig. 6.2. The wear indices (WI) developed to date for predicting wear of rails/wheels have been listed below (expected wear is proportional to WI):

$$WI = \mu_f F_f \theta \qquad \text{(Heumann, 1954)}$$

$$WI = \mu_f F_f [(a/r)^2 + (\theta \tan\beta)^2]^{1/2} \qquad \text{(Marcotte } et\ al.\text{, 1978)}$$

$$WI = P\theta/809,305 \text{ (approximate)} \qquad \text{(Marcotte } et\ al.\text{, 1978)}$$

$$WI = 2T_y\theta \qquad \text{(Ghonem and Gonsalves, 1978)}$$

$$WI = K_1\theta + K_2\theta^2 \quad \theta < 20° \qquad \text{(Ghonem and Kalousek, 1982)}$$

$$WI = K_3\theta \quad \theta > 20°$$

$$WI = (\mu_f/\sin\beta)(Y + \mu_a Q_1 \cos\delta)\theta \qquad \text{(Vogel and Kurek, 1981)}$$

$$WI = K(T_1\gamma_1 + T_2\gamma_2) \qquad \text{(Elkins and Eickhoff, 1979) (Derby Wear Index)}$$

where:
μ_f = adhesion coefficient at the flange contact
F_f = flange force
θ = angle of attack
a = radial distance from tread to flange
r = wheel radius
β = flange angle to horizontal
P = tie plate force (N)
T_y = total lateral force external force on wheel
K, K_1, K_2, K_3 = constants
Y = horizontal guide force on the high rail
μ_a = resultant adhesion coefficient on flanging wheel
Q_1 = wheel load
δ = superelevation angle
T_1, T_2 = longitudinal and lateral creep forces (tread and flange creep forces)
γ_1, γ_2 = longitudinal and lateral creepages (tread and flange creepages)

All but one of these indices are either linear or quadratic functions of angle of attack, θ, implying that wear tends to zero as θ tends to zero. Kumar and Rao (1984) showed that these indices were not fully valid as wear was apparent in tests run with an attack angle of zero. The 'Derby' WI, shown below, developed by Elkins and Eickhoff (1979), is based

on the proportionality of wear with work done at the wheel–rail contact. Various researchers have reported wheel–rail wear results using twin-disc test machines of varying geometries that support this approach, including Lewis and Dwyer Joyce (2004). They found that several wear coefficients were required, however, to characterise the different wear regimes (see Fig. 6.5). This approach has been used successfully in predicting wheel profile evolution using the process outlined in Fig. 6.2 (Braghin *et al.*, 2006), as will be discussed in Section 6.5. Another modelling approach has been to use the Archard Equation for wear (given in Eq. 6.2). Enblom has used this model in successfully predicting wheel wear (Enblom, 2006).

6.4 Wheel–rail contact mechanics and its effect on wheel wear

As discussed in Section 6.3, material removal from the wheel surface is a function of the contact stresses and sliding between the wheel and rail surfaces. Therefore, the detailed analysis of contact phenomena occurring at wheel–rail interface is a key ingredient in view of understanding and modelling wheel wear. In this section, the fundamental issues concerning wheel–rail contact mechanics will be introduced, confining the presentation to those areas of direct importance for the wheel wear process. It is not the purpose of this section to provide an extensive description of wheel–rail contact modelling approaches (the reader is referred to Chapter 4 for a comprehensive treatment of this subject), but instead to describe how the local shape of the contacting bodies and the global forces exchanged will affect material removal. To this end, a brief introduction of the most widely used approaches to contact modelling is given and will serve as an introduction to the state-of-the-art survey of wheel wear modelling reported in Section 6.5.

The wheel–rail contact problem presents a number of difficulties which make the solution hard without introducing reasonable assumptions. In particular, the following simplifications are often made:

1. **The contacting surfaces are assumed to be ideally smooth.** This is a simple idealisation of the real geometry. Attempts have been made to incorporate roughness effects into the analysis of wheel–rail contact (Knothe and Theiler, 1996). However, the basic properties of wheel–rail contact can be defined without the need to introduce rough surfaces.
2. **The constitutive law of the wheel and rail materials is assumed to be linear elastic.** Although in some cases calculations performed provide maximum stresses that are above the material yield strength, this assumption can be accepted in most cases.
3. **The dimensions of the contact patch are small with respect to wheel and rail dimensions so that the two contacting bodies can be**

approximated to half-spaces from the structural point of view. This assumption is fairly well satisfied when the contact patch is located on the wheel tread and on the rail head, but may become questionable when the contact occurs on the wheel flange and on the rail gauge corner.

The last of these three assumptions deserve a further comment: it has been demonstrated (Telliskivi and Olofsson, 2001) that when contact is located on the rail gauge corner the maximum normal pressure computed using the elastic half-space assumption may present large differences with respect to that computed by the finite element method considering the actual shape of the contacting bodies. However, Telliskivi and Olofsson obtained their results considering only the presence of normal contact stresses at wheel–rail contact, whereas no satisfactory alternative to the use of the elastic half-space assumption is known for cases where the effect of tangential contact stresses is added in the problem. Moreover, finite element techniques are highly expensive from the computational point of view. Thus, at present, the finite element method is used only to compute a small number of reference solutions on which simpler approaches are validated.

The wheel–rail contact problem can be divided into two parts: the 'normal problem' and the 'tangential problem'. In the normal problem, the contact patch (position, shape and dimension) and the distribution of normal contact stresses is obtained as a function of the local geometry of the contacting profiles and of the resulting normal force reacted by the bodies. In the tangential problem, tangential contact stresses and slip pages are defined inside the contact area as a function of the normal pressures and of the resulting frictional forces or, alternatively, as a function of the normal pressures and of the creepage components.

Given the simplifying assumptions listed in points (1) to (3) above, and assuming the wheel and rail materials to have the same Young's modulus and Poisson's ratio, the quasi-identity property applies (Kalker, 1990) with the implication that the tangential stresses will not affect the distribution of normal pressures, i.e. that the normal and tangential problems are decoupled and can be solved separately. In the next two subsections, some typical strategies used for solving the normal and tangential problems under the quasi-identity assumption are summarised and the effect of the local geometry of wheel and rail profiles on stress and slip distributions is discussed.

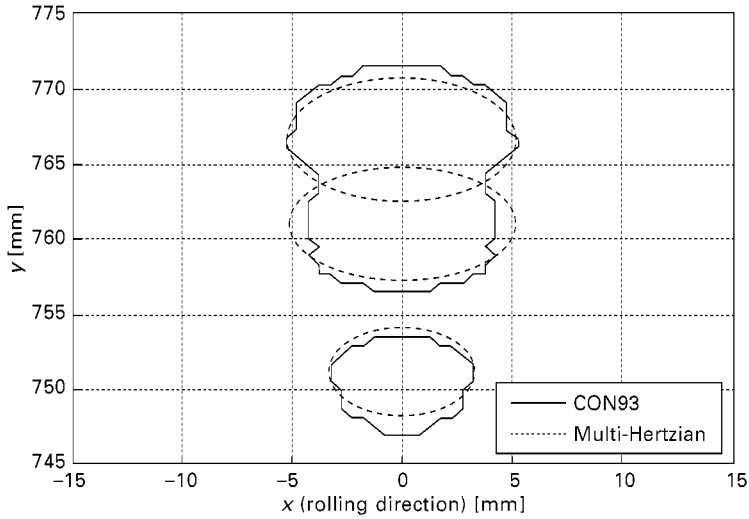
6.4.1 The normal problem

The simplest and most widely used method to solve the normal problem is represented by Hertz theory, where the two contacting bodies are assumed to have constant curvature over a region which is sufficiently wide to include the entire contact patch. Under this additional condition, and retaining assumptions

(1) to (3) listed above, an elliptic contact area and a semi-ellipsoidal pressure distribution are obtained. The dimensions and shape (represented by the a/b ratio of the semi-axes) of the contact patch and the maximum value of the contact pressure essentially depend on the curvature of the two bodies and on the resulting normal force and are expressed by analytical formulae based on elliptic integrals, thus making for the Hertzian model the most computationally efficient solution the normal problem.

Unfortunately, the Hertzian assumption of constant curvature may be quite far from reality in wheel–rail contact, especially when considering the contact of worn wheels and rails. Surface modifications produced by wear tend to modify the profile of wheels and rails so that they may become conformal. In this case, the Hertzian solution cannot be accepted, and more complicate strategies to solve the normal problem have to be implemented.

The most simple upgrade of the Hertzian approach is to approximate the undeformed distance function by piecewise constant curvature polynomials, to which the Hertzian theory is applied separately (Bruni *et al.*, 2000; Piotrowski and Chollet, 2005). The result of this procedure, known as 'multi-Hertzian' solution, is that the actual non-elliptical contact patch is approximated by the super position of more than one ellipse. As an example, Fig. 6.17 compares the non-elliptical contact patch obtained using the CONTACT algorithm (an accurate but time-consuming method to solve the normal problem described later in this subsection) with the multi-Hertzian approximation. In the case shown, which refers to conformal contact between an ORES1002 wheel



6.17 Approximation of a non-Hertzian contact patch by the multi-Hertzian approach. (Reference shape computed by CONTACT)

profile and a worn UIC60 profile with 1:20 inclination, the approximation of the contact patch by the multi-Hertzian method is good. In cases presenting sudden variations of profile curvature the result could be less satisfactory.

A more general and accurate approach to solve the normal problem is represented by the CONTACT algorithm (Kalker, 1990). In this case assumptions (1) to (3) are still retained, but the contacting bodies may have a generic transversal profile. In the longitudinal direction a constant curvature is assumed. This is acceptable for the wheel–rail couple where the longitudinal curvature of the wheel is almost constant and nominally equal to the inverse of the wheel radius and the rail surface has with good approximation zero curvature in the longitudinal direction. A rectangular area encompassing the expected contact patch must be provided by the user as well as the rectangular cell dimensions used to discretise such area. The two bodies are then pressed one against the other by the resulting normal force (or alternatively by a relative normal displacement) and the normal pressure at the nodes of the contact grid is computed as a function of the elastic displacements of the two bodies using the Boussinesq–Cerruti formulae for elastic half-spaces. With an iterative procedure, an approximation of the contact patch and the value of the normal pressure at the nodes of the contact grid is obtained. The disadvantage of this method is represented by its computational complexity that discourages its use for wheel wear prediction since repeated solution of the contact problem are required as shown in the feedback scheme of Fig. 6.2. Instead, this method is often used to compute reference solutions against which the validity of faster but simplified techniques is assessed.

With the aim of achieving an accuracy comparable to that of the CONTACT algorithm, but with much lower computational effort, approximate methods have been specifically developed for non-Hertzian contact conditions (Kik and Piotrowski, 1996; Ayasse and Chollet, 2006). In general, these methods are based on two simplifying assumptions:

- the contact patch is obtained as the intersection of the rigid wheel and rail surfaces when they are interpenetrated by a given amount;
- the normal stresses have a semi-elliptical distribution in the longitudinal direction.

The use of approximate non-Hertzian solutions of the contact problem is attractive from the perspective of wheel wear analysis, since these methods combine a good level of accuracy with an acceptable computational efficiency, allowing the repeated calculation of contact conditions which is required to reproduce the modification of the profile with service. At present, however, the use of these techniques for wheel wear prediction is confined to a small number of cases, of which Enblom and Berg (2008) represents an example.

6.4.2 The tangential problem

The tangential problem consists in determining the tangential tractions and the slip pages in the contact patch, given the creepages and the normal pressure distribution. As with the normal problem, the tangential problem can be approached at different levels of complexity. An approximate but fast solution is provided by the FASTSIM algorithm (Kalker, 1982, 1990). The contact patch is divided into longitudinal strips and for each strip a constant gradient of deformation is computed based on the creepage components. The tangential contact stresses are then computed at a discrete number of locations along each strip assuming each stress component to be proportional to the corresponding component of wheel–rail deformation in the same point through constant flexibility coefficients. The tangential stresses are then saturated to the traction bound corresponding to the local normal pressure times the friction coefficient and the slip is determined from the difference between the unsaturated and the saturated tangential stress using the flexibility coefficients.

The FASTSIM algorithm was originally developed for applications over Hertzian contact patches where the values of the flexibility coefficients can be chosen to match the analytical expression of the linearised creep force – creepage expression (Kalker, 1990). However, methods have been proposed (Kik and Piotrowski, 1996; Ayasse and Chollet, 2006) to adapt this method to non-elliptic contacts, the main difficulty being the identification of appropriate values for the flexibility parameters.

A more general and accurate but also more computationally demanding numerical solution to the tangential problem has been provided once more by Kalker, under the name CONTACT (Kalker, 1990). This technique is used in combination with the solution of the normal problem introduced under the same name in Section 6.4.1. In this method the assumption of elliptic contact patch is removed without the need to introduce simplifications additional to (1) to (3) listed at the beginning of the section. The contact patch is divided into rectangular cells (the same used to solve the normal problem as previously described). A relationship between the local deformation and the unsaturated tangential tractions is derived from Boussinesq–Cerruti analytical solution for elastic half-spaces. The saturation of the tangential stresses to the traction bound is introduced according to Coulomb's friction law applied to every discrete element in the contact patch. Under these assumptions, the problem is reduced to a constrained minimisation problem whose solution is found through an iterative procedure.

6.4.3 Effect of wheel and rail profiles on contact mechanics

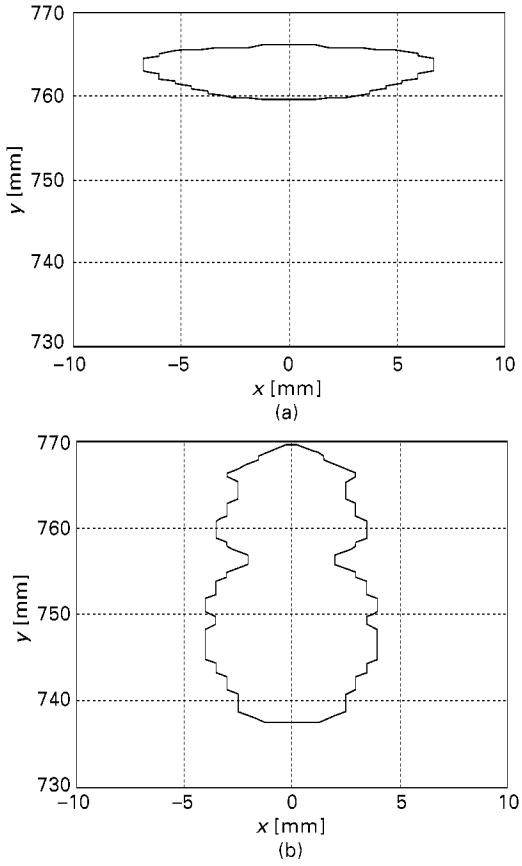
The local shape of wheel and rail surfaces in the contact region strongly affects the shape and dimensions of the contact patch, the contact stresses and hence the entire process of wheel wear. In general, with respect to coned profiles like the European ORES1002, worn wheel profiles or profiles designed to have a worn-like shape tend to produce much wider contact patches, which lower the sensitivity to wear, not only because of the decrease of contact pressures for the same normal load but also because of a reduction of the $T\gamma/A$ frictional power density. This has led to a number of attempts to design wear-resistant profiles, as summarised in Section 6.6.2.

To consider this subject in more detail, wheel–rail contact calculations performed using the CONTACT algorithm are compared for a new ORES1002 wheel coupled with a new UIC60 rail with 1:20 inclination and for worn ORES1002/UIC60 profiles. The global contact parameters considered are: zero lateral shift of the wheelset over the track; 60 kN vertical wheel load; combined longitudinal and lateral creepage condition with 0,003 creepage component in both directions; zero spin creepage and a friction coefficient of 0.36.

Figure 6.18 compares the contact patches obtained for the two cases considered: the new profiles produce a contact patch with a shape not far from elliptic, streamlined along the rolling direction, whereas the worn and conformal profiles produce a contact patch having a much larger extension in the transverse direction and a total area of contact approximately three times larger than that of the new wheel–rail profiles. Correspondingly, the normal contact stresses (compared in Fig. 6.19) are significantly higher for the new profiles than for the worn ones. A higher sensitivity to wear can then be expected for the new wheel–rail profiles on account of the higher contact pressure values, as shown by wear mapping results reported in Section 6.3.4.

As with the normal problem, the solution of the tangential problem is highly influenced by the shape and dimension of the contact patch and hence by wheel and rail profiles. Figure 6.20 compares the distribution of frictional power density $T\gamma/A$ evaluated for each cell in the discretised contact patch. It is observed that the maximum value of the frictional power density is approximately 3.5 times greater for the new profiles than for the conformal worn ones although the the overall value of the $T\gamma$ over the contact patch is almost the same for the two cases. Again, the new wheel–rail profiles are found to be more sensitive to wear effects, since larger values of $T\gamma/A$ give rise to higher wear rates.

From the examples shown, it can be concluded that wear effects are strongly affected by the amount of wear already accumulated on the wheel



6.18 Wheel–rail contact patch for zero lateral shift of the wheelset over the track, 60 kN vertical wheel load: (a) new ORES1002/UIC60 rail with inclination 1:20; (b) worn ORES1002/UIC60 rail.

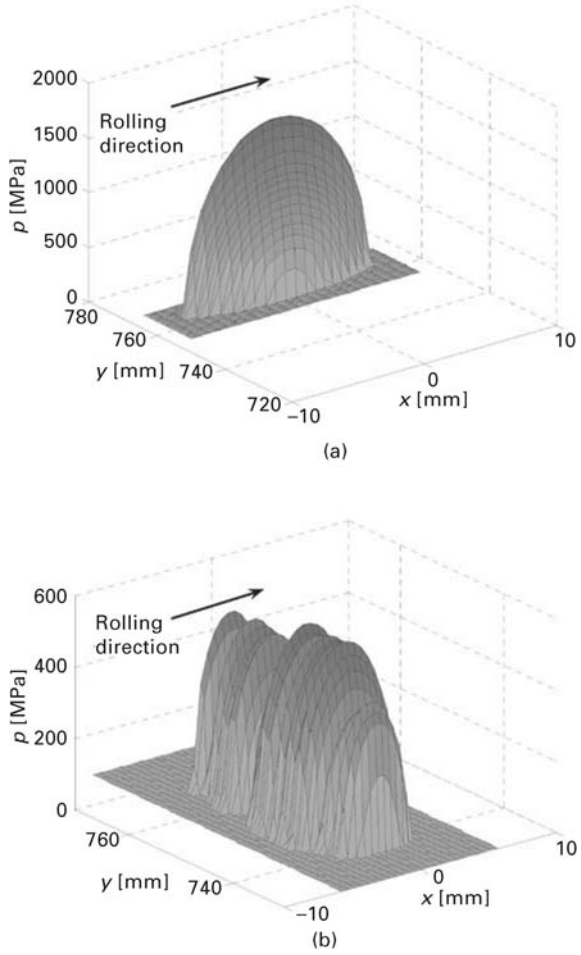
and rail surfaces, which confirms the 'closed loop' nature of the wheel wear process, as shown by Fig. 6.2. This has to be properly accounted for when addressing the study of wheel profile evolution due to wear.

6.5 State-of-the-art of uniform wheel wear modelling

6.5.1 Ingredients for regular wear modelling

A mathematical model for wheel wear prediction is a rather complex tool, combining several different tasks, as outlined in Fig. 6.2:

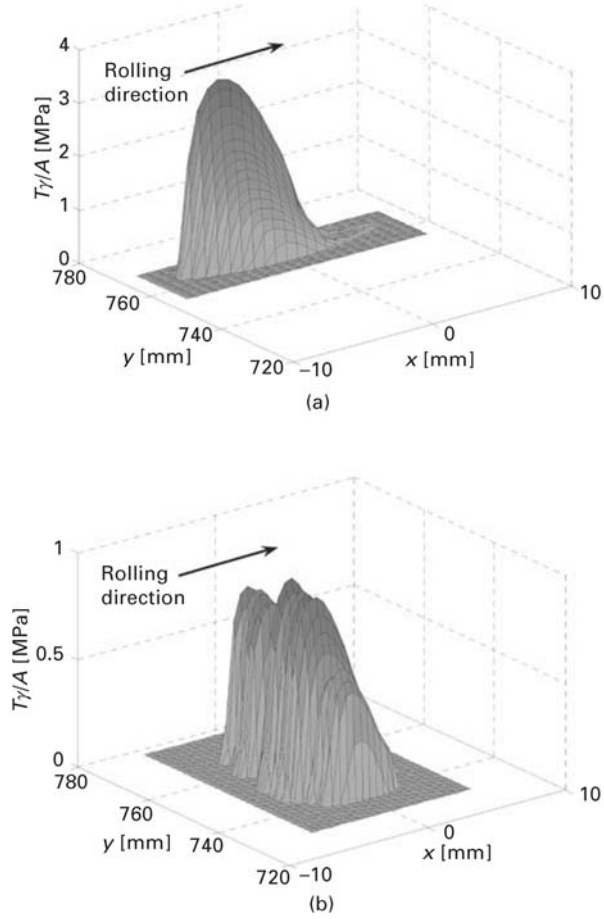
- Starting from the vehicle characteristics and initial wheel and rail profiles, a sequence of service conditions (e.g. tangent track running at different



6.19 Distribution of the normal pressure for the same contact conditions as shown in Fig. 6.18: (a) new ORES1002/UIC60 rail with inclination 1:20; (b) worn ORES1002/UIC60 rail.

speeds, curve negotiation at specified cant deficiencies, etc.) called the 'service profile' is simulated using a multibody model of vehicle/track interaction.

- At discrete time steps, a local contact analysis for each contact point is performed in order to determine stress as well as slip distributions inside the contact patch.
- These distributions are then converted into material loss using an appropriate wear model.
- Since the variation of wheel profiles produced by one single mission track



6.20 Distribution of the frictional power density $T\gamma/A$ for the same contact conditions as in Fig. 6.17. Combined 3 % creepage in longitudinal and lateral direction, no spin: (a) new ORES1002/UIC60 rail with inclination 1:20; (b) worn ORES1002/UIC60 rail.

is usually so small that it produces negligible variations on the results of the multi-body dynamic simulations and of the local contact analysis, in order to speed up the wheel profile wear prediction, the amount of wear is scaled to represent a larger distance run by the vehicle. This modification of the wheel profile has, however, to be kept small enough to produce small variations on the results of the dynamic simulation and of the local contact analysis, which opens the problem of choosing a profile updating strategy, as discussed below.

- The updated wheel profiles are finally fed back into the multibody

6.21 (a) Trackside and (b) onboard friction modifier applicators.

model of the railway vehicle and the whole procedure is started again. In some cases, smoothing of the updated profiles is performed to avoid unrealistic results of the multi-body simulation and of the local contact analysis caused by numerical perturbations.

As described in the previous paragraphs, each single task may be solved using different approaches that may be more or less accurate and, correspondingly, may require a higher or lower computational effort. The accuracy of the overall wear prediction procedure is not only determined by the modelling approaches applied to solve each task in the above list, but is also strongly affected by the choice of the service profile and of a suitable profile updating strategy.

The service profile has to include all relevant track geometry features such as curve radii and their occurrence, rail profiles, track irregularity along the line(s) serviced by the vehicle. Furthermore, the influence of service speed and environmental conditions on the adhesion coefficient has to be considered. The service profile is synthesized in a finite set of vehicle working conditions which are used for multi-body simulations.

As far as the profile updating strategy is considered, two main alternatives are available: in the 'fixed distance run' strategy, see e.g. Pearce and Sherratt (1991), each iteration of the loop procedure outlined above is considered to represent a fixed distance travelled by the vehicle (e.g. 1000 km), and wheel profile modification is scaled accordingly. In the 'fixed wear depth' strategy, see e.g. Jendel and Berg (2002), each iteration is scaled to produce a predefined maximum depth of wheel wear, and the corresponding travelled distance is computed. Both strategies have proved to work well, provided that reasonable values are chosen for the maximum distance run/maximum wear depth parameters, so that the service life of the wheel between two consecutive re-profiling operations is discretised appropriately.

6.5.2 The state of the art

Research activities aiming at the prediction of wheel uniform wear began to be published in the mid-1980s. It was soon clear to researchers that the solution to such a problem would be demanding both in terms of computer power and cross-discipline mathematical models.

One of the first attempts to simulate wear of railway wheels was proposed by Pearce and Sherratt (1991). Their model is very simple in order to achieve a reasonable simulation time: after the calculation of the global contact forces and creepages acting on the contact patch, the amount of material removed is calculated through the 'Derby wear index' (see Section 6.3). The considered track is made of a straight line plus an S curve. A 'fixed-distance run' updating strategy is used and the optimal route length for the updating is found to be 1100 km.

The same year, Kalker presented a method for the simulation of wheel tread wear achieving reasonable results with respect to measurements (Kalker, 1991). In this method, a very simple model is used to reproduce vehicle dynamics, whilst local contact analysis is carried out by the Hertzian theory and by the FASTSIM algorithm. Wear is assumed to be proportional to the frictional work at the wheel–rail interface and the proportionality coefficient is tuned so as to have a good correspondence with field tests. Kalker also introduces a statistical approach to wear prediction: instead of solving the contact problem at each occurrence, the probability of the contact being at a certain wheel–rail lateral co-ordinate is considered.

The first method to predict wheel wear using a non-Hertzian model to solve the normal problem was proposed by Linder and Brauchli (1996): the shape of the contact area is taken as the rigid body com penetration area. Similar to what is done by the FASTSIM algorithm for the tangential problem, the contact area is divided into strips parallel to the rolling direction and for each strip a Hertzian problem is solved. Again, wear is assumed to be proportional to the frictional power (two wear regimes are considered) taking the proportionality coefficients from the literature. Profiles are updated after a given fixed-distance run.

Zobory used Hertzian theory to solve the wheel–rail normal problem and FASTSIM to solve the tangential problem (Zobory, 1997). Different wear modelling approaches are discussed, mainly based on the proportionality of wear with the energy dissipated at the contact. Due to the very different wear regimes on wheel tread and wheel flange, Zobory introduced two proportionality constants, one for the 'mild' regime on the tread and one for the 'severe' regime on the flange. The transition between the two regimes depends on wheel and rail material properties. Simulation results are compared with various line measurements related to a vehicle running on the Gotthard line for a maximum mileage of 27 000 km. Wheel profiles are updated every 1000 km. To be able to compare experimental data and numerical results,

a smoothing procedure (based on a smoothing spline) was applied to the updated wheel profile.

In 2000 Liebig *et al.* analysed the influence of vehicle model complexity on the wheel wear results (Liebig *et al.*, 2000). Also in their work, wear was assumed to be proportional to the frictional work, The evolution of the wheel profile was predicted by using both a very simple vehicle model and a fully non-linear multi-body vehicle model. The results obtained show that, at increasing vehicle model complexity, the wear distribution along the wheel transversal profile becomes less uniform thus becoming more similar to measured wheel profiles.

In 2002 Jendel *et al.* developed a model similar to that of Zobory using the multi-body code GENSYS for the dynamic railway vehicle simulations (Jendel and Berg, 2002). Again, the local contact analysis was solved by applying the Hertzian theory and the FASTSIM algorithm for the normal and tangential problems, respectively. The contact model used within the multi-body simulation was able to find at most two contact points at the same time on a given wheel–rail pair. For the wear prediction, Archard's wear model was applied locally. The wheel profile was updated every time a maximum wear depth of 0.1 mm was reached or a maximum distance of 1500 km was run. A cubic spline was applied both on wear distribution and the updated wheel profile for smoothing purposes. The simulation results were compared with measurements of serviced wheels on the commuter railway network in Stockholm.

Braghin *et al.* (2002) developed a wheel wear prediction model using a fully non-linear multi-body model for the simulation of vehicle dynamics, and CONTACT algorithm to solve both the normal problem and the tangential problem under non-Hertzian assumptions. Material removal was assumed to be proportional to the local frictional power density ($T\gamma/A$ product), the proportionality constant being determined from full-scale experimental tests on the BU300 roller rig described in Section 6.3. Due to the complexity of the model used for local contact solution, the computational effort of this approach was extremely high. Thus, wheel profiles were updated only twice over a mileage of 10 500 km. Comparisons between calculated and measured wear showed good agreement for the tread while on the flange wear was under-estimated due to the fact that the wear coefficient was determined for the mild wear regime only.

The prediction of uniform wheel wear was improved both in terms of results and of computational time required by Braghin *et al.* by using a simpler multi-Hertzian–FASTSIM method to perform local contact analysis and a wear model based on experiments on a twin, disc machine (Braghin *et al.*, 2006). Numerical results were compared with full-scale wear experiments performed on the BU300 roller rig, showing satisfactory agreement under precisely measured testing conditions such as contact forces, angle of attack and friction coefficient.

Enblom achieved satisfactory prediction of wheel uniform wear by extending Jendel's model through the introduction of the tangential flexibility of contacting surfaces, thus allowing a more realistic calculation of relative sliding velocities and distances (Enblom, 2006). Enblom also included braking conditions in the vehicle service profile.

6.6 Means to reduce uniform wheel wear

Starting from the analysis of wheel wear causes and mechanisms, a number of means to reduce the effect of wear on wheel and rail surfaces have been identified, studied and partly put into practice. This final section of the chapter provides a review of these measures.

According to Fig. 6.2, wheel wear is the resulting effect of three interacting subprocesses: vehicle dynamics, contact mechanics and wheel–rail contact tribology. Measures to reduce wheel wear will accordingly be based on modifying at least one of these three subprocesses. Considering tribological issues first, an obvious way to reduce wheel wear is to use a material having improved wear-resistance properties. This is, however, limited by the fact that a relatively small number of standardised steel grades are allowed for wheel construction. Also, a material having high wear resistance properties is likely to show high sensitivity to RCF damage. A second possibility is represented by altering the frictional properties of the wheel–rail interface by the use of friction modifiers, as described in Section 6.6.1.

As far as contact mechanics is concerned, the use of appropriate wheel profiles allows for the reduction of local contact stresses and hence of wheel wear (see Section 6.6.2). Finally, wheel wear can be reduced by modifying the running behaviour of the entire vehicle, so as to lower the contact forces and slippages produced by vehicle service. This can be obtained in several ways, e.g. by the use of a special design for the vehicle architecture, by the choice of appropriate values for suspension parameters or by the use of advanced technologies like vehicle mechatronics.

Note that any means able to reduce wheel wear will have a significant impact on other fundamental issues related to train–track interaction, especially system performances, and other damaging effects on the wheel and on the rail. This explains why an optimal and universal solution to mitigate wear effects cannot be defined. Instead, best deal solutions, possibly consisting of the combined use of different means described hereafter, have to be sought, depending on the specific application considered.

6.6.1 Use of friction modifiers

Where wheel–rail contact conditions are most severe (high lateral loads and slips), typically in curves where the flange contacts the gauge corner,

mitigation is required to try and reduce the likelihood of wear occurring to both wheel and rails. Wheel tread–top of rail friction control is being increasingly used to provide this wear mitigation as well as minimise curve noise, lateral forces and corrugations (Eadie *et al.*, 2005, 2006). Application of friction modifiers (FMs), which are basically an engineered composite of inorganic and polymeric materials (in liquid or solid form), provides a thin dry film which, in combination with naturally occurring third body components (primarily iron oxides), gives an intermediate coefficient of friction (about 0.35 measured by a Salient Systems push tribometer) (Hooper *et al.*, 2003; Lu *et al.*, 2005). These friction level minimises lateral forces and wear without compromising braking or traction. In addition, the positive traction–creep relationship ('positive friction') can eliminate roll-slip oscillation in sharp curves under conditions of saturated lateral creepage (Eadie *et al.*, 2002; Suda *et al.*, 2003; Lu *et al.*, 2005). The two main methods used to apply friction modifiers to the wheel–rail contact are via *trackside devices* or by *on-board applicators* (as shown in Fig. 6.21). *Mobile lubricators*, which are essentially railway vehicles designed to apply lubricant to the gauge corner of the track, are also sometimes used.

Trackside lubricators are mounted next to the track and apply FMs to the rail gauge corner/top of rail. There are three types: mechanical, hydraulic and electronic. Onboard lubricators apply solid stick FMs (sticks are spring loaded against the wheel surface) or spray liquid FM onto the wheel flange or tread area, which is then transferred to the rail. Complex control systems are used in the application process to avoid the application of FMs at inappropriate locations.

6.6.2 Wear-resistant wheel profiles

The description of wear regimes reported in Section 6.3.4 shows that wear effects are governed by the normal and tangential components of the contact stresses and by the slip. An increase in the size of the contact patch for the same contact force components is therefore likely to produce a reduction of the overall wear effect since the contact stresses will be reduced. This consideration has led to the design of the so-called 'conformal' or 'wear-like' wheel profiles whose shape is designed to achieve conformal contact with the rail, thereby increasing the size of the contact region (Fröhling, 2003). One drawback of this solution is that conformal profiles usually give rise to a higher conicity than coned profiles and this may negatively affect vehicle stability (see Chapter 15 "Effect of damage on vehicle dynamics").

Some wheel profiles may present two-point contact with the rail at flange root. This increases the wear rate, on account of the large longitudinal slippages that are generated by the rolling radius difference between the two points of contact. Therefore, a second way to improve wear resistance is to design

the wheel profile so that the formation of two-point contact is avoided. The effectiveness of this strategy in improving wheel wear resistance is examined in Markine *et al.* (2007).

A third possible way to influence wear through the design of the wheel profile is to optimise the rolling radius difference function with a view to improving the vehicle dynamic behaviour. In this way, the lateral motion of the wheelset produced by track irregularities, hunting instability and curve negotiation is minimised, and the corresponding creepages and creep forces are reduced. This approach is proposed by (Shevtsov *et al.*, 2005) and by (Persson and Iwnicki, 2004). Both papers tackle the optimisation of the wheel profile from a more general point of view of vehicle dynamics but show some beneficial effects also in terms of wheel wear rate reduction.

The design of wear-resistant wheel profiles needs to consider that the shape of the profile will change with the progress of wear. To ensure that the wheel profile will preserve good wear resistance properties across the entire life before re-profiling, a thorough understanding of how the wear process will evolve is required. This can be achieved using a wear prediction model of the type described in Section 6.5.

6.6.3 Vehicle design

In railway vehicles, the natural tendency of solid wheelsets to give rise to large lateral oscillations known as hunting is avoided by joining wheelsets in pairs through a bogie frame and relatively stiff primary suspensions. Such a stabilising measure, however, has consequences for the vehicle curving behaviour, since the naturally good curving properties of the isolated wheelset are lost, and the vehicle usually gives rise to much larger longitudinal and lateral creep forces while running through a curve than those strictly required to balance non-compensated centrifugal effects and traction/braking actions (Wickens, 2003).

Therefore, a possibility to reduce wheel wear effects is to design the vehicle so as to reduce as much as possible creep forces arising in curve negotiation. This in turn can be obtained by controlling the angle of attack of the wheelsets, that is the rotation of the axle with respect to the radial direction in the curve. In particular, for vehicles negotiating the curve at low values of cant deficiency, the radial alignment of the wheelset (i.e. zero angle of attack) shall be considered the best curving condition.

With the exception of active vehicle control, which is treated in the next section, two means can be used to reduce the angle of attack of the wheelsets in curve and hence wear effects associated with curving: the first one consists in lowering the stiffness of the primary suspension and/or shortening the bogie wheelbase; the second is represented by the use of articulated bogies.

The decrease of primary suspension stiffness and of bogie wheelbase

both have a negative impact on vehicle stability, since the 'wheelset joining' effect provided by the bogie is lowered. Therefore, the mechanical design of the railway bogie can be viewed as the research of a trade-off between stability and curving requirements, the optimal solution depending on the type of vehicle considered (high-speed, conventional rail, freight transport) and on the service profile specified for the vehicle. As shown in Mazzola *et al.* (2007), the problem of defining optimal values for bogie parameters can be set in the form of a constrained optimisation problem, in which wear resistance is optimised under the constraint of the vehicle ensuring a required region of stable running.

Articulated bogies are characterised by a special design in which mechanical links introduce additional degrees of freedom to favour the radial alignment of the wheelsets in curve, whereas wheelset movements that do not have a role in the curving are constrained to avoid compromising too much vehicle stability. Examples of the use of articulated bogies are known especially in the field of freight trains, where this technology has been demonstrated to provide significant wear reduction. In the case of passenger transport, the decrease of the critical speed which is anyhow associated with the use of articulated bogies makes this solution less attractive, and hence a limited number of applications are known.

In a comprehensive study (Garcia *et al.*, 2000) that has considered high-speed, regional and mass transit service and three different concepts of articulated bogie, it is shown that a consistent reduction of wheel wear with respect to a conventional vehicle can be obtained using a 'forced-steering' bogie, where a mechanical link between the carbody and the bogie steers the two axles in the bogie. However, according to the results presented in the paper, the advantage provided by this concept is limited to curves having radius below 700 m, whereas for curves having larger radius the conventional bogie determines a lower wear index. The same analysis also shows that very stiff primary suspensions are needed to stabilise the forced-steering bogie. The other articulated bogie concepts considered in the same paper are shown to provide a higher wear index than the conventional bogie, except in the case of sharp curves having radius below 300 m, for which a small reduction of wheel wear is obtained.

6.6.4 Active control of vehicle dynamics

As pointed out in the previous section, the primary suspension design for a railway vehicle inherently imposes a trade-off between curving and stability, so that the final result of the design process will not provide the best possible performances with respect to curving and specifically to wheel wear. In recent years, there have been proposals to use active suspensions to improve vehicle dynamics, allowing not only an increase in vehicle running performances, but

also lower the aggresiveness of the vehicle with respect to the infrastructure and to the vehicle itself. Although these proposals for a 'mechatronic train' are still confined to the research field and have not yet found applications, the subject surely deserves a brief examination to underline the possible impact of new technologies in the mid-term future.

The different implementations to actively control the dynamics of a railway vehicle can be grouped into the following categories:

- **active primary suspension** (Shen and Goodall, 1997). In this concept, proposed with applications both to vehicles having solid wheelsets and independently rotating wheels (IRW), the passive primary suspension is replaced in part or entirely by an actuation system able to apply control forces on the axles. Depending on the different configurations (actuated solid wheelset/actuated IRW), active control can be used to improve stability, curving and/or to provide guidance.
- **independently driven wheels** (Gretzschel and Bose, 2002). The control action is represented by independent traction/braking torques applied to the two wheels of the axle, providing guidance to the IRW couple and improving curving performances.
- **active secondary suspension** (Diana *et al.*, 2002). This concept is based on applying a yaw torque on the bogie, rather than on the axles, which can be used to improve stability and/or curving performances.
- **active wheel steering** (Wickens, 1994). In this concept the wheels of the railway vehicle are independently rotated and steered by actuators to provide stability and guidance.

For most of these concepts, the possibility of achieving wheel wear reductions with respect to passive vehicles has been specifically studied. In Perez *et al.* (2002), using numerical simulations, the use of an active primary suspension is shown to reduce the wear index during curve negotiation by a factor of 5 with respect to a vehicle equipped with standard passive suspensions. The active primary suspension concept has also been shown to be beneficial in view of reducing rolling contact fatigue in rails (Perez *et al.*, 2006) and rail corrugation (González *et al.*, 2007), the study being confined to mathematical modelling and simulation. For the independently driven wheels concept, the reduction of frictional power was experimentally verified through a 1/5 scaled roller rig (Gretzschel and Bose, 1999). For active secondary suspensions actuation, full-scale tests performed on a bogie using the NTSEL roller rig test stand in Tokyo (Matsumoto *et al.*, 2004) demonstrated the possibility as reducing guiding forces in curves of small radius. Wear measurements were not specifically performed, but the reduction of the guiding forces is expected to produce reduced wear. Finally, for the actively steered wheels concept, no study specifically addressing the reduction of wear could be found; however, the concept clearly has a high potential

for wheel wear reduction, by almost eliminating longitudinal creep forces through the use of IRWs and by minimising transversal creep forces needed for curve negotiation.

6.7 Conclusions

This chapter has presented the basic principles governing the process of regular wear in railway wheels including a review of the modelling approaches proposed for this problem and a description of possible measures to reduce wheel wear and its effects.

Wheel wear effects represent one of the major sources of maintenance costs in the management of a fleet of railway vehicles and have direct implications on rolling stock performance as well as on vehicle impact on the track. This explains the large effort spent in the past to understand wheel wear mechanisms and to find mitigation measures. Nevertheless, wheel wear still needs to be considered as an open problem, where new materials, new testing techniques and more advanced modelling and simulation approaches can be expected to appear in the mid and long term.

From the point of view of scientific research, wheel wear represents a complex and multidisciplinary problem which needs to be addressed by combining knowledge coming from different fields, including tribology, behaviour of metallic materials, solid mechanics and vehicle dynamics. This places high demands on future research in this area, the challenge being not only to produce advancements in single pertinent branches of science and engineering, but also to integrate knowledge coming from different fields. Accordingly, it has been our effort, while writing this chapter, to outline the interactions between different competencies underlying a comprehensive understanding of the wheel wear process. It is our hope that the material contained in this chapter, together with other related chapters of this handbook, will form a comprehensive basis to address wear and damage problems at the wheel–rail interface.

6.8 References

Archard JF (1953), Contact of rubbing surfaces and flat surfaces, *Journal of Applied Physics*, **24**, 981–88.

Ayasse JB and Chollet H (2006), Wheel-rail contact in Iwnicki S (ed.), *Handbook of Railway Vehicle Dynamics*, CRC, Boca Raton, FL, 85–120.

Beagley TM (1976), Severe Wear of Rolling/Sliding Contacts, *Wear*, **36**, 317–35.

Blok H (1963), The flash temperature concept, *Wear*, **6**, 483–94.

Bolton PJ and Clayton P (1984), Rolling-sliding wear damage in rail and tyre steels, *Wear*, **93**, 145–65

Bolton PJ, Clayton P and McEwen IJ (1982), Wear of rail and tyre steels under rolling/sliding conditions, *ASLE Transactions*, **25**, 17–24.

Braghin F, Bruni S and Resta F (2002), Wear of railway wheel profiles: a comparison between experimental results and a mathematical model, *Vehicle System Dynamics Supplement*, **37**, 478–89.

Braghin F, Lewis R, Dwyer-Joyce RS and Bruni S (2006), A mathematical model to predict railway wheel profile evolution due to wear, *Wear*, **261**, 1253–64.

British Steel Makers Creep Committee (1973), *BSCC High Temperature Data*, The Iron and Steel Institute for the BSCC, London, UK.

Broster M, Pritchard C and Smith DA (1974), Wheel–rail adhesion: it's relation to rail contamination on British Railways, *Wear*, **29**, 309–21.

Bruni S, Collina A, Diana G and Vanolo P (2000), Lateral dynamics of a railway vehicle in tangent track and curve: tests and simulation, *Vehicle System Dynamics Supplement*, **33**, 464–77.

Carter FW (1926), On the action of a locomotive driving wheel, *Proceedings of the Royal Society*, **A112**, 151–7.

Dearden J (1960), The wear of steel rails and tyres in railway service, *Wear*, **3**, 43–9.

Diana G, Bruni S, Cheli F and Resta F (2002), Active control of the running behaviour of a railway vehicle: stability and curving performances, *Vehicle System Dynamics*, **37**(Supplement), 157–70.

Eadie DT, Kalousek J and Chiddick KS (2002), The role of high positive friction (HPF) modifier in the control of short pitch corrugations and related phenomena, *Wear*, **253**, 185–92.

Eadie D T, Vidler B, Hooper N E and Majowsky T W (2003), Top of rail friction control: lateral force and rail wear reduction in a freight application, *Proceedings International Heavy Haul Conference*, Dallas, TX, USA, May 5–9, 5.73–5.81.

Eadie DT, Santoro M and Kalousek J (2005), Railway noise and the effect of top of rail liquid friction modifiers: changes in sound and vibration spectral distributions in curves, *Wear*, **258**, 1148–55.

Eadie DT, Oldknow K, Maglalang L, Makowsky T, Reiff R, Sroba P and Powell W (2006), Implementation of wayside top of rail friction control on North American heavy haul railways, *Proceedings 7th World Congress on Railway Research*, Montreal, QC, Canada, 4–8 June, on CD.

Elkins JA and Eickhoff BM (1979), Advances in nonlinear wheel/rail force prediction methods and their validation *ASME Winter Annual Meeting*, New York, USA.

Enblom R (2006), *On simulation of uniform wear and profile evolution in the wheel-rail contact*, Doctoral Thesis KTH, Stockholm, Sweden.

Enblom R and Berg M (2008), Impact of non-elliptic contact modelling in wheel wear simulation, *Wear*, **265**, 1532–41.

Fröhling RD (2003), Strategies to control wheel profile wear, *Vehicle System Dynamics Supplement*, **37**, 490–501.

Gallardo-Hernandez EA, Lewis R and Dwyer-Joyce RS (2006), Temperature in a twin-disc wheel/rail contact simulation, *Tribology International*, **39**, 1653–63.

Garcia JF, Olaizola X, Martin LM and Gimenez JG (2000), Theoretical comparison between different configurations of radial and conventional bogies, *Vehicle System Dynamics*, **33**, 233–59.

Garnham JE and Beynon JH (1992), 'Dry rolling-sliding wear of bainitic and pearlitic steels, *Wear*, **157**, 81–109.

Ghonem H and Gonsalves R (1978), *Comparative performance of type II trucks*, Canadian Pacific Department of Research, Report No. S576–78, Calgary, AB, Canada.

Ghonem H and Kalousek J (1982), The use of angle of attack measurements to estimate wear under steady state rolling conditions in Tong P and Grief R (eds), *Wheel/Rail Load and Displacement Measurement Techniques*, US Transportation Development Center, Boston, MA, USA, 1310–34.

González F, Pérez J, Vinolas J and Alonso A (2007), Use of active steering in railway bogies to reduce rail corrugation on curves, *Journal of Rail and Rapid Transit*, **221**, 509–19.

Gretzschel M and Bose L (1999), A mechatronic approach for active influence on railway vehicle running behaviour, *Vehicle System Dynamics*, **33**(Supplement), 418–30.

Gretzschel M and Bose L (2002), A new concept for integrated guidance and drive of railway running gears, *Control Engineering Practice*, **10**, 1013–21.

Heumann H (1954), *Grundzüge der Führung der Schienenfahrzeuge*, Verlag R Oldenburg, Munich, Germany.

Jendel T (2002), Prediction of wheel profile wear – comparisons with field measurements, *Wear*, **253**, 89–99.

Jendel T and Berg M (2002), Prediction of wheel profile warp, *Vehicle System Dynamics Supplement*, **37**, 502–13.

Kalker JJ (1982), A fast algorithm for the simplified theory of rolling contact, *Vehicle System Dynamics*, **11**, 1–13.

Kalker JJ (1990), *Three-dimensional Elastic Bodies in Rolling Contact*, Kluwer, Dordrecht, the Netherlands.

Kalker JJ (1991), Wheel-rail rolling contact theory, *Wear*, **144**, 243–61.

Kik W and Piotrowski J (1996), A fast, approximate method to calculate normal load at contact between wheel and rail and creep forces during rolling, in Zobory I (ed.), *Proceedings 2nd Mini Conference on Contact Mechanics of Rail/Wheel Systems*, Budapest University of Technology, Budapest, Hungary, S52–61.

Knothe K and Theiler A (1996), Normal and tangential contact problem with rough surfaces, in Zobory, I (ed.), *Proceedings 2nd Mini Conference on Contact Mechanics of Rail/ Wheel Systems*, Budapest University of Technology, Budapest, Hungary, 34–43.

Krause H and Poll G (1986), Wear of wheel-rail surfaces, *Wear*, **113**, 103–22.

Kumar S and Rao DLP (1984), Wheel–rail contact wear, work, and lateral force for zero angle of attack – a laboratory study, *Transactions of the ASME, Journal of Dynamic Systems, Measurement, and Control*, **106**, 319–26.

Lewis R and Dwyer-Joyce RS (2004), Wear mechanisms and transitions in railway wheel steels, *Proceedings I of the MechE, Part J: Journal of Engineering Tribology*, **218**, 467–78.

Lewis R and Olofsson U (2004), Mapping rail wear regimes and transitions, *Wear*, **257**, 721–9.

Lewis, R, Olofsson, U, Dwyer-Joyce, RS, Hallam, RI (2004), 'Wheel material wear mechanisms and transitions, *Proceedings 14th International Wheelset Congress*, Orlando, FL, 17–21 October on CD.

Liebig S, Kiss Z, Quarz V (2000) Prediction of wheel-wear by means of multibody system simulation, in Zobary I (ed.), *Proceedings 7th Mini Conference on Vehicle System Dynamics, Identification and Anomalies*, Budapest University of Technology, Budapest Hungaryt, 179–84.

Lim SC and Ashby MF (1987), Wear mechanism maps, *Acta Metallica*, **35**, 1–24.

Linder C and Brauchli H (1996), Prediction of wheel wear, in Zobory I (ed.), *Proceedings 2nd Mini Conference on Contact Mechanics and Wear of Rail/Wheel Systems*, Budapest University of Technology, Budapest, Hungary, 215–20.

Lu X, Cotter J and Eadie DT (2005), Laboratory study of the tribological properties of friction modifier thin films for friction control at the wheel-rail interface, *Wear*, **259**, 1262–9.

Matsumoto A, Sato Y, Ohno H, Mizuma T, Suda Y, Michitsuji Y, Tanimoto M, Miyauchi E and Sato Y (2004), Research on high curving performance trucks – concept and basic characteristics of active-bogie-steering-truck, *Vehicle System Dynamics*, **41**(Supplement), 33–42.

Marcotte PP, Caldwell WN and List HA (1978), Performance analysis and testing of a conventional three-piece freight car truck retrofitted to provide axle steering, *ASME Winter Annual Meeting*, San Francisco, CA, USA.

Markine VL, Shevtsov IV and Esveld C (2007), An inverse shape design method for railway wheel profiles, *Structural and Multidisciplinary Optimization*, **33**, 243–53.

Mazzola L, Alfi S and Belforte P (2007), Numerical simulations for the virtual assessment of technological options in high speed rail vehicles design, *Proceedings Multibody Dynamics 2007*, ECCOMAS Thematic Conference, Milan, Italy, 25–28 June, on CD.

McEwen IJ and Harvey RF (1985), Full scale wheel-on-rail wear testing: comparisons with service wear and a developing theoretical predictive method, *Lubrication Engineering*, **1**, 80–88.

Nilsson R (2000), Wheel and rail wear – measured profiles and hardness changes during 2.5 years for Stockholm commuter traffic, *Proceedings Railway Engineering – 2000*, London, UK, 5–6 July.

Olofsson U and Nilsson R (2002), Surface cracks and wear of rail: a full scale test and laboratory study, *Journal of Rail and Rapid Transit*, **216**, 249–64.

Pearce TG and Sherratt ND (1991), Prediction of wheel profile wear, *Wear*, **144**, 343–51.

Perez J, Busturia J and Goodall RM (2002), Control strategies for active steering of bogie-based railway vehicles, *Control Engineering Practice*, **10**, 1005–12.

Perez J, Stow J and Iwnicki S (2006), Application of active steering systems for the reduction of rolling contact fatigue on rails, *Vehicle System Dynamics*, **44**(Supplement), 730–40.

Persson I and Iwnicki SD (2004), Optimisation of railway profiles using a genetic algorithm, *Vehicle System Dynamics*, **41**(Supplement), 517–527.

Piotrowski J and Chollet H (2005), Wheel-rail contact models for vehicle system dynamics including multi-point contact, *Vehicle System Dynamics*, **43**, 455–84.

Roberti R and Bruni S (2001) Development of operations of tilting train on Italian network, *Proceedings 5th World Congress on Railway Research*, Köln, Germany, 25–29 November, on CD.

Shen G and Goodall RM (1997), Active yaw relaxation for improved bogie performance, *Vehicle System Dynamics*, **28**, 273–89.

Shevtsov IY, Markine VL and Esveld C (2005), Optimal design of wheel profile for railway vehicles, *Wear*, **258**, 1022–30.

Steele RK (1982), Observations of in-service wear of railroad wheels and rails under conditions of widely varying lubrication, *ASLE Transactions*, **25**, 400–409.

Suda Y, Iwasa T, Komine H, Fuji T, Matsumoto K, Ubukata N, Nakai T, Tanimoto M and Kishimoto Y (2003), The basic study on friction control between wheel and rail (experiments by test machine and scale model vehicle), *Proceedings 6th International Conference on Contact Mechanics and Wear of Rail/Wheel Systems*, Gothenburg, Sweden, 10–13 June, 343–8.

Telliskivi T and Olofsson U (2001), Contact mechanics analysis of measured wheel–rail profiles using the finite element method, *Journal of Rail and Rapid Transit*, **215**, 65–72.

Vogel R and Kurek EG (1981), *Railroad Track; Theory and Practice: Material, Properties, Cross Sections, Welding and Treatment*, Fastenrath F (ed.), Frederick Ungar, New York.

Wickens AH (1994), Dynamic stability of articulated and steered railway vehicles guided by lateral displacement feedback, *Vehicle System Dynamics*, **23**(Supplement), 541–53.

Wickens AH (2003), *Fundamentals of Rail Vehicle Dynamics*, Swets and Zeitlinger, Lisse, the Netherlands.

Zobory I (1997), Prediction of wheel/rail profile wear, *Vehicle System Dynamics*, **28**, 221–59.

Copyrighted Materials
Copyright © 2009 Woodhead Publishing Retrieved from www.knovel.com

7
Fatigue of railway wheels

A. EKBERG, Chalmers University of Technology, Sweden

Abstract: Wheel fatigue – the formation of cracks on wheel tread, in wheel rim or on the wheel disc – is a crucial phenomenon in assuring safe and reliable rail transportation. The current chapter gives a background to the phenomenon. The appearance and mechanisms of different forms of wheel fatigue are discussed and predictive models are presented. The influence of some operational factors and interacting deterioration phenomena are then discussed before the chapter ends with concluding remarks and an outlook on future trends.

Key words: railway wheel fatigue, rolling contact fatigue, thermal loading, damage, phenomenology, predictive models.

7.1 Introduction

7.1.1 Background

There are roughly 25–50 million railway wheels in operation in the world today. Allowing an annual failure rate of one in 1000 would mean 25 000–50 000 failed wheels every year. It is quite clear that if 'failure' here refers to a complete fracture of the wheel causing the train to be inoperative, railways would not be an efficient means of transportation.

In reality most 'failures' due to wheel fatigue are of a less dramatic nature. In the vast majority of cases cracks are mitigated before they have become sufficiently large to cause catastrophic failures. Still, wheel fatigue may have major operational and economical consequences. The fact that wheel fatigue failures tend to come in 'epidemics' causing major operational disturbances further adds to this.

As we increase axle loads and train speeds together with the overall utilisation of train fleets, wheel fatigue issues are bound to increase if countermeasures are not taken. Due to the massive number of operational wheels and the decreasing margins for operational errors, there is a vast potential in cost savings from improved wheel fatigue management. Even marginal savings gained through planned maintenance, improved fatigue lives and enhanced performance of the wheels will add up to significant cost savings.

As indicated above, accidents caused by wheel fractures are rare. Though fatal accidents are even more rare they are often highly publicised, which may cause customers to abandon train transportation. This has a highly negative consequence since a decrease in train transportation, whether due to

questionable safety, decreased operational reliability or reduced economical competitiveness, draws passengers and freight to cars, trucks and airplanes. All of these are inferior means of transportation from a socioeconomic, environmental and safety point of view.

So, from an operational point of view, an understanding and implementation of knowledge on wheel fatigue is of significant importance. However, the subject is also interesting from a scientific point of view. Historically, railways and fatigue research have been closely linked. With the introduction of railways, dynamic loads of sufficient magnitude and duration to cause fatigue failures were introduced, and the first systematic study of metal fatigue was carried out on railway axles in the 1850s. Since then a lot of knowledge has been gained, but fatigue in general, and fatigue of railway wheels in particular, is a problem that is still far from fully solved. Some reasons for this are discussed below.

7.1.2 Description of the problem

Material fatigue is the process in which small cracks in a component initiate and grow under the influence of a repeated (normally mechanical) loading, possibly to a final failure. Fatigue analysis aims at quantifying the effect of changes in loading and component geometry, etc. on the resulting fatigue life. Fatigue design is the process in which fatigue analysis is employed to keep cracking within acceptable limits and to predict effects of changes in operational conditions. This also includes the identification of causes and efficient mitigating actions when failures have occurred. Naturally fatigue design needs to be accomplished within economical and operational frames. This chapter deals with fatigue design of railway wheels.

7.1.3 Description of the chapter

The text begins with a discussion on material fatigue in general and on wheel fatigue in particular. Here it is important to note that 'wheel fatigue' is not a single phenomenon. Plain fatigue can occur in the wheel disc, rolling contact fatigue can occur in the wheel rim and on the wheel tread. Further, thermal loading can add to the fatigue impact.

After describing wheel fatigue mechanisms, the focus turns to means of numerical analysis and prediction of wheel fatigue. This section does not aim to cover all predictive models, but rather to highlight general approaches by means of some selected models. A section that discusses the interaction between wheel fatigue and other wheel–rail interface-related phenomena follows. Finally, an attempt is made to indicate in which areas knowledge is lacking today and in which direction future trends should go.

Any attempt to describe such a wide and complex field as wheel fatigue

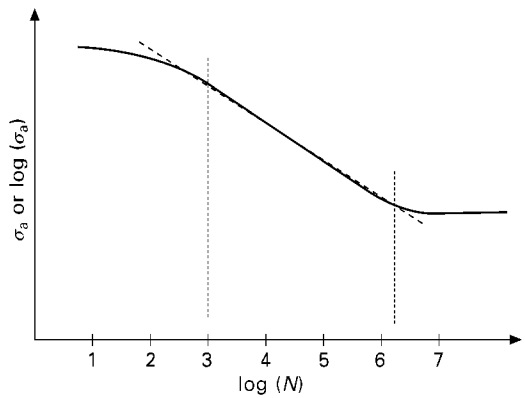
in a single chapter is more or less futile. Significant areas will be missing or glossed over, the text will be coloured by its author's (mis)perception of the subject, etc. To some extent this is mitigated by other chapters in the book that deal with related subjects. In addition, the chapter concludes with some recommendations for further reading to broaden the perspective.

7.2 Appearance and mechanisms of wheel fatigue

7.2.1 Fatigue

In classic fatigue analysis of components subjected to alternating stresses, the fatigue life is often related to the stress amplitude. Graphically this relation can be plotted in a stress–number (S–N) curve, also referred to as a Wöhler-curve,[1] as outlined in Fig. 7.1. The fatigue life, in numbers N of load cycles, is here related to the stress amplitude σ_a as a straight line in a lin–log or log–log diagram between fatigue lives of roughly 10^3 and 10^6 load cycles. Note the asymptote at a fatigue life of 10^6–10^7 load cycles. The corresponding stress amplitude is denoted the fatigue limit. For fatigue lives shorter than roughly 10^4 cycles, plastic deformation will occur and the corresponding fatigue phenomenon is referred to as low-cycle fatigue (LCF). The LCF life is normally predicted by the magnitude of the strain amplitude.

In fatigue design, it is important to recall that fatigue is essentially a threshold phenomenon. As seen in Fig. 7.1, the fatigue life is (theoretically) infinite at stress (or strain) magnitudes below the fatigue limit. At higher loads



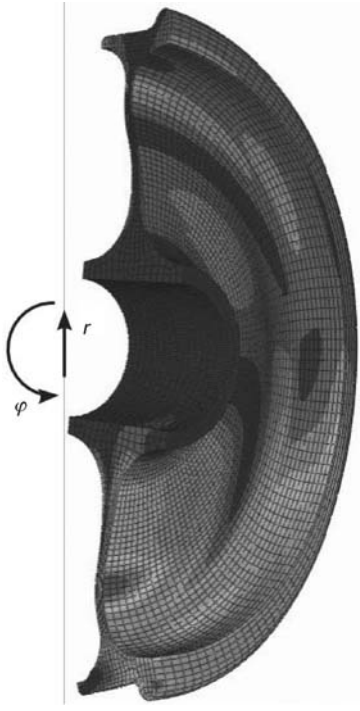
7.1 The Wöhler (*S–N*) curve.

[1]Named after August Wöhler (1819–1914) who did his pioneering work on fatigue life analysis of railway axles in the 1850–60s.

the fatigue life decreases exponentially with the stress (strain) magnitude. In the case of a large number of components loaded to roughly the same level and when this level increases (or the fatigue resistance of the material decreases) slightly, the result may be a rapid increase in the number of fatigue failures for the component studied. 'Load' and 'failure' should here be interpreted in a very broad sense.

7.2.2 Different kinds of wheel fatigue

The stress field in a loaded railway wheel is sketched in Fig. 7.2. Consider a material point in the wheel disc (properly, a small material volume around a point fixed in the rotating wheel). This point will be in a fairly unloaded state when it is positioned in the upper part of the wheel. During a revolution the load will be increased when the material point is located in the highly-loaded portion of the wheel disc and then decrease as the point is moving out of the highly-loaded portion. Consequently, a wheel revolution will correspond to a load cycle.



7.2 Stress field in a fully worn wheel. Curving load case according to EN 13979-1 (CEN/TC 256, 2003). (Picture Tore Vernersson/Copyright Lucchini Sweden)

Depending on the point studied in the wheel disc, the stress field is more or less biaxial (i.e. only the stress components σ_φ, σ_r and $\tau_{r\varphi}$ are significant). For three simultaneously acting stress components, the definition of a 'stress amplitude' σ_a is not straightforward. Normally an equivalent scalar stress measure (similar to the use of von Mises or Tresca effective stresses in plasticity analysis) is employed to characterise the fatigue impact, see e.g. Socie and Marquis (2000). In principle, the derived equivalent stress could be introduced in a Wöhler curve to give the resulting fatigue life. Alternatively, it could be ensured that the equivalent stress magnitude is below the fatigue limit, in which case the fatigue life will (theoretically) be infinite. Although such an approach is allowed by current codes, a more simplistic approach is generally adopted (CEN/TC 256, 2003).

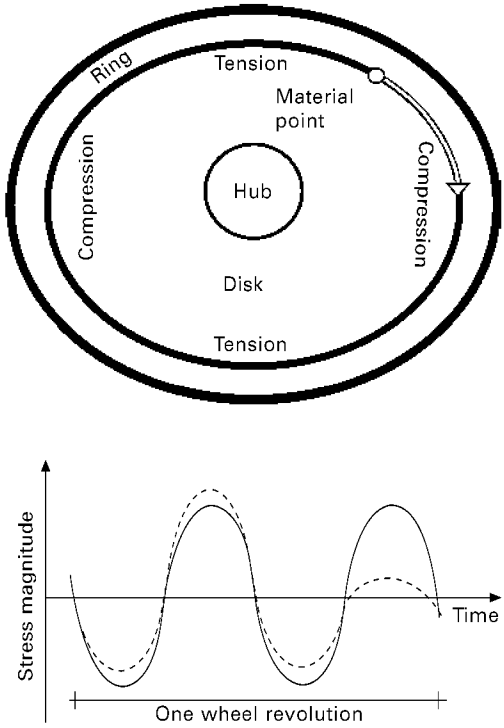
Of particular interest for wheel fatigue is the occurrence of holes in the wheel disc, e.g. for mounting of brake discs. A hole will cause a stress concentration that will locally increase the stress magnitude, potentially to above the fatigue limit. It is therefore vital that holes are placed in positions where stress magnitudes are low.

Now consider a resilient wheel with an elastic layer between the solid wheel disc and the tyre. When loaded the tyre will be ovalised. A material point on the inside of the tyre will thus, during a revolution of the wheel, be subjected to a stress history according to Fig. 7.3, i.e., two load cycles per revolution. If the ovalisation is allowed to be large, the state of stress in the material point will be approximately uniaxial. Also, fatigue cracking here is largely a consequence of stress concentrations, in this case normally due to material defects or scratches on the inside of the wheel rim. More details on this kind of fatigue may be found in Esslinger *et al.* (2004).

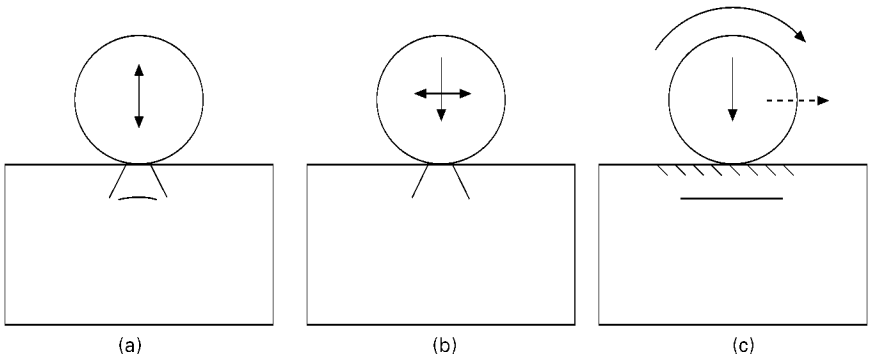
Fatigue in the wheel disc or on the inside of a resilient tyre is not related to the topic of this book, the wheel–rail interface, other than that the risk of fatigue increases with an increased total contact load on the wheel. In contrast, fatigue initiated in the wheel rim is closely related to the contact stress field. This fatigue phenomenon is generally referred to as rolling contact fatigue (RCF).

Consider a material point in the rim of a solid railway wheel. A load cycle is easily identified since it will correspond to the passage through the rail–wheel contact stress field, i.e. one load cycle equals one revolution of the wheel. However, since the contact stress field is multiaxial and has six independent components at each material point, we will have to adopt an equivalent stress criterion as described above. In addition, the overall loading is essentially compressive. This will increase the resulting fatigue life, but to what extent depends on the load configuration.

Consider two components in mechanical contact that are subjected to pulsating compression, as in Fig. 7.4a. If the load magnitude is sufficiently high and a sufficiently large number of load cycles are applied, cone cracks



7.3 Ovalisation of a wheel ring and sketched evolution of stress for a material point on the inside of a wheel ring. The solid line represents an idealised case, whereas the dashed line represent a more realistic case where the effect is more pronounced at the bottom of the wheel close to the wheel–rail contact.



7.4 (a) 'Plain' contact fatigue. (b) Fretting fatigue. (c) Rolling contact fatigue (RCF).

and/or subsurface cracks will appear. However, the cracks will only grow when they are within the contact stress zone. Once they have grown out of this stress field they will be arrested.

If the compression is combined with a tangential motion the result will be so-called fretting fatigue. The interfacial friction between the bodies will cause wear and create a stress concentration at the stick–slip interface (see Chapter 3). This will significantly promote crack initiation. Due to the interaction between wear and fatigue there are two possible scenarios. (i) If the tangential motion is sufficiently high, wear will dominate and initiated cracks will be worn off. The result will be a long fatigue life, but severe wear. (ii) If the tangential motion is less, fatigue cracking will dominate. Also in this case, fatigue cracks will be arrested once they have grown out of the contact stress zone unless there is an additional loading of the bulk material.[2]

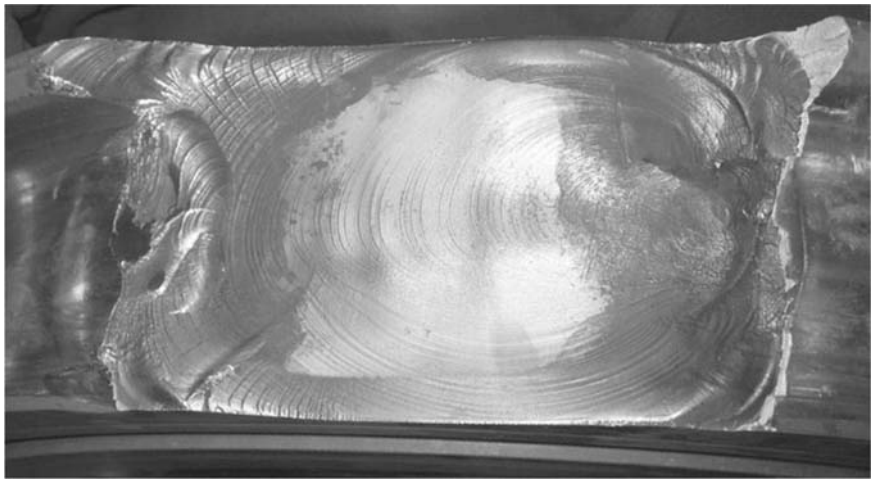
The third case is that one component is rolling (with or without friction) on the other. The contact stress zone will then move along with the rolling component. It will therefore be possible for initiated cracks to propagate to a significant length. Depending on the characteristics of the loading and the material resistance, RCF cracks may initiate at or below the surface. We will study these cases separately below.

7.2.3 Subsurface-initiated rolling contact fatigue

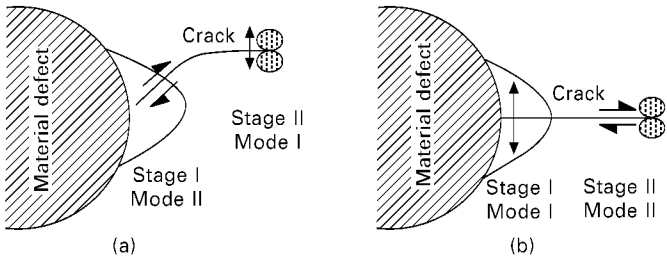
An example of a failure due to subsurface-initiated RCF is given in Fig. 7.5. Such failures are rare, but the consequences are often severe. In a worst case scenario the train may derail, but also when the train does not derail the impact of the cracked wheel may cause significant damage to both rail and bogie.

To understand the mechanisms behind subsurface-initiated RCF we need to understand both the contact mechanics and the mechanisms behind fatigue crack initiation and growth. In plain fatigue (i.e., fatigue due to repeated tensile or tensile/compressive uniaxial loading), the crack initiation phase is related to shear deformation at pre-existing micro-cracks in the material. The initiation stage includes the formation and growth of small macro-cracks and is often referred to as Stage I, whereas the shear deformation that drives crack growth is referred to as Mode II. The growth of large cracks (sometimes referred to as Stage II) is instead driven by an opening mode deformation of the crack (Mode I) (see Fig. 7.6a). Plain fatigue cracks are normally initiated at the surface of the material in connection to surface defects and scratches.

[2]This is generally the case of wheel–axle assemblies where the axle is subjected to bending in addition to the fretting loading of the assembly.



7.5 Wheel fracture caused by subsurface-initiated RCF.



7.6 Short crack growth at a material defect with global loading direction being perpendicular to the long crack. A grey colour indicates plastic zones; initially at the side of the defect and at the tip of the long crack. (a) Plain uniaxial fatigue. (b) Rolling contact fatigue.

However, they can also initiate just below the surface in the presence of material defects. Figure 7.6a shows such a case.

Now consider a material volume with a defect (a pore or an inclusion of a soft foreign material) that is subjected to a compressive (rolling contact) loading. Due to stress concentration, the largest compressive stress will occur at the side of the defect. If sufficiently large, the material will yield in compression and, upon unloading, a tensile residual stress field will form close to the defect. Consequently, a crack in this location will be subjected to an opening mode deformation. However, this effect is very localised to the side of the defect. Once the crack has grown out of this zone of tensile residual stress, it is propagated by the global shear stress (see Fig. 7.6b). This is exactly the opposite to plain fatigue. Naturally, the discussion above

assumes idealised conditions; in reality the two crack driving mechanisms are likely to interact.

In pure rolling under Hertzian contact conditions (see Chapter 3), the largest shear stress in a railway wheel occurs some few millimetres below the surface. Consequently, this is where we would expect the initiation of cracks. However, the reality is more complicated, as initiation and growth of subsurface-initiated RCF is influenced by the occurrence of residual stresses, distorted microstructure (due to operations, forging, casting, etc.) and material defects. In practice, the initiation of subsurface RCF will therefore occur at depths of some 4–25 mm below the surface. In general, shallow initiation is highly influenced by the contact geometry, whereas deep initiation is more related to vertical load magnitudes and material defect sizes. A more shallow subsurface initiation than at 4 mm may be possible. Theoretically, this should relate to small local contact radii. However, at these depths subsurface cracks are likely to interact closely with surface-initiated cracks and it would be difficult to distinguish root causes (cf 'squats' on rails as discussed in Chapter 13).

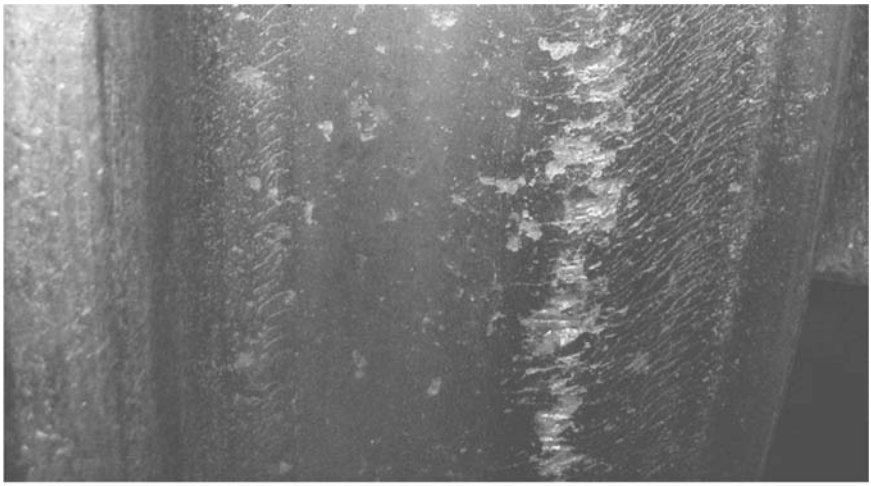
The growth direction of larger shallowly initiated subsurface cracks typically deviates towards the wheel hub. The reason is likely to be the influence of crack face friction. The larger the crack, the higher will be the interfacial friction between the crack faces. Thus, even if a deviation towards larger depths will decrease the magnitude of the shear stress acting on the crack, it will also decrease crack face friction so that the total effect will be an increased fatigue impact. The deviation may also be promoted by the residual stress field.

Final failure due to subsurface-initiated RCF typically occurs as a branching towards the surface. At failure, the crack often has a diameter larger than a decimetre. However, there is a large scatter and there does not seem to exist a well-defined critical crack size as is typically the case for plain fatigue.

7.2.4 Surface-initiated rolling contact fatigue

Surface-initiated RCF cracks are more common, but less severe than subsurface-initiated cracks. In fact, if there are no signs of surface-initiated RCF on a wheel fleet, the loading is likely to be lower than what is optimum from an economical point of view. Still, this is a delicate balance. When the loading (in a broad sense) is increased above a certain level, an epidemic of surface initiated cracks will occur with cracking frequencies and severity exceeding acceptable levels. The remedy to surface-initiated RCF is re-profiling, which is costly and calls for taking trains out of traffic. The typical appearance of severe surface-initiated RCF is shown in Fig. 7.7a,b.

Surface-initiated RCF is related to the frictional loading between the wheel and the rail. If the frictional stress in the interface exceeds the yield



(a)

(b)

7.7 (a) Overview of severe surface-initiated rolling contact fatigue. Picture: Roger Deuce/Copyright Bombardier Inc. (b) Magnification of a deep surface cavity, roughly 2 cm wide, caused by a combination of surface-initiated RCF and thermal loading.

limit in shear of the wheel material, the result will be plastic deformation of the surface material. This will result in a deformation ('mangling out') of a thin surface layer, which eventually may result in fatigue crack initiation and growth of small cracks at a shallow angle into the wheel. The crack mouth is (ideally) orientated perpendicularly to the resultant direction of the frictional force. This may provide a good indication of whether curve negotiation or braking friction is the main cause of the frictional loading (Magel and Kalousek, 1996).

As the cracks grow deeper, the propagation will be influenced by a possible penetration of fluid. Fluid trapped in a crack will be pressurised during the rail contact. This results in decreased crack face friction and an opening loading (Mode I) of the crack that promotes growth (Bower, 1988). From an operational point of view, this should not be interpreted as if lubrication should be avoided; a proper lubrication will delay the initiation of surface cracks. However, if liquid lubrication is adopted, it must be consistently used. As is well known from laboratory testing (Tyfour *et al.*, 1996), the most detrimental combination is a dry phase to initiate cracks followed by a lubricated phase. A natural version of such conditions occurs during winter when the conditions generally are dry, but snow, melting ice and frost may lubricate initiated cracks. This seasonal effect is clearly manifested in wheel failure statistics (Kalousek *et al.*, 1996).

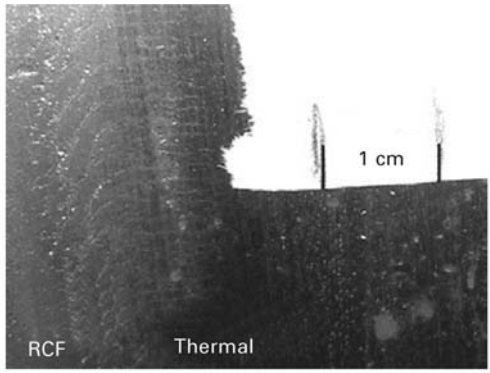
The typical morphology of surface-initiated cracks is an initial shallow growth followed by a more radial growth. At a depth of some millimetres, branching and deviation to a growth orientation perpendicular to the wheel rim tends to occur. By the joining of cracks, pieces of the wheel rim may break loose. The result is a cavity as shown in Fig. 7.7b.

7.2.5 Thermal damage

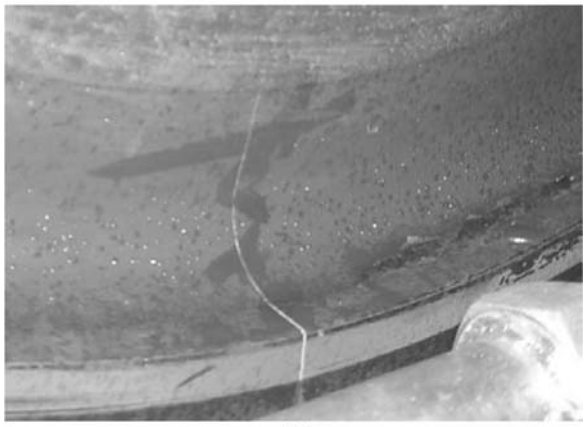
In this section, thermal damage relates both to the cracking of a wheel resulting from thermal loading and to material transformation due to elevated temperatures in the wheel. In practice these phenomena normally appear in parallel with one dominating the damage pattern.

The typical appearance of mild thermal fatigue is a 'brick' or 'crocodile skin' pattern on the wheel rim as seen in Fig. 7.8a. A more severe thermal loading will give rise to transverse cracks that may extend several millimetres radially into the wheel tread. In extreme cases the result may even be a complete failure of the wheel when the crack propagates into the wheel disc, as seen in Fig. 7.8b.

The cause of thermal cracking is a too severe heating of the outer layer of the wheel tread. This is normally related to tread braking although other mechanisms of friction generation (such as wheel–rail friction) also add to the heating. The suppressed thermal expansion in the circumferential direction



(a)



(b)

7.8 (a) Mild thermal fatigue resulting in a 'brick pattern' just left of the marking tape. Further left, the thermal cracking is overtaken by rolling contact fatigue due to curving (note the inclined cracks). (b) A thermal crack that has propagated through the rim and up into the wheel disk, seen from the field side of the wheel.

will cause compressive yielding in the outermost layer of the wheel. During subsequent cooling, tensile circumferential stresses are induced and cause cracking. In mild cases, the cracks are only micrometers deep. Often, but not always, the cracking is accompanied by heat transformation of the material. The severity can range from a slight shading that is only visible after etching to the formation of wheelflats with a large martensitic zone (see Fig. 7.9 and Chapter 8).

Apart from being a problem in its own right, thermal fatigue will also interact with mechanical fatigue. Initiated thermal cracks may continue to propagate under the influence of mechanical loading. Similarly, the propagation

7.9 Wheel flat formed when the wheel has been sliding on the rail.

of mainly mechanically induced cracks is promoted by thermal loading. The dominating cause can normally be identified from the inclination of the crack into the wheel material: the steeper the inclination, the larger the thermal influence. As an example, the steep inclination of the surface crack in Fig. 7.7b indicates that it has formed under the combined influence of rolling contact and a significant thermal loading.

7.3 Prediction of wheel fatigue

There are a multitude of different numerical models to analyse wheel fatigue reported in the literature. Here, the presentation will focus on fatigue initiation in the wheel rim. Only some selected models are presented. These have the benefit of being described in analytical form, which facilitates identification of the relative influence of the parameters included. Some additional methods are briefly described in this section and in Section 7.4. For further information, the reader is directed to the references.

7.3.1 Subsurface-initiated rolling contact fatigue

To facilitate the prediction of subsurface-initiated RCF, we will distinguish between the cases of shallow initiation (say, from some 4 mm to some 10 mm below the wheel tread) and deep initiation (say, at some 10–25 mm below the tread).

For the shallow cracks we presume fatigue initiation to be related to the magnitude of the shear stress. Assuming Hertzian contact conditions the maximum shear stress will occur some millimetres below the tread surface and have an approximate magnitude of:

$$\tau_{\max} \approx \frac{p_0}{3} = \frac{F}{2\pi ab} \qquad [7.1]$$

Here p_0 is the peak normal contact pressure, F the total normal contact load (the vertical wheel load including dynamic contributions) and a and b the semi-axes of the Hertzian contact patch.

Since the state of stress in the wheel rim is multiaxial, an equivalent stress criterion will be employed to quantify the fatigue impact. Here we use a multiaxial fatigue criterion (Dang Van *et al.*, 1989), which can be expressed as:

$$\sigma_{EQ} = \tau_a + a_{DV}\sigma_h \qquad [7.2]$$

where a_{DV} is a positive material parameter and $\sigma_h = (\sigma_x + \sigma_y + \sigma_z)/3$ is the hydrostatic stress (positive in tension).

For wheel–rail contact conditions there is a pulsating shear stress evolution on the shear plane experiencing the highest 'amplitude'. The shear stress 'amplitude' τ_a is for this case half of the maximum shear stress:[3]

$$\tau_a = \frac{F}{4\pi ab} \qquad [7.3]$$

If additional frictional loading of moderate magnitude is applied, the maximum shear stress may be estimated at:

$$\tau_{\max} = \frac{F}{2\pi ab}(1 + f^2) \qquad [7.4]$$

Here $f = \sqrt{F_x^2 + F_y^2}/F$ where F_x and F_y are tangential forces.

Combining Eqs (7.3) and (7.4), the equivalent stress can be expressed as:

$$\sigma_{EQ} = \frac{F}{4\pi ab}(1 + f^2) + a_{DV}\sigma_{h,res} \qquad [7.5]$$

where $\sigma_{h,res}$ is the hydrostatic part of the residual stress. Equation (7.5) is valid for wheel–rail contact. For line contact (e.g. in a twin-disc test) a similar approximation is given in Ekberg *et al.* (2004) following a discussion in Ciavarella and Maitournam (2004).

Close to the wheel rim, the residual stress is compressive which should result in a negative term $a_{DV}\sigma_{h,res}$ and consequently a reduction in the equivalent stress magnitude. It has been shown, however, that under conditions resembling those in a railway wheel rim, such a beneficial effect does not

[3]This is a simplistic reasoning. In reality the shear stress vector will follow a path in the shear plane (see Ekberg, 1997).

exist (Desimone *et al.*, 2006). A pragmatic approach to account for this is to put $\sigma_{h,res}$ equal to zero.

Fatigue is predicted to occur if $\sigma_{EQ} > \tau_{FL}$ where τ_{FL} is the fatigue limit in shear, which is roughly $\sigma_u/3$ where σ_u is the tensile fracture stress of the material. However, as mentioned above, subsurface-initiated RCF is closely connected to the presence of material defects. Due to the large number of operating railway wheels and the large loaded material volume of each wheel, statistics of extremes should preferably be used to define a 'worst case' defect size (Beretta *et al.*, 2006). As a rough estimation, a defect with a diameter of 1 mm (typical allowed size for high-speed applications) in a critical position will decrease the fatigue limit by 50 % (Ekberg and Sotkovszki, 2001). It should be noted that only defects larger than about 1 mm can be reliably detected in today's ultrasonic inspections.

As seen above, the size of the contact patch (πab) will influence the magnitude of the fatigue impact (σ_{EQ}). In addition, the shape of the contact patch will have an influence. In wheel–rail contact, the patch is normally elongated ($a > b$) in the rolling direction, and the orientation of the material plane subjected to the largest σ_{EQ} will then be inclined by 45° to the wheel rim. In line contact, as for example in twin–disc testing, the critical plane is parallel to the contacting surface (Ekberg *et al.*, 2004). This is of practical interest since the material in a forged wheel is highly anisotropic (Ekberg and Sotkovszki, 2001). Anisotropy and the variation of material properties with the depth below the wheel tread surface must also be considered when wheel test samples are to be extracted from the wheel rim.

If we look at RCF initiation deeper into the wheel rim, interfacial wheel–rail friction and the size and shape of the contact patch will have less influence. Instead, the magnitude of the total contact load and the occurrence of material defects will be the main influential factors (Kabo, 2002; Kabo and Ekberg, 2005).

For the analysis of long cracks at large depths (i.e. 10–25 mm), linear elastic fracture mechanics (LEFM) is often used. Consider an elastic material with a subsurface crack oriented in parallel to the contact surface where a point load is acting. For the two-dimensional (2D) case, the Mode II stress intensity range ($\Delta K = K_{II}^{max} - K_{II}^{min}$) can be derived under the presumptions of friction-free crack face and wheel–rail interfaces (Hearle and Johnson, 1985; Ekberg *et al.*, 2007) as:

$$\Delta K_{II} = \frac{3F}{4h}\sqrt{\frac{3L}{2\pi}} \qquad [7.6]$$

Here F is the magnitude of the line load, h the depth of the crack below the surface and L the length of the crack.

If the same simplification with friction-free interfaces is employed in a 3D analysis of a penny-shaped crack loaded by a uniform shear stress due

to a point load at the wheel surface, a stress intensity range can be derived (Ekberg *et al.*, 2007) as:

$$K_{II} = \frac{192}{25\sqrt{5}(2-\nu)\pi\sqrt{\pi}} \cdot \frac{\sqrt{r}}{h^2} F \qquad [7.7]$$

Here r is the radius of the crack, ν Poisson's ratio and F the contact load. It is seen that Eq. (7.7) breaks down if the crack is located too close to the wheel tread ($h \rightarrow 0$).

Crack growth rates can be predicted from ΔK_{II} magnitudes using a Paris law type of equation, i.e.

$$\frac{da}{dN} = C_{II}(\Delta K_{II})^{m_{II}} \qquad [7.8]$$

Here da/dN is the crack growth increment per cycle, and C_{II} and m_{II} are material parameters. An example of such a prediction is given in Liu *et al.* (2007).

If one should generalise, the main problem with a prediction based on fracture mechanics is how to account for the compressive loading manifested as crack face friction. On the other hand, it has been shown that the presumption of limited plasticity inherent in the LEFM theory is fulfilled (Lansler *et al.*, 2006). For analysis based on an equivalent stress approach, the main stumbling block is how to quantify the influence of material defects. It should also be mentioned that testing is very cumbersome. As an example, it is not obvious how to interpret results from scaled testing since the microstructure of the material is not scaled, whereas the stress gradients are.

7.3.2 Surface-initiated rolling contact fatigue

As described above, initiation of surface cracks is related to plastic deformation of the material in the surface layer of the wheel rim. In addition, irregularities on the contact surface (e.g. due to surface roughness or cavities) will give rise to locally high stress magnitudes, which may promote further damage (Kapoor, 1994).

The main cause of global plastic deformation of the surface material is the applied interfacial shear stress between the wheel tread and the rail.[4] As the frictional loading increases from zero, the depth of the location of the maximum shear stress decreases (and the magnitude increases) slightly as compared to pure rolling. At a traction coefficient of roughly $f \approx 0.3$ the location of maximum shear stress will jump to the surface and its magnitude

[4]High shear magnitudes also exist in the interfaces between wheel and brake blocks. However, here the main effect is normally thermal loading and related damage.

increases strongly with increasing interfacial friction. In addition, the shear stress magnitude decreases rapidly with depth.

To predict surface-initiated RCF, a shakedown map (Johnson, 1989) is commonly adopted. Here full slip is assumed, i.e. the shear stress (traction) $q(x,y)$ in the wheel–rail interface is presumed to be proportional to the contact pressure via the traction coefficient f:

$$q(x, y) = f \cdot p(x, y) \qquad [7.9]$$

The peak traction will thus be:

$$q_0 = f \cdot p_0 \qquad [7.10]$$

where p_0 is the peak contact pressure, which according to Hertzian theory is expressed as:

$$p_0 = \frac{3F}{2\pi ab} \qquad [7.11]$$

A criterion for surface plasticity may be expressed as:

$$q_0 = k \qquad [7.12]$$

where k is the yield limit in shear of the wheel material.

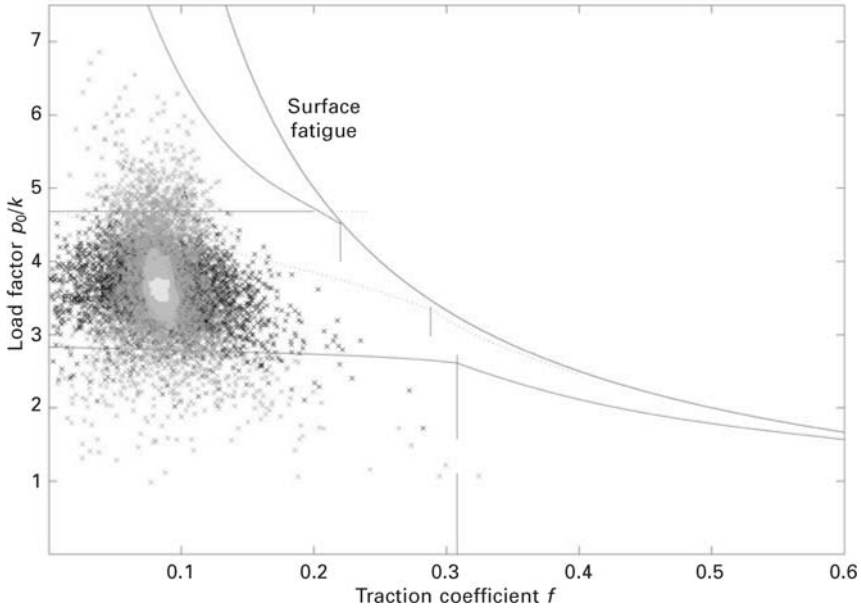
A combination of Eqs (7.10)–(7.12) results in the yield criterion:

$$f \cdot p_0 = k \Leftrightarrow \frac{3F}{2\pi abk} = \frac{1}{f} \qquad [7.13]$$

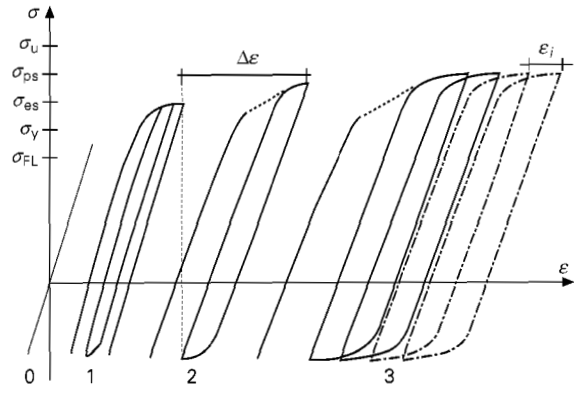
This equation corresponds to the top right curve in the shakedown map in Fig. 7.10. If our operational conditions correspond to a combination of f and p_0 outside this line, we will exceed the material's yield limit. Plastic deformation and fatigue cracking are then presumed to follow.

Another approach to predicting surface-initiated RCF would be to evaluate fatigue life from explicitly evaluated plastic deformations. To this end elastoplastic finite element (FE) simulations are needed. It is important here to understand how the material responds to high load magnitudes. There are basically three scenarios (see Fig. 7.11):

1. Residual stress formation and related stress re-distribution in the wheel rim may be sufficient to eventually result in an elastic response (elastic shakedown). We will then expect cracking to be a relatively long-term process.
2. Plastic deformation is not suppressed, but stress re-distribution and plastic hardening causes it to be stabilised (plastic shakedown). In other words, even if the loading acts in one direction, there will *eventually* be no net flow in this direction.
3. The plastic deformation will remain unstable and there will be an incremental strain growth at each load cycle (ratcheting).



7.10 Shakedown map. Low-pass filtered response from simulated operation on corrugated track with cut-off frequency 90 Hz (light grey), 200 Hz (grey) and 1000 Hz (dark grey) compared to the non-filtered response (black). (Adopted from Nielsen *et al.*, 2005)



7.11 Types of material responses in uniaxial loading: elastic (0), elastic shakedown (1), plastic shakedown (2) and ratcheting (3). The denoted stress amplitudes are the fatigue limit σ_{FL}, the yield limit σ_Y, the elastic shakedown limit σ_{es}, the plastic shakedown limit σ_{ps}, and the fracture stress σ_u.

Predictive models can set out from any of these mechanisms although elastic shakedown conditions are normally of little practical interest for the case of surface initiated RCF cracking of wheels.

If the assumed main material response is in the plastic shakedown stage, a LCF criterion that is able to account for the multiaxial state of stress and strain should be used. An example is given in Jiang and Sehitoglu (1999) where a fatigue parameter FP is defined as

$$FP = \Delta\varepsilon \cdot \langle\sigma_{\max}\rangle + J \cdot \Delta\gamma\,\Delta\tau \qquad [7.14]$$

The criterion is evaluated for a material plane where $\Delta\varepsilon$ is the range of the normal strain acting on the plane, $\langle\sigma_{\max}\rangle = \max(\sigma_{\max}, 0)$ with $\sigma_{\max}$ being the largest normal stress on the plane. Further, J is a material parameter, $\Delta\gamma$ the range of the (engineering) shear strain and $\Delta\tau$ the range of the shear stress. Note that the evaluation of $\Delta\gamma$ and $\Delta\tau$ needs to account for the fact that these are vector quantities.

The fatigue life N can then be evaluated by an extension of traditional (uniaxial) LCF predictive models as:

$$FP = \frac{(\sigma'_{\mathrm{f}})^2}{E}(2N)^{2b} + \sigma'_{\mathrm{f}}\varepsilon'_{\mathrm{f}}(2N)^{b+c} \qquad [7.15]$$

where E is the elasticity modulus, and σ'_{f}, $\varepsilon'_{\mathrm{f}}$, b and c are additional material parameters.

If, instead, ratcheting is considered as the dominating damage mechanism, a ratcheting criterion (Kapoor, 1994) may be employed:

$$\sum_i \varepsilon_i = \varepsilon_{\mathrm{c}} \qquad [7.16]$$

Here ε_i is the current strain increment and ε_{c} the fracture strain. When applied to surface-initiated RCF in railway wheels, ε_i needs to account for the multiaxial state of strain. An approach to this is to define an effective strain increment in line with the use of equivalent stresses discussed above (Kapoor, 1994).

Regardless of which approach is adopted, it is vital that the constitutive model employed can correctly predict the evolution of the plastic deformation. Consider Fig. 7.11. A LCF prediction according to Eqs (7.14) and (7.15) needs a precise estimation of the strain range $\Delta\varepsilon$, whereas a ratcheting prediction according to Eq. (7.16) needs a proper evaluation of the strain increment ε_i. The latter, especially, is very cumbersome since it calls for both sophisticated constitutive models and the simulation of a large number of load passages to obtain a stabilised response (Ringsberg, 2000).

Analyses based on both LCF and ratcheting predict the overall fatigue life. If the aim instead is to explicitly predict the growth of surface-initiated cracks, fracture mechanics methods need to be adopted. This is more complicated than for the analysis of subsurface-initiated RCF. Firstly, there is the interaction

with lubricants that needs to be dealt with if relatively large cracks are studied (for smaller cracks this should be less of an issue for wheels than for rails due to the rotational speed of the wheels). This issue is covered in Chapter 9. Secondly, the cracks are propagating in a plastically deformed material. This implies that material anisotropy may need to be considered. It also implies that methods based on linear elastic fracture mechanics are questionable. There are alternative methods presented in the literature, such as the use of material forces or the J-integral (Bergkvist, 2005). However, these methods need to be supported by more research before becoming operational. There are also approaches that combine the analysis of wear and RCF. These will be discussed in Section 7.4.

7.3.3 Thermal fatigue

Thermal loading may result in different types of damage as discussed in Section 7.2. It may cause material phase transformations with martensite formation (which causes embrittlement of the material, an increase in material volume and a subsequent formation of tensile residual stresses) as an extreme example. This thermomechanical problem has been numerically and experimentally studied (see e.g. Jergeus, 1998).

Significant heating may also decrease the yield limit, thus promoting surface-initiated RCF (and wear). This has been studied in the literature (Böhmer *et al.*, 2003) by modifying the shakedown map to account for thermal stresses.

Finally, the thermal loading may cause cracking of its own accord. To numerically predict this phenomenon, a simulation of the thermal loading and the corresponding yielding followed by the cooling and formation of tensile residual stresses can be carried out (Vernersson, 2006). A fracture mechanics analysis can then be performed to evaluate to which depth cracking is likely to occur.

7.4 Wheel fatigue put in context

In this section, interactions between wheel fatigue and other phenomena and parameters are discussed. It should be noted that this discussion is not conclusive and that the focus is on primary interactions. Secondary interactions, such as the influence of surface cavities due to RCF on noise generation, are not considered.

7.4.1 Track geometry

The influence of track geometry on wheel fatigue can be divided into its effect on the induced loads and that on the contact geometry. Poor track

geometry will result in higher vertical and lateral track forces. This will increase the risk of subsurface and surface-initiated RCF, respectively. The other factor of importance is the contact geometry. As seen in eqs (7.5) and (7.13), a decreased contact patch size will increase the risk of RCF. Such a decrease may be the result of worn rail profiles or a mismatch between wheel and rail profiles.

In curves with rail gauge that is too wide, the wheel–rail contact may be located far out on the field side of the wheel tread. The close vicinity to the free edge of the wheel will promote plastic flow, especially in heavy-haul applications. This may be manifested in tread rollover ('lipping') (see Fig. 7.12a). In addition, a high vertical load close to the field side will result in high subsurface shear stress magnitudes. The consequence may be subsurface crack initiation leading to 'spread rim' failures (see Fig. 7.12b). Such cracks propagate in parallel to the wheel rim at a shallow depth and may reach a considerable size before a portion of the wheel tread is detached.

7.4.2 Corrugation and out-of-round wheels

Operations on corrugated rails and/or with out-of-round wheels will introduce an additional dynamic loading. For the case of moderate to high-speed operations, there are large contributions in the frequency range 200–1000 Hz. An illustration of this is given in Fig. 7.10 where predicted vertical and lateral load magnitudes are presented in a shakedown map (see Section 7.3.2). Low-pass filtering at 200 Hz and 90 Hz, respectively, results in marked decreases in the registered load scatter. The consequence is that the effect of corrugation and out-of-roundness on increased load magnitudes can be analysed neither by using conventional multi-body dynamics simulations nor by employing traditional wheel–rail contact force measurements, since these do not capture high enough frequencies. More details on these phenomena are given in Chapter 8.

Operations on a stretch of corrugated track will result in a fatigue impact distribution. Numerical simulations of high-frequency vertical train–track dynamics (Ekberg *et al.*, 2007) with integrated fatigue analysis show that an increased axle load increases the fatigue impact by increasing the mean magnitude of the vertical load and fatigue impact. The amount of undamped mass is found to have practically no influence. An increased speed and an increased corrugation magnitude will increase the scatter and thus the magnitude of the peak impact. Interestingly, there seems to be a saturation effect regarding the influence of speed, as seen in Fig. 7.13.

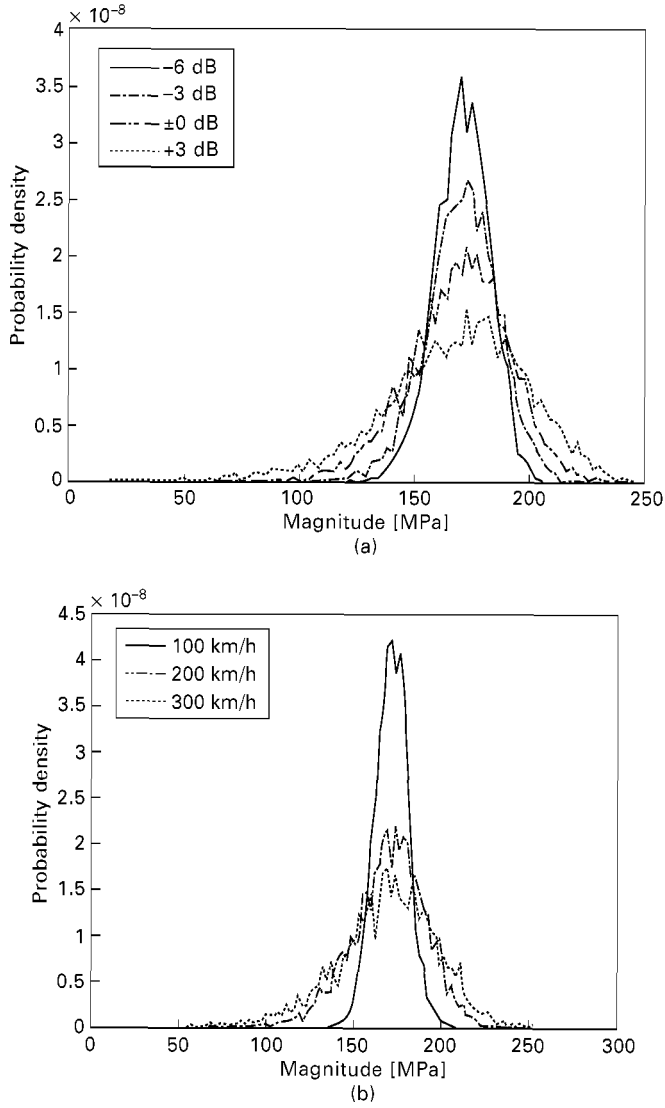
A detailed analysis showed that the increased vertical load magnitude was the main cause of the increased fatigue impact. For low-speed heavy-haul operations, the deteriorated contact geometry due to the corrugation will also have an effect. However, in this case the fatigue impact is lower.

(a)

(b)

7.12 (a) Moderate tread rollover due to plastic flow (Fröhling *et al.*, 2006). (b) 'Spread rim' failure (Ekberg and Marais, 2000).

The above discussion focuses on the effect of corrugated rails. The conclusions are largely valid also for operations with out-of-round wheels provided that the corrugated profiles are similar. One should note that an out-of-round wheel will be more severe with respect to wheel fatigue since



7.13 Frequency distribution of the Dang Van equivalent stress (Eq. 7.5) as a function of varying corrugation roughness (a) and speed (b). The base value of roughness comes from field measurements on a corrugated stretch. (From Ekberg *et al.*, 2007)

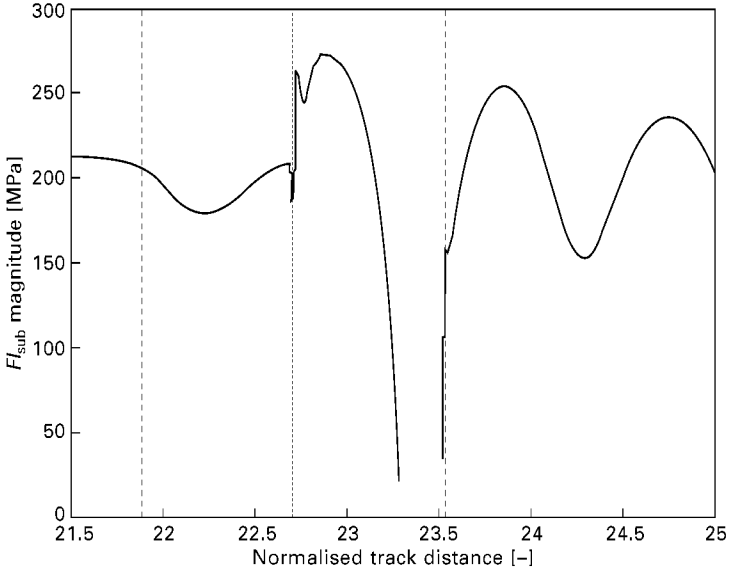
the wheel will experience increased loading at every revolution, whereas the influence of rail corrugation is confined to the corrugated stretches. In this context it should be remembered, however, that the roughness of a wheel tread is normally lower than that of a corrugated rail.

7.4.3 Single rail irregularities

The influence of single rail irregularities on wheel fatigue is normally rather limited. The reason is that a fixed material point in the wheel experiences relatively few loads due to single rail irregularities. Consequently, the accumulated damage due to these events will be limited. They could, however, have an effect in promoting final failure, which may range from detachment of a piece of the surface material up to complete wheel failure.

Figure 7.14 presents the evaluated fatigue impact (with $FI_{sub} = \sigma_{EQ}$ according to Eq. 7.5) during a wheel negotiation of an insulated joint. It is seen that the passage of the insulating pad between the ends of the two rails (at normalised track distance 22.7) results in a load peak. However, the irregularity also introduces a transient vertical vibration, which causes a loss of contact between the wheel and rail (normalised track distance 23.3–23.5).

Switches and crossings constitute another kind of rail 'irregularity'. The loading on a wheel during switch negotiation is very complex with poor contact conditions, high transient loads and multiple point contacts (Kassa, 2007). In general, wheels passing through switches are subjected to a high



7.14 Fatigue impact (according to Eq. 7.5) on a wheel passing an insulated joint. The wheel is travelling from left to right at a speed of 125 km/h. The axle load is 25 tonnes. The joint section of length 1 m and depth 3 mm (error in alignment between rail ends) is within the dashed lines. The dotted line gives the position of the insulating pad. (From Kabo *et al.*, 2006)

fatigue impact both regarding subsurface crack initiation (high vertical loads with major contributions in the high-frequency spectrum combined with poor contact geometries) and surface initiation (high lateral forces combined with poor contact geometries).

An important aspect here regarding wheel maintenance is the influence of hollow worn wheels (see below). These wheels generally cause poor steering. In addition, during switch negotiation, there is a risk that the wheel tread close to the field side impacts the crossing nose. Although the main fatigue damage will be on the crossing nose and not on the wheel, such a loading may promote 'spread rim' failures as discussed above.

A phenomenon that may appear in switches is flange back contact with check or wing rails. The result will be flange back wear, as shown in Fig. 7.15. This is in itself mainly a cosmetic issue, but the lateral impact loading at flange back contact will cause a lateral shift in the wheel position accompanied by a high interfacial wheel–rail friction.

7.4.4 Wheel wear

As the wheel tread wears during operation, small initiated cracks will wear away. This is especially apparent for cracks that are formed close to the flange root: if the wear-in process is too slow, these may grow to a non-acceptable size, which calls for re-profiling.

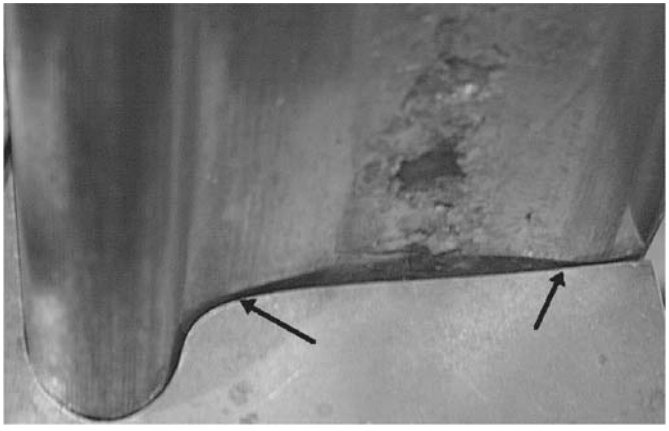
7.15 Flange back wear.

This balance between wear and rolling contact fatigue is the background to the concept of a 'magic wear rate' where crack formation is exactly balanced by wheel wear. The problem is that the wheel is not uniformly worn. Consequently, vehicle dynamics and contact conditions will change as the wheel wear progresses, a process that may even accelerate crack growth.

A very detrimental form of wheel wear is the so-called hollow wear, where the tread surface is worn into a concave form (see Fig. 7.16). This reduces or even eliminates the wheelset's ability to steer by a radius difference between the two wheels. The resulting high lateral forces in combination with the poor contact geometry result in a rapid initiation of RCF cracks at the surface. Hollow wear is commonly counteracted by strict limits on allowed wear depth (with 2 mm being a common value). However, numerical simulations (Fröhling *et al.*, 2008) indicate that this is not a conclusive measure of the impact of hollow wear.

A second wear-related phenomenon may occur if the diameter difference between the two wheels of a wheelset is too large. This may cause erroneous steering, resulting in very severe flange wear on one wheel and surface-initiated RCF on the other (Fröhling, 2006).

In some sense, wear and surface-initiated RCF can be seen as two manifestations of the same phenomenon. Both are caused by high frictional stresses in the wheel–rail interface and both result in the detachment of material from the wheel tread. What differs is the size of the detached metal chips. There are a number of predictive models based on these similarities that attempt to predict wear and surface-initiated RCF in a unified model. One example is the T–γ-model where the product of lateral force T and the contact patch creepage γ are employed in an empirical damage function



7.16 Hollow wear (between the arrows). Also note the severe surface-initiated RCF damage in the worn section of the wheel tread.

(Burstow, 2004). A second example is the brick model (Franklin and Kapoor, 2007) where the contacting body is discretised in bricks. Every load passage that induces a shear strain in the brick above the yield limit will add to the accumulated shear strain in that brick. Once a critical shear strain is exceeded, the brick is considered failed and is removed. Depending on the pattern of the failed material, the result may be characterised as 'crack-like' or 'wear-like'. These models have mainly been employed in the prediction of rail fatigue and are further described in Chapter 9.

7.4.5 Material characteristics

For avoiding surface-initiated RCF, the choice of material in the wheel seems to be quite straightforward. Judging from Eq. (7.13), the higher the yield limit, the better. Sadly though, reality has proven to be much more complex. It should first be noted that the yield limit k in Eq. (7.13) is not the monotonic yield limit, but rather the cyclic yield limit during operation. These two limits may be very different since residual stresses will arise and the material will transform at the onset of yielding. Normally, the material will experience a cyclic softening, but this effect is counteracted by the formation of compressive residual stresses. The result tends to be component hardening. However, in order for this process to be possible, the material needs a high ductility. Since harder steels normally have less ductility, the chances are that the needed deformation cannot occur without crack formation in the material.

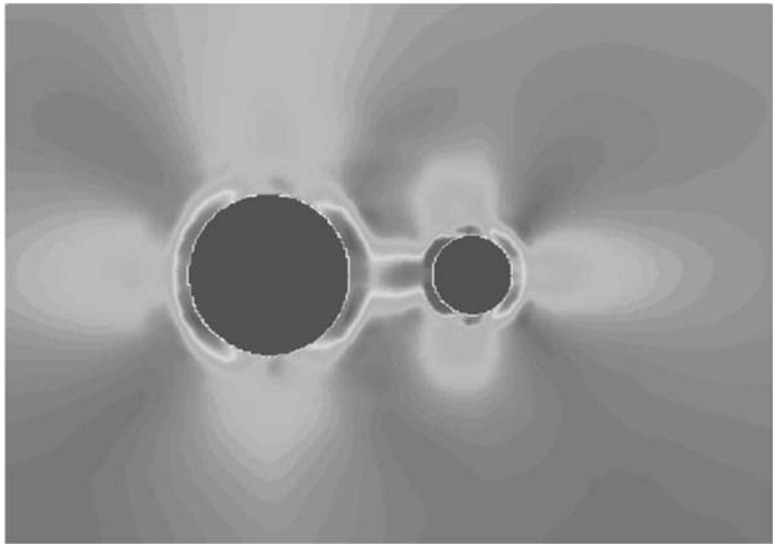
A second effect is the wear-in during which the wheel profile is being adapted to the rail profile.[5] This wear-in normally tends to decrease the contact pressure by increasing the contact patch size, although exceptions occur, such as for hollow worn wheels. A harder material tends to be less prone to wear, and consequently the wear-in process is delayed, during which time cracks may form and grow. Thirdly, since the carbon content of harder wheel steels is normally higher, these are more prone to thermal damage and, in particular, to martensite formation. This is an issue especially for tread braked wheels.

In conclusion, a hard wheel steel may very well give a higher resistance against surface-initiated RCF, but it comes at the expense of the wheel being much more sensitive to wheel–rail mismatches and thermal loading. In practice, a slightly softer material therefore often has a better performance. However, this tends to come at the expense of a poorer wear resistance. Consequently wheels of a soft material are not suited for conditions of hard tread braking or poor curving.

[5]In reality, it is more complicated than that since the wheel traverses a large number of different rail profiles with contact at different lateral position.

Promising alternatives include pearlitic steels with increased silicone and manganese content and wheels of bainitic steel. These target different operational conditions, with bainitic steels shown to achieve very good operational rolling contact and thermal fatigue resistance under heavy-haul conditions (Gianni *et al.*, 2007). To counteract the inherently low wear resistance of bainitic materials, a rigorously controlled manufacturing process, regarding the chemical analysis in particular, is needed. In addition, the allowed diameter difference between the two wheels of the same wheelset needs to be restricted in order to prevent the combination of flange wear and surface-initiated RCF as discussed above.

For subsurface-initiated RCF, the material cleanliness is a crucial factor. When the material close to a material defect yields in compression, tensile residual stresses form at unloading as discussed in Section 7.1. These cause crack initiation at the side of the defect. As the crack grows, the influence of the material defect will decrease. However, if a nearby defect exists, a highly stressed zone between the two defects will form through which the crack may propagate (see Fig. 7.17) (Kabo, 2002; Kabo and Ekberg, 2005). As an example, interaction between several defects in a streak of manganese sulphide may result in a large crack although the size of each of the sulphide defects is in the order of 10 μm. To avoid wheels with large material defects, manufactured wheels are commonly tested ultrasonically before delivery. In addition the oxygen content is normally restricted in order to prevent the formation of oxides.



7.17 Interaction between two nearby subsurface defects subjected to a passing contact load. The impact is represented in terms of the Jiang fatigue parameter. (Picture courtesy Elena Kabo)

In the context of material cleanliness, also the material's hydrogen content must be considered. During repeated fatigue loading trapped hydrogen atoms diffuse into pores where they form H_2 molecules. This causes a high pressure in the pores and subsequent cracking, a phenomenon commonly referred to as hydrogen embrittlement. The longer the operational lives of the wheels, the more important this phenomenon becomes since the repeated loading promotes hydrogen diffusion. As a preventive action, the hydrogen content of wheel materials is normally highly restricted.

7.5 Conclusions

7.5.1 What is known and what is not

Although research in wheel–rail contact phenomena has a long history, there are still areas where knowledge is severely lacking. A major stumbling-block in gaining deeper insight and providing more accurate predictions is the stochastic nature of the fatigue phenomenon in general, and of RCF in particular. Consider the Wöhler curve in Fig. 7.1. Here the fatigue life (the number of loading cycles to failure) is indicated as being a deterministic function of the magnitude of the stress amplitude. In reality there will be a significant scatter in fatigue life also under controlled laboratory conditions. This scatter is due to scatter in the applied loading; recall that small deviations in stress amplitude will result in large differences in the resulting fatigue life owing to the logarithmic nature of the stress–life relationship. A second source of scatter is the deviations in material strength at both micro-level and macro-level. Recall that fatigue is a local phenomenon; fatigue cracks will initiate and grow where the local stress/strength ratio is the highest. The stochastic scatter is most significant at long fatigue lives since fatigue cracks are here on the border between arrest and continued growth.

In the case of RCF, deviations in contact geometry and roughness add to the imposed scatter. This influence is both on the acting load magnitudes (with the influence of corrugation and out-of-round wheels as extreme examples) and on the contact geometry, and consequently on the local contact stresses.

Accounting for statistical scatter and thereby making proper risk analyses is a major challenge that requires not only proper statistical tools but also more reliable predictive models and input data. This need includes fundamental studies to improve the understanding of basic mechanisms and influencing parameters. It also includes improved engineering models and, not least, the adoption of these in practice. A broader adoption will not only improve fatigue management but also promote the collection and analysis of relevant operational data.

Traditionally RCF analysis consisted of applying a Hertzian contact pressure, in some cases together with an interfacial friction corresponding to

full slip, to a 2D semi-infinite body followed by an analysis of the resulting elastic stress field and a prediction of resulting fatigue impact. In recent years, this process has been expanded. The integration of RCF analysis with the analysis of wheel–rail interaction has provided the means of evaluating operational load spectra and their resulting fatigue impact. Further, integration of wear simulations into these models provides a means to evaluate profile evolutions and crack truncations. The introduction of FE simulations and the development of more sophisticated constitutive models have given the possibility of a better analysis of the elastoplastic material response. More sophisticated fatigue analysis models have refined the fatigue life predictions. Numerical simulations and measurements have furthered the insight into operational temperatures in wheels. In conclusion, there has been progress in many fields. Still, there is a lack of knowledge in some separate areas but, even more, regarding the integration of different phenomena. As a simple example, even a slight indentation on a wheel tread will distort the contact geometry. Is this a phenomenon that has a significant influence and needs to be considered in the fatigue analysis? There are many similar questions that can be raised.

Design against, and analysis of, wheel disc fatigue is in practice very different as compared to RCF since it is carried out according to design codes. However, here also there are large uncertainties and the load cases considered are so crude that the possibilities for optimisations are very limited. A design based more on actual acting loads and operational conditions of the wheels could improve both economy and safety of the wheels.

Another area where knowledge is lacking is material characteristics. Today, a number of material parameters are evaluated for wheel steels. However, there is still no commonly accepted agreement as to which are the key parameters in defining RCF strength (Ghidini *et al.*, 2003). Some parameters are known to have a significant influence (e.g., material cleanliness and hydrogen content), some are likely to have an influence (e.g., yield strength and fracture and impact toughness) although it is not obvious how these parameters should be rated (e.g. are higher values always better?). Adding to the frustration is the mismatch between material parameters employed in numerical simulations and the material parameters that are experimentally evaluated.

Last but not least, there are problems with the validation through operational tests. Fatigue is, as mentioned above, a threshold problem. Consequently, continuously monitoring of fatigue deterioration is not easily done. A typical operational scenario is a period of very limited damage followed by an epidemic of RCF cracks. To monitor the evolution of damage and systematically alter operational parameters one-by-one in controlled steps is then normally not an option.

For subsurface-initiated cracks the complications are even worse since these

failures (luckily) occur very rarely. It is normally not possible afterwards to track down the operational history and, more importantly, how this differs from what other wheels in the fleet are experiencing. In addition, the crack surface is often damaged, which makes it hard to identify initiating material defects.

7.5.2 Likely future trends

The trend in railway operations is very clear: higher axle loads and higher train speeds are being introduced throughout the world. In addition, both railway lines and train fleets are utilised to a higher degree. The consequence is that margins for errors in terms of failures and/or unplanned maintenance are constantly shrinking. The implication for railway operators and infrastructure managers is that performance and operational reliability must improve.

In this evolutionary process, the railways are likely to take the same route as, for example the aerospace and car industries: design and development will be more and more computer-aided; the adoption of new designs, models and solutions will rely much more on simulations; the components will be more 'intelligent' through the use of sensors and actuators together with added computational power and data storage.

In this trend, the area of wheel fatigue will be crucial. Higher operational demands with shrinking margins for failures raise the need for reliable predictions combined with rapid and accurate preventive actions. To this end, there is a need to consider the entire railway system. If higher speeds and more frequent operations increase revenues, it may be worth the cost of a higher wheel consumption and more frequent wheel re-profiling to avoid traffic disturbances. It may even be worth the cost of, say, additional detectors to measure the impact on each wheel, thereby making it possible to plan the maintenance on a bogie basis. The foundation for these future improvements is, however, a thorough operational understanding and knowledge of wheel fatigue mechanisms and available means of prediction and counteraction.

7.6 Sources of further information and advice

Compact overviews of wheel fatigue are given in the state-of-the art studies (Ekberg and Kabo, 2005; Tunna *et al.*, 2007) and in the references therein. Of these reviews the former deals with mechanisms and predictive models whereas the latter focuses on the interaction with vehicle dynamics.

To interpret observed wheel damages, the UIC atlas of wheel and rail defects (Stone, 2004) and the Bombardier overview of wheel tread damages (Deuce, 2007) are excellent guides.

Rich sources of additional information are the proceedings of the conference series Contact Mechanics and Wear of Rail/Wheel Systems and

the Wheelset Congress. Selected papers of the former are published in *Wear*. More specialised conferences with bearing on the subject are the conference series of the International Heavy Haul Association (IHHA) and of the International Association of Vehicle Systems Dynamics (IAVSD). Selected papers from the latest IHHA technical session (Kiruna, 2007) are published in the *Journal of Rail and Rapid Transit (IMechE)*. Selected papers of the IAVSD conferences are published in *Vehicle Systems Dynamics*.

Scientific papers on wheel fatigue may be found, in addition to the journals mentioned above, in *International Journal of Fatigue* and in *Fatigue & Fracture of Engineering Materials & Structures*. More operationally focused articles may occasionally be found in railway journals such as *Railway Gazette*, *International Railway Journal* and the *European Railway Review*.

7.7 Acknowledgements

The author is deeply thankful for the day-to-day support of colleagues at CHARMEC and for the valuable comments on draft versions that have been received from Robert Fröhling of Transnet, Andrea Gianni and Lennart Nordhall of Lucchini, Roger Deuce of Bombardier and Bengt Åkesson, Tore Vernersson, Jens Nielsen and Elena Kabo of Chalmers. Special thanks to Roger Deuce of Bombardier and Even Bergsengstuen of NSB, Elena Kabo of CHARMEC and to Johan Marais for picture material.

Finally I wish to express my sincere appreciation to all researchers and engineers working in the field of wheel fatigue. Without your efforts this chapter would have been very bland indeed.

7.8 References

Beretta, S, Anderson, C and Murakami Y (2006), Extreme value models for the assessment of steels containing multiple types of inclusion, *Acta Materialia*, 54(8), 2277–89.

Bergkvist, A (2005), *On the Crack Driving Force in Elastic–Plastic Fracture Mechanics With Application to Rolling Contact Fatigue in Rails*, Licentiate thesis, Chalmers University of Technology, Gothenburg, Sweden.

Böhmer, A, Ertz, M and Knothe K (2003), Shakedown limit of rail surfaces including material hardening and thermal stresses, *Fatigue & Fracture of Engineering Materials & Structures*, **26**(10), 985–98.

Bower, A F (1988), The influence of crack face friction and trapped fluid on surface initiated rolling contact fatigue cracks, *ASME Journal of Tribology*, **110**, 704–11.

Burstow, M C (2004), *Whole life rail model application and development for RSSB – continued development of an RCF damage parameter*, Rail Safety & Standards Board, report AEATR-ES 2004–880, Issue 2.

CEN/TC 256 (2003), *Railway applications–wheelsets and bogies–monobloc wheels–technical approval procedure–part 1: Forged and rolled wheels*, Technical Report EN 13979-1, European Committee for Standardization, Brussels, Belgium.

Ciavarella, M and Maitournam H (2004), On the Ekberg, Kabo and Andersson calculation

of the Dang Van high cycle fatigue limit for rolling contact fatigue, *Fatigue & Fracture of Engineering Materials & Structures*, **27**(6), 523–26.

Dahlberg, T and Ekberg A (2002), *Failure Fracture Fatigue*, Lund, Sweden, Studentlitteratur.

Dang Van, K, Cailletaud, G, Flavenot, J F, Le Douaron, A and Lieurade H P (1989), Criterion for high cycle fatigue failure under multiaxial loading, in Brown M W and Miller K J (eds), *Biaxial and Multiaxial Fatigue*, Mechanical Engineering Publications, London, UK, 459–78.

Desimone, H, Bernasconi, A and Beretta S (2006), On the application of Dang Van criterion to rolling contact fatigue, *Wear*, **260**(4–5), 567–72.

Deuce, R (2007), *Wheel Tread Damage – an Elementary Guide*, Report 100115000, Bombardier Transportation, Germany.

Ekberg, A (1997), Rolling contact fatigue of railway wheels – a parametric study, *Wear*, **211**(2), 280–8.

Ekberg, A and Kabo E (2005), Fatigue of railway wheels and rails under rolling contact and thermal loading – an overview, *Wear*, **258**(7–8), 1288–1300.

Ekberg, A and Marais J (2000), Effects of imperfections on fatigue initiation in railway wheels, *Proceedings Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, **214**(1), 45–54.

Ekberg, A and Sotkovszki P (2001), Anisotropy and rolling contact fatigue of railway wheels, *International Journal of Fatigue*, **23**(1), 29–43.

Ekberg, A, Kabo, E and Andersson H (2004), Answer to the letter to the editor by M Ciavarella and H Maitournam, *Fatigue and Fracture of Engineering Materials and Structures*, **27**(6), 527–8.

Ekberg, A, Kabo, E, Nielsen, J C O and Lunden R (2007), Subsurface initiated rolling contact fatigue of railway wheels as generated by rail corrugation, *International Journal of Solids and Structures*, **44**(24), 7975–87.

Esslinger, V, Kieselbach, R, Koller, R and Bernhard Weisse (2004), The railway accident of eschede–technical background, *Engineering Failure Analysis*, **11**(4), 515–35.

Franklin, F J and Kapoor A (2007), Modelling wear and crack initiation in rails, *Proceedings Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, **221**(1), 23–33.

Fröhling, R D (2006), Analysis of asymmetric wheel profile wear and its consequences, *Vehicle System Dynamics*, **44**(S1), 590–600.

Fröhling, R, Ekberg, A and Kabo E (2008), Developing hollow wear limits based on field experience and numerical simulations, *Wear*, **265**(9–10), 1283–91.

Ghidini, A, Cantini, S and Roberti R (2003), Mechanical behaviour of materials for railways solid wheels: a simplified criterion to estimate and compare rcf resistance, In *Proceedings 6th International Conference on Contact Mechanics and Wear of Rail/Wheel Systems*, Gothenburg, Sweden, 10–13 June, 403–11.

Gianni, A, Karlsson, T, Ghidini, A and Ekberg A (2007), Bainitic steel grade for solid wheels: metallurgical, mechanical and in-service testing, *Proceedings International Heavy Haul Association Specialist Technical Session*, Kiruna, Sweden, 11–13 June, 701–11.

Hearle, A D and Johnson K L (1985), Mode II stress intensity factors for a crack parallel to the surface of an elastic half-space subjected to a moving point load, *Journal of the Mechanics and Physics of Solids*, **33**, 61–81.

Jergeus, J (1998), *Railway Wheel Flats: Martensite Formation, Residual Stresses and Crack Propagation*, PhD Dissertation, Chalmers University of Technology, Gothenburg, Sweden.

Jiang, Y and Sehitoglu H (1999), A model for rolling contact fatigue, *Wear*, **224**, 38–49.

Johnson, K L (1989), The strength of surfaces in rolling contact, *Proceedings Institution of Mechanical Engineers*, **203**, 151–63.

Kabo, E (2002), Material defects in rolling contact fatigue – influence of overloads and defect clusters, *International Journal of Fatigue*, **24**(8), 887–94.

Kabo, E and Ekberg A (2005), Material defects in rolling contact fatigue of railway wheels – the influence of defect size, *Wear*, **258**(7–8), 1194–200.

Kabo, E, Nielsen, J C O and Ekberg A (2006), Prediction of dynamic train-track interaction and subsequent material deterioration in the presence of insulated rail joints, *Vehicle System Dynamics*, **44**, 718–29.

Kalousek, J, Magel, E Strasser, J, Caldwell, W N, Kanevsky, G and Blevins B (1996), Tribological interrelationship of seasonal fluctuations of freight car wheel wear, contact fatigue shelling and composition brakeshoe consumption, *Wear*, **191**, 210–18.

Kapoor, A (1994), Re-evaluation of the life to rupture of ductile metals by cyclic plastic strain, *Fatigue & Fracture of Engineering Materials & Structures*, **17**(2), 201–19.

Kapoor, A, Franklin, F J Wong, S K and Ishida M (2002), Surface roughness and plastic flow in rail wheel contact, *Wear*, **253**(1–2), 257–64.

Kassa, E (2007), *Dynamic Train–turnout Interaction Mathematical Modelling, Numerical Simulation and Field Testing*, PhD Dissertation, Chalmers University of Technology, Gothenburg, Sweden.

Lansler, E, Ekberg, A, Kabo, E and Andersson H (2006), Influence of plastic deformations on growth of subsurface rolling contact fatigue cracks in railway wheels, *Proceedings of the IMechE, Part F: Journal of Rail and Rapid Transit*, **220**(4), 461–73.

Liu, Y, Liu, L and Mahadevan S (2007), Analysis of subsurface crack propagation under rolling contact loading in railroad wheels using FEM, *Engineering Fracture Mechanics*, **74**(17), 2659–74.

Magel, E and Kalousek J (1996), Identifying and interpreting railway wheel defects, *Proceedings International Heavy Haul Association Specialist Technical Session 'Running heavy, running fast into the 21st century'*, Montreal, Qc, Canada 9–12 June, 5.7–5.21.

Nielsen, J C O, Ekberg, A and Lunden R (2005), Influence of short-pitch wheel/rail corrugation on rolling contact fatigue of railway wheels, *Proceedings of the IMechE Part F: Journal of Rail and Rapid Transit*, **219**(F3), 177–87.

Ringsberg, J (2000), Cyclic ratchetting and failure of a pearlitic rail steel, *Fatigue & Fracture of Engineering Materials & Structures*, **23**(9), 747–58.

Socie, D F and Marquis G B (2000), *Multiaxial Fatigue*, Society of Automotive Engineers, Warrendale, PA, USA.

Tunna, Sinclair och Perez (2007), A review of wheel wear and rolling contact fatigue, *Proceedings of the IMechE, Part F: Journal of Rail and Rapid Transit*, **221**, 271–89.

Tyfour, W R, Beynon, J H and Kapoor A (1996), Deterioration of rolling contact fatigue life of pearlitic rail steel due to dry-wet rolling-sliding line contact, *Wear*, **197**(1–2), 255–65.

Stone, D (ed.) (2004), *Atlas of Wheel and Rail Defects*, UIC, Paris, France.

Vernersson, T (2006), Tread Braking of Railway Wheels–Noise-Related Tread Roughness and Dimensioning Wheel Temperatures: Tests, Rig Measurements and Numerical Simulations, PhD Dissertation, Chalmers University of Technology, Gothenburg, Sweden.

Copyrighted Materials
Copyright © 2009 Woodhead Publishing Retrieved from www.knovel.com

8

Out-of-round railway wheels

J. NIELSEN, Chalmers University of Technology, Sweden

Abstract: Out-of-round railway wheels can have a detrimental influence on track and vehicle components, contributing to increased risks of rail breaks, sleeper cracking, high-cycle fatigue of wheels and axles and bearing damage. Environmental consequences are rolling noise and impact noise, ground vibrations and passenger discomfort. Causes of different types of out-of-round wheels are surveyed. Experimental equipment and assessment procedures for quantification of out-of-roundness are explained. Results from field tests illustrating consequences of out-of-round wheels are presented. Wheel impact load detectors and examples of alarm limits used in different countries for detection of detrimental wheel defects are listed. Mathematical models and computer programs for simulation of the influence of out-of-round wheels on wheel–rail contact forces and vehicle/track responses are described.

Key words: railway wheel out-of-roundness, wheel flat, polygonalisation, tread braking, roughness, detectors; alarm limit, dynamic vehicle–track interaction.

8.1 Introduction

Out-of-round railway wheels can have a detrimental influence on track and vehicle components, contributing to increased risks of rail breaks, sleeper cracking, high-cycle fatigue of wheels and axles and bearing damage. Examples of wheel out-of-roundness (roughness, waviness, irregularities) are local tread damage such as wheel flats causing severe repeated impact loads, and polygonal wheels containing a periodic deviation from the nominal wheel radius that is dominated by a few wavelengths (orders, harmonics) around the wheel circumference. A polygonal wheel leads to increased components of the dynamic vertical wheel–rail contact force at certain excitation frequencies that are determined by train speed and the irregularity wavelengths, whereas wheel flats generate impact forces with significant contributions in a wide frequency range. Impact noise and rolling noise are other consequences of wheel out-of-roundness.

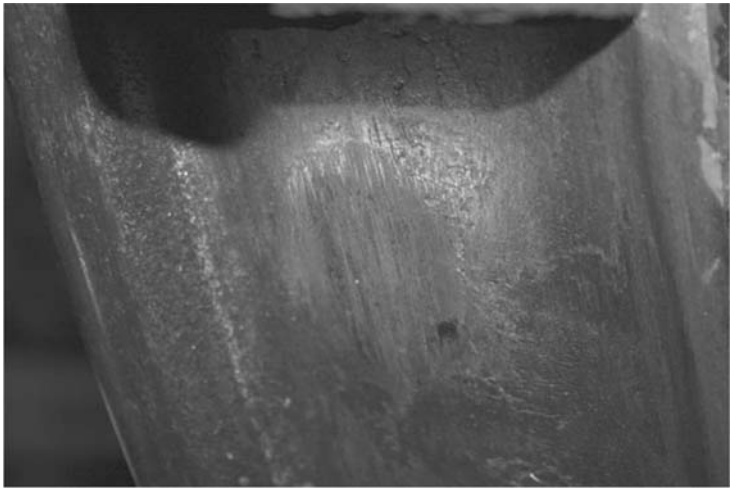
There are several mechanisms or issues that may result in out-of-round wheels. Examples are irregular wear around the wheel circumference, brake system failures, wheel machining issues, misaligned axle bore holes, surface- or subsurface-initiated fatigue cracking, local variations in material microstructure, plastic deformation and buildup of brake block material on the wheel tread. In this chapter, causes and consequences of different types of wheel out-of-roundness are explained.

8.2 Classification and quantification of wheel out-of-roundness

In this section, brief descriptions of some common types of wheel out-of-roundness are given. Experimental equipment and assessment procedures for quantification of out-of-roundness are explained.

8.2.1 Introduction to discrete wheel tread defects – wheel flats

A discrete wheel tread defect is a deviation from the nominal wheel radius on a small section of the wheel tread that may generate an impact load in the wheel–rail contact for each wheel revolution. One of the most common discrete tread defects, the wheel flat (see Fig. 8.1), is developed due to unintentional sliding (without rolling) of the wheel along the rail. The reason for the sliding may be that the brakes are poorly adjusted, frozen or defective, or that the braking force is too high in relation to the available wheel/rail friction.[1] Contaminations on the rail surface, such as leaves, grease, frost and snow, aggravate the problem. As a consequence, a part of the wheel tread is worn off and locally the wheel temperature is raised significantly due to the dissipated friction energy. When the wheel starts rolling again, this is followed by a rapid cooling due to conduction into the large steel volume surrounding the flat. This may lead to material phase transformations (formation of martensite) and residual stresses. The residual stresses are predominantly compressive in the martensitic region and tensile in the region surrounding the martensite.



8.1 Wheel flat. (Photo courtesy Anders Ekberg)

The initial flat with sharp edges will soon be transformed into a longer flat with rounded edges because of plastic deformation of the wheel material at subsequent impacts with the rail.[2] Further, if martensite is formed, cracks will initiate and propagate in the brittle material due to the rolling contact loading and the repeated impacts. Due to the tensile residual stresses in the surrounding material, cracks may grow to considerable depths and relatively large parts of the wheel tread may detach. In severe cases, the initial wheel flat with a typical length of 40–60 mm (and depth in the order of 1 mm) can deteriorate into a discrete tread defect that could be as long as 300–500 mm and 5 mm deep.

Other reasons for discrete wheel defects may be local deviations from the nominal material properties, plastic deformation and tread metal buildup where wheel, brake block and rail debris are welded to the wheel tread due to heavy tread braking.[3] The consequences of this type of wheel out-of-roundness are discussed in more detail in Section 8.3.

8.2.2 Discrete wheel tread defects – rolling contact fatigue

Rolling contact fatigue (RCF) defects can be distinguished by whether the cracks are initiated at the surface or at some millimetres below the surface.[4,5] The surface-initiated cracks are often initiated due to repeated cyclic tangential loading of the wheel–rail contact leading to accumulated plastic deformation of the surface material. Due to subsequent rolling contact loading and hydraulic pressure from fluids (water, grease) trapped in the cracks, the cracks propagate into the wheel bulk material typically down to a depth of about 5 mm. Other initiating mechanisms may involve surface defects and asperities. Further, thermal loading (due to tread braking or wheel sliding) and possible subsequent material transformation promote the initiation and growth of surface-initiated cracks. Fracture of surface-initiated fatigue cracks typically result in the detachment of a shallow part of the wheel tread. The occurrence of surface-initiated cracks can be reduced by using a well-adopted wheel material, by applying a surface coating or by using rim quenching to introduce beneficial compressive residual stresses in the wheel tread. The rim quenching technique is today used on most railway wheels. Reduction of tangential wheel–rail contact loads is another efficient remedy to suppress the formation of surface cracks. Also, a controlled amount of wheel tread wear may counteract the crack propagation growth rate.

Initiation of subsurface-initiated cracks is promoted by the presence of material defects, high vertical contact load magnitudes and by poor wheel–rail contact geometry leading to small contact patches and high contact stresses. Cracks may initiate at a depth of 5–25 mm, propagate along the wheel circumference and finally fracture towards the wheel tread, breaking off a

large part of the wheel rim. Here, preventive measures are avoiding overloads from, for example, rail corrugation and poorly aligned rail joints, management of wheel–rail contact geometry to avoid contacts with excessive normal contact stresses, and using a material with appropriate fatigue and fracture resistance. Defects related to RCF are sometimes referred to as spalling and shelling.[3] RCF in wheels is treated in more detail in Chapter 7.

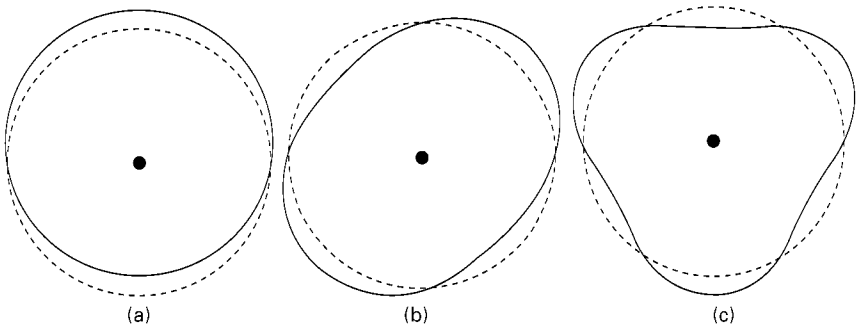
8.2.3 Wheel polygonalisation

Polygonal wheels exhibit a periodic radial tread irregularity from the mean wheel radius.[6] Often the wheel out-of-roundness is dominated by a few waves that correspond to 1–5 wavelengths (orders, harmonics) around the wheel circumference. Amplitudes are in the order of 1 mm.

The shape of the out-of-roundness can be classified by conducting a Fourier analysis of the measured radial tread irregularity to determine the contributions of different harmonics. The corresponding wavelengths λ [m] are defined by:

$$\lambda = \frac{2\pi R}{\theta}, \theta = 1, 2, 3, \ldots \quad [8.1]$$

where θ is the order and R is the wheel radius. The case $\theta = 1$ corresponds to an eccentricity which is caused by misaligned axle bore holes or by a misalignment in the fixation of the wheel during profiling. The case $\theta = 2$ corresponds to a wheel ovality. The first three orders are illustrated in Fig. 8.2. Normally, several different out-of-roundness orders exist simultaneously. When a polygonal wheel is rolling on the track at speed v [m/s], the excitation frequency f is given by:



8.2 Examples of different orders of wheel out-of-roundness: (a) first out-of-roundness order ($\theta = 1$, eccentricity), (b) second out-of-roundness order ($\theta = 2$, ovality); (c) third out-of-roundness order ($\theta = 3$).[6] (Figure courtesy Anders Johansson)

$$f = v/\lambda \qquad [8.2]$$

For a wheel with 1–5 wavelengths, this means that the excitation frequencies are in the range 5–125 Hz when train speeds are in the interval 50–250 km/h. Such a periodic wheel out-of-roundness may lead to annoying vibration levels for the passengers at certain train speeds, and to ground vibrations affecting residents living close to the railway lines.

Wheels with a periodic out-of-roundness may be caused by a system interaction issue or by a wheel machining issue.[7] Examples of wheel polygonalisation with one, three and four harmonics around the circumference have been found on disc-braked wheels on inter-city express high-speed trains in Germany.[8] Fixation (claw clamping in a three-jaw chuck) of the wheel during profiling may be the cause of an initial polygon with three harmonics.[9] In an investigation of ICE wheelsets, it was concluded that the third harmonic dominated for solid steel wheels, while the second harmonic was common for rubber sprung wheels.[10]

Vertical wheel–rail contact forces with high magnitudes are generated when the excitation frequency caused by the wheel out-of-roundness interacts with the fundamental vertical system resonance where wheelset, rails and sleepers are vibrating in phase on the track support stiffness. Increasing track support stiffness leads to higher force magnitudes at this resonance. The lower bending and torsional modes of the wheelset, and the fundamental vertical system resonance, have been suggested as wavelength-fixing mechanisms leading to polygonal wheels with 1–5 harmonics. The most common strategy to reduce the effects of wheel polygonalisation seems to be a shortening of the interval between wheel re-profilings.

Out-of-round wheels with an irregularity that is not dominated by a particular wavelength are sometimes referred to as stochastic out-of-round wheels. This can be caused by the presence of a mixed microstructure of the wheel surface material as a result of heat treatment issues during wheel manufacture.[7] After a few re-profilings, the mixed microstructure is removed leading to that the development of stochastic out-of-roundness stops.

8.2.4 Introduction to wheel roughness induced by tread braking

Rail vehicles equipped with tread brakes using cast iron brake blocks generate higher rolling noise levels than vehicles with disc brakes at the same train speed. The reason is that such tread braking leads to surface irregularities with higher magnitudes (noise-related tread roughness).[11] The dominating wavelengths of the roughness are in the interval 50–70 mm, while amplitudes are of the order of 10 μm.

A key phenomenon in the generation of wheel roughness by tread braking

is the hot spotting of the tread caused by the thermomechanical interaction between block and wheel tread.[11,12] The heat generated in tread braking is not uniformly distributed over the wheel and brake block contact surfaces. Instead, hotter areas (hot spots) may develop. Hot spotting, along with material transfer from block to wheel tread, seems to generate the high roughness levels. Composite and sinter material blocks, contrary to cast iron blocks, generate low roughness levels on the wheel treads. They may even polish the tread to an extent that it needs roughening in order to allow for sufficient traction levels in the wheel–rail contact. This type of wheel out-of-roundness is discussed in more detail in Section 8.4.

8.2.5 Measurement of wheel out-of-roundness

The instrument in Fig. 8.3 is one example of equipment for measurement of wheel radius as a function of a co-ordinate around the wheel circumference. The stationary wheelset is lifted by hydraulic jacks, and the brakes are released so that the wheels can be rotated freely by hand. Three probes in mechanical contact with the wheel tread measure the deviation from the mean wheel

8.3 Measurement of out-of-roundness on a stationary wheel.[13] Three probes are in mechanical contact with the wheel tread. (Photo courtesy Ødegaard and Danneskiold-Samsøe A/S)

radius. The centre probe is typically positioned at the nominal contact point, while the two other probes can be positioned at a selected distance on each side of the centre probe. The present equipment has a sampling distance of 0.5 mm and an amplitude resolution of 0.06 μm.[13]

Another approach that instead measures out-of-roundness on wheels while in traffic is based on the measurement of how far the wheel flange projects below the upper surface of the rail.[14] This can be accomplished by mounting a resiliently supported rigid bar alongside the rail. The displacement of the bar when displaced by the wheel flange is recorded. A survey on roughness measurement is given in Verheijen.[15]

8.2.6 Roughness level spectrum

Based on a measured profile of deviation from the mean wheel radius, a roughness level spectrum can be calculated. The roughness level spectrum quantifies roughness level as a function of roughness wavelength. Such spectra are important as input in the prediction of rolling noise and in the simulation of vertical dynamic vehicle–track interaction.

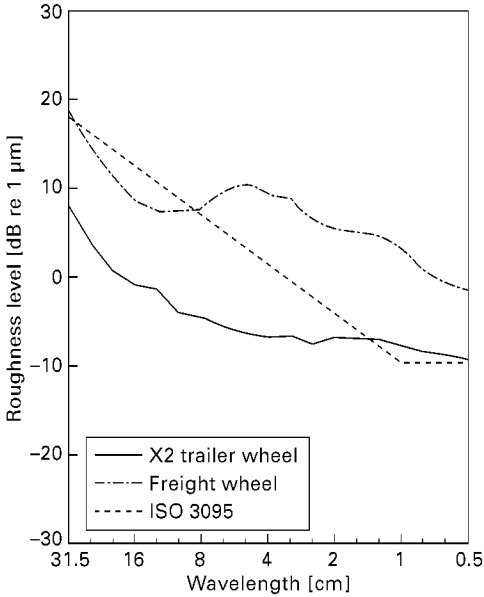
Based on the calculated power spectral density of the measured profile, the root mean square values $\tilde{r}_k$ [m] of the profile are evaluated in one-third octave bands k with centre wavelengths λ_k. Typical wavelengths λ_k are in the range from 31.5 cm down to 0.5 cm or less. Roughness level L_r^k is defined in decibels relative to a reference value r_{ref} = 1 μm according to:

$$L_r^k = 20 \log_{10}\left(\frac{\tilde{r}_k}{r_{ref}}\right) \text{ [dB re 1 μm]} \qquad [8.3]$$

Two examples of wheel roughness level spectra are compared in Fig. 8.4. Using the equipment in Fig. 8.2, the spectra were determined based on measurements on five passenger train (X2) trailer wheels and 14 freight wagon wheels.[16] It is observed that freight wagon wheels that are braked with cast iron brake blocks are rough with large magnitudes for wavelengths between 40 and 80 mm, whereas the disc braked X2 trailer wheels are smooth. At measurement, all measured wheels had a travelled distance of at least 100 000 km since their previous re-profiling. Roughness level spectra based on various types of Dutch railway traffic are presented in the paper by Dings and Dittrich.[17]

8.2.7 Order spectrum

In Fig. 8.5a, one example of measured deviation in wheel radius from the nominal radius versus a circumferential coordinate is illustrated. Using the equipment in Fig. 8.3, the wheel radius was measured on a subway train



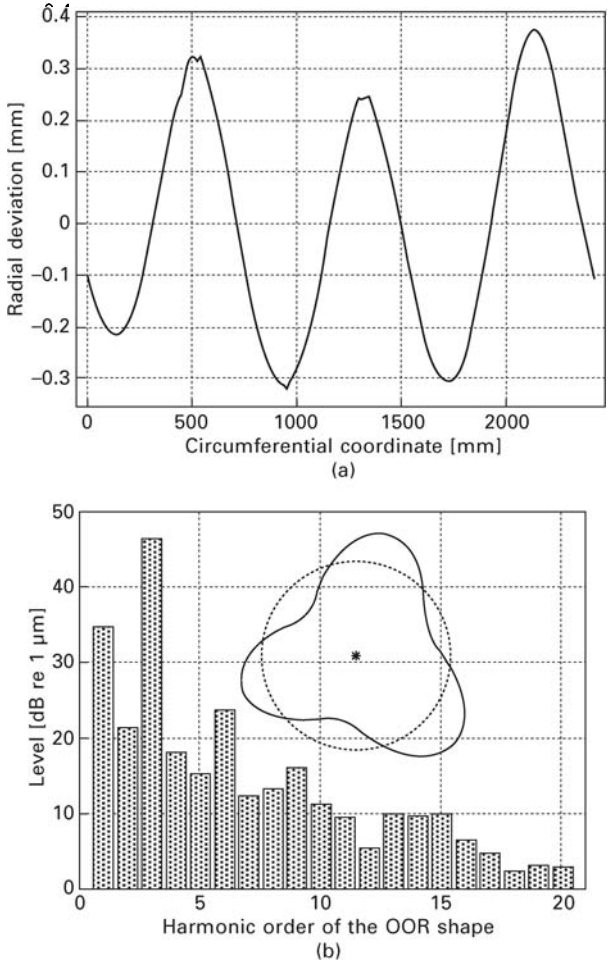
8.4 Roughness level spectra in one-third octave bands based on measurements of five (disc-braked) X2 trailer wheels and 14 (tread-braked with cast iron brake blocks) freight wagon wheels.

wheel after 300 000 km of travelled distance. It is observed that the measured deviation is dominated by the third-order (harmonic), i.e. a periodic radial irregularity with three wavelengths around the wheel circumference. The different orders that are contained in the wheel out-of-roundness can be displayed by performing a Fourier decomposition of the measured data, see Fig. 8.5b. It is concluded that the first (eccentricity) and third orders are dominating, and that the out-of-roundness contains contributions also from several other orders.

8.3 Discrete wheel tread defects

A discrete wheel tread defect can be a major safety hazard as it may be the triggering factor that leads to rail fracture and a possible derailment. A severe and common type of discrete defect generating repeated wheel–rail impact loads with high magnitudes is the wheel flat.

One scenario that can lead to rail fracture is that the magnitude of the impact load is sufficiently high to generate a stress state that initiates crack growth from a defect in the rail. Examples of rail defects are inclusions and pores in a weld, or surface cracks in the railhead caused by RCF (head checks). Subsequent loading of the crack from passing wheels leads to crack growth and eventually the stress intensity at the crack tip may exceed



8.5 Radial deviation from mean wheel radius versus a circumferential coordinate. (a) Measurement on a subway wheel after 300 000 km travelled distance. (b) Corresponding order spectrum.[19] [OOR = out-of-roundness]

the fracture toughness of the rail material. A large impact load could also generate final fracture in a poorly maintained rail if, for example, RCF cracks are not detected and ground in time. As mentioned in Section 8.1, repeated impact loads may also lead to bearing damage and high-cycle fatigue of the wheelset.

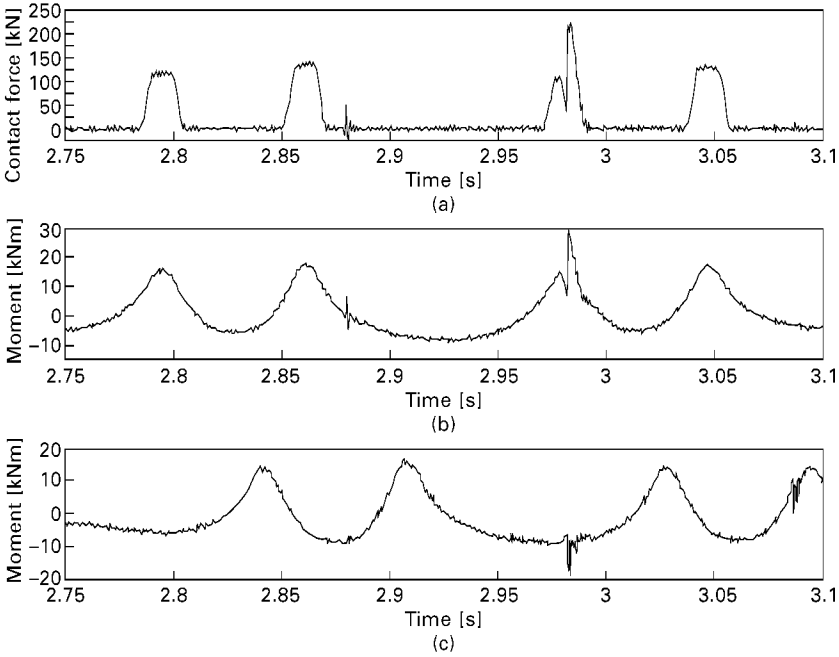
The growth of transverse rail cracks in modus I is governed by the combination of tensional stresses generated at low temperatures* (deviation

* Low temperature also leads to a reduction of fracture toughness.

from the stress-free rail temperature) and tensional stresses due to rail bending.[18] The rail bending moment is generated by the quasistatic loading from each passing wheel, but also from dynamic load increments due to out-of-round wheels and track irregularities. The magnitudes of the impact loads from an out-of-round wheel are dependent on the size and shape (mainly the depth and length) of the tread defect, train speed, track properties, unsprung wheelset mass and axle load. The magnitude of the bending moment in the rail near a crack is also dependent on axle distance and the distance between the crack and the positions of impact.

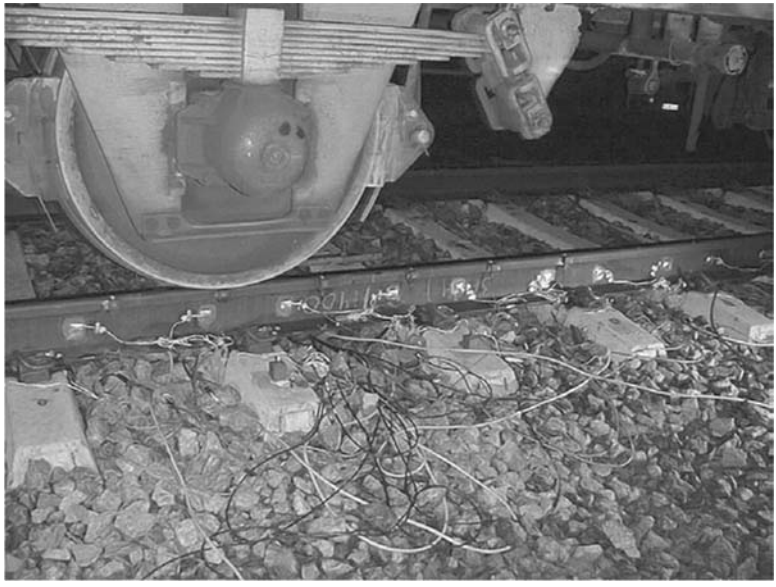
8.3.1 Field tests

Field tests have been performed to investigate the influence of wheel flat dimensions and train speed on impact loads and track response. Figure 8.6 illustrates results from a field test[20] where the wheel–rail contact force was measured by a strain gauge-based wheel impact load detector (see Section



8.6 Measured influence of a wheel flat with length 60 mm and depth 1.0 mm: (a) vertical wheel–rail contact forces measured between two sleepers; (b) rail bending moment above the sleeper that was closest to the impact position; (c) rail bending moment above a sleeper that was located two sleeper distances away from the impact position. Train with axle loads 24–30 tonnes and speed 100 km/h. Track with UIC60 rails, resilient rail pads and concrete monobloc sleepers on ballast.

8.3.2), and bending moments in the rail were measured by strain gauges mounted on the rail foot above the sleepers (see Fig. 8.7). The field test was performed on a track with UIC60 rails, resilient rail pads and concrete monobloc sleepers on ballast. A freight train with axle loads in the interval 24–30 tonnes and speed 100 km/h was used to provide the loading of the track. In the upper figure, the wheel–rail contact force in one sleeper bay of the impact load detector is displayed. Each of the four wheels passing the instrumented section is observed as a measured deviation in load from 0 kN. It is observed that the third wheel generated an impact load of 230 kN at time = 2.98 s. This was due to a wheel flat with length 60 mm and depth 1.0 mm (axle load 26 tonnes). The middle figure illustrates the resulting bending moment in the rail above the sleeper that was closest to the studied impact position. Positive bending moments (leading to tensional stresses in the rail foot) in the order of 15 kNm were generated by each passing wheel. In between the wheels, negative bending moments (leading to tensional stresses in the railhead) in the order of –5 kNm are observed. The impact from the wheel flat led to a maximum bending moment of about 29 kNm, which could be detrimental if there was a crack in the rail foot in the near vicinity of the impact position. In the lower figure, the same impact generated a significant contribution to the negative bending moment (–19 kNm) in the

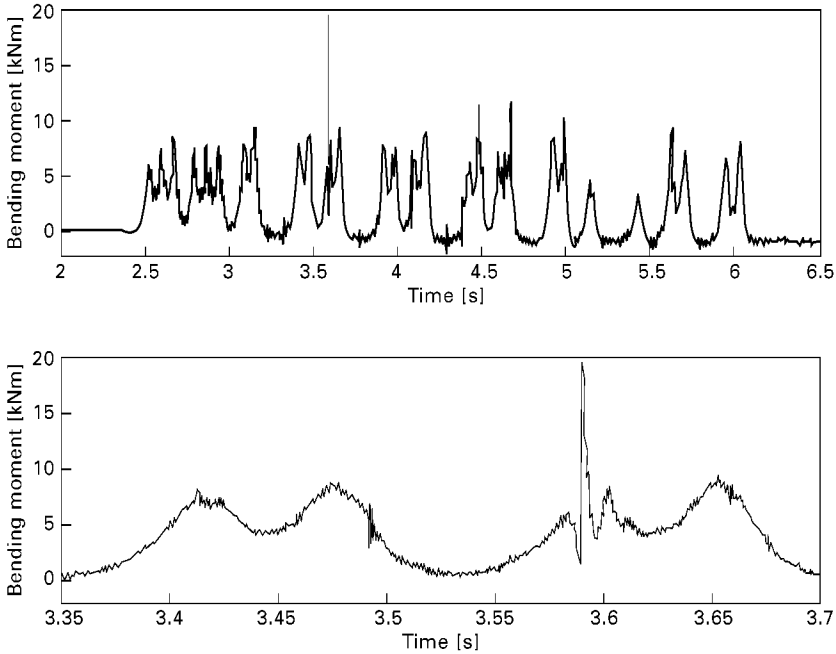


8.7 Track instrumented with strain gauges on the neutral axis of the rail to measure the vertical wheel–rail contact force and on the railfoot to measure bending moments in the rail. Accelerometers are mounted on the sleepers to measure sleeper vibrations.

rail two sleeper distances away from the impact. In this case, a crack in the railhead could be critical.

In the same field test, sleeper bending moments at rail seat were measured. The magnitudes of sleeper bending moments due to wheel loading are significantly influenced by, for example, rail pad stiffness, and the stiffness and distribution of the ballast support along the sleeper and along the track. In this particular test, the track was in good condition after having been recently tamped. In Fig. 8.8, it is observed that a maximum rail seat moment of 19 kNm occurred at time 3.6 s. This was again caused by the 60 mm wheel flat on a wheelset with axle load 26 tonnes and train speed 100 km/h.

Studies have been carried out which contribute to the field and which have had a significant influence on the early understanding of the effects of out-of-round wheels.[21,22] In experiments involving wheel tread irregularities, difficulties may occur in locating the position of impact near to the instrumentation of the track. Therefore, an indentation equivalent to a wheel flat was ground into the railhead.[22] The influence of train speed on the magnitude of impact loads have been investigated in a series of wheel



8.8 Measured sleeper bending moment at rail seat. The maximum bending moment 19 kNm at 3.6 s was caused by a 60 mm wheelflat. Train with axle loads 24–30 tonnes and speed 100 km/h. Track with UIC60 rails, resilient rail pads and concrete sleepers on ballast. The lower figure is a zoomed-in section of the upper figure.

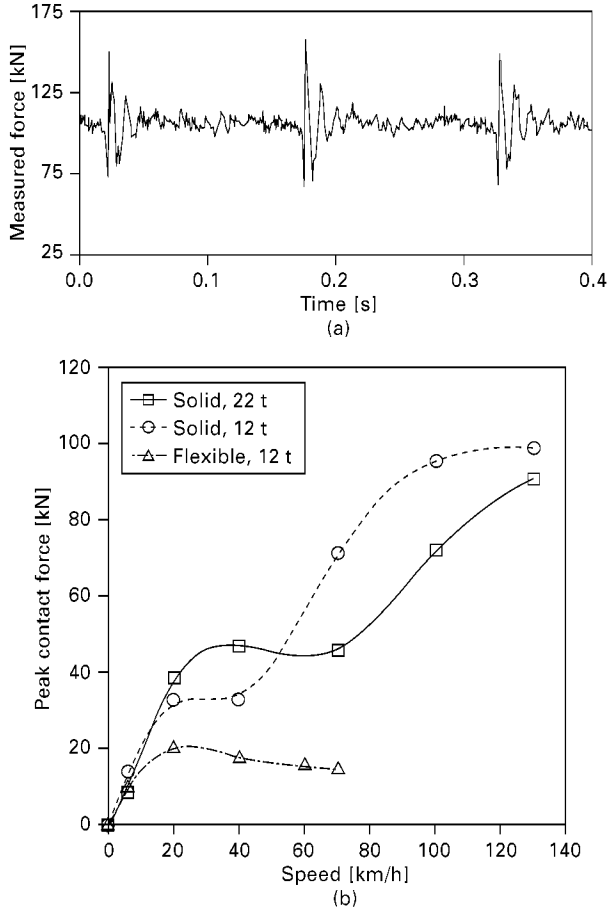
impact tests.[23–25] It has been concluded that impact loads increase with the length (and depth) of the wheel flat.

The influence of axle load and train speed on impact load magnitudes due to wheel flats was investigated using instrumented wheelsets.[26] Vertical wheel–rail contact forces generated by a wheelset with standard solid disc wheels were compared to those for a wheelset with radially flexible wheels. Measurement wheels instrumented with strain gauge bridges on each side of the wheel disc were employed.[27] Wheel flats with length 40 mm were manually ground on site, but unfortunately resulting in somewhat different depths (0.35 mm for the solid wheel and 0.25 mm for the flexible wheel). Emphasis was put on having the experiments as controlled as possible. Therefore: (i) frequency response functions of the track were measured for a characterisation of the track; (ii) surface irregularities on railhead and wheel tread were measured; (iii) the same test train was used in all test runs (only switching between measurement wheelsets with standard or flexible wheels); (iv) transient vertical contact forces (obtained indirectly via strain gauges on the instrumented wheelsets) and strains and accelerations of sleepers and rails were measured; (v) the location of the instrumented wheelset in relation to the instrumented section of the track was determined at each instant of time using optical sensors; (vi) each revolution of the instrumented wheels was registered in the recorded data in order to determine the angular position of the wheel flat.

Three test runs with each combination of train speed and axle load were performed. One example of a measured contact force time history is shown in Fig. 8.9a. Starting at the time instant when the front end of the wheel flat reaches contact with the rail, the wheel will move downwards and the rail upwards to compensate for the deviation in wheel radius. This leads to an initial reduction in contact force. Towards the back end of the wheel flat where the wheel radius is increasing, the wheel will continue downwards due to its greater inertia forcing the rail to do the same. This results in an increase of the contact force. The peak contact force deviation from the nominal static wheel load is plotted versus train speed in Fig. 8.9b. Despite the difference in wheel flat depth, it is concluded that the lower unsprung mass of the flexible wheel leads to lower impact loads.

8.3.2 Wheel impact load detectors

Efficient monitoring of out-of-round wheels and early detection of detrimental wheel defects can be achieved by a proper distribution of wheel impact load detectors in the railway network. In order to minimise damage, the defective wheels need to be detected as early as possible and taken out of service for maintenance. Several types of detectors based on different types of sensors exist on the market. In addition to identifying defective wheels



8.9 (a) Example of measured vertical contact force time history for a standard solid disc wheel with a 40 mm long wheelflat, train speed 70 km/h and axle load 22 tonnes. The contact force was band pass filtered in the interval 10–1000 Hz. (b) Measured peak contact force deviations from the nominal static wheel load due to a 40 mm long wheelflat (average of 11 peaks for each combination of train speed and wheel type). (From Fermér and Nielsen)[26]

by indirectly measuring the static and dynamic components of the vertical wheel–rail contact force, the detectors can be used to collect information on, for example, total train weight (tonnage), unbalanced vehicles, train speed, direction of travel, number of axles, axle distances and train length.

The original and most common type of detector is based on a series of strain gauge load circuits mounted on the neutral axis of the rail in between two adjacent sleepers in several consecutive sleeper bays. The measuring zone can be up to 16 m long to cover the complete wheel circumference of wheels with various wheel diameters. The strains (shear forces) are converted

to time histories of vertical wheel–rail contact force exerted by each passing wheel. The resolution of the system has been reported to be 50 N for train speeds in the interval 50–300 km/h.[28]

A second system for detection of defective wheels is based on rail foot mounted fibre-optic sensors.[29,30] A standard detection station uses eight or 12 sensors to measure the bending of the rail during a train passage. After analysis and processing of the signals, the combination of the calculated maximum value and the root mean square value (RMS-value) of the dynamic force for each wheel are used in a diagnosis of the overall wheel quality and to determine the type of wheel defect. The system is calibrated using vehicles with known weights. The accuracy of the output is mainly dependent on the number of sensors that are used in the system and on the track quality. After the passage of around 100 recognised calibration vehicles, the accuracy is reported to be around 5 %.

A third system is based on the measurement and analysis of vertical forces transmitted between rail and sleepers. Certain weighing sleepers similar to conventional concrete monobloc sleepers are installed in a track section with sufficient length to continuously scan at least one wheel circumference.[31] Wheel out-of-roundness detectors based on measured rail accelerations are also available but, in this case, the quantification of the wheel–rail impact load is difficult.

8.3.3 Alarm limits

The output from a wheel impact load detector can be reported in the form of various numerical values that have been evaluated based on the measured vertical wheel–rail contact force Q. Examples of such values are the maximum load $Q_{max} = \max(Q)$, the maximum dynamic load increment $Q_{dyn} = \max(Q) - \text{mean}(Q) = Q_{max} - Q_{mean}$ and the dynamic load factor $1 + Q_{dyn}/Q_{mean} = Q_{max}/Q_{mean}$. To identify defective wheels, the measured values are compared to maximum allowed values, alarm limits, set by the railway administration or by the railway operator.

To ensure operational safety and asset life optimisation, different railway administrations use different alarm limits and corresponding action rules in their assessment of outputs from their wheel impact load detectors. Reporting criteria are flexible; for example, a low-level alarm can be used to identify wagons/coaches with wheels that need re-profiling at the next scheduled maintenance activity, a medium-level alarm to limit the maximum speed of the train until the wagon/coach can be set out and a high-level alarm that directs the train to stop as quickly and safely as possible to avoid a potential derailment.[28] Examples of current low-level alarm limits are summarised in Table 8.1. Tracks with mixed traffic generally have lower alarm limits to meet passenger safety.

Table 8.1 Low-level alarm limits used in different countries

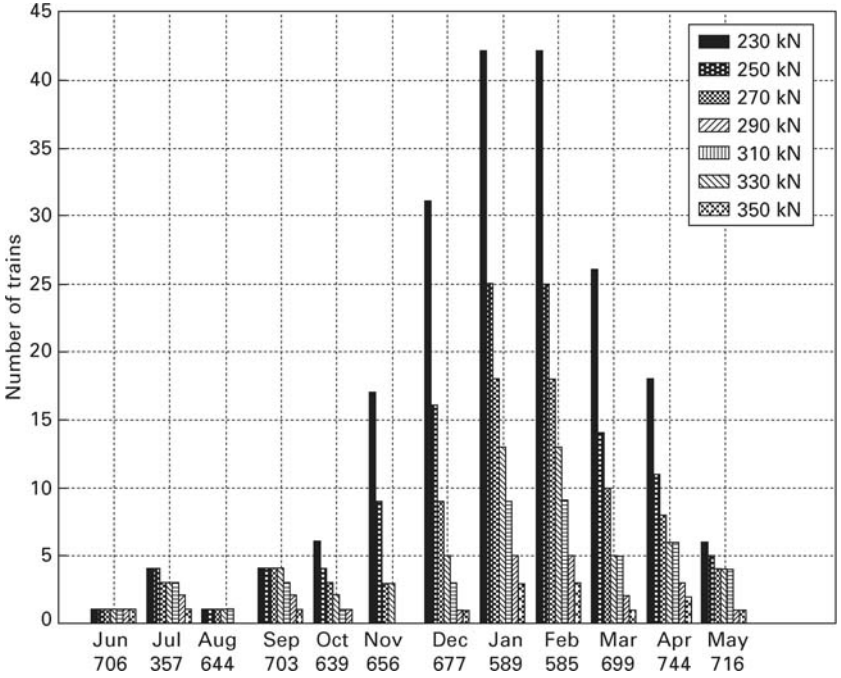
Country	Low alarm limit [kN]	Axle load [tonnes]	Requirement
Australia/QR	400	30	
Australia/BHP-IO	450	40	
Germany/ICE trains	–	13	$1 + Q_{dyn}/Q_{mean} = 1.8$
South Africa/Transnet Freight	–	–	$1 + Q_{dyn}/Q_{mean} > 2.1$[a]
Sweden	290	30	
UK	350[b]	25	
USA/AAR	400	–	

[a]The maintenance limit $Q_{max}/Q_{mean} > 2.1$ is independent of axle load, the stop train limit is $Q_{max}/Q_{mean} > 2.7$.
[b]The limits are: 200 kN = level 1 (warning), 350 kN = level 2 (alarm and 50 mph speed restriction), 400 kN = level 3 (alarm and 20 mph speed restriction) and 500 kN = level 4 (alarm and 10 mph speed restriction).

The strain gauge-based wheel impact load detector described in Section 8.3.2 is used on the heavy-haul traffic line in the northern part of Sweden. Currently on this line, most loaded trains have axle loads of 25 tonnes, although the newer iron-ore wagons carry axle loads of 30 tonnes. The influence of the alarm limit level on the number of trains with at least one alarming wheel is illustrated in Fig. 8.10.[32] The statistics from this detector clearly show that wheel damage is essentially a winter-related problem. It is observed that with the alarm limit set to 290 kN, the number of trains in January with at least one alarming wheel was approximately 1 in 45 (13 in 589). The same ratio in June was 1 in 700.

The use of an impact load detecting system offers the benefit of being able to define a criterion for removal of wheels that is based on the actual loads that the vehicle and track components are subjected to. On some railways, the wheel removal criterion is instead based on visual inspection and the size of the wheel defect. In this case, the maximum size of a certain wheel defect is said to correspond to a certain maximum impact load. However, it has been shown that criteria based either on measured impact load magnitudes or on visual inspections of defect dimensions often do not harmonise.[25] Changes in the North American criteria for removal of out-of-round wheels have been reviewed.[33]

A conceptual framework for investigating the economic consequences of high-impact wheels has been proposed.[33] Based on a life-cycle cost analysis, the objective was to determine at which impact load level it is economically beneficial to remove a defective wheel. It is concluded that, for North American conditions, wheels should be removed from service when they cause impact loads greater than 380 kN. The tests, which led to the current Association of American Railroads (AAR) wheel removal criterion, are described by Kalay *et al.*[25,34]



8.10 Monthly (June 2002–May 2003) statistics of the number of trains with at least one wheel exceeding a certain load magnitude. The total number of trains passing the detector per month is listed on the abscissa. Wheel impact load detector in Harrträsk on the southern route of Malmbanan.[32] (Figure courtesy Eric Berggren)

More recently, fracture mechanics have been used to study the influence of wheel defects on rail life.[35–37] A rail break risk assessment has been performed based on the size of a transverse crack in the railhead (head check) or at the edge of the rail foot.[18,36] The influences of impact load magnitude, distance between the crack and the location of wheel impact, and rail temperature on stress intensity and crack growth are studied. Depending on the statistically low number of wheel impacts occurring within a critical distance of any given crack, it was concluded that rail life measured from the time of the latest rail inspection depends more on the size and location of the cracks that passed the inspection, rail temperature and the number of passing nominal wheel loads than on the magnitudes of the occasional wheel impacts.

8.4 Wheel roughness induced by tread braking

The most common braking system used on European freight wagons is tread (block) braking. The braking action on a wheel is obtained by pressing one, two or four brake blocks, consisting of cast iron, composite or sinter materials,

against the tread of the wheel. Tread braking transforms the kinetic energy of the train into heat, which is partitioned between the wheel and the blocks and is further conducted to the rail and radiated to the surrounding air. Tread braking can also promote improved adhesion between wheel and rail by the scrubbing of the wheel tread.

A major environmental concern is that higher rolling noise levels are generated by wheels with tread brakes, and in particular by those wheels that are equipped with cast iron brake blocks. This is because wheels with cast iron tread brakes develop a tread surface irregularity (roughness) with dominating wavelengths in the interval 50–70 mm (see Fig. 8.4). The amplitude of the periodic irregularity is in the order of 10 μm. Another disadvantage with tread braking is that excessive heating may cause damage to the wheel.

8.4.1 Thermoelastic instability and tread braking

A key phenomenon leading to the generation of wheel roughness induced by tread braking is the hot spotting of the tread that is caused by the thermomechanical interaction between block and tread. For brake blocks made from cast iron material, hot spotting, along with material transfer, seems to generate the high roughness levels. The transfer of block material to hot spot areas on the wheel, seen as blue spots on the tread after braking, has been confirmed through a metallographic analysis.[38] An increased hardness of the blue spots was also detected. Composite and sinter material blocks, in contrast to cast iron blocks, generate low roughness levels on the wheel treads. The generation of wheel roughness induced by tread braking has been studied extensively.[11,12]

At tread braking, the heat is not uniformly distributed over the contact surfaces of wheel and brake block. Due to the thermoelastic interaction between brake block and wheel tread, hotter areas (hot spots) may develop (see Fig. 8.11.)[39] The development of hot spots is caused by a thermoelastic instability (TEI) that may occur because of the friction in the sliding contact. The hot spots are generated on the tread when the variations of the pressure in the sliding contact between the tread and block result in locally increased

8.11 Hot spots visualised with thermography showing part of wheel tread during braking as tread moves out of contact with organic composite brake block. The bright areas have elevated temperature, i.e. hot spots, while the dark areas are colder. The shown part of the tread is 7 cm in axial direction and 70 cm in circumferential direction. From inertia dynamometer tests during brake experiments.[39] (Figure courtesy Martin Petersson)

frictional heating and increased surface temperature in the areas of high contact pressure. Thermal expansion of the material then results in a spatial concentration of the contact and, eventually, to a few contact patches. The hot spots can be either stationary or slowly moving on the surfaces. The proneness to hot spotting on the wheel tread is determined by the block material and the block design. The process is accelerated by increased rate of frictional heat generation, high thermal expansion and high elastic modulus.[40] It is slowed down by high thermal conductivity and high wear rate.

The correlation between surface temperatures of tread braked railway wheels and wheel roughness has ben studied by Vernersson.[41,42] Full-scale braking tests in an inertia dynamometer showed that the wheel–block interaction induced hot spots on the wheel tread. The behaviour for different kinds of brake block material (cast iron, composite and sinter material) was compared. Cast iron brake blocks were found to generate strong hot spot patterns almost instantly after the application of the brake blocks on the tread. The roughness profile of the wheel measured after the test was strongly related to the hot spot pattern during the brake cycle. Furthermore, it was shown that the roughness generated on the wheels in the inertia dynamometer was dominated by the same wavelength interval as those measured on wheels in revenue service. Transient finite element simulations of the thermomechanical interaction between tread and brake block showed the same trends of the propensity of hot spotting as shown by the experimental study. However, the reason why roughness generated by cast iron blocks is dominated by the wavelength interval 50–70 mm is still an open question. Another consequence of hot spots is the local damage that is induced at the spots. Due to RCF and thermal fatigue, this may in the long run give rise to local detachments of tread material.[43,44]

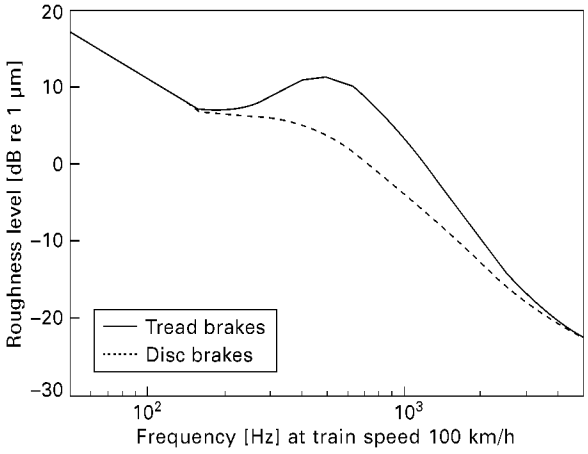
An extensive experimental study on 18 prototypes of brake blocks has been reported by Petersson.[39] The prototype composite brake blocks that were tested showed varying behaviour. Some of the block prototypes generated no hot spots on the treads during braking and thus no severe wheel roughness was developed, while others induced some hot spotting followed by roughness development. However, the roughness levels were much lower than those generated by the cast iron blocks. The sinter brake blocks generated hot spots, but they were not spatially fixed, and the developed roughness levels were low. The fundamental difference between cast iron blocks and the other blocks is the process of material transfer from block to wheel occurring in the previous case. Note also that the hot spot patterns generated by composite and sinter brake blocks generally take much longer to build up than is the case for cast iron blocks, and that material in this case is worn away at the hot spots, creating troughs in the tread upon cooling.

8.4.2 Rolling noise

As for the other types of wheel out-of-roundness, wheel roughness generated by tread braking with cast iron brake blocks leads to a prescribed relative displacement excitation of wheel and rail. The consequences are higher dynamic wheel–rail contact forces, vibrations and noise. Due to the increased roughness levels in the wavelength interval 50–70 mm, it has been concluded that the increased roughness on wheels with cast iron brake blocks may lead to up to 10 dB higher noise levels compared to disc braked wheels at the same speed.[45]

Based on measurements on wheels in the Netherlands,[17] two representative spectra for combined wheel–rail roughness have been defined corresponding to either a tread braked wheel (with cast iron brake blocks) or a disc braked wheel on a rail with low roughness.[46] In Fig. 8.12, these two spectra have been transformed to the frequency domain by assuming constant train speed 100 km/h. The tread braked spectrum is up to 9 dB (re 1 μm) higher than the corresponding spectrum for the disc braked wheelset in the frequency range 200–2000 Hz. This corresponds to around a three times higher root mean square value of the roughness profile for the tread braked wheel compared to the disc braked one.

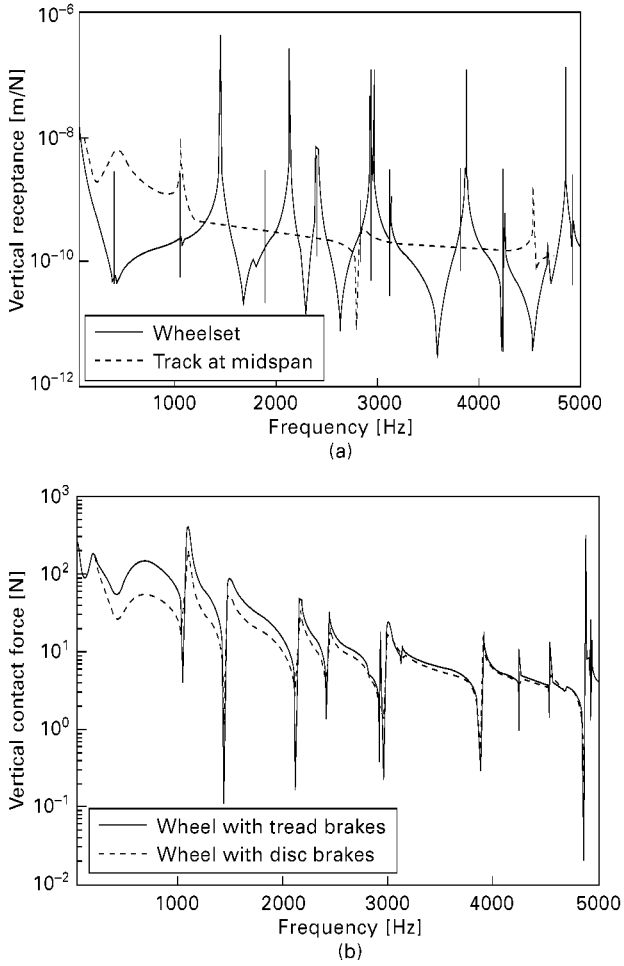
To illustrate the influence of different shapes of roughness spectra on vertical wheel–rail contact force and rolling noise, results calculated using the software TWINS[47] (Track–Wheel Interaction Noise Software) are compared in Fig. 8.13. In the present example, the wheelset is a standard freight wheelset (SJ57H) and the static wheel load is 100 kN. The track design includes



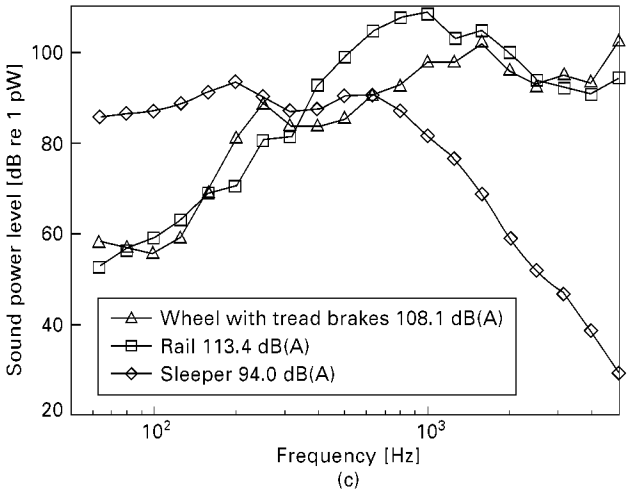
8.12 Examples of combined wheel–rail roughness level spectra for wheel with tread brakes (solid line) and wheel with disc brakes (dashed line).[19, 46] Spectra were calculated for train speed 100 km/h and include the effect of contact patch filtering.

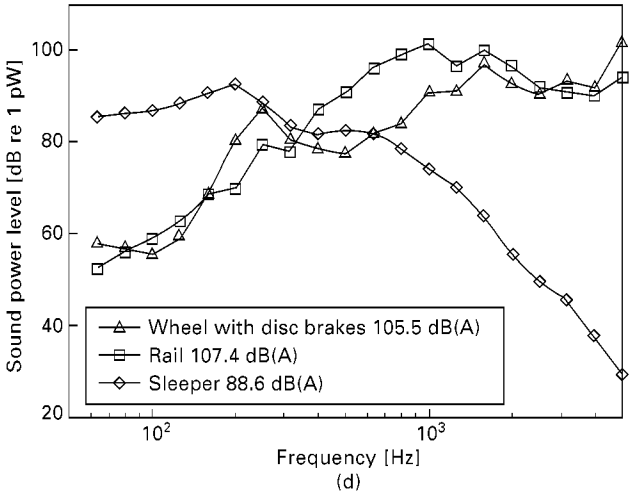
UIC60 rails, resilient rail pads and concrete monobloc sleepers with sleeper distance 0.6 m.[48] The vertical wheelset and track receptances are displayed in Fig. 8.13a.

Figure 8.13b illustrates the vertical wheel–rail contact force. Minima in the contact force correspond to regions of high levels in the wheelset or track receptance. Note the increase in contact force when the wheelset is tread



8.13 Examples of results calculated with TWINS for two different wheel–rail roughness spectra: (a) vertical wheel and track receptances (wheel rotation leading to a splitting of wheel resonance frequencies is neglected here); (b) vertical wheel–rail contact force; (c) sound power level spectra for wheel with tread brakes; (d) sound power level spectra for wheel with disc brakes. Sound power levels for wheel, rail and sleeper, respectively, are reported in the legends. Freight traffic with train speed 100 km/h is studied.[19]





8.13 Cont'd

braked. Finally, Figs 8.13c and 8.13d show the sound power level spectra for wheel, rail and sleeper. In this example, sound power level per wheelset is increased by around 6 dB(A) when the tread braked wheelset is used. For frequencies in the interval 400–2000 Hz, the dominating noise component is radiated by the rails even if the higher roughness levels are present on the wheels. Rolling noise is treated in more detail in Chapter 16.

8.4.3 Strategies to reduce wheel roughness

In most cases, a change to disc brakes would significantly reduce rolling noise levels from freight traffic. However, since the number of freight wagons in

Europe is about 600 000 and the life of the wagons is up to 30 years, such a replacement would be expensive. Therefore, efforts are made to replace the cast iron brake blocks with blocks of composite or sinter material. Since the physical properties of composite and sinter materials can vary significantly, the design of tailor-made blocks for a specific operation is possible.

One approach is to replace the present cast iron blocks with other types of brake blocks that, in addition, require a modification of the complete braking system. The use of high friction level materials, such as in the type K blocks, would lead to significant noise reductions but, again, this requires large investments by the railway operators. Thus, the introduction of the K blocks is slow, and the focus is instead on the development of new retrofit blocks with low life-cycle costs. The aim is to design new brake blocks, LL blocks, that are completely interchangeable with the existing cast iron brake blocks without changing the braking system. The material cost for K/LL blocks is higher than for cast iron brake blocks, but their life is also longer.[49] When composite material blocks are used, a larger part of the generated heat is entered into the wheel than is the case for conventional cast iron brake blocks, and this has implications for the wheel design.

8.5 Simulation of consequences of out-of-round wheels

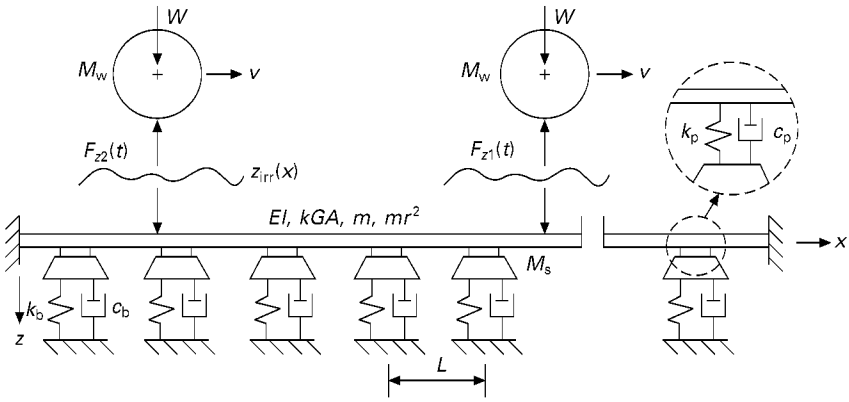
The interaction between a vehicle with out-of-round wheels and a railway track may involve dynamic excitation with a wide range of frequencies. To predict the resulting dynamic responses, such as wheel–rail contact forces, wheel and rail vibrations and rail and sleeper bending moments, appropriate models of both wheelset and track dynamics that are accurate in the excited frequency range are required. For a polygonal wheel with three harmonics on a train travelling at 300 km/h, this means frequencies up to about 100 Hz, cf Eq. (8.2). Repeated wheel–rail impacts due to a severe wheel flat may excite loads with significant contributions at frequencies up to and above 1 kHz, while rolling noise generated by rough wheels covers frequencies up to 5 kHz.

8.5.1 Mathematical models and computer programs

Simulation of dynamic vehicle–track interaction and consequences of out-of-round wheels can either be studied in the frequency domain or in the time domain. Each approach offers both advantages and disadvantages when compared to the other. For example, when interaction is solved in the frequency domain, the included models of vehicle, track and wheel–rail rolling contact need to be linear. This implies that the use of frequency domain models is limited to investigations of wheel out-of-roundness with small amplitudes

that result in maintained wheel–rail contact. Time domain models, on the other hand, can account for non-linear rolling contact mechanics. Further, they may be used to study consequences of severe wheel flats leading to temporary loss of wheel–rail contact followed by the impact when contact is recovered. However, solution times are generally significantly longer compared to frequency-domain models. Surveys of modelling and simulation of dynamic vehicle–track interaction are presented.[18,50,51] An illustration of a mathematical model containing two wheelsets on a discretely supported track is given in Fig. 8.14.

For wheel–rail contact excitation at frequencies above about 20 Hz, the dynamic behaviour of carbody and bogies is decoupled from the wheelsets by the secondary and primary suspensions. Therefore, for investigations of causes and consequences of most types of wheel out-of-roundness, only the gravity loads of carbody and bogies need to be accounted for. Further, simulation of vehicle–track interaction can often be limited to the vertical xz-plane, and lumped models of the unsprung mass of the wheelsets are sufficient. However, track dynamics need to be modelled in more detail, including the properties of rails, rail pads, sleepers and ballast/subgrade, for example by applying the finite element method. An exception requiring more detail in the wheel model occurs when vibration amplitudes and rolling



8.14 Illustration of a mathematical model of a coupled vehicle–track system. Vehicle model with two half wheelsets with unsprung mass M_w, half static axle load W and train speed v. Discretely supported rail modelled as a Rayleigh-Timoshenko beam with bending stiffness EI, shear stiffness kGA, mass per unit beam length m and rotary inertia per unit beam length mr^2. Rail pads with stiffness k_p and viscous damping c_p. Rigid half sleepers with mass M_s and sleeper distance L on supports with stiffness k_b and viscous damping c_b. Prescribed wheel/rail irregularity and unknown vertical wheel–rail contact forces F_{z1} and F_{z2}.

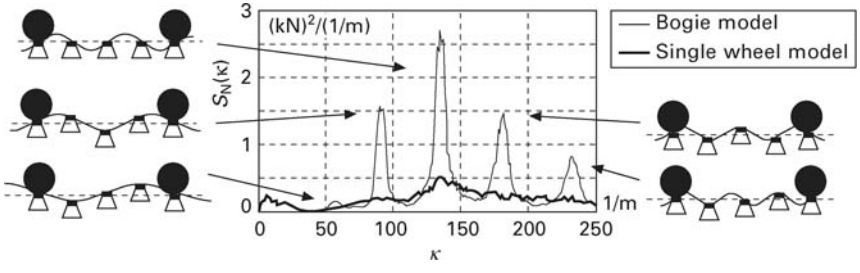
noise of the wheel are to be calculated. In general, the equations of motion for vehicle (index v) and track (index t) can be formulated in discretised form as:

$$\left.\begin{aligned}\mathbf{M}_v\boldsymbol{x}_v + \mathbf{C}_v\boldsymbol{x}_v + \mathbf{K}_v\boldsymbol{x}_v &= \boldsymbol{F}_v + \boldsymbol{F}_z\\ \mathbf{M}_t\boldsymbol{x}_t + \mathbf{C}_t\boldsymbol{x}_t + \mathbf{K}_t\boldsymbol{x}_t &= \boldsymbol{F}_z\end{aligned}\right\} \quad [8.4]$$

where $\mathbf{M}$, $\mathbf{C}$ and $\mathbf{K}$ are the mass, damping and stiffness matrices of vehicle and track, and $\boldsymbol{x}$ is the displacement vector containing the corresponding degrees of freedom. The vector $\boldsymbol{F}_z$ contains the unknown wheel–rail contact forces, while the given vehicle gravity loads are assembled in $\boldsymbol{F}_v$.

At higher frequencies, the effects of rail bending wave reflections between adjacent wheelsets (in a bogie or in adjacent coaches) on wheel–rail contact forces and track vibrations are not negligible, especially when rail pad stiffness is low. The mean spectral density of the calculated normal contact force when using either one or two rigid wheel set masses to model the vehicle is compared in Fig. 8.15.[52] The spectral density of the normal contact force between the front wheel (in a model with two wheelsets) and the rail versus wave number κ is illustrated. Five peaks appear in the investigated wave number interval ($0 < \kappa < 250$ m^{-1}, $\kappa = 2\pi/\lambda$ where λ is the irregularity wavelength). At the corresponding frequencies, see Eq. (8.2), the track vibrates with two, three, four, five and six half-wavelengths respectively between the two wheels. Note that the influence of these modes on the contact force is excluded when a single wheelset model is used. The influence of multiple wheelsets on the wheel–rail contact force has also been investigated.[53,54]

A non-linear compressive stiffness of the wheel–rail contact can be determined by assuming three-dimensional contact between two elastic half-spaces according to Hertz. Some of the assumptions in Hertz theory are that



8.15 Example of calculated spectral density $S_N(\kappa)$ of normal wheel–rail contact force versus wave number $\kappa = 2\pi/\lambda$ of rail roughness for a vehicle travelling on a track with a random rail irregularity. Vehicle speed is 150 km/h. Calculated vibration modes at resonances in coupled vehicle–track system are sketched and related to the corresponding peaks of the contact force spectrum.[52] (Figure courtesy Annika Lundberg)

the curvatures of the two bodies in contact are constant within the contact patch and that they can be described by second-order polynomials. Hertz theory results in an elliptical contact patch, and that the wheel–rail contact force $F_{z,i}$ and the elastic approach δ_i of wheel and rail are related according to:

$$F_{z,i} = k_{\mathrm{H}}\, \delta_i^a \qquad [8.5]$$

where $a = 3/2$ and k_{H} is a constant that is dependent on the material properties and the undeformed wheel–rail contact geometry. However, Hertz assumptions are violated in the contact between a wheel flat and a rail. To overcome this, the elastic non-Hertzian normal contact theory proposed by Kalker[55] can be used.[56] For different combinations of contact force and positions (angles) of the wheel flat relative to the rail, the approach δ is calculated prior to the dynamic simulation to save computational cost. By curve fitting, the resulting relationships can be formulated as in Eq. (8.5) but with different values for k_{H} and a. Further, for wheel/rail irregularities with short wavelengths, a contact filter[57] is necessary to avoid an over-estimation of the normal contact force. The contact filter leads to an attenuation of wheel–rail roughness amplitudes. The effect of the filter increases when the roughness wavelength approaches the dimensions of the contact patch.

The most realistic and accurate type of model in simulation of vehicle–track interaction takes into account that the wheels are moving along the rail, and that the wheel/rail irregularities are treated as prescribed relative wheel–rail displacements along the path of the wheels. This approach, sometimes referred to as the 'moving mass' model, is normally used in time-domain models. In the alternative 'moving irregularity model', the vehicle model remains at a fixed position along the rail, while an imaginary strip containing the combined effect of wheel and rail irregularities is pulled at a steady speed between the models of vehicle and track. The latter model is the common approach when vehicle–track interaction is solved in the frequency domain.

To investigate the influence of out-of-round wheels on wheel–rail contact forces and vehicle/track response, a function prescribing the wheel irregularity is required to constrain the relative motion of wheel and rail. If the order spectrum (see Fig. 8.5) of a wheel with radius R is known, the prescribed displacement z_{irr} can be written as:

$$z_{\mathrm{irr,polygonal}} = \sum_i A_i \sin\left(\frac{i}{R}\, x + \phi_i\right) \text{[m]} \qquad [8.6]$$

where i is the order of the wheel out-of-roundness, A_i is the amplitude [m] that is taken from the order spectrum, ϕ_i is a phase angle and x is a (circumferential) coordinate.

Based on a roughness level spectrum L_{r} with M one-third octave bands, samples of wheel irregularity profiles can be determined by using the formula:

$$z_{\text{irr,roughness}} = \sum_{k=1}^{M} \sum_{l=1}^{N} A_k \sin\left(\frac{2\pi}{\lambda_{kl}} x + \phi_{kl}\right) \text{[m]} \quad [8.7]$$

where the amplitudes of the N sines in each band k are determined as:

$$A_k = \sqrt{\frac{2}{N}} \times 10^{L_r^k} \text{[m]} \quad [8.8]$$

The wavelengths λ_{kl} are distributed within each band, for example by assuming a constant wave number increment, while the phase angles ϕ_{kl} can be uniformly randomly distributed between 0 and 2π.

A commonly used irregularity function to describe a wheel flat with length L, depth d and rounded edges is:[35, 58]

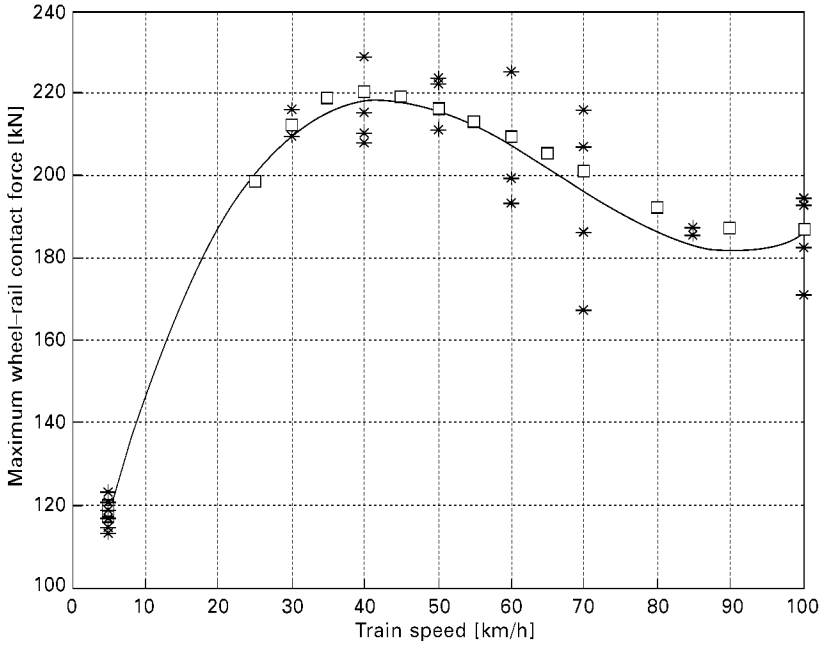
$$z_{\text{irr,flat}} = \frac{d}{2}\left(1 - \cos\left(\frac{2\pi}{L} x\right)\right), 0 \le x \le L \quad \text{[m]} \quad [8.9]$$

However, this function is simplified in the sense that it does not represent the true wheel trajectory as it neglects the circular shape of the wheel. Alternative irregularity functions for fresh and rounded wheel flats are given by Wu and Thompson[59] and Steenbergen.[60]

Examples of computer programs for simulation of dynamic vehicle–track interaction that are based on the frequency-domain approach in combination with the moving irregularity model are TWINS[47] and TRACK,[61] whereas DIFF,[62] DIFF3D,[63] VIA[64] and VICT[65] are examples of time-domain models. Several high-frequency vehicle–track interaction models were compared in a benchmark test.[66] Calculated maxima of wheel–rail contact forces, railhead accelerations and sleeper bending moments were compared. Significant differences in results from different models were observed in some cases. A more recent benchmark test has been summarised by Steffens and Murray.[67] Figure 8.16 illustrates an example from a study where a time-domain model was validated versus field measurements of the maximum wheel–rail contact force at different train speeds due to a 100 mm long wheel flat.[35] The irregularity function in Eq. (8.9) was used to prescribe the relative wheel–rail displacement. In another validation exercise,[68] good agreement was obtained between measured and calculated wheel–rail contact forces for a passenger train on a corrugated rail, where the combined wheel–rail roughness was described by Eqs (8.7) and (8.8).

8.5.2 Simulation of long-term irregular wheel wear

Several different damage mechanisms may be responsible for the long-term development of out-of-round wheels. Examples are wear, plastic deformation and RCF. In addition, different wear mechanisms (abrasive



8.16 Influence of train speed on maximum wheel–rail contact force due to a wheel flat with length 100 mm and depth 0.9 mm. Axle load $2W$ = 24 tonnes, unsprung wheelset mass $2M_w$ = 1185 kg. Measured data: *, calculated data: □.[35] The solid line is a curve fit (third-order polynomial) to visualise the trend in measured data.

wear, adhesive wear, delamination wear and oxidation or corrosive wear) can occur simultaneously.[69]

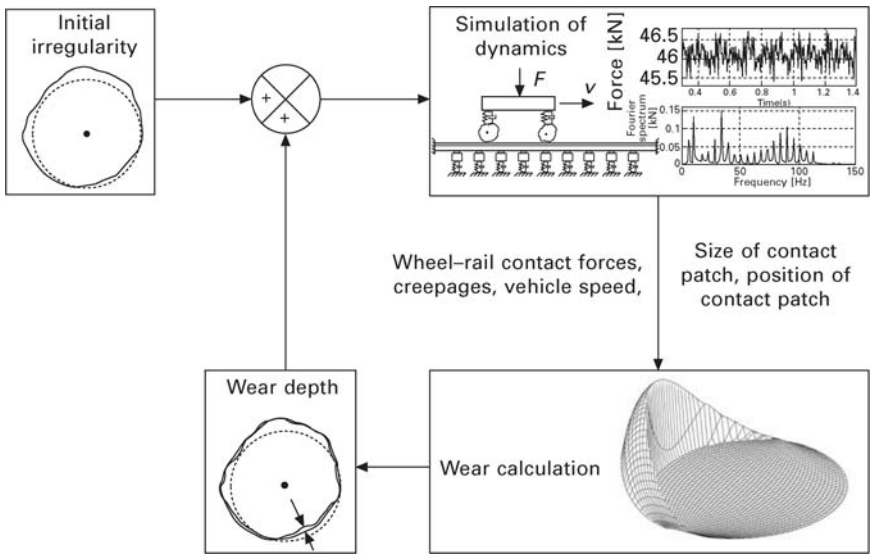
Several empirical wear models or 'wear indices' have been proposed. According to Archard,[70] the volume V_{wear} of removed material due to wear is formulated as:

$$V_{wear} = k \frac{F_z s}{H} \text{ [m}^3\text{]} \qquad [8.12]$$

Here F_z[N] is the normal wheel–rail contact force, s [m] the sliding distance, H [N/m^2] the hardness of the softer material and k a non-dimensional wear coefficient. Archard's wear model has been used in wheel–rail applications although the principal cause of wheel/rail wear is rolling and not pure sliding.[71] Another common approach is based on the assumption that the removed material volume is proportional to the frictional work in the contact patch. In this case, the frictional work is calculated as a sum of products containing the tangential forces and spin moment multiplied with the corresponding creepages in the wheel–rail contact.[72] The wear coefficient k may be determined from wear maps.[73]

Most mathematical models for prediction of irregular wheel wear include (i) a model of dynamic vehicle–track interaction in order to determine the distribution of forces and creepages in the wheel-rail contact patch, and (ii) a wear model to account for the long-term wear process of the wheel surface. To determine the wear distribution, the contact patch is discretised into an element grid. Hertz contact theory and the FASTSIM algorithm by Kalker[74] can be used to determine the normal contact pressure and the sliding distance for each element in the grid. Note that the Archard wear model is valid only in the slip region of the contact patch. The boundary between the stick and slip regions is determined from the calculations with FASTSIM. One basic approach in most models is the use of different time scales. In the dynamic interaction model, the timescale of the vibrations can be expressed in the order of seconds, while a time interval corresponding to an order of 10^6 wheel revolutions is considered in the wear model. For the dynamic interaction model, this assumption means that the wheel irregularities can be treated as constant in each simulation with given conditions on vehicle speed, axle load and track properties.

The coupling of the two models is often illustrated by a feedback loop, such as the one in Fig. 8.17. The initial wheel irregularity can be selected to contain contributions from a wide range of wavelengths, for example the result of a wheel re-profiling. Forces and creepages in the contact patch are calculated by the dynamic interaction model. Irregular wear versus location



8.17 Integration of dynamic vehicle-track interaction and damage process in a long-term feedback loop.[6] (Figure courtesy Anders Johansson)

around the wheel circumference is then determined based upon the chosen wear hypothesis. After each wheel revolution, the running surface is modified slightly by the wear. By this procedure, the long-term wear can be monitored in an iteration process. Certain model parameters, such as vehicle speed, axle load and track properties, can be varied from one iteration to the next in order to simulate more realistic operating conditions.

The initial wheel surface irregularity can be assumed in the form of a spectrum with wavelengths in a relevant interval, cf Eq (8.7). The wheel irregularity acts as input to the coupled vehicle–track system, which results in fluctuating (normal and tangential) contact forces, creepages and contact patch dimensions. Simulation of dynamic vehicle–track interaction will reveal that wheel–rail contact forces and wear are higher at certain frequencies than at others. For a certain vehicle speed, this means that wear will be higher at certain wavelengths. After millions of wheel revolutions, an irregular wear pattern may form. Common wavelength-fixing mechanisms are resonances of the coupled vehicle–track system, e.g. the P2 resonance, the vertical pinned–pinned resonance of the discretely supported rail or a torsional mode of the wheelset.[19]

As mentioned earlier, wear may not be the only active damage mechanism. Repeated normal and tangential loading may lead to accumulated plastic deformation of the wheel tread. The variation in wear rate due to hardening of the wheel surface has been accounted for in a model for simulation of wheel polygonalisation.[75]

8.6 Sources of further information and advice

Several literature surveys and text books dealing with the causes and consequences of out-of-round wheels, modelling of dynamic vehicle–track interaction and prediction of wheel wear have been published.

A classification of wheel defects is given in the *Atlas of Wheel and Rail Defects*.[76] Modelling of railway track and vehicle–track interaction at 'high frequencies' is surveyed by Knothe and Grassie.[50] The frequency range of interest is 20–5000 Hz, because at lower frequencies the dynamic behaviour of the track is not so significant. The upper frequency limit was selected because then wheelset and track response spectra contain the most dominating components leading to railway noise. The survey[51] discusses vehicle–track dynamics in the 'mid-frequency range' 50–500 Hz. An overview of modelling of wear phenomena leading to out-of-round wheels is also given. Research on out-of-round wheels has been the subject of surveys.[6,19,77]

Wheel–rail noise generation is one consequence of dynamic vehicle–track interaction with out-of-round wheels. A survey on this subject has been compiled by Thompson and Jones.[78] Overviews of RCF and wear are given in the survey[79] and in the text book.[69] Literature surveys on wheel flats, RCF

in wheels, wheel polygonalisation and wheel roughness on tread braked wheels are included in the dissertations.[1,5,6,11,12]

8.7 Acknowledgements

Large parts of this text stem from work performed within research projects at the Centre of Excellence CHARMEC (CHAlmers Railway MEChanics) at the Department of Applied Mechanics, Chalmers University of Technology in Gothenburg, Sweden. In particular, the contributions made by Drs Anders Ekberg, Johan Jergéus, Anders Johansson and Tore Vernersson are acknowledged.

8.8 References

1 Jergéus J (1998), *Railway Wheel Flats – Martensite Formation, Residual Stresses, and Crack Propagation*, PhD Thesis, Chalmers University of Technology, Gothenburg, Sweden.
2 Snyder T, Stone D H and Kristan J (2003), Wheel flat and out-of-round formation and growth, *Proceedings 2003 IEEE/ASME Joint Rail Conference*, Chicago, IL, USA, 22–24 April 143–8.
3 Fröhling R D (2007), 'Wheel/rail interface management in heavy haul railway operations – applying science and technology, *Vehicle System Dynamics*, **45**(7–8), 649–77.
4 Ekberg A and Kabo E (2005), Fatigue of railway wheels and rails under rolling contact and thermal loading – an overview, *Wear*, **258**, 1288–300.
5 Ekberg A (2000), *Rolling Contact Fatigue of Railway Wheels*, PhD Thesis, Chalmers University of Technology, Gothenburg, Sweden.
6 Johansson A (2003), *Out-of-Round Railway Wheels – Causes and Consequences*, PhD Thesis, Chalmers University of Technology, Gothenburg, Sweden.
7 Deuce R, 'Wheel tread damage – an elementary guide', Bombardier Transportation, 2007, 38 pp.
8 Zacher M (1990), Unrunde Räder und Oberbausteifigkeit, *Eisenbahntechnische Rundschau*, **45**(10), 605–10.
9 Rode W, Müller D and Villman J (1997), Results of DB AG investigations – out-of-round wheels, *Proceedings Corrugation Symposium – Extended Abstracts*, IFV Bahntechnik, Technische Universität Berlin, Berlin, Germany.
10 Pallgen G (1998), Unrunde Räder an Eisenbahnfahrzeugen, *Eisenbahningenieur*, **49**(1), 56–60.
11 Vernersson T (2006), *Tread Braking of Railway Wheels – Noise-Related Tread Roughness and Dimensioning Wheel Temperatures*, PhD Thesis, Chalmers University of Technology, Gothenburg, Sweden.
12 Petersson M (1999), *Noise-Related Roughness of Railway Wheels – Testing of Thermomechanical Interaction Between Brake Block and Wheel Tread*, Licentiate Thesis Chalmers University of Technology, Gothenburg, Sweden.
13 Ødegaard and Dannieskiold-Samsøe (2003) *Wheel and Rail Roughness Measuring System*, Copenhagen, Denmark, available at: http://www.odegaard.se/Files/Web_railway_roughness.pdf, accessed March 2009.

14 Kay C J, Rasaiah W G and Villanueva-leal A (2003), *Railway wheel monitoring*, US patent 6564467.
15 Verheijen E (2006), A survey on roughness measurements, *Journal of Sound and Vibration*, **293**(3–5), 784–94.
16 Johansson A (2006), Out-of-round railway wheels – assessment of wheel tread irregularities in train traffic, *Journal of Sound and Vibration*, **293**(3–5), 795–806.
17 Dings P C and Dittrich M G (1996), Roughness on Dutch railway wheels and rails, *Journal of Sound and Vibration*, **193**(1), 103–12.
18 Nielsen J C O, Kabo E and Ekberg A (2007), *Larmgräns för hjulskadedetektorer – en utredning av risk för rälbrott på Malmbanan (Alarm limit for wheel impact load detectors – an assessment of risk for rail fracture on Malmbanan)*, Research report 2007:05, Department of Applied Mechanics, Chalmers University of Technology, Gothenburg, Sweden (in Swedish).
19 Nielsen J C O, Lundén R, Johansson A and Vernersson T (2003), Train–track interaction and mechanisms of irregular wear on wheel and rail surfaces, *Vehicle System Dynamics*, **40**(1–3), 3–54.
20 Johansson A and Nielsen J C O (2003), Out-of-round railway wheels – wheel–rail contact forces and track response derived from field tests and numerical simulations, *Proceedings of the ImechE, Part F, Journal of Rail and Rapid Transit*, **217**, 135–46.
21 Jenkins H H, Stephenson J E, Clayton G A, Morland G W and Lyon D (1974), The effect of track and vehicle parameters on wheel/rail vertical dynamic forces, *Railway Engineering Journal*, January, 2–16.
22 Newton S G and Clark R A (1979), An investigation into the dynamic effects on the track of wheel flats on railway vehicles, *Journal of Mechanical Engineering Science*, **21**(4), 287–97.
23 Kalay S, Tajaddini A and Stone D H (1992), Detecting wheel tread anomalies, *ASME Rail Transportation Division (Publication)*, **5**, 165–74.
24 Stone D H, Kalay S and Tajaddini A (1992), Statistical behaviour of wheel impact load detectors to various wheel defects, *Proceedings 10th International Wheelset Congress*, Sydney, NSW, Australia, 27 September–October, 9–13.
25 Kalay S, Tajaddini A, Reinschmidt A and Guins A (1995), Development of a performance-based wheel removal criteria for North American Railroads, *Proceedings 11th International Wheelset Congress*, Paris, France, 18–22 June, 227–33.
26 Fermér M and Nielsen J C O (1994), Wheel/rail contact forces for flexible versus solid wheels due to tread irregularities, *Vehicle System Dynamics*, **23**(Supplement), 142–57.
27 Gullers P, Andersson L and Lundén R, High-frequency wheel–rail contact forces – field measurements and influence of track irregularities, Wear, 265 (9–10), 2008, pp 1472–8.
28 Salient Systems, Inc. (2005) *Wheel Impact Load System*, Dublin, Ireland, available at; http://www.salientsystems.com/prod_wild.html, accessed March 2009.
29 Lloyds Register (2007), *GOTCHA® Wayside monioring platform to measure the quality of trains and the utilisation of the track*, London, UK available at: http://gotchamonitoringsystems.com/index.html, accessed March 2009.
30 Bontekoe T, 'Beneficial monitoring: not whether, but when?', *European Railway Review*, 2007(5) 69–73.
31 Schenck Process Group (2007), *MULTIRAIL® WheelScan – Diagnosing and Detecting Precise in-transit wheel diagnosis*, Prague, Czech Republic, available at: http://www.

schenchk.cz/en/products/product/wheel-diagnosis-wheelscan.html, accessed March 2009.
32 Berggren E (2003), *Statistik från hjulskadedetektorer 2002-01–2003-09 (Statistics from wheel impact load detectors 2002-01–2003-09*, The Swedish National Rail Administration (Banverket), BBS report 2003-10-20, Borlänge, Sweden (in Swedish).
33 Zeta-Tech Associates (1997), *Wheel impact detection systems – the North American experience*, Suite Cherry Hill, NJ, USA.
34 Kalay S and Hargrove M B (1994), Research, test & development – examining the economics of high-impact wheel loads, *Railway Track & Structures*, **90**, 29–33.
35 Nielsen J C O, Ringsberg J and Baeza L (2005), Influence of railway wheel flat impact on crack growth in rails, *Proceedings 8th International Heavy Haul Conference*, Rio de Janeiro, Brazil, **14–16**, June, 789–97.
36 Sandström J and Ekberg A (2007), Predicting crack growth and risks of rail breaks due to wheel flat impacts in heavy haul operations, *Proceedings International Heavy Haul Association Specialist Technical Session*, Kiruna Sweden, 11–13 June, 379–88.
37 Mok H, Chiu W K, Peng D, Sowden M and Jones R (2007), Rail wheel removal and its implication on track life: a fracture mechanics approach, *Theoretical and Applied Fracture Mechanics*, **48**, 21–31.
38 Vernersson T, Petersson M and Hiensch M (1998), Thermally induced roughness of tread braked railway wheels, *Proceedings 12th International Wheelset Congress*, Qingdao, China, 21–25 September, 68–75.
39 Petersson M (2000), Noise-related roughness of railway wheel treads – full-scale testing of brake blocks, *Proceedings of the ImechE: Part F, Journal of Rail and Rapid Transit*, **214**, 63–77.
40 Kennedy F E Jr (1984), Thermal and thermomechanical effects in dry sliding, *Wear*, **100**, 453–76.
41 Vernersson T (1999), Thermally induced roughness of tread braked railway wheels, part 1: brake rig experiments, *Wear*, **236**, 96–105.
42 Vernersson T (1999), Thermally induced roughness of tread braked railway wheels, part 2: modelling and field measurements, *Wear*, **236**, 106–16.
43 Fec M C and Sehitoglu H (1985), Thermal-mechanical damage in railroad wheels due to hot spotting, *Wear*, **102**, 31–42.
44 Cole I S (1992), The effects of localized heating on railway wheel steels – an investigation into effects of tread braking, *Proceedings 10th International Wheelset Congress*, Sydney, NSW, Australia, 27 September–1 October, 319–24.
45 Hemsworth B and Jones R R K (2000), *Brite/Euram III Silent Freight – final report*, Brite/Euram III Silent Freight Technical Report 5E0U15T1.DB, AEA Technology, Derby, UK.
46 Thompson D J (1997), *Definition of the reference roughness*, Brite/Euram III Silent Freight Technical Document 3S7J28T1.DA, Institute of Sound and Vibration Research, University of Southampton, Southampton, UK.
47 *TWINS (Track Wheel Interaction Noise Software) theoretical and user's manual (version 3.0)*, TNO Institute of Applied Physics, Delft, the Netherlands, 1999.
48 Jones C J C and Thompson D J (2003), Extended validation of a theoretical model for railway rolling noise using novel wheel and track designs, *Journal of Sound and Vibration*, **267**(3), 509–22.
49 Oertli J (2008), Railway noise abatement: the case for retrofitting freight vehicles

with composite brake blocks, in Schulte-Werning B, Thompson D and Gautier P (eds), *Noise and Vibration Mitigation for Rail Transportation Systems, Notes on Numerical Fluid Mechanics and Multidisciplinary Design 99*, Springer-Verlag, Berlin Heidelberg, Germany, 236–42.

50 Knothe K and Grassie S L (1993), Modelling of railway track and vehicle/track interaction at high frequencies, *Vehicle System Dynamics*, **22**(3–4), 209–62.

51 Popp K, Kruse H and Kaiser I (1999), Vehicle-track dynamics in the mid-frequency range, *Vehicle System Dynamics*, **31**, 423–64.

52 Igeland A (1996), Railhead corrugation growth explained by dynamic interaction between track and bogie wheelsets, *Proceedings of the ImechE: Part F, Journal of Rail and Rapid Transit*, **210**(1), 11–20.

53 Wu T X and Thompson D J (2002), Behaviour of the normal contact force under multiple wheel/rail interaction, *Vehicle System Dynamics*, **37**(3), 157–74.

54 Johansson A and Nielsen J C O (2007), Rail roughness growth – influence of powered wheelsets with wheel tread irregularities, *Wear*, **262**, 1296–307.

55 Kalker J J (1990), *Three-Dimensional Elastic Bodies in Rolling Contact*, Kluwer, Dordrecht, the Netherlands.

56 Baeza L, Roda A, Carballeira J and Giner E (2006), Railway train–track dynamics for wheel flats with improved contact models, *Nonlinear Dynamics*, **45**, 385–97.

57 Remington P J (1987), Wheel/rail rolling noise I: theoretical analysis, *Journal of the Acoustical Society of America*, **81**, 1805–23.

58 Wu T X and Thompson D J (2004), The effects of track non-linearity on wheel/rail impact, *Proceedings of the ImechE: Part F, Journal of Rail and Rapid Transit*, **218**, 1–15.

59 Wu T X and Thompson D J (2002), A hybrid model for the noise generation due to railway wheel flats, *Journal of Sound and Vibration*, **251**(1), 115–39.

60 Steenbergen M (2007), The role of the contact geometry in wheel–rail impact due to wheel flats, *Vehicle System Dynamics*, **45**(12), 1097–116.

61 Grassie S L, Gregory R W, Harrison D and Johnson K L (1982), The dynamic response of railway track to high frequency vertical excitation, *Journal of Mechanical Engineering Science*, **24**(2), 77–90.

62 Nielsen J C O and Igeland A (1995), Vertical dynamic interaction between train and track – influence of wheel and track imperfections, *Journal of Sound and Vibration*, **187**(5), 825–39.

63 Andersson C and Abrahamsson T (2002), Simulation of interaction between a train in general motion and a track, *Vehicle System Dynamics*, **38**(6), 433–55.

64 Baeza L, Roda A and Nielsen J C O (2006), Railway vehicle/track interaction analysis using a modal substructuring approach, *Journal of Sound and Vibration*, **293**(1–2), 112–24.

65 Zhai W and Sun X (1994), A detailed model for investigating vertical interactions between railway vehicle and track, *Vehicle System Dynamics*, **23**(Supplement), 603–15.

66 Grassie S L (1996), Models of railway track and vehicle/track interaction at high frequencies: results of benchmark test, *Vehicle System Dynamics*, **25**(Supplement), 243–62.

67 Steffens D and Murray M (2005), Establishing meaningful results from models of railway track dynamic behaviour, *Proceedings 8th International Heavy Haul Conference*, Rio de Janeiro, Brazil, 14–16 June, 41–9.

68 Nielsen J C O, High-frequency wheel–rail contact forces – validation of a prediction model by field testing, Wear, 265 (9–10), 2008, pp 1465–71.
69 Bayer R G (2002), *Wear Analysis for Engineers*, HNB, New York, USA.
70 Archard J F (1953), Contact and rubbing of flat surfaces, *Journal of Applied Physics*, **24**, 981–8.
71 Jendel T (2002), Prediction of wheel profile wear – comparisons with field measurements, *Wear*, **253**, 89–99.
72 Elkins J A and Eickhoff B M (1979), Advances in non-linear wheel–rail force prediction methods and their validation, *ASME Winter Meeting*, New York, USA, 25–30 November.
73 Lewis R, Dwyer-Joyce R S, Olofsson U and Hallam R I (2004), Wheel material wear mechanisms and transitions, *Proceedings 14th International Wheelset Congress*, Orlando, FL, USA, 17–21 October, on CD.
74 Kalker J J (1982), A fast algorithm for the simplified theory of rolling contact, *Vehicle System Dynamics*, **11**(1), 1–13.
75 Meywerk M (1999), Polygonalization of railway wheels, *Archive of Applied Mechanics*, **69**(2), 105–20.
76 UIC (2004), Atlas of Wheel and Rail Defects, International Union of Railways, France.
77 Nielsen J C O and Johansson A (2000), Out-of-round railway wheels – a literature survey, *Proceedings of the ImechE: Part F, Journal of Rail and Rapid Transit*, **214**, 79–91.
78 Thompson D J and Jones C J C (2000), A review of the modelling of wheel/rail noise generation, *Journal of Sound and Vibration*, **231**(3), 519–36.
79 Tunna J, Sinclair J and Perez J (2007), A review of wheel wear and rolling contact fatigue, *Proceedings of the ImechE: Part F, Journal of Rail and Rapid Transit*, **221**(2), 271–89.

Copyrighted Materials
Copyright © 2009 Woodhead Publishing Retrieved from www.knovel.com

9

Rail surface fatigue and wear

D. I. FLETCHER, University of Sheffield, UK; F. J. FRANKLIN, University of Newcastle, UK; A. KAPOOR, Swinburne University of Technology, Australia

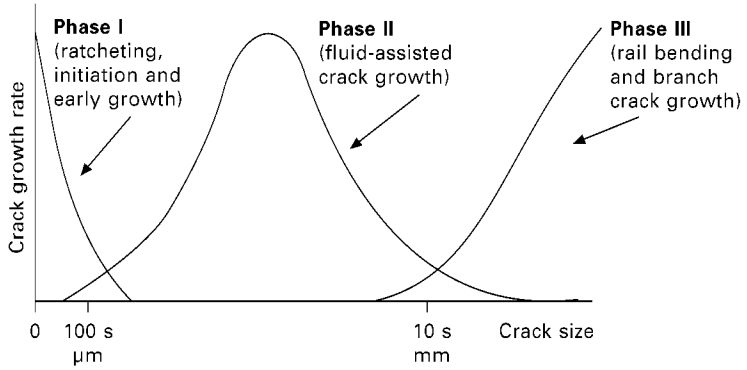
Abstract: Rail fatigue and wear are degradation processes which drive rail maintenance and replacement, and hence have economic importance in railway operation. The underlying mechanisms of these processes are described, and the importance of their interaction in determining the life of the rail is discussed. Methods for detailed prediction of crack growth rate and rail wear rate are introduced, providing a link between vehicle loads and the rail damage produced.

Key words: wear, fatigue, fracture mechanics, head checks, rolling contact fatigue.

9.1 Introduction

This chapter deals with wear and fatigue failures at the surface of the rail, although failure by fatigue may also occur in the rail web (Sih and Tzou, 1985) or foot (Cannon, 2003). Rail fatigue and wear depend crucially on the repeating contact loads to which the rail surface is subjected, and earlier chapters on contact mechanics have covered the calculation of contact stress in detail. The main factors differentiating fatigue and wear failure at the rail–wheel contact from failures in the majority of engineering components are that the repeatedly applied loads are highly compressive. Most fatigue failures in general engineering components are produced by tensile loading, and the mechanisms by which compressive loads produce failure are explored in this chapter. Cracks very similar to those that develop in rails can form in gears (Townsend *et al.*, 1986; Rieger and Din, 2003; Kahraman, 2004) and the rolls use in steel production (Frolish *et al.*, 2002). These share the characteristics of the rail–wheel contact in that they include highly concentrated non-conformal contacts, in which the surfaces are in combined rolling and sliding motion relative to one another.

Since rail fatigue and wear are complex subjects in themselves, this chapter uses the Hertzian approach to describe the contact stress, although most of the crack growth and wear mechanisms could equally be described and investigated through calculation of the real stresses present at the rail–wheel interface. In terms of crack size, the life of a surface breaking rolling contact fatigue (RCF) crack in a rail can be split into three main phases (Fig. 9.1).



9.1 Crack growth can be divided into a series of phases, each characterised by a different dominant crack growth mechanism, and corresponding model.

In phase I the rail is either-crack free, or undergoing the process of crack initiation and very early crack growth, with defects up to hundreds of microns in size. This stage of crack growth is dominated by development of defects in the severely plastically deformed near-surface layers of the rail, and the defect sizes correlate with the depth of this layer. Transition to the second phase of crack life takes place as defects initiated in the near surface layers of the rail grow large enough for a range of additional crack growth mechanisms to begin driving their growth. These mechanisms are applicable to cracks of millimetres to tens of millimetres in size, and are the main subject of this chapter. It is at this stage that mechanisms must describe how a compressive load (which would usually be expected to close cracks and restrict their growth) can lead to crack propagation.

The final stage of crack growth is usually rapid, as the crack becomes large, and is driven by bulk stresses in the rail, including rail bending stress and tensile residual stresses. Models exist to describe this stage of life for cracks in rails, and these are touched on below. However, cracks of this size are too large to be left in the rail and managed, and the only solution is urgent replacement of the rail.

During the process of crack growth the contact stress at the rail–wheel interface is almost always driving wear of the surfaces as well as initiating and possibly propagating cracks. This is not always disadvantageous, since natural wear is able to remove material containing very small cracks before these become dangerous. However, wear can also lead to loss of rail profile, thereby changing the contact mechanics of the rail–wheel interface. Rail wear can also produce changes in the gauge of the track, affecting the vehicle dynamics, and excessive loss of rail cross-section is a criterion for rail replacement.

The introduction above has mainly discussed crack size and life rather

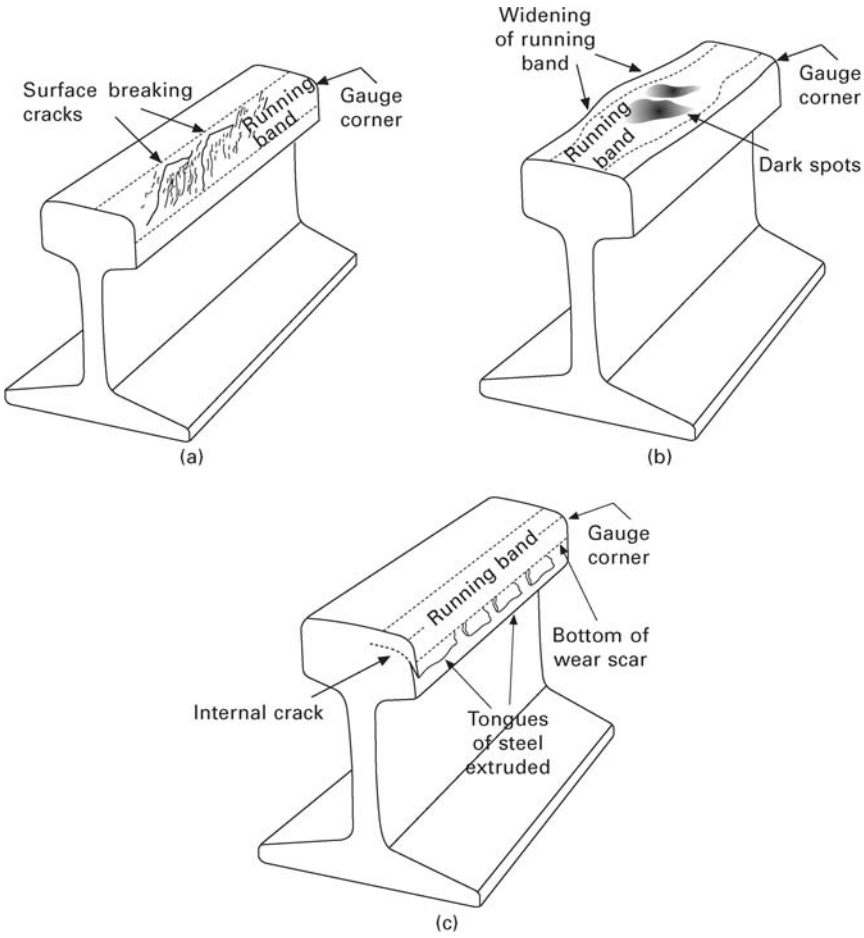
than overall rail life. This is because a rail may contain a range of different defect sizes at any time, since large cracks or defects do little to prevent the initiation and development of new smaller defects in neighbouring sections of the rail. In such cases overall rail life would be determined by the largest defect present. An additional complication is that two or more mechanisms may be driving crack growth at certain stages of growth. A common approach to modelling this has been to consider the fastest individual mechanism among any set of competing crack growth mechanisms, and to assume that this mechanism dominates (Hourlier and Pinean, 1981).

The stress to which the rail is subjected in service varies widely even for a single location. A rail will be exposed to a wide range of different wheels, making contact in slightly different places on the railhead and with different conformity of contact, and hence different pressures. The actual vehicles and their dynamic performance will also vary so, even if the wheel–rail geometry was identical, the contact loads will differ. Track forces are also dependent on the support structures, ballast thickness, sleeper type and the presence of welds or joints which can excite dynamic loading. This means that the forces driving fatigue and wear are highly variable and their quantification is complex.

9.2 Rail rolling contact fatigue

There are a wide range of terms used to describe different types of rail fatigue cracking, and these often vary between different regions of the world. In fact, the type of cracking experienced varies too, since different rail grades, traffic types, vehicle characteristics and maintenance regimes are in place in different areas. In the UK the main categories of railhead fatigue failure are 'head checks', 'squats' and 'tongue lipping', all shown in Fig. 9.2, with the following broad characteristics:

- **Head checks**. Also known as 'gauge corner cracking' when it occurs at the gauge corner, these are initially small and fine cracks on the rail surface, which grow down into the rail at a shallow angle below the surface. As they enlarge they may branch down causing a rail break, or branch up leading to spalling of the rail surface. This type of cracking is typically found on highly canted curves, with the curve radius and cant deficiency helping to determine whether cracks develop on the rail crown or gauge face. These cracks are discussed in detail below.
- **Squats**. Identified by a darkening of the rail surface, combined with a widening of the running band. This appearance is due to a horizontal crack below the rail surface which allows the near surface material to flow sideways, thereby producing widening of the running band, and produces a surface depression which collects dirt and becomes corroded leading to the characteristic shadow. Branch cracks can form from the



9.2 Schematic illustrations of common rail fatigue damage types: (a) head checks; (b) squats; (c) tongue lipping.

original horizontal crack, leading to spalling of the rail surface, or a rail break.

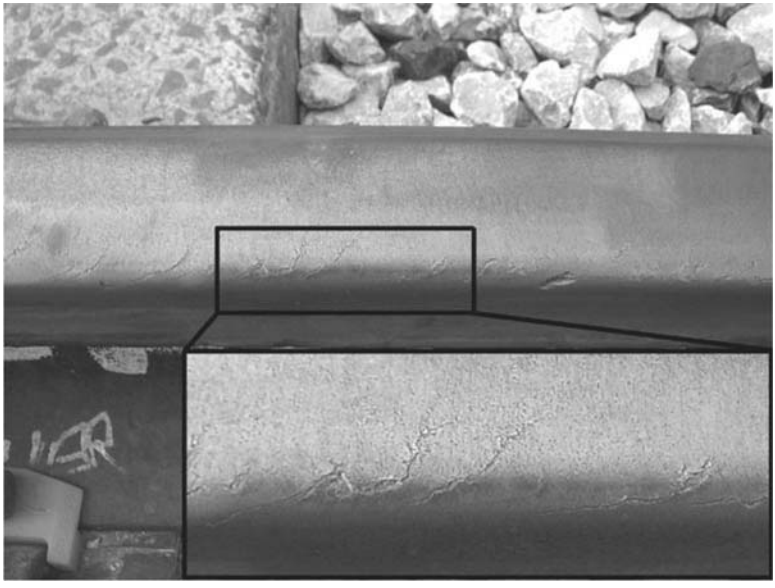
- **Tongue lipping**. Identified by the extrusion of thin slivers or tongues of material from the running band of the rail, often extending down the gauge face of the rail by several millimetres. Cracks can form below the extruded steel and grow into the railhead in a near horizontal plane. Branches can then form running either up or down, leading to a rail break.

Fatigue failures such as tache ovale (Cannon *et al.*, 2003) which occur due to internal defects in the steel of the rail head are now very rare due to improved steel cleanliness, and these are not discussed here.

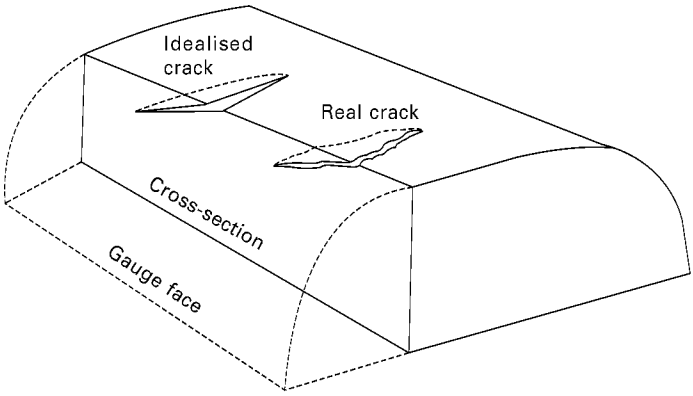
9.2.1 Head checks – detailed description

Head check cracking often leads to a series of closely spaced cracks of similar appearance, and Fig. 9.3 shows the typical surface appearance of a rail in which this has happened. On straight (or tangent) track, train acceleration or braking are dependent on longitudinal traction, and some lateral and spin tractions will also exist to steer the vehicle. In a curve, the proportion of lateral traction increases to supply the force required for cornering, and cracks which develop are aligned closely with the direction of resultant traction. This can be seen in Fig. 9.3, where the angle of the crack mouths relative to the longitudinal direction of the rail reflects the high proportion of lateral traction in the rail–wheel contact.

Internal geometry of head check cracks is shown schematically in Fig. 9.4. At very small crack sizes (phase I of growth) the shape of the crack will be largely microstructurally determined, but during phase II growth the shape is primarily controlled by the applied stress. Crack shapes can be revealed through destructive examination of cracked rails, for example by grinding back incrementally through the rail and mapping points to define the crack at each stage. The laborious nature of such crack mapping means that it has been conducted on only a small number of cracks, and in most cases in which destructive examination is used crack size is assessed from a single cross-section through the rail. Crack shape varies greatly depending on where the crack begins to grow relative to the rail gauge and crown,



9.3 Head check visual appearance.



9.4 Schematic illustration of internal head check geometry.

how large it has become, and whether any branches have developed. Rail maintenance and changes in vehicles can mean that existing cracks lie to the side of the current running band, so non-uniform growth around the crack front can be expected. To avoid over complexity in modelling these cracks simplifications of the real crack shapes are often required. Much of the modelling described below makes the assumption that the crack shape is semi-circular or semi-elliptical and that it lies in a single plane at a shallow angle below the rail surface. In addition, the assumption is often made that the crack lies in a semi-infinite half-space or quarter-space. Comparison with computer models of rails containing similar cracks has shown that this is a reasonable assumption for most phase II cracks, particularly at smaller crack sizes for which the railhead geometry has less of an influence.

The orientation of the cracks below the surface is typically at a shallow angle. The crack growth mechanisms of phase II lead to longitudinal crack growth in the opposite direction to the traction at the rail surface. Although most phase II crack growth takes place beyond the depth of plastic flow in the near surface layers of the rail, the crack growth direction can easily be visualised if this plastic flow is considered. Effectively, the near-surface parts of the crack are being 'pushed back' by traction forces and plastic flow produced by driving wheels, moving them away from the crack tip, while the crack tip moves forward away from the crack mouth. The distinction that the damage is produced by a driving wheel rather than a braking wheel is significant. The underlying mechanisms of crack growth depend on traction acting in the opposite direction to wheel motion (i.e. a driving wheel) and do not always apply to braking wheels.

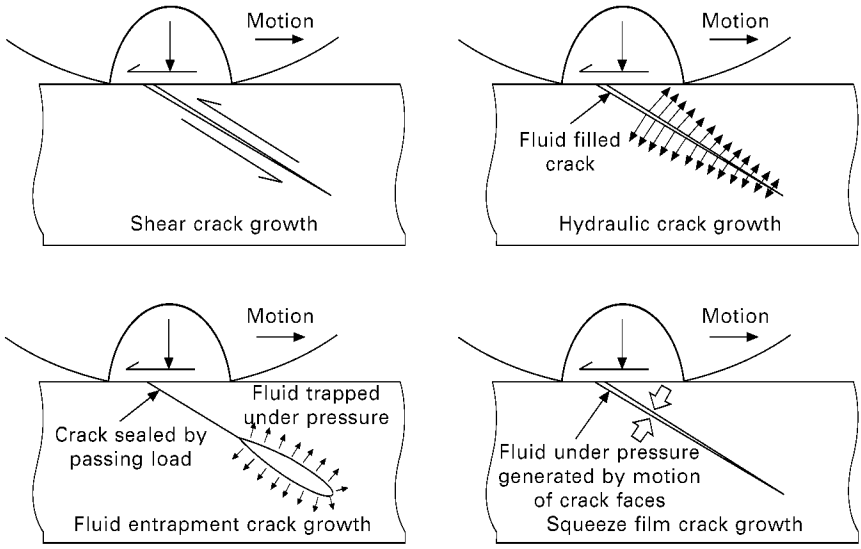
9.2.2 Underlying mechanisms

Assuming that small phase I cracks have already developed, two primary mechanisms exist for propagation of phase II cracks, which develop in a

compressive stress regime, and are not subject to the bulk tensile stress which drives crack growth in most situations. First, even in a compressive stress regime, shear stress can act leading to crack growth through sliding of the crack faces over one another. Second, since head check cracks are surface breaking rather than internal, fluids (e.g. rain water) can enter and assist growth through a hydraulic mechanism. Fluids can also lubricate the crack faces, reducing friction which would otherwise restrict shear mode growth. The role of fluids is supported by experimental and field evidence that fluids are required for the propagation into the railhead of head checks and gauge corner cracks in most cases. However, fluids are not required for crack initiation and, since plastic flow can extend to several millimetres in depth, shallow cracking can develop without fluids.

Figure 9.5 illustrates four fluid-assisted crack growth mechanisms applicable to head check cracks using a cross-sectional view of a rail containing a crack. The action of these mechanisms is summarised below and in Fig. 9.5:

1. **Shear driven crack growth**. This mechanism is based on the cyclic shear stress to which passing wheels subject the rail. Bower (1988) found the crack growth predictions for this mechanism to be consistent with both the experimental and railway experience of RCF cracking, but that propagation by cyclic shear was unlikely unless the crack face friction coefficient for typical cracks was below 0.2. At the date of Bower's work, the experimental evidence was also that cyclic shear loading always led rapidly to crack branch formation, when examination



9.5 Fluid-assisted crack growth mechanisms.

of cracked rails shows that head checks actually propagate up to tens of millimetres in their original plane before branching. Only subsequently did Bold *et al.* (1991) demonstrate that coplanar shear-driven growth is possible under the type of sequential mixed mode loading experienced in the railhead.

The shear of crack faces over one another (an assumption of the shear crack growth mechanism) is particularly difficult to observe in controlled conditions because of the requirement for a contact load and simultaneous shearing load. However, through construction of a special test stand Stupnicki and Morbarigazzi (2004) have observed the deformation using holographic interferometry. Tests were performed in 900A rail steel, and showed the microslip of the faces of a subsurface fatigue crack.

2. **Hydraulic pressure transmission.** Direct hydraulic transmission of contact pressure to the crack faces, without entrapment of fluid in the crack, was found by Bower (1988) to be unable to account for the effect of load motion or traction direction on crack growth. However, the stress intensity factor ranges predicted for this mechanism were sufficient to produce crack growth.
3. **Fluid entrapment.** Fluid sealed and compressed in a crack by the passing wheel load was found by Bower (1988) to predict sufficiently high stress intensities (and therefore crack growth rates) for crack growth to take place, and the growth predicted was sensitive to the direction of the motion of the load, as it should be to explain observed crack growth. However, the complex combination of tensile and shear stress predicted at the crack tip made it difficult to predict the crack growth direction, so this could not positively be correlated with observed crack growth.
4. **Squeeze film fluid action.** The 'squeeze film' (Walker *et al.*, 1969; Bogdanski, 2002a) lubrication mechanism was originally used to predict the pressure developed between components of rolling element bearings and also human synovial joints such as the knee, in which lubricant is alternately squeezed out and drawn into a gap. Modelling this mechanism predicts that high pressures will be generated in a crack filled with fluid when crossed by a rolling/sliding contact similar to a rail–wheel contact, even if the crack is not sealed and the fluid is not 'entrapped'. The possibility that high pressures (and consequently high crack growth rates) can be produced even when a crack is not sealed is important because surface breaking cracks in rails continue to grow even when their dimensions are too large for them to be covered and sealed by passing wheel contacts. Under these conditions, entrapment is implausible, so mechanisms 2 and 3 would be unlikely to account for crack growth.

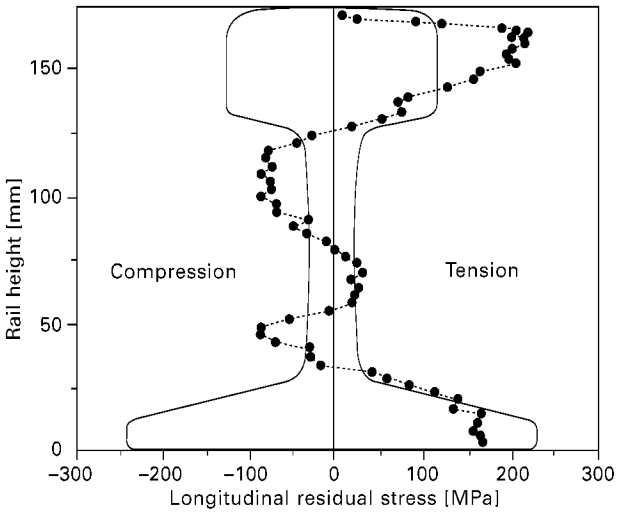
Early experimental investigations focusing on rail materials identified the importance of fluids in crack growth. Stress analysis was conducted to investigate whether application of fluid to a contact modified the friction and other contact conditions enough for crack growth to occur, or whether the fluid itself was responsible for crack growth. This work (Dowson *et al.*, 1963) took place to accompany tests undertaken by Ollerton and Morey (1963) in which a BS11 rail steel disc was tested against a wheel steel in pure rolling at a contact pressure of 1960 MPa. With no lubrication the disc wore, but there was no spalling or pit formation at 40 million contact cycles; however, pitting developed rapidly when a lubricant was applied.

More recent work with solid lubricants has allowed further investigation of whether it is the change of friction brought about by lubrication or the entry of lubricant into cracks which is the primary trigger for crack growth (Fletcher and Beynon, 1999a, 2000a). It was found that residue of dry solid lubricant present in surface breaking cracks (applied using a liquid carrier medium, but dried prior to testing) was correlated with a dramatic acceleration in crack growth, indicating that reduction of crack face friction rather than direct hydraulic action was an important factor controlling crack growth.

The importance of fluid in the propagation of RCF cracks has been further emphasised in a recent study by Burris and Kudish (2004), in which it was found that in the presence of fluid, and depending on crack orientation and size, the stress intensity factors for surface cracks may be two or more orders of magnitude higher than for similar subsurface ones, which are not fluid filled. It should be noted that this study was concerned with cracks of similar size to the contact patch, and these crack growth rate predictions may not apply to the later stages of RCF crack growth in which rail bending becomes important, or to very large internal cracks such as tache ovales.

In addition to contact stress, there are several other influences on crack growth in phase II. These include:

1. **Thermal stress.** Longitudinal expansion or contraction is restricted in long lengths of continuously welded rail, so longitudinal stress can develop as temperature changes. In general, compressive longitudinal stress must be avoided to prevent rail buckling in hot weather, although excessive tensile stress must also be avoided since this drives small cracks to become rail breaks (Hooper, 2008). The rail 'neutral temperature' is the temperature at which the rail is stress-free, and pre-tensioning can be used to set this point during installation. Rail longitudinal stress will change by 2.5MPa per degree Celsius change away from the neutral temperature (Szelazek, 1992; Jeong *et al.*, 1998).
2. **Residual stress.** Two sources of residual stress affect rails. First, stress is locked into the rail during manufacture, particularly through 'roller straightening' of the rail. A typical longitudinal residual stress distribution is shown in Fig. 9.6 (Webster *et al.*, 1992), although some rails have now

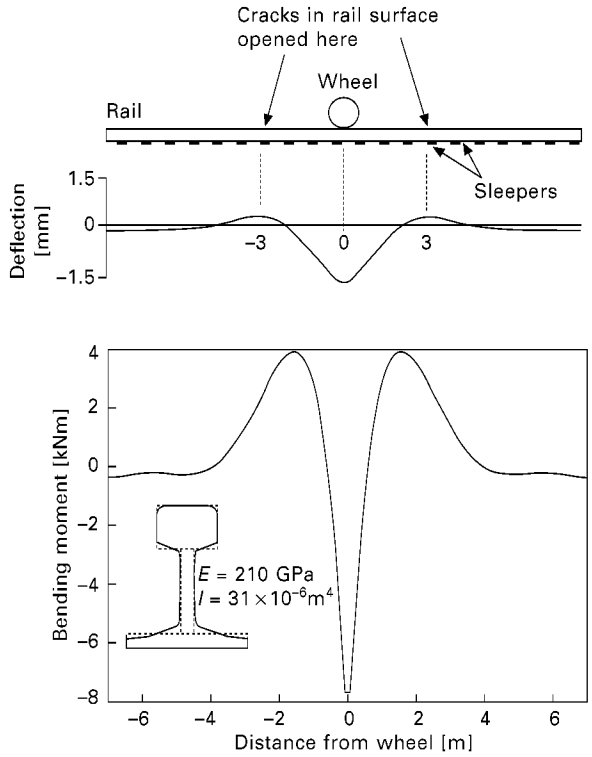


9.6 Typical internal longitudinal residual stress distribution.

been developed to give a more favourable compressive stress distribution in the railhead. The second source of residual stress is plastic flow of the steel near the contact surface during the passage of trains. The stresses develop as part of the 'shakedown' process, and are usually protective against further plastic damage. The rate of plastic flow therefore falls rapidly as traffic uses the rails (Williams, 2005).

3. **Rail bending.** The simplest rail bending models assume that the rail is supported on a continuous elastic foundation, and are known as 'Winkler' models. More complex models consider multiple wheels running on a rail which is supported intermittently on a (relatively) flexible sleeper and ballast system. Such models show the bending of the rail and accompanying longitudinal stress which can be expected to develop, and these are shown schematically in Fig. 9.7. Longitudinal bending moments in the rail of approximately 4000 Nm have been found during modelling work (Dukkipati and Dong, 1999; Fletcher and Kapoor, 2006a and Kapoor and Fletcher, 2006b). Bending in the lateral direction is also possible (Hay, 1982), and models can be used to predict the stress produced in the rail by this lateral bending.

Although in most cases these additional stresses are low relative to the contact stress, they can have a large effect when acting in combination with contact stress. For example, a shallow crack under compressive load can be restricted from growing in shear by friction between the crack faces (mechanism 1 above). Modelling has shown (Kapoor and Fletcher, 2006a) that even a modest additional stress (such as a longitudinal thermal stress



9.7 Schematic representation of rail bending.

which can develop in a long continuous rail) can be sufficient to allow the crack faces to slide, allowing the contact load to drive crack growth. The minor longitudinal stress effectively 'unlocks' the system, allowing the contact stress to do more damage. Such interaction between applied stresses through crack closure effects means that, when considering the combined effect of contact and addition stresses, a resultant stress should be found from which to calculate crack growth rate. Superposition of growth rates predicted for each stress applied individually will give unrealistic results, even though most models of Phase II crack growth treat the rail as an elastic system.

9.3 Experimental investigations

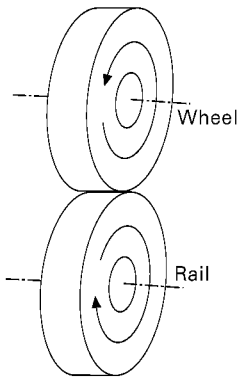
9.3.1 Laboratory investigations

The complexity of rail–wheel contact fatigue failure makes it difficult to investigate through observation of crack growth on rails in service, where any piece of rail will be subject to many different vehicle types and wheel profiles. Over the majority of crack development, crack propagation rate

is of the order of a few nano-metres during each wheel pass, so on-site observation can also require years of monitoring.

Laboratory tests have therefore been developed to look at RCF crack growth, making possible rapid application of large numbers of contact passes to a rail material. In addition, laboratory-based tests offer much greater control of the applied stress, so it becomes possible to understand the individual influence of normal load, traction load, lubrication and other variables. In many cases, the simplification is made of representing the rail–wheel contact by a 'twin-disc' system (Fig. 9.8) which uses two parallel cylinders of rail and wheel material to produce rolling-sliding loads characteristic of rail–wheel contact. With this system it is usually possible to generate only a single tangential load representing longitudinal traction (driving/braking) in addition to the normal contact load. Lateral traction and spin which affect a real contact usually cannot be simulated, although approximation can be made with specially profiled discs, or by operating with the disc axes at an angle to one another. Factors such as residual or bending stresses cannot be considered.

A common issue across all laboratory approaches to fatigue investigation is the need to define the 'failure' point of the specimen. Depending on the apparatus, this has ranged from detection of well-developed cracks and spalls using accelerometers to sense machine vibration, through to early detection approaches allowing repeatable detection of RCF cracks of only a few hundred microns in size (Garnham and Beynon, 1991). For small-scale machines, consideration must be given to how well the test represents a full-size rail–wheel contact. Hertzian stress analysis allows calculation of the load required to produce equal bulk contact pressure for small- and large-scale cases, but asperity (roughness) contacts and stresses cannot be scaled in this way and may be disproportionately large relative to the nominal contact area when conducting small-scale tests. Successful correlation of small-scale test



9.8 Twin-disc representation

results with models and field performance is still possible, provided these scaling issues are considered.

The first use of a twin-disc approach to investigate rolling contacts was by Way (1935). This was not conducted for a railway application, but it did reveal the crucial involvement of fluids in the generation of cracks for this contact configuration. The effects of changes between lubricated and unlubricated contact were also investigated, and this has continued to receive attention since it is an inevitable aspect of rail operation (Tyfour *et al.*, 1996; Fletcher and Beynon, 2000a). Way suggested that the lubricant acted as a hydraulic pressure transmitter in the manner of mechanism 2 in Fig. 9.5, although his work predated the identification of the other crack growth mechanisms now considered.

Several twin-disc test machines have been developed including that by Garnham and Beynon (1991), enhanced by Fletcher and Beynon (2000b), and the Amsler design, first reported by Amsler (1922) and used in many other tests, for example by Clayton and Su (1996). The Amsler machine has a simple design, although this can restrict the range of contact pressure and slip combinations which can be achieved without special test specimen manufacture. Alternative machines often standardise test specimen design, and use more complex control systems to achieve the required combination of pressure and slip at the contact.

A compromise between small-scale laboratory work and full-scale vehicle tests is found in a small number of laboratory test machines which work with full size rails and wheels. A machine of this type was developed by British Rail Research in the UK, capable of replicating the rolling-sliding contact of a wheel on a rail, including the angle of attack of the wheel, (see for example, McEwen and Marvey, 1985). This machine ran a wheel on a short straight piece of rail and lifted the wheel off for its return movement, so unidirectional traffic was simulated, although this design meant that only slow wheel speeds were possible. The alternative approach of using a fully rolling design with a large 'rail' roller (effectively a large-scale twin-disc machine, although retaining wheel and rail profile rather than using cylindrical specimens) has been adopted in machines developed at Deutsch Bahn (Ullrich and Luke, 2001; Ullrich *et al.*, 2005) and Voest Alpine (Eadie *et al.*, 2006). The approach taken by Nishida *et al.* (1985) differed in using a 6 m diameter circular track on which wheels ran upright. Several mid-scale simulators using scaled down wheels have also been produced (Kalousek *et al.*, 1982; Iwnicki and Wickens, 1998; Chen *et al.*, 2006), although these have in many cases been focused on rail–wheel dynamics rather than wear and fatigue.

9.3.2 Field and full-scale test track investigations

Although full-scale or field tests are limited by uncertainty about the stress which will exist between the rail and wheel, investigation using full-scale test

tracks or lines in commercial service does provide real-world test conditions which are very difficult to replicate in the laboratory. Some of the earliest work undertaken in this way was in measurement of the contact patch which develops between the rail and wheel, by placing carbon paper on the rail and observing the print produced by the wheel (Andrews, 1958). Photoelastic techniques have also been used to accomplish this task, well before finite element (FE) or other numerical techniques were available (Hetènyi and McDonald, 1958; Haines and Ollerton, 1963). Recent similar work has used ultrasonics (Marshall *et al.*, 2003).

The Facility for Accelerated Service Testing (FAST) developed by the Association of American Railroads (AAR) uses a large closed loop of track with close to continuous running, able to expose rails and wheels to high volumes of repeatable and well known traffic, producing results such as those by Steele (1991). Many tests of track behaviour in commercial service have also been undertaken, such as those by Grohmann *et al.* (2002), Cantera (1993) Olofsson and Nilsson (2002) and Olofsson and Telliskivi (2003). This type of experimental work has been particularly successful when a change is made to the system (e.g. addition of lubrication on curves) producing a measurable effect on wear or crack growth, but it is less useful in understanding the detailed mechanisms of rail degradation. A recent full scale test track investigation has, however, been able to demonstrate the entry of fluids into cracks (Fletcher *et al.*, 2006), which has been highlighted by laboratory work and modelling as a driver of RCF crack growth.

In addition to running specific experiments, many railways collect data on traffic and correlate this with maintenance requirements, producing a direct link between traffic and system failure without the need to understand failure mechanisms. This can be very successful for a well understood and 'steady-state' system, but it cannot provide predictions of how the system will respond to changes such as new rail steel types or a different traffic mix. An important use for the data collected is therefore in validating predictive models for current conditions, thereby increasing confidence in their output for new conditions.

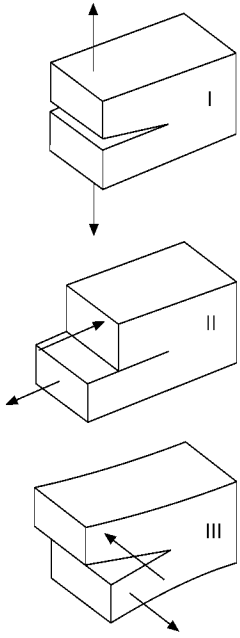
9.4 Calculating crack growth rate

Calculation of crack growth rates for RCF cracks remains a highly complex field, and there is currently no simple way to view a cracked rail and to know how fast the cracks it contains will grow. Experience has shown that cracks of a certain visible size can be considered 'mild' while longer cracks are classified as 'severe', and in the UK there are guidelines for the actions which must be taken if cracks of these sizes are found (Railtrack, 2001). Such guidelines are usually based on previous experience of similar cracks in well known rail steels. The models described below provide fracture

mechanics methods for calculating stress intensity factor and hence crack growth rate and, as described above, they offer a method by which changes to a rail system can be examined, which could not be predicted through guidelines based on previous system performance.

9.4.1 Fracture mechanics

For medium length cracks which have moved away from the severe plastic deformation of material at the rail surface, linear elastic fracture mechanics (LEFM) can be applied to calculate crack growth rates. Research has been undertaken into the application of LEFM to shaken-down materials (Huang and Stein, 1995, 1996) so, even though the near-surface material has been previously plastically deformed, it can be argued that LEFM can still be applied, just as it is to other components whose materials undergo plastic deformation during production, but whose later behaviour is elastic. To describe the way a crack is advancing, fracture mechanics refers to different 'modes' to represent cracks being opened, sheared or torn, as illustrated in Fig. 9.9. A crack in a rail will undergo different combinations of these failure modes at different stages of its growth, although most modelling has been confined to cases combining mode I and II failure.



9.9 Fracture mechanics Mode I (opening), II (shearing) and III (tearing).

Fracture mechanics is based around the calculation of a 'stress intensity factor' (SIF) which takes account of the stress applied to the body considered, its external shape (particularly important when the crack is approaching the boundaries of the body) and the crack shape. Applied stress can produce a combination of SIFs for modes I to III, and these must be combined and converted to a crack growth rate using a crack growth law. This law is usually specific to a material, and is developed through laboratory testing in which loads are applied to simple specimens for which the SIF can be easily calculated. The assumption is made that growth rate will be the same for similar cracks in the same material when they have the same SIF, even though loading conditions and external geometry of the bodies can differ. Further information on fracture mechanics can be found in standard reference books on the subject (Anderson, 1994; Suresh, 1998).

A range of methods has been developed to calculate SIFs for inclined cracks beneath a rolling contact load, including those by Kaneta, Murakami and co-workers (Kaneta *at al.*, 1985, 1986, 1988; Kulkarni *et al.*, 1991; Bower (1988), Fletcher and Beynon, (1999b,c), Fletcher and Kapoor (2006b) and Bogdański *et al.* (Bogdański and Brown, 1997, 2002; Bogdański *et al.*, 1997). Each method is relevant to a specific mechanism of crack growth, and makes certain assumptions about factors such as crack shape and crack inclination angle below the surface. Some of these are discussed in detail below.

9.4.2 Stress intensity factor calculation methods

Body force method

One of the earliest and most widely respected methods of calculating SIFs for RCF cracks is the 'Body force method' developed by Kaneta and Murakami *et al.* (Kaneta *et al.*, 1985, 1986; Kaneta and Murakami, 1991). This was developed for semi-circular and semi-elliptical cracks, and later applied in the modelling of more realistic crack shapes seen in twin-disc rolling contact tests and in rails (Murakami *et al.*, 1994; Kaneta *et al.*, 1998).

The work by Kaneta and Murakami *et al.* considered fluid in the cracks as a transmitter of contact pressure to the crack faces by a hydraulic mechanism. In this mechanism, the areas of the crack in which the crack faces were open (apart) and closed (crack faces in contact) were predicted for every position of the wheel passing the crack, and the pressure acting on the faces modified accordingly. The lubrication of the crack faces, with consequent increase in their mode II shear stress intensity factors, was also considered.

The pressure transmitted by fluid in the crack was assumed to decrease linearly, dropping from the rail–wheel contact pressure at the crack mouth to zero at the crack tip. In a static system such a variation would not exist, but it was chosen to represent the dynamic nature of a fluid briefly pressurised

at the crack mouth, and then allowed to depressurise after the wheel has passed. The sliding of the crack faces was modified by the friction between the faces in regions where the crack was closed. Where the applied shear stress was insufficient to overcome this friction force, that region of the crack was assumed to be locked. These pressure and friction distributions were also applied in the later work of Fletcher and Beynon (1999b,c). A more recent application of the body force method is in examination of the combined effect of crack face fluid pressure and frictional heating on a crack under rolling-sliding contact, conducted by Goshima *et al.* (2002). It is found that, with fluid pressure present, the crack growth is dominated by tensile mode growth (mode I), with limited thermal and frictional effects. Without pressurisation the growth is by a shear mechanism (mode II), and thermal and frictional effects are larger. The change of crack shape at the mouth and deeper below the surface as the crack grows is also considered.

Finite element and dislocation based methods

The study undertaken by Bower (1988) is one of the most comprehensive examinations of the mechanisms of RCF, and has been discussed above. Bower's work followed work by Hearle and Johnson (1985) and Sheppard *et al.* (1985) who calculated stress intensities for subsurface cracks. It used dislocation distribution methods similar to those of Keer and Bryant (1983). Bower presented results for a crack of length to half width ratio of 0.5, equivalent to around a 3 mm crack in a rail or a 0.1 mm crack in a twin-disc test. The crack was taken to lie at 25° below the contact surface and have a crack face friction coefficient (i.e. friction coefficient inside the crack) of 0.1. Considerable care was necessary in stepping the wheel position forward during the modelling to ensure that the positions of crack face contact were sensibly predicted. More recently the work of Bogdański *et al.* (Bogdański *et al.*, 1997; Bogdański and Brown, 1997; Bogdański, 2002), using FE techniques has been significant in increasing the understanding of mixed mode SIFs generated during rolling-sliding contact such as that between a rail and wheel. Most importantly, Bogdański *et al.* have moved from two-dimensional modelling to full three-dimensional cases, with the ability to predict the full cycle of SIF to which the crack will be subjected during the passage of a contact. This includes modes I, II and III.

Extensive FE models of the rail–wheel contact have also been produced by Chalmers University in Sweden (Ringsberg 2001; Ringsberg and Josefson, 2001; Ringsberg and Lindback, 2003). These models have concentrated on crack initiation and the very early phases of propagation, using such approaches as cycle by cycle modelling of ratcheting strain accumulation, low-cycle fatigue and a critical plane approach in which a search algorithm is used to find the most damaging crack plane, using the stress and strain tensors determined from FE analyses.

Nowell and Hills (1987) and Hills and Nowell (1989) presented methods for the calculation of SIFs for both open and closed surface breaking cracks using a technique of edge dislocation distribution along the line of a crack within a body, a technique that is similar to that used by Bower. More recently Seed (2001) has obtained modes I and II SIFs for an inclined crack under various contact loadings.

Boundary element modelling

Boundary element analysis has proved to be particularly well suited to examining cracks in rails. Its strength lies in its mathematical formulation which enables the high stress gradient ahead of the crack to be modelled without the requirement for very fine (and hence time-consuming) meshing in the region ahead of the crack (Goldberg, 1998). This is possible because the problem is formulated in terms of the integration of stress across the boundary surface of the rail rather than through its volume, as would be the case for FE modelling. Successful modelling of cracks under contact loading using this method has been carried out in the UK (Kapoor *et al.*, 2004; Mellings *et al.*, 2005) and Japan (Akama and Moni, 2002). Modelling of single and multiple cracks in a rail in bending (Fletcher *et al.*, 2004) has also been carried out using boundary element software developed at Cornell University in the USA.

Although very comprehensive, FE and to a lesser extent boundary element modelling has the major drawback that its complexity makes model generation very time-consuming, and solving the models is slower than alternative methods such as using the Green's functions described in the next section.

Influence and Green's functions

Many calculations of SIF are specific to a single loading condition (i.e. rail surface contact pressure and friction) and are difficult to generalise to alternative loadings, but this difficulty can be overcome using a Green's function or influence function approach. Calculations of SIF are conducted for point loads at a range of positions on the contact surface or crack face. Then, if the full applied contact load can be represented as the summation of an array of point loads, an integration process can then be used to find the total SIF for a contact load or even for the combination of contact, residual and bending loads. For example, if boundary element analysis provides the SIF produced by individual point forces at positions along a crack face, the Green's functions provide a means to treat an arbitrary crack face load as a series of point forces, and to sum up the combined SIF for the entire arbitrary load.

Based on Green's functions developed by Rooke *et al.* (1992) for calculation

of SIF for an inclined surface breaking crack, methods have been developed to consider cracks containing pressurised fluid (hydraulic crack growth, Fig. 9.5) and cases in which contact and shear of the crack faces is considered (shear mechanism) (Fletcher and Beynon, 1999b,c). Since use of this method requires only the summation of existing results rather than complete SIF calculation it is typically very fast to run, allowing a large range of contact conditions to be investigated. However, it has the disadvantage that it may only be applied to cases for which the Green's functions have been previously calculated.

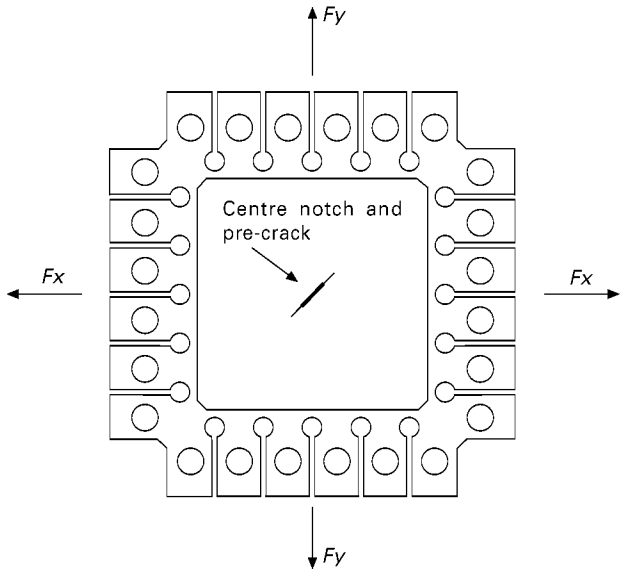
The Fletcher and Beynon (1999b,c) method of SIF calculation is implemented for the contact stresses due to wheel loads, residual, continuously welded and rail bending stresses. Results in good agreement with those from the body force model developed by Kaneta *et al.* (1985) have been produced. Both models are based on very similar assumptions about the stresses driving crack growth and the ways in which lubricants affect crack growth, crack closure and sliding.

9.4.3 Crack growth laws

Crack growth laws developed for simple push-pull cyclic loading conditions have been found inadequate to represent the growth of a crack in a rail because of the simultaneous presence of mode I, II and III crack growth driving stresses. Crack growth laws have therefore been developed using loading conditions closer to those experienced in the rail, helping to take account of the 'non-proportional' changes of mode I (opening), II (shearing) and III (tearing) loading which take place during the passage of the wheel (Bower, 1988). The variation in relative proportions of mode I, II and III SIFs means that at some stages in the passage of a wheel the mode I crack opening failure mechanism will dominate, whereas mode II or III failure will dominate at other stages during the passage of the wheel.

Bold *et al.* (1991) developed crack growth laws for RCF cracks after extensive testing using biaxial fatigue specimens of normal grade rail steel cut from the rail web (Fig. 9.10). Equations (9.1) and (9.2) summarise the work, showing stress intensity factor ranges in mode I (opening) and II (shearing) measured in MPa $m^{0.5}$ and crack growth rate da/dN in nm/cycle. ΔK_{eq} is an equivalent stress intensity factor, used to combine the mode I and II factors, even though the peaks in the SIFs for different modes of failure occur out of phase with one another during the passage of the contact. Crack length is a and ΔK_{th} is the threshold stress intensity factor range.

$$\Delta K_{eq} = \sqrt{\Delta K_I^2 + \left[\left(\frac{614}{507}\right)\Delta K_{II}^{3.21}\right]^{\frac{2}{3.74}}} \qquad [9.1]$$



9.10 Biaxial specimen.

$$\frac{da}{dN} = 0.000507\ (\Delta K_{eq}^{3.74} - \Delta K_{th}^{3.74}) \qquad [9.2]$$

The stress cycles applied during the development of this law were based on predictions by Bower (1988) of the stresses to which inclined RCF cracks in rail are subjected. An alternative crack growth law for rail steel is available from the work of Akama and Mori (2002) which is also based on biaxial fatigue testing of rail steel samples.

9.5 Crack branching predictions

Real cracks in rails have a complex three-dimensional geometry, and analysis to predict branching and further growth must be carried out by finite or boundary element modelling of specific crack geometries. Despite this, some understanding of crack branching can be gained through calculations based on the mode I (opening) and mode II (shearing) SIFs for a simplified crack, and the discussion here relates only to a simplified crack in a two-dimensional approximation to the rail–wheel contact.

To examine crack branching in rails it is important to consider both mode I and II loading, since these almost always occur together, although non-proportionally, for inclined cracks. To do this Otsuka *et al.* (1975, 1981, 1984a,b) used the theories of maximum shear stress and maximum tangential stress, Eqs (9.3) and (9.4), originally developed by Erdogan and Sih (1963),

making the assumption that crack growth should occur by whichever mode first gave a resultant SIF exceeding threshold for the material. The directions of propagation are θ_σ and θ_τ for tensile and shear mode growth respectively. K_σ and K_τ are the tensile and shear mode equivalent stress intensity factors respectively.

$$K_\sigma = \cos\frac{\theta_\sigma}{2}\left[K_\mathrm{I}\cos^2\frac{\theta_\sigma}{2} - \frac{3}{2}K_\mathrm{II}\sin\theta_\sigma\right] \quad [9.3]$$

$$K_\tau = \frac{1}{2}\cos\frac{\theta_\tau}{2}\left[K_\mathrm{I}\sin\theta_\tau + K_\mathrm{II}(3\cos\theta_\tau - 1)\right] \quad [9.4]$$

Kaneta *et al.* (1985) applied these theories for cases of non-proportional loading usually present in RCF situations, i.e. the mode I and II stress intensities varying out of proportion to one another during the passage of the contact. Kaneta *et al.* showed that the value of θ_σ which corresponds to the maximum value of K_σ is given by one of the roots of the derivative of K_σ with respect to θ_σ, given by Equation (9.5).

$$\tan\frac{\theta_\sigma}{2} = \frac{(1 \pm \sqrt{1 + 8\gamma^2})}{4\gamma},\ \gamma = \frac{K_\mathrm{II}}{K_\mathrm{I}} \quad [9.5]$$

Similarly the value of θ_τ which gives the maximum value of K_τ is found using one of the roots of Eq. (9.6).

$$\tan^3\frac{\theta_\tau}{2} - \frac{1}{\gamma}\tan^2\frac{\theta_\tau}{2} - \frac{7}{2}\tan\frac{\theta_\tau}{2} + \frac{1}{2\gamma} = 0,\ \gamma = \frac{K_\mathrm{II}}{K_\mathrm{I}} \quad [9.6]$$

The values of K_I and K_II at instants throughout the passage of the wheel are used to calculate crack growth direction from Eqs (9.5) and (9.6) evaluated at each instant, producing a range of resultant K_σ, K_τ, θ_σ and θ_τ values. Kaneta *et al.* (1985) suggest restrictions be placed on the values of K_I and K_II to ensure the model is physically realistic. For example, if the resultant value K_σ is found to be negative it should be set to zero since the faces of the crack cannot inter-penetrate one another. The question of which of these resultant values actually governs crack growth direction is not resolved by the method, but it is assumed that the crack will propagate in the direction corresponding to the maximum resultant K_σ or K_τ. Results of modelling based on this method of branch prediction and also taking into account residual stress have shown that crack branching can be particularly sensitive to residual stress acting in addition to contact stress (Fletcher and Kapoor, 2006b).

9.6 Rail wear

9.6.1 Overview

Wear of rails leads to problems including a widening of the track gauge, loss of rail profile leading to high rail–wheel forces and poor dynamics and, in extreme cases, a loss of rail cross-section so significant that the rigidity and load-bearing ability of the rails deteriorates. Measurement of wear can be conducted using a cross-level gauge able to measure the separation of the rails and their elevation relative to one another. Sophisticated vehicle-mounted laser based systems can provide full details of the rail profile and gauge. The MiniProf® profile measuring device can also be used to provide a rail cross-sectional profile by moving a small pivoted stylus around the rail cross-section by hand.

Having said that wear leads to problems, a small degree of wear can be beneficial in removing damaged material from the surface of the rail. For newly installed rails, wear can improve the conformity of contact between rail and wheel, thereby reducing rail stress and subsequent damage. This increase in conformity can happen faster with less wear-resistant (softer) rails than with harder more wear-resistant grades. Grinding can be used to manage and maintain rail profile, even for newly installed rails. However, excessive wear or grinding will limit rail life and require costly maintenance or rail replacement.

Rail wear is often highest where the wheel flange or flange root comes into contact with the rail gauge face. On lines with fast (typically passenger) traffic, the outer (high) rail in curves can wear excessively since this is supplying cornering forces. For slower (typically freight) traffic, excessive wear is more often found on the inner (low) rail of the curve since the train is effectively being dragged round the curve on this rail rather than being forced to corner by the high rail.

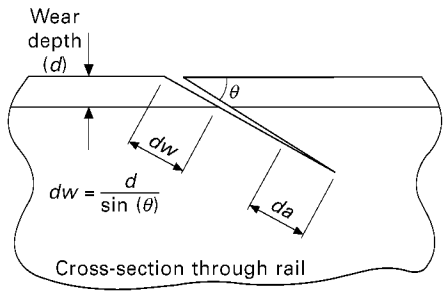
Because of the geometry of the rail and wheel, a contact between the wheel flange to rail gauge face is a predominantly sliding as opposed to a predominantly rolling contact when the rail and wheel touch at the rail crown. High amounts of sliding typically produce high wear rates, so the high degree of sliding which often occurs in curves usually leads to high wear at these locations. Lubrication is often applied to the rail to reduce wear in these locations, either using track-based lubricators which release lubricant when triggered by passing wheels, or using train-based systems. Migration of the lubricant from the rail gauge face to the railhead is problematic since braking and acceleration depend on rail–wheel friction. Friction modifiers can also be applied, being capable of reducing high friction without producing the very low friction characteristic of grease-or oil-based lubricants. Solid lubricants such as graphite or molybdenum disulphide can also be employed in cases where fire resistance or avoidance of lubricant migration are important

factors. In some cases, liquid lubricants can accelerate crack growth if they enter cracks (Fletcher and Beynon, 1999a), just as water does, whereas dry solid lubricants do not lead to this problem.

9.6.2 Rail wear and wear–fatigue interaction

Wear mechanisms are covered elsewhere in the book, and this information will not be repeated here. Rail wear depends on the location of the point of contact between wheel and rail (i.e. gauge corner or crown of the rail; tread or flange of the wheel) which will influence the degree of sliding on the contact. Large differences in contact position can be seen between different fleets of vehicles due to (for example) differences in suspension characteristics, but also between different wheels on the same vehicle. This means that wear is spread across the width of the rail and wheel, and affects the whole profile. Profile change as a result of wear is in fact a major cause of the variation of the point of contact, that is, wear and contact location form a feedback system. The spread of wear makes it difficult to relate profile change accurately to rail-wheel contact since the latter is not constant (and generally not known precisely). Railhead profile wear has been studied for two different rail steels by Olofsson and Telliskivi (2003), and in many other studies worldwide.

Wear of the surface of a rail can be beneficial in removing small cracks, and by slowing the effective crack growth rate for larger cracks. Figure 9.11 shows a two-dimensional representation of a crack growing into a railhead, for which the predicted crack growth at the tip (i.e. the crack tip advance rate) during the passage of a wheel is labelled *da*. However, during the passage of the wheel, the rail surface is wearing away and this produces a reduction in crack length of *dw* at the crack mouth (i.e. the crack mouth truncation rate). The net crack growth is *da* – *dw*, as shown in the figure. The relationship

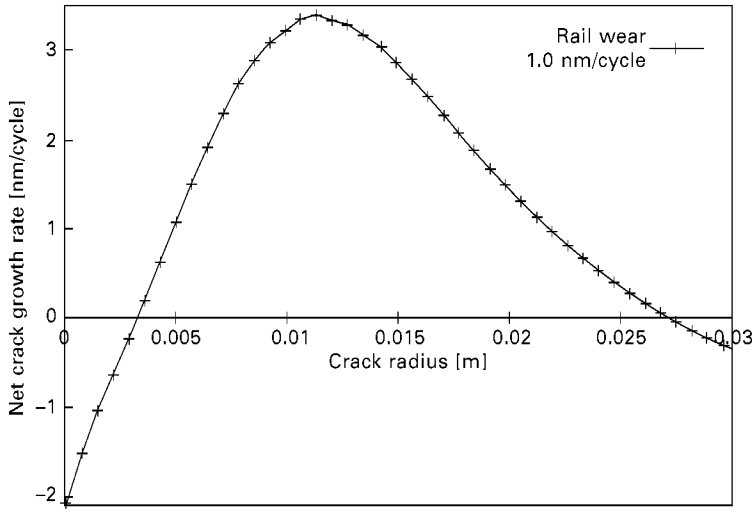


9.11 Schematic representation of a rail cross-section illustrating wear–fatigue interaction during phase II crack growth. Cracks can only grow if they are driven fast enough to 'keep ahead' of rail surface wear.

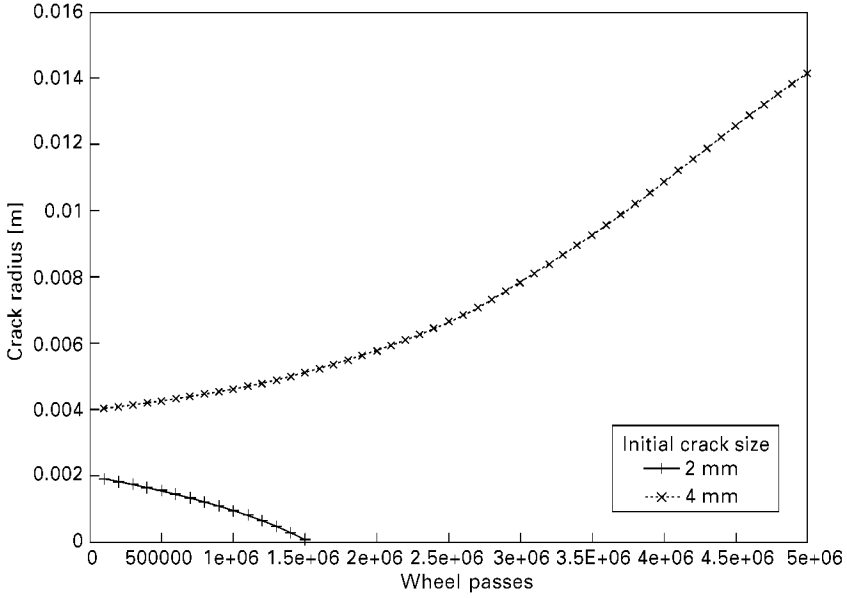
between vertical wear rate and crack length reduction depends on the angle of the crack below the rail surface. Cracks very close to the surface are often at angles below 30°, and this amplifies the vertical wear rate to give a large reduction in crack length with each wheel passage.

If wear rate is sufficient, the net crack growth rate can be reduced to zero, indicating no progression of existing cracks, or even to a negative rate, indicating that existing cracks are being worn away. Figure 9.12 shows a typical plot of net crack growth rate against crack size, based on the shear (mode II) crack growth of a semi-circular inclined crack at a specific contact pressure and friction level. Using this example, for a wear rate of 1 nm per wheel pass it can be seen that cracks below approximately 3.3 mm and above approximately 27 mm in radius have negative or zero crack growth rates, while cracks between these sizes would be predicted to undergo sustained growth. For the crack-growth rate and wear rate in Fig. 9.12, Fig 9.13 shows the evolution of crack size over a number of wheel passes for two initial crack sizes. A 2 mm initial radius is insufficient for sustained crack growth, and its size is reduced by passing wheels (negative net growth rate shown in Fig. 9.12). The 4 mm initial crack size has a positive net growth rate, and therefore increases in size with each wheel pass.

Under the contact conditions for which Fig. 9.12 is calculated, small cracks below 3.3 mm are not predicted to grow, but are in an unstable situation,



9.12 Net crack growth rate, for the case of a 600 mm crown radius rail with 75 kN vertical load (approximately 900 MPa contact pressure), at a friction coefficient of 0.20. Surface wear rates 1.0 nm/cycle. Contact semi-axis lengths 5.6 mm × 7 mm. Negative regions of the curve indicate wear dominance and crack size reduction.



9.13 Crack growth predictions based on the crack growth rate data in Fig. 9.12 for initial crack sizes of 2 and 4 mm radius. The 2 mm initial radius is insufficient for sustained crack growth, and its size is reduced by passing wheels (negative net growth rate shown in Fig. 9.12. The 4 mm initial crack size has a positive net growth rate, and therefore increases in size with each wheel pass.

since any deviation in contact pressure or traction may be enough to increase their size sufficiently for sustained growth to begin. Larger cracks of 27 mm or above with a zero net growth rate are in a stable equilibrium because of the negative slope of the crack growth rate curve; however, their size at this stage is typically too large for the rail to remain in service. Additional factors such as loss of rail strength through reduction of cross-section and increased growth through rail bending (phase III in Fig. 9.1) mean that large cracks of these sizes could not safely be left in track even if they are predicted to have a zero net crack growth rate by the phase II crack growth models. It should be noted that the crack sizes stated above are examples for a specific contact condition, and will vary depending on these conditions.

Use of grinding as a rail maintenance technique is able to exploit the effect of wear on crack growth. Grinding can be used as an artificial wear process, removing cracks and reducing the size of remaining cracks to below the minimum size at which phase II growth is sustainable (Fig. 9.12). Grinding can also be used to re-profile the rail, giving an improved rail–wheel contact geometry and reduced stress. In addition to removal of very small cracks, the driving stress for any remaining (or later new) defects is therefore reduced.

9.7 References

Akama, M. and Mori, T. (2002), Boundary element analysis of surface initiated rolling contact fatigue cracks in wheel/rail contact systems, *Wear*, **253**(1–2), 35–41.

Amsler, A. J. (1922), Abnützungsmaschine für metalle (wear machine for metals), *Z. VDI*, **66**(15), 377–8.

Anderson T. L. (1994), *Fracture Mechanics: Fundamentals and Applications*, (2nd edn), CRC, Boca Raton, FL, USA.

Andrews, H. I. (1958), The contact between a locomotive driving wheel and the rail, *Wear*, **59**(2), 458–484.

Bogdański, S. (2002), A rolling contact fatigue crack driven by squeeze fluid film, *Fatigue and Fracture of Engineering Materials and Structures*, **25**, 1061–71.

Bogdański, S. and Brown, M. W. (1997), Modelling of surface crack growth in ehd contact, *Proceedings International Conference on Engineering Against Fatigue*, University of Sheffield, Sheffield, UK, 17–21 March.

Bogdański, S. and Brown, M. W. (2002), Modelling the three-dimensional behaviour of shallow rolling contact fatigue cracks in rails, *Wear*, **253**, 17–25.

Bogdański, S., Stupnicki, J., Brown, M.W. and Cannon, D.F. (1997), A two dimensional analysis of mixed-mode rolling contact fatigue crack growth in rails, *Proceedings 5th International Conference on Biaxial/Multiaxial Fatigue and Fracture*, Krakow, Poland, 8–12 September, **2**, 189–206.

Bold, P. E., Brown, M. W. and Allen, R. J. (1991), Shear mode crack growth and rolling contact fatigue, *Wear*, **144**, 307–17.

Bower, A. F. (1988), The influence of crack face friction and trapped fluid on rolling contact fatigue cracks, *ASME Journal of Tribology*, **110**, 704–11.

Burris, I. and Kudish, K. W. (2004), Modelling of surface and subsurface crack behaviour under contact load in the presence of lubricant, *International Journal of Fracture*, **125**, 125–47.

Cannon, D. F. (2003), *An International Cross Reference of Rail Defects*, report commissioned by the Steering Group of UIC/WEC Joint Research Project 1 – Rail Defect Management, UIC, Paris, France.

Cannon D. F., Edel, K.-O., Grassie, S. L. and Sawley K. (2003), Rail defects: an overview, *Fatigue and Fracture of Engineering Materials and Structures*, **26**(10), 865–86.

Cantera, F. (1993), Investigation of wheel flange wear in Santander FEVE rail – a case study, *Wear*, **162**, 975–9.

Clayton, P. and Su, X. (1996), Surface initiated fatigue of pearlitic and bainitic steels under water lubricated rolling/sliding contacts, Wear, **200**, 63–73.

Chen, H., Ban, T., Ishida, M. and T Nakahara, T. (2006), Effect of water temperature on the adhesion between rail and wheel, *Proceedings of the IMechE, Part J: Journal of Engineering Tribology*, **220**(7), 571–79.

Dowson D., Higginson G. R. and Whitaker A. V. (1963), Stress distribution in lubricated rolling contacts, *Proceedings IMechE Symposium on Fatigue in Rolling Contact*, London, UK, paper 6.66, 66–75.

Dukkipati, R. V. and Dong, R. (1999), Idealized steady state interaction between railway vehicle and track, *Proceedings of the IMechE, Part F; Journal of Rail and Rapid Transit*, **213**, 1529.

Erdogan, F. and Sih, G. C. (1963), On the crack extension in plates under plane loading and transverse shear, *ASME Journal of Basic Engineering*, **85**, 519–25.

Eadie, D., Elvidge, D., Oldknow, K., Stock, R., Pointner, P., Kalousek, J. and Klauser, P.

(2006) The effects of top of rail friction modifier on wear and rolling contact fatigue: full scale rail-wheel test rig evaluation, analysis and modelling, *Proceedings 7th International Conference on Contact Mechanics and Wear of Rail/Wheel Systems*, Brisbane, QLD, 24–27 October, 411–19.

Fletcher, D. I. and Beynon, J. H. (1999a), The influence of lubricant type on rolling contact fatigue of pearlitic rail steel, in Dowson, D., Priest M., Taylor, A., Ehret, P, Childs, T. H. C., Dalmaz, G., Verthier, Y., Flamand, L., Georges J. M. and Lubrechgt, A. A. (eds), *Lubrication at the Frontier, Proceedings 25th Leeds-Lyon Conference on Tribology, Paris 1998*, Elsevier, Amsterdam, the Netherlands, 299–310.

Fletcher, D. I. and Beynon, J. H. (2000a), The effect of intermittent lubrication on the fatigue life of pearlitic rail steel in rolling sliding contact, *Proceedings of the IMechE, Part F: Journal of Rail and Rapid Gransit*, **214**, 145–58.

Fletcher, D. I., and Beynon, J. H. (2000b) Development of a machine for closely controlled rolling contact fatigue and wear testing, *ASME Journal of Testing and Evaluation*, **28**(4), 267–75.

Fletcher, D. I. and Beynon, J. H. (1999b), A simple method of stress intensity factor calculation for inclined fluid-filled surface-breaking cracks under contact loading, *Proceedings of the IMechE, Part J: Journal of Engineering Tribology*, **213**, 299–304.

Fletcher, D. I. and Beynon, J. H. (1990c), A simple method of stress intensity factor calculation for inclined surface breaking cracks with crack face friction under contact loading, *Proceedings of the IMechE, Part J: Journal of Engineering Tribology*, **213**, 481–6.

Fletcher, D. I. and Kapoor, A. (2006a), A rapid method of stress intensity factor calculation for semi-elliptical surface breaking cracks under three-dimensional contact loading, *Proceedings of the IMechE, Part F: Journal of Rail and Rapid Transit*, **220**, 219–34.

Fletcher, D. I. and Kapoor, A. (2006b), Post Hatfield rolling contact fatigue – The effect of residual stress on contact stress driven crack growth in rail: Combination of bending and contact stresses, Nov 2006, Office of Rail Regulation, London, UK, http://www.rail-reg.gov.uk/server/show/nav.1184.

Fletcher, D. I., Hyde, P. and Kapoor, A. (2004), Growth of multiple rolling contact fatigue cracks driven by rail bending modelled using a boundary element technique, *Proceedings of the IMechE, Part F: Journal of Rail and Rapid Transit*, **218**(3), 243–53.

Fletcher, D. I., Hyde, P. and Kapoor, A. (2006), Modelling and full scale trials to investigate fluid pressurisation of rolling contact fatigue cracks, *Proceedings 6th International Conference on Contact Mechanics and Wear of Rail/Wheel Systems*, Brisbane, QLD, Australia, September 24–26 463–71.

Frolish M. F., Fletcher D. I. and Beynon J. H. (2002), A quantitative model for predicting the morphology of surface initiated rolling contact fatigue cracks in back-up roll steels, *Fatigue and Fracture Of Engineering Materials and Structures*, **25**(11), 1073–86.

Garnham, J. E. and Beynon, J. H. (1991), The early detection of rolling-sliding contact fatigue cracks, *Wear*, **144**, 103–16.

Goldberg, M. (ed.) (1998), *Boundary Integral Methods Numerical and Mathematical Aspects. Computational Engineering*, Vol 1, WIT Press, Southampton, UK.

Goshima, T., Ishihara, S. and Shimizu, M. (2002), The influence of crack-face fluid pressure on the fatigue crack propagation due to rolling contact with frictional heat, *Journal of Thermal Stresses*, **25**(4), 373–88.

Grohmann, H. D., Hempalmann, K. and Gross-Thebing, A. (2002), A new type of RCF, experimental investigations and theoretical modelling, *Wear*, **253**(1–2), 67–74.

Haines, J. and Ollerton, E. (1963), Contact stress distribution on elliptical contact surfaces subjected to radial and tangential forces, *Proc. Institution of Mechanical Engineers*, **177**(4), 95–114.

Hay, W. W. (1982), *Railroad engineering*, Wiley.

Hearle, A. D. and Johnson, K. L. (1985), Mode II stress intensities for a crack parallel to the surface of an elastic half-space subjected to a moving point load, *Journal of the Mechanics of Physics of Solids*, **33**, 66–81.

Hetènyi, M. and McDonald, P. H. Jr (1958), Contact stresses under combined pressure and twist, *Journal of Applied Mechanics*, **25**, 396–401.

Hills, D. A. and Nowell, D. (1989), Stress intensity calibrations for closed cracks, *Journal of Strain Analysis*, **24**(1), 37–43.

Hooper, N. (2008), Reducing broken rail derailments in dark territory, *Parts 1 and 2 Interface Journal*, available at: http://www.interfacejournal.com/features/01-08/brokenrail1/1.html and http://www.interfacejournal.com/features/04-08/brokenrail2/1.html, accessed March 2009.

Hourlier, F., and Pineau, A. (1981), Fatigue crack propagation behaviour under complex mode loading, *Proceedings 5th International Conference on Fracture, Advances in Fracture Research*, **4**, 1841–49.

Huang, Y. J. and Stein, E. (1995), Prediction of the fatigue threshold for a cracked body using shakedown theory, *Fatigue and Fracture of Engineering Materials and Structures*, **18**(3), 363–70.

Huang, Y. J. and Stein, E. (1996), Shakedown of a cracked body consisting of kinematic hardening material, *Engineering Fracture Mechanics*, **54**(1), 107–12.

Iwnicki S. D., and Wickens A. H. (1998), Validation of a MATLAB railway vehicle simulation using a scale roller rig, *Vehicle System Dynamics*, **30**(3), 257–70.

Jeong, D. Y., Tang Y. H., Orringer, O. and Perlman, A. B. (1998), *Propagation Analysis of Transverse Defects Originating at the Lower Gage Corner of Rail*, Technical Report DOT/FRA/ORD-98/06, US Department of Transportation, Washington DC, USA.

Kahraman, T. L. and Krantz, A. (2004), An experimental investigation of the influence of the lubricant viscosity and additives on gear wear, *Tribology Transactions*, **47**, 138–148.

Kalousek, J., Rosval, G. and Ghonem, H. (1983), Lateral creepage and its effect on wear in rail wheel interface, in Kalousek, J., Dukkipati, R. V. and Gladwell, G. M. L. (eds), *Proceedings of the Conference on Contact Mechanics and Wear of Rail/wheel Systems*, Vancouver, BC, Canada, July 6–9, 1982, University of Waterloo Press, Waterloo, ON, Canada.

Kaneta, M. and Murakami, Y. (1991), Propagation of semi-elliptical surface cracks in lubricated rolling/sliding elliptical contacts, *ASME Journal of Tribology*, **113**, 270–75.

Kaneta, M., Yatsuzuka, H. and Murakami, Y. (1985), Mechanism of crack growth in lubricated rolling/sliding contact, *ASLE Transactions*, **28**(3), 407–14.

Kaneta, M., Suetsugu, M. and Murakami, Y. (1986), Mechanism of surface crack growth in lubricated rolling/sliding spherical contact, *ASME Journal of Applied Mechanics*, **53**, 354–60.

Kaneta, M., Matsuda, K., Murakami, K. and Nishikawa, H. (1998), A possible mechanism for rail dark spot defects, *Trans ASME, Journal of Tribology*, **120**, 304–9.

Kapoor, A. (1994), A re-evaluation of the life to rupture of ductile metals by cyclic

plastic strain, *Fatigue and Fracture of Engineering Materials and Structures*, **17**(2), 201–19.

Kapoor, A. and Fletcher, D. I. (2006a), Post Hatfield rolling contact fatigue – The effect of residual stress on contact stress driven crack growth in rail Part 2: Data, Nov 2006, Office of Rail Regulation, London, UK, see http://www.rail-reg.gov.uk/server/show/nav.1184, Acessed March 2009.

Kapoor, A. and Fletcher, D. I. (2006b), Post Hatfield rolling contact fatigue – The effect of residual stress on contact stress driven crack growth in rail. Literature Review, Nov 2006b, Office of Rail Regulation, London, UK, see http://www.rail-reg.gov.uk/server/show/nav.1184, accessed March, 2009.

Kapoor, A., Fletcher, D. I. and Smith, L. (2004), *Modelling 3D Cracks in Rail for the Whole Life Rail Model: Progress and Status Report*, Technical Report MEC/AK/RSSB/August04, University of Newcastle-upon-Tyne, Newcastle-upon-Tyne, USA.

Keer, L. M. and Bryant, M. D. (1983), A pitting model for rolling contact fatigue, *ASME Journal of Lubrication Technology*, **105**, 198–205.

Kulkarni, S., Hahn, G. T., Rubin, C. A. and Bhargava, V. (1991), Elasto-plastic finite element analysis of three-dimensional pure rolling contact above the shakedown limit, *Journal of Applied Mechanics,* **58**, 347–53.

Marshall, M. B., Lewis, R., Drinkwater, B. W., and R. S. and Dwyer-Joyce, R. S. (2003), An approach for contact stress mapping in machine joints and concentrated contacts, *Journal of Strain Analysis for Engineering Design*, **39**(4), 339–50.

McEwen, I. J. and Harvey, R. F. (1985), Full-scale wheel-on-rail testing: comparisons with service wear and a developing theoretical predictive model, *Lubrication Engineering*, **41**(2), 80–88.

Mellings, S., Baynham, J. and Adey, R. A. (2005), Automatic crack growth prediction in rails with BEM, *Engineering Fracture Mechanics*, **72**(2), 309–18.

Murakami, Y., Sakae, C. and Ichimaru, K. (1994), Three-dimensional fracture mechanics analysis of pit formation mechanism under lubricated rolling-sliding contact loading. *STLE Tribology Transactions*, **37**(3), 445–51.

Nishida, S., Sugino, K., Vrashima, C. and Masumoto, H. (1985), Study on contact rolling fatigue of rails. 1. Development of high speed rail testing machine, *Bulletin of the JSME*, **28**(243), 1814–18.

Nowell, D. and Hills, D. A. (1984), Open cracks at or near free edges, *Journal of Strain Analysis*, **22**(3), 177–85.

Ollerton, E. and Morey, J. W. W. (1963), Fatigue strength of rail steel in rolling contact, *Proceedings IMechE Symposium on Fatigue in Rolling Contact*, London, UK, paper 2, 11.

Olofsson, U. and Nilsson, R. (2002), Surface cracks and wear of rail: a full-scale test on commuter train track, *Proceedings of the IMechE, Part F: Journal Of Rail and Rapid Transit*, **216**(4), 249–64.

Olofsson U. and Telliskivi T. (2003), Wear, plastic deformation and friction of two rail steels: a full-scale test and a laboratory study, *Wear*, **254**(1–2), 80–93.

Otsuka, A., Mori, K., Oshima, T. and Tsuyama, S. (1981), Mode II fatigue crack propagation in aluminium alloys and mild steel, in Francais, D. (ed.) *Advances in Fracture Research*; *Proceedings 5th International Conference on Fracture*, Pergamon, Oxford, UK, **4**, 1851–9.

Otsuka, A., Mori, K. and Miyata, T. (1975), The condition of fatigue crack growth in mixed mode condition, *Engineering Fracture Mechanics*, **7**, 429–39.

Otsuka, A., Tohgo, K., Kiba, T. and Yamada, S. (1984a), Mode II fatigue crack growth

characteristics and mechanisms I: aluminium alloy 7N01–T4 weldments under mode II loading, in *Advances in Fracture Research: Proceedings 6th International Conference on Fracture*, Pergamon, Oxford, UK, **4**, 1671 Valluni, S. R. *et al.* (eds),

Otsuka, A., Mori, K. and Tohgo, K. (1984b) Mode II fatigue crack growth in aluminium alloy, Tanaka, T., Jonom and Komai, K. (eds), *Current Research on Fatigue Cracks*, The Society of Materials Science, Japan, 127.

Railtrack plc. (2001), Rolling contact fatigue in rails; a guide to current practice, *RT/PWG/001*, Issue 1, February.

Rieger, Y. and Ding, N. F. (2003), Spalling formation mechanism for gears, *Wear*, **254**, 1307–17.

Ringsberg, J. W. (2001), Life prediction of rolling contact fatigue crack initiation, *International Journal of Fatigue*, **23**, 575–86.

Ringsberg, J. W. and Josefson, B. L. (2001), Finite element analyses of rolling contact fatigue crack initiation in railheads, *Proceedings of the IMechE, Part F: Journal of Rail and Rapid Transit*, **215**, 243–259.

Ringsberg, J. W. and Lindback, T. (2003), Rolling contact fatigue analysis of rails including numerical simulations of the rail manufacturing process and repeated wheel-rail contact loads, *International Journal of Fatigue*, **25**, 547–58.

Rooke, D. P., Rayaprolu, D. B. and Aliabadi, M. H. (1992), Crack line and edge Green's functions for stress intensity factors of inclined edge cracks, *Fatigue and Fracture of Engineering Materials and Structures*, **15**(5), 441–61.

Seed, G. M. (2001), Stress intensity factors for a surface-breaking crack in a half-plane subject to contact loading, *Fatigue and Fracture of Engineering Materials and Structures*, **24**, 69–79.

Sheppard, S., Barber, J. R. and Comninou, M. (1985), Short subsurface cracks under conditions of stick and slip caused by a moving compressive load, *ASME Journal of Applied Mechanics*, **52**(4), 811–17.

Sih G. C. and Tzou Dy (1985), Rail-end bolt hole fatigue crack in 3 dimensions, *Theoretical and Applied Fracture Mechanics*, **3**(2) 97–111.

Steele, R. K. (1991), The effect of metal removal, steel cleanliness and wheel load on the fatigue life of rail, *Wear*, **144**, 71–87.

Stupnicki, C. and Morbarigazzi, J. (2004), Measurement of microslips of subsurface fatigue crack faces due to rolling loads, *Journal of Strain Analysis for Engineering Design*, **39**, 161–71.

Suresh, S. (1998), *Fatigue of Materials* (2nd edn), Cambridge University Press, Cambridge, UK.

Szelazek, J. (1992), Ultrasonic measurement of thermal stresses in continuously welded rails, *NDT & E International*, **25**(2), 77–85.

Townsend, D. P., Zaretsky, E. V. and Scibbe, H. W. (1986), Lubricant and additive effects on spur gear fatigue life, *ASME Journal of Tribology*, **108**, 468–77.

Tyfour, W. R., Beynon, J. H. and Kapoor, A. (1996), Deterioration of rolling contact fatigue life of pearlitic rail steel due to dry-wet rolling–sliding line contact, *Wear*, **197**, 255–65.

Ullrich D. and Luke M. (2001), Simulating rolling-contact fatigue and wear on a wheel/rail simulation test rig, *Proceedings 5th World Congress on Railway Research*, Cologne, Germany, 25–29 November, on CD.

Ullrich D., Maedler K. and Zoll A. (2005), *Testing of wheel/rail technologies on test rigs and in operational trials*, RTR – Railway Technical Review 04/2005, 29–33.

Walker, P. S., Dowson, D., Longfield M. D. and Wright, V. (1989), Rheological behaviour of human joints, *Rheologica Acta*, **8**(2), 234–39.

Way, S. (1935), Pitting due to rolling contact, Trans ASME, *Journal of Applied Mechanics*, **2**, A49–A58.

Webster P. J., Wang, X., Mills, G. and Webster, G. A. (1992), Residual stress changes in railway rails, *Physica B*, **180/181**, 1029–31.

Williams, J. A. (2005), The influence of repeated loading, residual stresses and shakedown on the behaviour of tribological contacts, *Tribology International*, **38**(9), 786–97.

Copyrighted Materials
Copyright © 2009 Woodhead Publishing Retrieved from www.knovel.com

10

The evolution and failure of pearlitic microstructure in rail steel – observations and modelling

F. J. FRANKLIN, University of Newcastle, UK;
J. E. GARNHAM and C. L. DAVIS,
University of Birmingham, UK;
D. I. FLETCHER, University of Sheffield, UK;
A. KAPOOR, Swinburne University of Technology, Australia

Abstract: Detailed metallurgical investigations of worn rails and test discs show that rolling contact fatigue (RCF) crack initiation and early propagation are strongly influenced by microstructure. Also, normally benign, ductile inclusions near and at the rail surface can weaken the microstructure when transversely strain-flattened. The effect of different rail steel grades has been shown and modelled in 2D. Early surface wear 'flaking' and RCF cannot be fully modelled in 2D, given the granular nature of the formative and final microstructures, so 3D aspects of this early crack growth are examined and 3D models of the initial and strained metallurgical structures are being developed.

Key words: wear, rolling contact fatigue, shakedown, ratcheting, pearlitic rail steel.

10.1 Introduction

In Chapter 5, the history of the development of rail steels, particularly those with a pearlitic structure, was outlined. The production processes and metallurgical characteristics of standard pearlitic rail steel were described together with how its microstructure is altered by the highly compressive rolling and sliding loads and tractions found in wheel–rail contact. These alterations make this inexpensive low-alloy, medium- to high-carbon steel surprisingly suitable for rails. However, as cost, safety, load, speed and traffic demands on rail networks increase, further refinements of pearlitic rail microstructures will be demanded together with explorations of cost-effective, alternative steel microstructures.

In this chapter, pearlitic rail steel will be further explored with regard to wear and rolling contact fatigue (RCF) failure, with observations from rail removed from track and from laboratory twin-disc test samples. The advantages and limitations of this test method will be briefly discussed, as will the development of modelling techniques from this metallurgical analysis for the prediction of rail life.

From this combined practical and modelling approach, the behaviour of rail steel microstructures under compressive rolling-sliding conditions can be analysed with regard to basic mechanical properties, and the results can then be fed into fundamental models of rail steel and used to predict rail life. The aim is to give a rapid and economical method of evaluating new or alternative rail steels, and so to select the most promising prior to track trials; track trials are expensive, but will always be needed. To support modelling, a detailed understanding of the mechanisms of microstructural evolution and failure is needed.

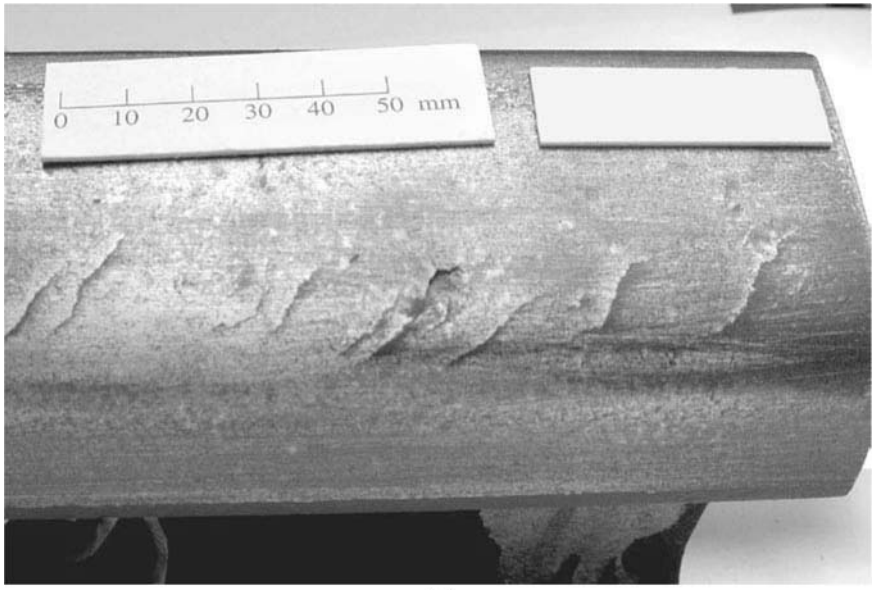
10.2 Observations of microstructural evolution and failure

10.2.1 Observations from rail

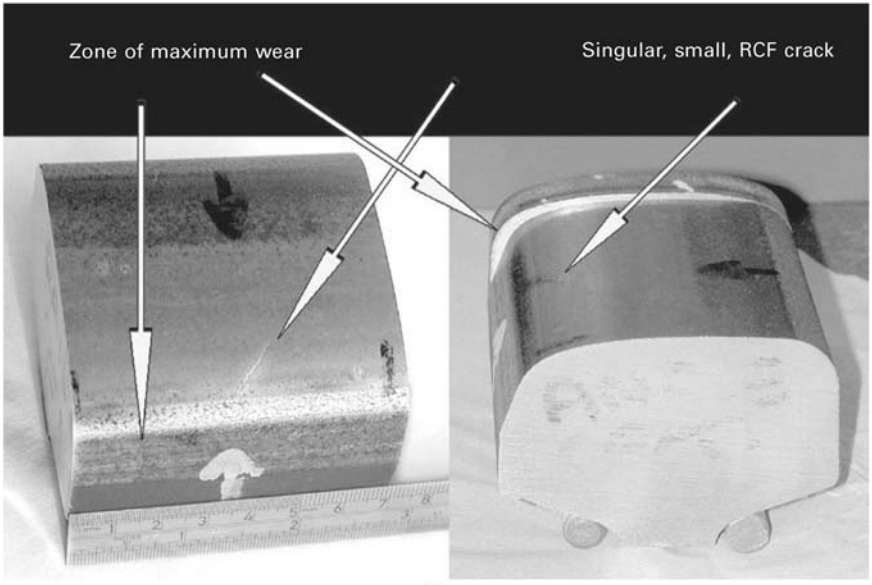
Figure 10.1 shows an old, low-carbon BS11 grade rail with multiple RCF cracks together with a more recent 260 grade rail with carbon levels at the high end of the range and which had a singular visual crack. On both, there were numerous surface micro-cracks as well as the visual cracks. On such high rails from curved track, RCF cracks form at a location *just above* the gauge corner and gauge face where there is maximum wear, i.e. at the location where there is a combination of high contact stress and relatively high creepage (Fig. 10.1b). In the gauge corner zone and below, there is high creepage or slip and so high wear (profile loss), and any initiating RCF cracks are worn away before they can propagate.

These and other sections of worn rail, mostly from the high rail of a curve, with both developed and just-initiating RCF cracks, have been surface-examined by optical and scanning electron microscopy (SEM) and multi-sectioned to provide microstructural samples for optical and SEM microscopy and for macro-, micro- and nano-hardness tests (Eden *et al.*, 2005; Garnham *et al.*, 2007; Garnham and Davis, 2008). The section orientations have been axial along the track, transverse and some angled at 90° to the centre of the RCF cracks, i.e. in the direction of maximum strain at that track profile location.

Surface SEM examinations of the steels from Fig. 10.1 showed surface micro-flakes, indicative of the ratcheting mode of wear (Kapoor and Franklin, 2000). The high matrix strain at the RCF crack location for BS11 'Rail D' is shown in Fig. 10.2, where it can be compared to the unstrained, core microstructure. Micro-examination of the RCF failure zones of UK standard grade rails (past BS11/now 220) and higher grade, higher carbon (260) rails showed that RCF cracks primarily initiated along strain-flattened, pro-eutectoid (PE) ferrite, located at the prior austenite (PA) grain boundaries. The cracks were mainly at the edges of these zones, at the boundary with

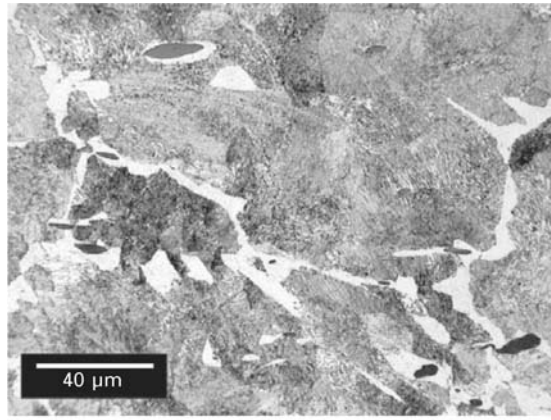


(a)

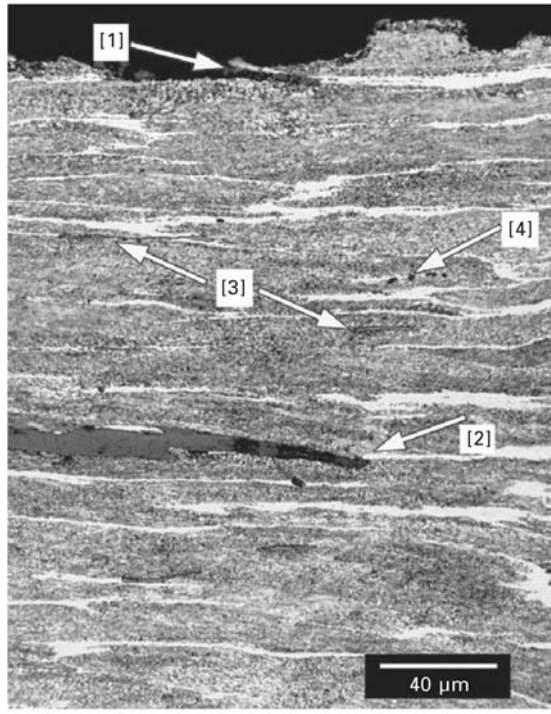


(b)

10.1 Rolling contact fatigue cracks in worn high rails from curves. (a) Developed RCF cracks in a BS11 grade rail ('Rail D'). (b) Two views of a singular, small, RCF crack on a section cut from a higher carbon, 260 grade rail (Rail 'F'). The crack is located above the region of maximum wear, as shown by comparing the section profile against a section of unused rail placed behind.



(a)



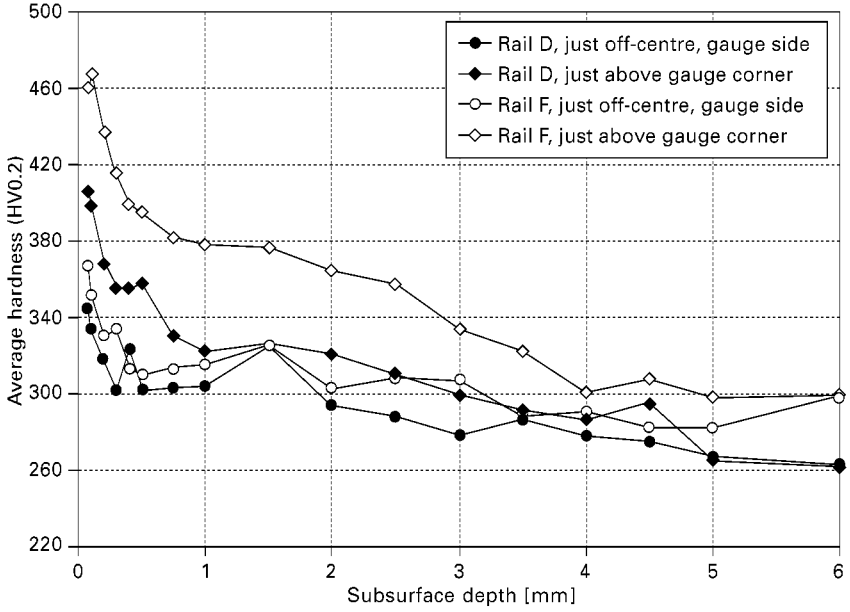
(b)

10.2 Optical microscope images of BS11 'Rail D' microstructure from a microsection angled normal to the maximum creepage vector and RCF cracks shown in Fig. 10.1a. (a) Unstrained core matrix. (b) Highly strained matrix at surface with RCF crack initiating along PE ferrite at the surface [arrow 1], an oxidised, established RCF crack propagating along PE ferrite band subsurface [arrow 2], highly strain-flattened MnS-based ductile inclusions [arrows 3] and non-distorted brittle inclusions [arrow 4]. (Surface pitting is due to post-removal corrosion.)

flanking pearlite (Fig. 10.2). In the 260 grade 'Rail F', where the amount of PE ferrite was far lower than in the BS11 Rail D, some cracks also initiated at pearlite colony boundaries.

Micro-hardness surveys of the two rails in Fig. 10.1, Rail D (BS11 grade) and Rail F (260 grade), showed the increased work-hardening at the RCF crack zones just above the gauge corner compared to the crown of the rail (Fig. 10.3). The chemical compositions of these rails are shown in Table 10.1. These show that the nitrogen and oxygen levels of the old BS11 rail are higher than modern specifications. Complementary work on worn rails (Eden *et al.*, 2005), including electron microscopic observations and nano-hardness tests of the strained phases, showed that there was strain-partitioning between the PE ferrite and adjacent pearlite zones during this hardening, with a higher percentage hardening in the PE ferrite. (It should be noted that micro- and nano-hardness values will be progressively higher than the macro-hardness value, the lesser the load, as they reflect sample preparation; they are, therefore, comparative not absolute.)

To explore crack initiation in PE ferrite, a programme of twin-disc testing was undertaken in which, for a given pearlitic rail material, the amount of PE ferrite was varied (see below). In the rails examined, it was noted that some very near-surface cracks initiated on a few strain-flattened, ductile



10.3 Micro-hardness plots showing work-hardening approaching the surface of 'Rails D and F' just gauge side of the track centre and at the RCF crack zone locations just above the gauge corner.

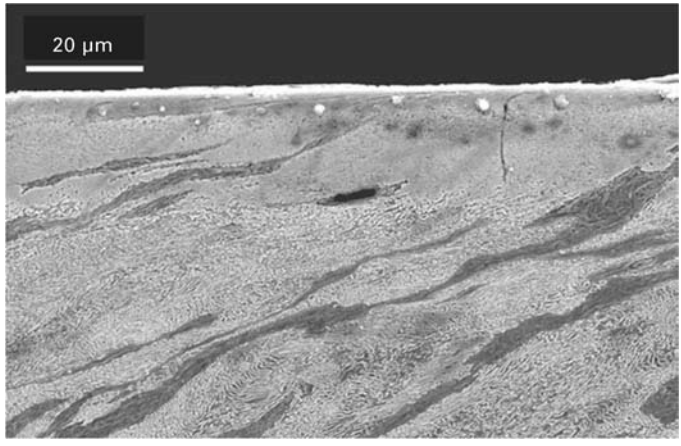
Table 10.1 Chemical compositions of *Rail D* and *Rail F* and respective standards at the time of production, together with the compositions and specifications of the 220 grade rail and R8 grade wheel used for the twin disc tests. Nitrogen and Oxygen were measured by the LECO (sample melt–gas analysis) method, all other elements by optical emission spectroscopy (OES)

Element (wt%)	**BS11 (1959-Acid Bessemer)**	***Rail 'D'* (to BS11, 1959)**	**Railtrack RT/CE/S/061 260 Grade (1996)**	***Rail 'F'* (260 grade)**	**Railtrack RT/CE/S/061 220 Grade (1996)**	***SUROS rail top discs* (220 grade)**	**BS5892-3 R8 monobloc wheel (1992)**	***SUROS wheel bottom discs* (R8)**
C	0.40–0.50	0.456	0.60–0.82	0.814	0.50–0.60	0.55	0.56 max	0.53
Si	0.05–0.20	0.129	0.13 – 0.60	0.197	0.20–0.60	0.24	0.40 max	0.29
Mn	0.95–1.25	1.078	0.65 – 1.25	0.955	1.00–1.25	1.10	0.80 max	0.72
S	0.060 max	0.027	0.008 – 0.030	0.029	0.008–0.030	0.020	0.040 max	0.007
P	0.060 max	0.012	0.030 max	0.021	0.030 max	0.022	0.040 max	0.016
Cr	–	0.013	0.15 max	< 0.60	0.15 max	0.03	0.30 max	0.20
Mo	–	0.024	0.02 max	< 0.002	0.02 max	0.005	0.08 max	0.005
Ni	–	0.094	0.10 max	< 0.031	0.10 max	0.02	0.30 max	0.02
Cu	–	0.405	0.15 max	0.022	0.15 max	0.01	0.30 max	0.09
Al	–	0.015	0.004 max	< 0.009	0.004 max	0.002	–	0.034
V	–	–	0.03 max	< 0.017	0.03 max	0.001	0.05 max	0.002
Ti	–	–	0.025 max	–	0.025 max	0.0002	–	0.0015
N	–	0.0648	0.0100 max	0.0012	0.0080 max	0.0040	–	0.0045
O	–	0.0358	0.0020 max	0.0041	0.0020 max	0.0009	–	0.0006

inclusions, although most had no associated cracks (see Fig. 5.14 in Chapter 5). Another observation from some of the several rails examined was that 'white etching constituent' (WEC) formed intermittently at the surface, at low wear/low creepage areas of the track profile such as the track centre on tangent track and near-centre on the high rails from curved track (Fig. 10.4, and Fig. 5.13 in Chapter 5). Here, the structure has been thermomechanically transformed by long-term, multi-cycle asperity contacts so that carbide is progressively taken back into solution, resulting in a form of martensite. Eventually, brittle cracks can form in this structure (Fig. 10.4). In the areas susceptible to RCF cracking, the wear rate is too high for this to happen. WEC can also be generated by singular, mostly thermal (quick heat and cool) events, such as pure sliding generated by wheel locking.

10.2.2 Observations from experiments (twin-disc tests, etc.)

To simulate and assess the formation of RCF cracks in rail steels, without the need for track trials which are necessarily long-term and expensive, an established method is twin-disc testing, with discs of rail steel and wheel steel under rolling-sliding, compressive contact. In twin-disc testing, the contact (maximum contact stress, stress distribution, creepage and environment) is



10.4 Axial section through a WEC zone situated on the track width profile crown of tangent track (0.44 wt% C BS11 grade). This SEM micrograph (ferrite dark, carbide light) shows: (i) The progressive adsorption of pearlitic lamellar carbide in the WEC zone approaching the surface; (ii) strain aligned, angled PE ferrite extending into the WEC zone; (iii) a perpendicular, brittle surface crack extending down into the WEC zone; (iv) a micro-crack initiating on a MnS-based inclusion located at the base of the WEC layer.

closely controlled, whereas on rail there is a mix of traffic contact conditions (variations in wheel profiles, loads, etc.) and environments (weather, track contaminants, etc.).

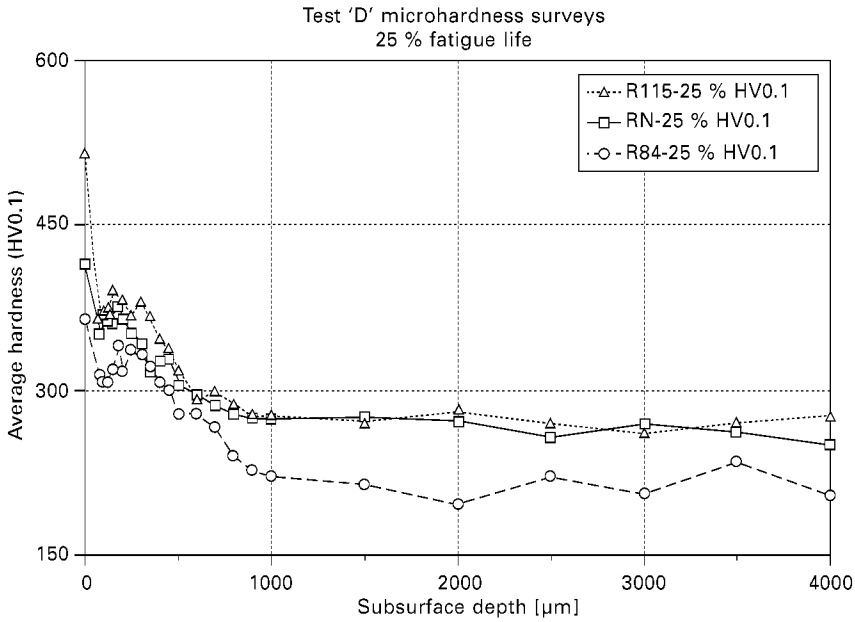
Twin-disc tests for this work were carried out on the SUROS test machine. Details of the equipment and test method can be found in Garnham and Beynon (1991, 1992), Garnham (1995) and Fletcher and Beynon (2000a,b). To simulate rain, clean, distilled water is used to lubricate the contact, which has the effect of reducing wear and facilitating the propagation of RCF cracks. The discs meet with a flat line contact of 10 mm track width (alternative twin-disc tests where there have been crowned discs giving elliptical contact zones, or top and bottom discs with different track widths, will not result in even contact conditions throughout the test due to uneven disc wear). The SUROS test machine has been found highly suitable by many researchers since the early 1990s for determining rail steel material behaviour under rolling-sliding, high-contact stress conditions that can be linked to the corresponding behaviour of those materials on track.

To simulate a driving wheel, the disc of rail material should always be the slower of the two discs; this is the situation on rail except for vehicle braking. On the slower disc, rolling and sliding are in opposite directions such that the mouths of early forming RCF cracks are closed by the contact thus trapping lubricant in the crack. This reduces crack face friction and hydraulically loads the crack tip thereby stimulating propagation. In all rolling-sliding lubricated contacts of similar materials, RCF cracks are always first observed on the slower moving surface. This aspect for rail has been further explored in depth by Tyfour *et al.* (1996), Fletcher and Beynon (1998, 2000c) and Fletcher *et al.* (2008).

The first major difference between rail and twin-disc is that, while the scale of microstructural features is common for both rail and test discs, contact dimensions and thus stress field dimensions are far larger with actual wheel–rail contact than with twin-disc contact. Surface micro-roughness is also similar for rail and twin-disc contact, with similar dimensions of subsurface stresses; these are very large, and concentrated within about 50 microns of the surface.

SUROS disc diameter is typically 47 mm. Comparing rail and disc micro-hardness plots (Figs 10.3 and 10.5), rail hardening extends to ~ 4 mm whereas disc hardening does not extend beyond 1 mm. With the water-lubricated SUROS disc hardness plots there is a distinct strain 'relaxation zone' between the high surface strain from asperity contact and maximum Hertzian strain subsurface. This is not so distinct with the rail hardness plots, because rails experience a wide range of different contacts as well as a mixture of wet and dry contacts if exposed to the weather, and the net effect is to smooth the subsurface distribution of strain.

For this test programme, heat treatment was used to examine the effect



10.5 Surface and subsurface micro-hardness results (HV0.1) for all three rail disc conditions at 25 % RCF life.

of different microstructures on test outcomes. All 'rail' discs were sourced from the head of a single length of unused 220 grade rail and 'wheel' discs from a the rim of an R8 grade wheel (composition details are given in Table 10.1). Rail discs were cut transversely across the railhead (in order to give similar grainflow around a test disc circumference). This meant that the rolling and sliding direction was normal to the grainflow and, in particular, to elongated, ductile MnS-based inclusions, which is the situation with transverse creepage on rail, the most significant creepage component with regard to RCF cracking on rail curves.

Wheel discs were cut from across the wheel rim. These were identified as 'WN'. The rail discs were divided into three sets. The set produced from the rail without further heat treatment are identified as 'RN'. Two sets were heat-treated in a vacuum furnace, with controlled inert gas flows to control cooling. Two heat-treatment conditions were used to maximise and minimise the PE ferrite content, respectively. These were:

- To maximise the PE ferrite content at PA grain boundaries, discs were heated to 840 °C in 1.1 hours, austenitised for 2.5 hours, slow cooled over 4 hours to 610 °C, held for 0.5 hours, then fast cooled to room temperature in 0.3 hours. These discs were identified as 'R84'.
- To minimise PE ferrite content, discs were heated to 1150 °C in 1.75 hours,

austenitised for 2.5 hours, then rapid-gas cooled to room temperature in 0.5 hours. These discs were identified as 'R115'.

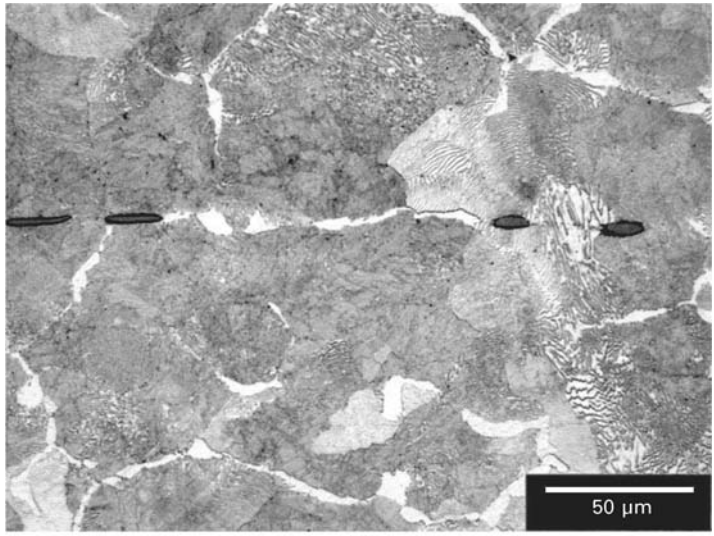
Prior austenite grainsize, as delineated by PE ferrite, was assessed on microsections taken from one disc of each set; it was determined by comparative charts from British Standard BS4490 [1989]. PE ferrite content was determined by image analysis of representative micrographs.

RCF tests were run-in dry for an initial 500 revolutions and then (filtered, distilled) water was added to the contact at one drip per second. The 500 cycle dry period has been found to reduce fatigue life (Fletcher and Beynon, 2000a) through generation of high shear stress at the disc surface, thereby increasing the rate of ratcheting strain accumulation, and thus the likelihood of crack initiation, at the surface. All tests were at a creepage of –1 % and mainly at a maximum contact stress of 1500 MPa, a stress typical for moderate curving wheel–rail contact (Allery, 1986). A few singular, comparative tests were carried out at the same creepage and a higher maximum contact stress of 1800 MPa. For the three rail disc microstructures (RN, R84, R115), at 1500 MPa, at least two tests were run to RCF failure and additional single tests were then carried out at 25 % and 10 % of average RCF life, and at 250 and 500 dry cycles.

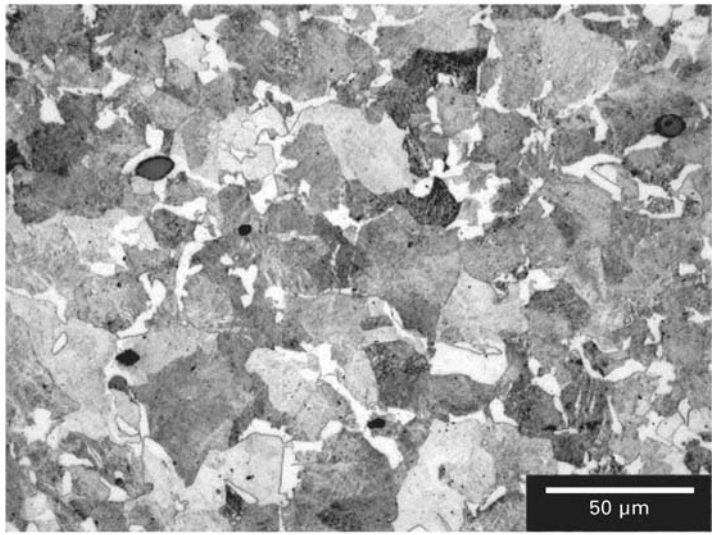
Disc pre-test microstructures are shown in Fig. 10.6 and pre-test properties (macro-hardness, PA grainsize, volume % PE ferrite, pearlite nodule micro-hardness) in Table 10.2. Note that for all three rail microstructures, PE ferrite at PA grain boundaries did not completely surround pearlitic areas, but there can be a considerable extent of ferrite along a boundary which can be aligned and thinned with strain to facilitate crack growth.

RCF lives are given in Table 10.3. This shows that the R115 discs had a longer mean fatigue life than the RN and R84 discs; however, there was also greater variability between the fatigue lives for the four R115 discs. More detailed test results, including dimensions of the largest cracks per disc, associated matrix strain and traction coefficients, are shown in Table 10.4. Surveys of all significantly large cracks in the 100 % RCF life and 25 % RCF life tests, with regard to crack length-depth ratios and crack density, are shown in Table 10.5. Also included in this table are comparative data for the target RCF cracks from the angled microsections, cut normal to the cracks, taken through BS11 grade Rail D and 260 grade Rail F (*cf* Fig. 10.1). Both sets of data indicate that RCF cracks are angled steeper in microstructures with lower PE ferrite (R115 and higher grade Rail F) reflecting the lower strain (Table 10.4), even after a higher number of test cycles. However, note that shallow-angled cracks in highly strained zones of weaker steel microstructures can turn more steeply down once propagated through the strained zone to the lesser strained region below (Fig. 10.7a).

The fatigue lives for the three microstructures were reflected in these

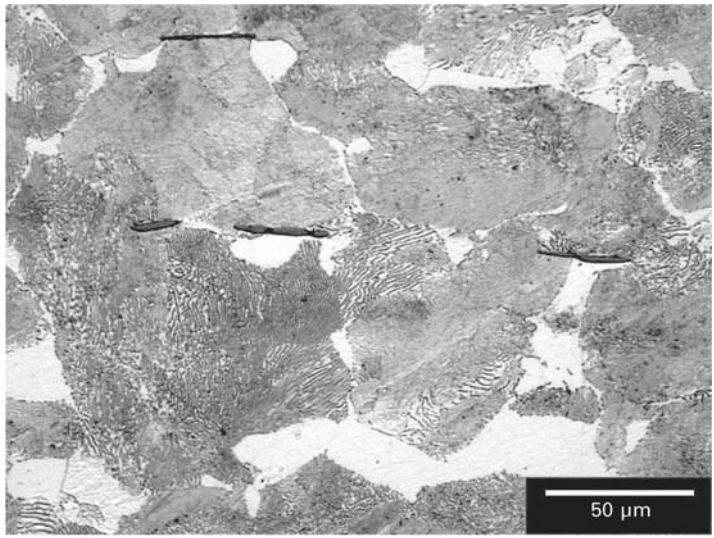


(a)

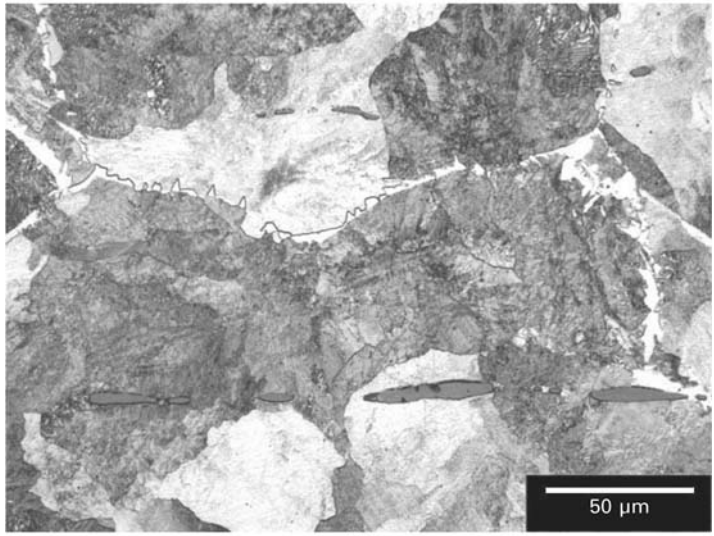


(b)

10.6 SUROS disc microstructures prior to testing. These are disc transverse sections (i.e. rail axial and wheel rim circumferential sections), which also show MnS-based inclusions, elongated along the rail axial direction. (a) 'RN' – disc from as-received 220 grade rail. (b) 'WN' – disc from as received R8 wheel. (c) 'R84' – rail disc heat-treated to maximize PE ferrite. (d) 'R115' – rail disc heat-treated to minimise grain PE ferrite.



(c)



(d)

10.6 Cont'd

'crack growth rates'. These cracks were predominantly in PE ferrite, clearly indicating that the different microstructures, i.e. amounts of PE ferrite, had the most significant effect on fatigue life. The difference in (core) hardness is not fully reflected in the RCF results, as will be discussed later.

Surface RCF crack initiation features were similar to those seen in rail; the circumferential, disc microsections indicate that numerous fatigue cracks in the detected 'defect zone' had initiated along highly strained, PE ferrite

Table 10.2 Disc material properties for SUROS tests

Material condition	**R84**	**RN**	**R115**	**WN**
Pre-test hardness (HV10)	204 ± 5	239 ± 6	250 ± 2	239 ± 4
Prior austenite grainsize (GA/ASTM)	4–6	3–4	2–4	6–7
Pro-eutectoid ferrite (%)	11.0 ± 2.0	5.8 ± 1.2	2.0 ± 1.6	–
Pearlite nodule hardness (HV1)	234 ± 11	264 ± 12	283 ± 9	270 ± 8

Table 10.3 SUROS disc fatigue results (average and individual) and the effect of higher maximum contact stress

Test programme **All –1 % creepage**		**RCF life (top disc revolutions)**		
		R84	**RN**	**R115**
1500 MPa		13 462 average 12 911 14 013	16 563 average 14 013 19 113	21 586 average 26 711 12 209 15 908 31 514
1800 MPa		4612	5414	5813
RCF life ratio:	1800 MPa / 1500 MPa	34 %	33 %	27 %

band boundaries. In the R115 discs, with the lowest PE ferrite, some cracks had also initiated on pearlite colony boundaries. The part life tests showed the exact location of crack initiation and initial propagation (Fig. 10.8a). Figure 10.7a shows an overview of a complete crack and the disc strain field. It shows that once a RCF crack has propagated through the angled strain field (with these concentrated line contact, twin-disc tests) it can turn toward a steeper angle in the less strained matrix below. After initiation and initial propagation (Fig. 10.7a), the direction of further crack propagation appeared primarily determined by the strain field; however, many crack path diversions were observed along the edges of strained PE ferrite bands (Fig. 10.7b), with occasional branching to strain-flattened MnS-based inclusions (Fig. 10.8b). With regard to Table 10.3, comparing RCF life for the three microstructures, R84, RN and R115, at maximum contact stresses of 1800 MPa and 1500 MPa, material ranking was the same, with RCF life at 1800 MPa approximately one third of that at 1500 MPa. This highlights the damaging effect of higher contact stresses found with curving, with track and wheel profile inconsistencies and with general track and vehicle inconsistencies which generate high dynamic loads.

Table 10.4 SUROS disc crack details and respective test conditions and traction coefficients. For RCF failure, discs were circumferentially sectioned at the location indicated by the eddy current probe. For 25 % and 10 % RCF life, discs were circumferentially sectioned track centre

Disc (all 1 % creepage)	Top disc revs	Length of longest crack [μm]	Length divided by top disc revs [μm/rev]	Depth of longest crack [μm]	Depth divided by top disc revs [μm/rev]	RCF life tests. Matrix strain angle (x)/crack propagation angle to surface (in zone of maximum subsurface hardening [a])	Maximum matrix strain subsurface tan(90°–angle *x*)	Strain divided by top disc revs [10^{-5} $revs^{-1}$]	Coefficient of traction near test end
1800 MPa p_o to RCF failure									
R84	4612	2816	0.611	745	0.162	14°/14°	4.0	86.7	0.165
RN	5414	2580	0.477	700	0.129	13°/12°	4.3	79.4	0.180
R115	5813	1495	0.257	565	0.097	18°/18°	3.1	53.3	0.195
1500 MPa p_o to RCF failure									
R84	12911	2142	0.166	750	0.058	15°/15°	3.7	28.7	0.160
R84	14013	2015	0.148	658	0.047	16°/16°	3.5	25.0	0.180
RN	14013	1227	0.088	437	0.031	20°/22°	2.7	19.3	0.175
RN	19113	1681	0.088	643	0.034	17°/20°	3.3	17.3	0.185
R115	26711	1370	0.051	511	0.019	17°/25°	3.3	12.4	0.200
R115[b]	12209	288	0.024	83	0.007	–	–	–	0.275
R115	31514	1124	0.036	525	0.017	19°/19°	2.9	9.2	0.195
1500 MPa p_o to 25 % RCF life									
R84	3363	97.2	0.029	36.2	0.011	–	–	–	0.210

RN	4112	158	0.038	41	0.010	–	–	–	0.245
R115	5412	92.2	0.017	19.8	0.004	–	–	–	0.275
1500 MPa p_o to 10 % RCF life									
R84	1352	51.4	0.038	9.3	0.007	–	–	–	0.220
RN	1657	56.4	0.034	21	0.013	–	–	–	0.240
R115	2173	47.5	0.022	10	0.005	–	–	–	0.250

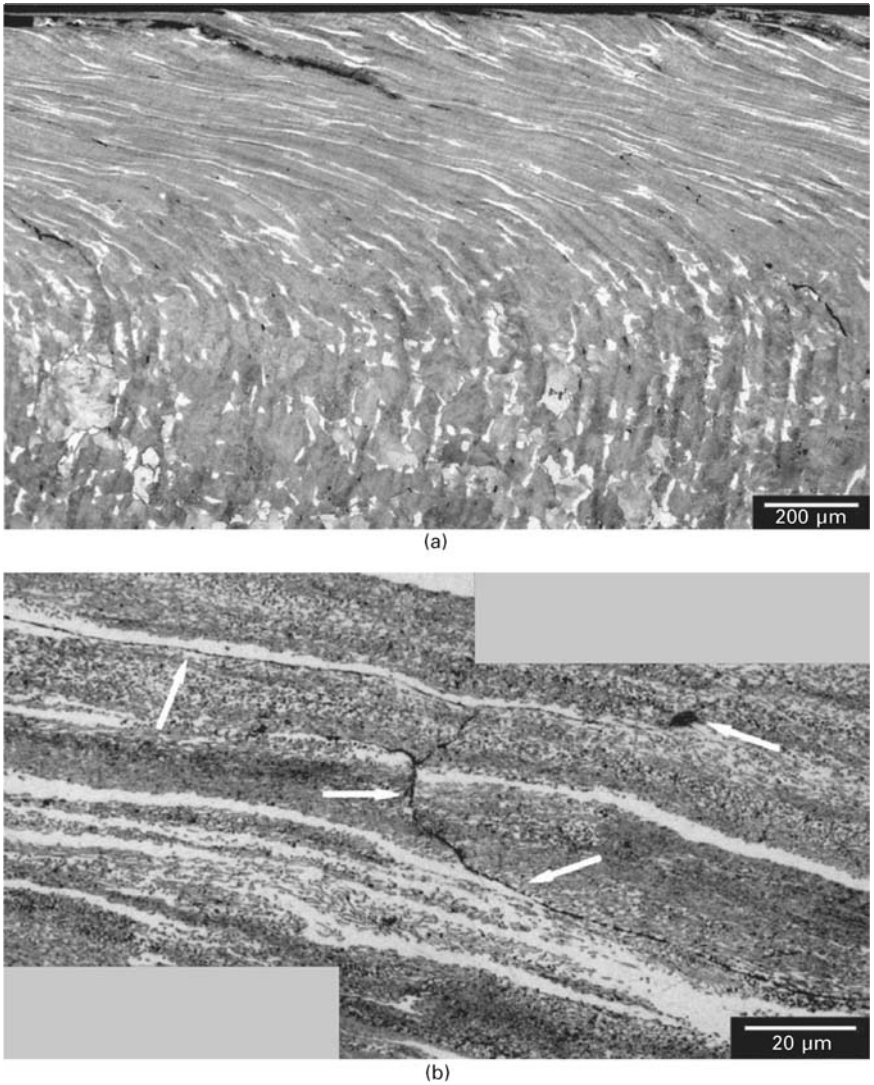
[a] Crack shapes generally following strain field. Maximum subsurface hardening at ~ 180 μm at 1500 p_o and at ~ 200 μm at 1800 p_o. The matrix strain angle is difficult to assess as it is usually higher above a crack than below the crack; in this instance, an average of the two is indicated below.

[b] Low life R115 result; high traction coefficient with flake tips lifting from surface. Appears to have triggered eddy-current gate before RCF cracks fully developed.

Table 10.5 SUROS disc RCF crack length/depth ratio statistics and comparative rail data

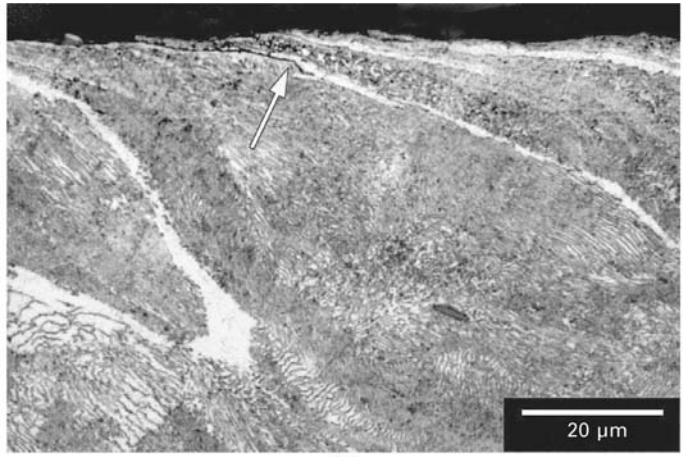
Disc	Top disc revs	Significant* crack length/depth ratio		[a] Normalised data: number of significant* cracks around circumference'	Significant cracks in section arc
		Mean	Standard deviation		
Tests at 1800 MPa to 100 % life (* 100 % life, significant cracks nominally chosen as deeper than 200 μm)					
R84 Test A	4612	3.70	0.56	90	(11)
RN Test A	5414	3.86	0.45	59	(7)
R115 Test A	5813	2.71	0.53	22	(2)
Tests at 1500 MPa to 100 % life (* 100 % life, significant cracks nominally chosen as deeper than 200 μm)					
R84 Test B	12911	3.20	0.52	126	(15)
R84 Test C	14013	2.98	0.33	151	(25)
RN Test B	14013	2.47	0.18	97	(12)
RN Test C	19113	2.78	0.29	167	(26)
R115 Test B	26711	2.26	0.26	160	(20)
R115 Test C	12209	–	–	–	Nil
R115 Test S	31514	1.94	0.12	74	(9)
Tests at 1500 MPa to 25 % life (* 25 % life, significant cracks nominally chosen as deeper than 10 μm)					
R84 Test D	3363	3.54	0.92	74	(7)
RN Test D	4112	6.31	1.62	64	(4)
R115 Test D	5412	4.30	–	9	(1)

For comparison, rail data:
Angled microsections, normal to RCF cracks, through BS11 grade 'Rail D' and 260 grade 'Rail F'.
Rail D, centre of targeted crack: length 7680 μm depth 2353 μm ratio 3.27
Rail F, centre of singular large crack: length 4460 μm depth 2130 μm ratio 2.09
[a] As the arc lengths of circumferential sections varied, the number of significant cracks in each arc has been normalised thus: (Number of sig. cracks in arc) x (circumference / arc length).
Note: The disc from R115 Test K, 1500 MPa to 100 % RCF life at 15 908 top disc revs, was kept intact for reference and not micro-sectioned

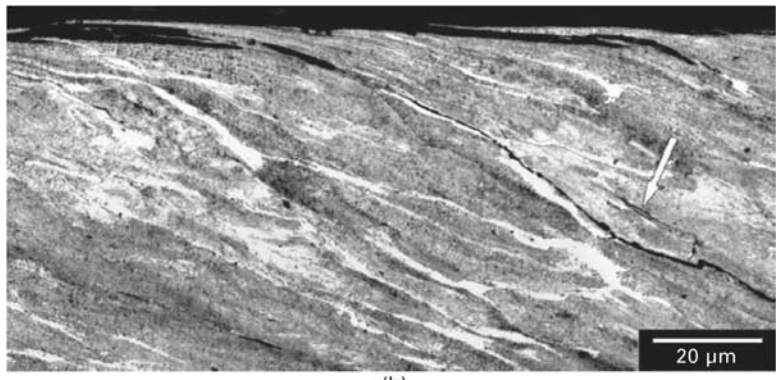


10.7 Circumferential section through a 'R84' rail test disc at 100 % RCF life. (a) Overview of the largest crack. (b) Detail of crack propagation in subsurface zone of maximum strain, showing crack branching and 'jumping' (arrowed) across pearlite zone from one PE ferrite zone to another.

In Section 5.5 of Chapter 5, the effect of highly strain-flattened, ductile, MnS-based inclusions is discussed and the associated figure shows cracking on such an inclusion near-surface. In all three microstructures tested here, crack initiation was observed on a few (of the many) ductile inclusions, together with some main cracks branching to ductile inclusions (Fig. 10.7b)



(a)



(b)

10.8 Circumferential sections through 'RN' rail test discs. (a) Test disc at 10 % average RCF life; surface crack initiation on border of strain flattened PE ferrite zone and flanking pearlite (arrowed). (b) Test disc at 100 % RCF life; near-surface crack propagation down the edge of a strain-flattened PE ferrite zone, with one branch to a strain-flattened, MnS-based inclusion (arrowed).

and ductile inclusions lying in the main crack paths. Because the R84 and RN discs were more comprehensively cracked than the R115 discs, a definitive linkage between microstructure and cracked MnS-based inclusion density could not be meaningfully determined. It was also observed that very near-surface, nano-cracks had initiated on some very small, angular brittle inclusions.

Fatigue lives reflect pro-eutectoid ferrite content and/or core hardness (*cf* Tables 10.2 and 10.3). However, micro-hardness plots for the full-life tests showed that the hardness difference between the three microstructures diminishes approaching the strain field at the disc surface, with all results

Previous Page
Copyrighted Materials
Copyright © 2009 Woodhead Publishing Retrieved from www.knovel.com

falling within a common 'envelope' (Garnham *et al.*, 2007; Garnham and Davis, 2008). Micro-hardness results from the transverse disc sections, for the 25 % RCF life tests for all three conditions, are shown in Fig. 10.5. In the region of the fatigued surface and near-surface, all except one of the micro-hardness curves again fall within a common 'envelope'. There is no longer a clear differentiation between microstructures, i.e. the 'lower hardness/higher PE ferrite' material has strain-hardened to a similar level to the higher hardness material, indicating that the PE ferrite component has shown greater percentage hardening. This pattern is nearly established in the 25 % RCF life tests.

This figure additionally shows data from micro-hardness testing on the worn disc surfaces (disc surfaces are too rough to do this at 100 % RCF life, areas between developing flakes were chosen on the 25 % life disc surfaces). The results reflect asperity hardening at the immediate surface and, interestingly, follow more closely the pattern of RCF life per condition. On the disc surfaces and 10 mm subsurface on the microsections (core hardness), hardness indents were made at loads of 0.1 kg and 1.0 kg (HV0.1, HV1) for comparison. Whereas subsurface HV0.1 hardness readings were on average 3.4 % higher than the HV1 readings, reflecting sample polishing, on the disc surface HV0.1 readings were on average 20 % higher than the HV1 readings, reflecting the localised hardness gradients within a few microns of the surface due to asperity effects on the contact. In other words, no definitive determination can be made for disc surface hardness, as each testing load/indent depth reflects part of the hardness gradient approaching the surface at a micron scale. However, the method does allow comparative determinations.

Nano-hardness results for the R84 and RN samples are shown in Table 10.6. Even with grids of 800 indents per sample, only a few indents fell fully within PE ferrite zones, particularly near-surface where the ferrite grain boundaries are most highly strained, and this restricts the data available. These limited results indicate clearly that the percentage mean PE ferrite hardness increase, in the zone of maximum strain-hardening at around 200 μm subsurface and near-surface, is higher than that of the pearlite phase, as observed with worn rail examinations (Eden *et al.*, 2005). For example, in the RN sample tested to 10 % fatigue life the PE ferrite 200 μm subsurface increased in hardness by 34 % compared to its undeformed value, whereas the pearlite hardened by 15 %. This indicates that there was significant preferential straining in the PE ferrite compared to the pearlite. This is more fully discussed by Garnham *et al.* (2007).

10.2.3 Summary of key microstructural features

Worn rail taken from track and rail material tested under near-track contact conditions has been examined at a macro-, micro- and nano-scale. In standard

Table 10.6 Percentage increase in mean nanohardness approaching strained surface over mean nanohardness of core (5000 μm) for RN and R84 material conditions at RCF failure and percentages of RCF life

RN (standard rail)	**% Increase in mean pearlite hardness *{Number of nano-indents fully in phase}***				**% Increase in mean PE ferrite hardness *{Number of nano-indents fully in phase}***			
Depth (μm) >	**10**	**200**	**500**	**5000**	**10**	**200**	**500**	**5000**
500 cycles (F)	27	20	5	0	55	41	8	0
(dry)	*{174}*	*{179}*	*{177}*	*{168}*	*{3}*	*{3}*	*{4}*	*{1}*
10 % life (E)	8	15	6	0	13	34	6	0
	{176}	*{176}*	*{176}*	*{185}*	*{4}*	*{2}*	*{3}*	*{2}*
25 % life (D)	21	25	6	0	–	–	–	0
	{174}	*{178}*	*{189}*	*{193}*	*{0}*	*{2}*	*{1}*	*{0}*
Full life (B)	19	27	1	0	–	–	–	0
	{31}	*{182}*	*{174}*	*{173}*	*{0}*	*{0}*	*{4}*	*{0}*
R84 (maximised PE ferrite)	**% Increase in mean pearlite hardness *{Number of nano-indents fully in phase}***				**% Increase in mean PE ferrite hardness *{Number of nano-indents fully in phase}***			
Depth (μm) >	**10**	**200**	**500**	**5000**	**10**	**200**	**500**	**5000**
500 cycles (F)	52	36	14	0	85	57	30	0
(dry)	*{178}*	*{156}*	*{179}*	*{167}*	*{4}*	*{8}*	*{2}*	*{15}*
10 % life (E)	13	17	9	0	–	16	16	0
	{159}	*{149}*	*{162}*	*{155}*	*{1}*	*{4}*	*{7}*	*{8}*
25 % life (D)	20	31	11	0	29	45	13	0
	{157}	*{164}*	*{176}*	*{173}*	*{14}*	*{9}*	*{6}*	*{5}*
Full life(B)	23	43	19	0	–	–	21	0
	{43}	*{159}*	*{164}*	*{166}*	*{1}*	*{1}*	*{6}*	*{5}*

and higher grade rail (where PE ferrite is present), RCF cracks were observed to initiate at, and propagate along, the surface boundary of the most highly strained PE ferrite zones and flanking pearlite. With twin-disc tests this was further explored where discs cut from rail were compared in three conditions: as-machined and after two heat-treatments which maximised and minimised the PE ferrite content.

PE ferrite strain-hardens more than pearlite during cyclic, rolling-sliding, compressive contact. The degree of strain of a PE ferrite zone, situated along PA grain boundaries, was found to be dependent upon its thickness and orientation to the strain direction (Garnham and Davis, 2008), and that is affected by the PA grainsize. With a fine grainsize, there would be a less direct route for propagation.

In the most highly strained, near-surface region, propagating cracks were observed in (two-dimensional–2D) cross-sections to have 'jumped' between parallel PE ferrite zones to accommodate the strain field. In the higher carbon rail and discs heat-treated to minimise PE ferrite, some crack initiation was also observed at pearlite colony boundaries.

Microstructures containing less PE ferrite exhibit long lives until fatigue crack initiation. However, a larger degree of scatter in life is observed. An increase in the amount of PE ferrite in a rail steel reduces RCF life, probably because the increased strain-partitioning/preferential straining in the PE ferrite leads to earlier fatigue crack initiation. In these microstructures, some crack initiation was also observed at pearlite colony boundaries; it could not be clearly determined whether these were situated at PA grain boundaries. Similar crack initiation mechanisms were observed at both 1500 and 1800 MPa, although RCF life was significantly reduced at the higher contact pressure.

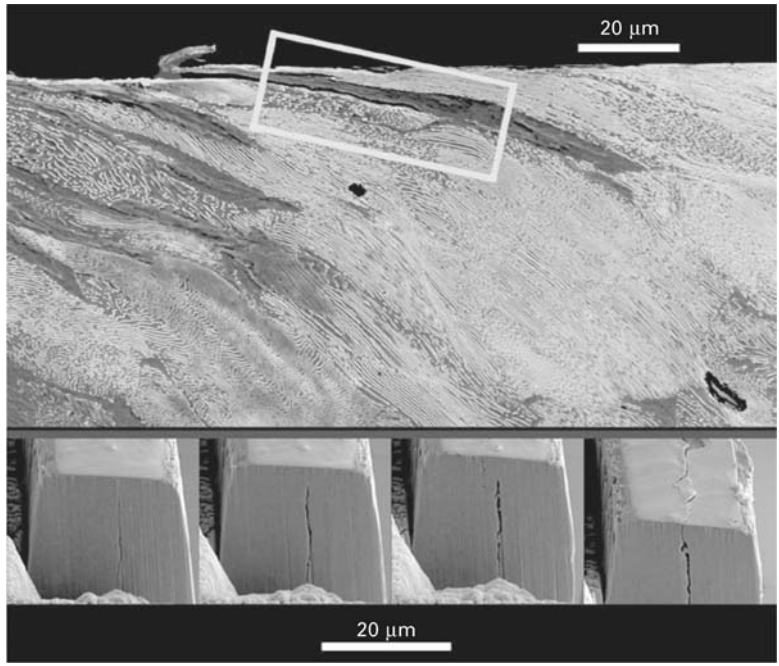
Rod-shaped, ductile inclusions can become highly strain-flattened to a 'disc' shape in rail near-surface microstructures deformed by a high vector of transverse creepage; such flattened inclusions near-surface can influence crack initiation and propagation. The more strain-flattened a ductile inclusion, the more likely it is to present a 'plane of weakness' facilitating crack propagation, i.e. inclusions located within the most highly strained, PE ferrite zones. Flattened inclusions also present a weak point within the pearlite matrix in the more resilient higher grade rail with far less PE ferrite. Micro-cracks were observed on some of these inclusions, although crack growth was limited and very near-surface. Nano-cracks were also observed on some very fine brittle inclusions near-surface. The effect of inclusion contents and distributions in rail steels will be further investigated.

10.2.4 Ongoing metallurgical work

One limitation of the examination methods described so far is that the effect of the three-dimensional (3D) aspects of the grain distribution on crack

propagation, particularly initial propagation, has not been fully examined. For example, how far will the early propagation of a surface RCF crack follow a PE ferrite boundary as it curves away from the strain field. Such information is required to fully calibrate and develop models of short crack growth. This is being investigated by three methods:

(1) Multiple polishing and microscopic examination of the same micro-crack within a microsection. This has the advantage that etched sections can be examined so that small shape changes in RCF cracks, PE ferrite zones, pearlite zones and inclusions can be seen.
(2) Focused ion beam ('FIB') milling and SEM-imaging through single initiating micro-cracks (Fig. 10.9). With this method a micro-crack is FIB milled out into a block of material. This block can then be progressively FIB eroded from the side, thus mapping the approach to the crack tip and the complete crack shape, including branches, that exists below the



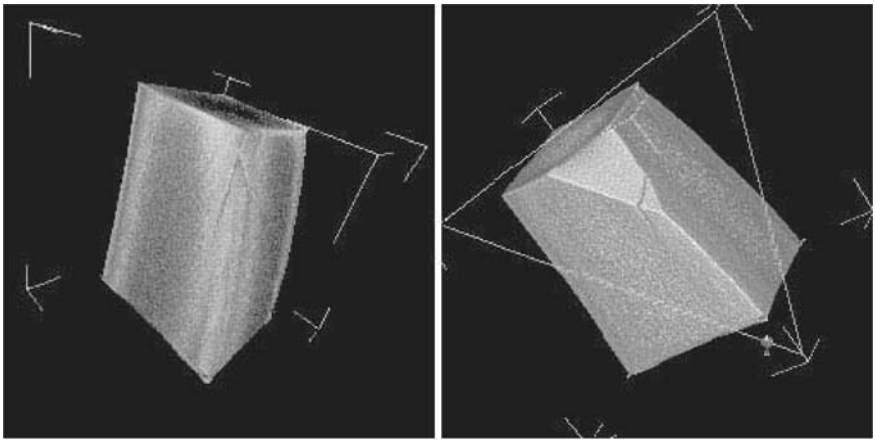
10.9 The area delineated by the white rectangle was focused iron beam (FIB) milled out of this 25 % life, R84, SUROS disc, circumferential microsection. This 'block' was then side-milled by FIB, from beyond the crack tip toward the crack mouth, to reveal the 3D subsurface crack shape. It was milled in 160 (¼ micron) steps with each step captured by scanning electron microscopy imaging; four of these images are shown above; at the 9¼ , 12, 15 and 20 μm stages.

plane of the microsection. At each stage a SEM image can be taken and a 3D construct, or video, of the micro-crack shape made.

(3) Digital computed radiography (also called computer-aided tomography, or 'CAT' scanning). The way this method has been used to date has been to extract a small block of rail material containing a small macro-crack and to rotate the block to different orientations before a focused X-ray source. The multiple scans are digitally computed to construct a 3D model of the volume and its defects. This model can then be 'sliced' (by computer) at any plane to examine crack shape characteristics (Fig. 10.10). The method examines cracks of a few millimetres visual surface length, and is being further developed and used by non-destructive evaluation companies for non-destructive use on track, particularly for weld examinations.

10.3 Modelling

Chapter 9 focuses on propagation of long cracks (i.e. 2–3 mm or longer) in rails. Here the earliest stages of material deformation and crack formation are discussed. The term crack initiation is ambiguous, but for rail inspectors a crack with surface length less than about 2 mm can be ignored; this corresponds to a crack depth of about 0.5 mm. For a model to predict rail life, it is necessary to understand how plastic deformation and damage



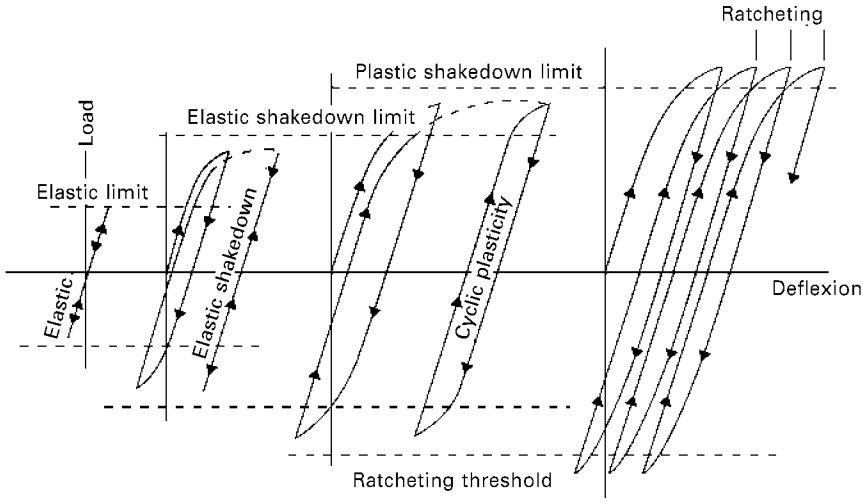
10.10 X-ray, computer aided tomography (CAT) scan of a small (visual) RCF crack on worn rail, similar to that shown in Fig. 10.1b. A cube of material has been extracted around the crack and then rotated before a scanning, focused X-ray. A 3D image has been digitally constructed (left image) which can then be 'sliced' at any plane to examine the crack features.

accumulate, and how microstructure influences formation and propagation of cracks, in this region close to the surface.

10.3.1 Material response to cyclic loading

In general, material subject to repeated loading can respond in four ways (see Fig. 10.11). If the cyclic stress is less than the material's elastic limit then the response will be purely elastic and failure, if any, will occur by high-cycle fatigue (HCF). If the cyclic stress is above the elastic limit then there will be some plastic flow, which will lead to the formation of protective residual stresses and may also cause the material to strain-harden; if the cyclic stress is below the 'elastic shakedown limit', the system will shakedown to a state where, after a few load cycles, the response is again purely elastic and failure, if any, will occur by HCF.

If the cyclic stress is above the elastic shakedown limit, then the system will not shakedown to a purely elastic state, and there will be plastic deformation each cycle. For certain stress cycles the resulting plastic strain may be fully reversing, and eventual failure by low-cycle fatigue (LCF) is likely; such closed cycles of plastic strain are only possible below the 'plastic shakedown limit'. If the cycle of plastic strain is open, then there is a net change in plastic strain after each cycle. The directional accumulation of plastic strain is called 'plastic ratcheting', and the plastic shakedown limit is therefore also called the 'ratcheting threshold'. Ideally, the train–track system and,



10.11 Material response to cyclic loading. Beneath the elastic shakedown limit, the system will be in, or settle into, a purely elastic state; above the elastic shakedown limit, there will be plastic deformation each cycle.

in particular, the wheel and rail materials should be designed or selected so that the wheel–rail contact runs at a stress beneath the elastic shakedown limits of the wheel and rail.

10.3.2 Shakedown limits

There are two ways of obtaining the shakedown limit. One is to run a finite element (FE) model to simulate rolling or sliding of the two bodies. The material properties evolve with time and are accounted for in the model. The loading case against the steady-state cycle achieved is analysed to determine if the behaviour was elastic shakedown, plastic shakedown or ratcheting. This approach suffers from the drawback that the computing resources required are large and the results are generally non-parametric. The Hahn group established shakedown limits by running FE simulations for a number of cycles until a steady-state was achieved (Ham *et al.*, 1988; Kulkarni, 1991). Another approach relies on the theories of plasticity to obtain lower and upper bounds to the shakedown limit. The actual shakedown limit lies between the lower and upper bounds and, by bringing the two bounds closer, an accurate value is obtained. This approach requires elastic stresses and a postulated residual stress field or failure mechanism to obtain respectively lower or upper bounds to the shakedown limit. The results are quick and parametric so it is easy to see the effect of material properties such as hardness on the shakedown limit. This approach has been extensively used by the Johnson group (Johnson, 1962; Ponter *et al.*, 1985; Bower and Johnson, 1991; Johnson, 1992) for determining lower bounds and the Kapoor group (Kapoor and Williams, 1994, 1996; Wong *et al.*, 1997a; Dyson *et al.*, 1999) for determining upper bounds. Additional results for shakedown limits for line contact (such as obtained in two contacting parallel cylinders), point contact (two contacting spheres) and elliptic contacts (two general smooth bodies such as eggs) are available (Kapoor *et al.*, 1996; Wong *et al.*, 1996, 1997b; Jones *et al.*, 1997; Williams *et al.*, 1999).

The shakedown limit is proportional to the shear yield stress of the material, and depends on the distribution of normal pressure and shear traction across the wheel–rail contact. The shear yield stress, k, is proportional to the hardness of the material, but depends on the hardening response to cyclic loading (Ponter *et al.*, 2004). An overview of material response to cyclic loading is provided by Fleck *et al.* (1994). Finite element techniques can easily incorporate the complex hardening response of the material. The upper and lower bounds approach can also estimate the effect of material hardening on the shakedown limit, and the publications mentioned above contain results for both kinematic hardening (where the yield locus moves in stress space without changing its radius) and isotropic hardening (where the radius of the yield locus changes).

To calculate shakedown limits, the wheel–rail contact can be approximated as an ellipse with a Hertz pressure distribution, and traction, q, proportional to the pressure p, i.e., $q = \mu p$ where μ is the coefficient of friction. For a given stretch of track, wheel–rail contact data can be determined from vehicle dynamics simulations, stresses in the wheel and rail can be determined from FE analysis and shakedown maps can be constructed on the basis of material data. This approach has been used, for example, to produce shakedown maps which can be used for selecting hard, metallic coatings for rails (Ringsberg *et al.*, 2005).

Real engineering surfaces have roughness on microscopic scale. When such bodies are brought together, the contact is made at asperities (higher bumps on the surface) and the pressure profile becomes peaky, causing plastic flow at the asperity contact even though the nominal contact pressure is beneath the shakedown limit. Analyses by Kapoor *et al.* (2002) and Morales-Espejel *et al.* (1999) demonstrate the important roles roughness and lubricant viscosity play in shakedown limit for a real contact.

10.3.3 Failure of material subjected to ratcheting

Above the elastic shakedown limit, there will be plastic deformation with each load cycle and failure by LCF is possible. Above the plastic shakedown limit, ratcheting failure (RF) is also possible once the plastic deformation accumulates to the point where the material's ductility is exhausted. For LCF, the number of cycles to failure, N_{LCF}, is related to the amplitude of the plastic strain cycle, $\Delta\varepsilon_{\mathrm{P}}/2$:

$$\frac{\Delta\varepsilon_{\mathrm{P}}}{2} = \varepsilon_{\mathrm{f}}' \, (2N_{\mathrm{LCF}})^{c} \qquad [10.1]$$

where $\varepsilon_{\mathrm{f}}'$ is the fracture ductility and $c \approx 2$. For RF, the number of cycles to failure, N_{RF}, is related to the plastic strain increment, $\Delta\varepsilon_{\mathrm{r}}$:

$$N_{\mathrm{RF}} = \frac{\varepsilon_{\mathrm{r}}}{\Delta\varepsilon_{r}} \qquad [10.2]$$

where ε_{c} is the critical strain at which ductility is exhausted and failure occurs. These failure mechanisms may be independent or additive (Kapoor, 1994, 1997). If independent, the real number of cycles to failure, N_{f}, is given simply by:

$$N_{\mathrm{f}} = \min \{N_{\mathrm{LCF}}, N_{\mathrm{RF}}\} \qquad [10.3]$$

i.e. by whichever mechanism predicts failure soonest. If additive, the real number of cycles to failure may be given by Miner's rule:

$$\frac{1}{N_f} = \frac{1}{N_{LCF}} + \frac{1}{N_{RF}} \quad [10.4]$$

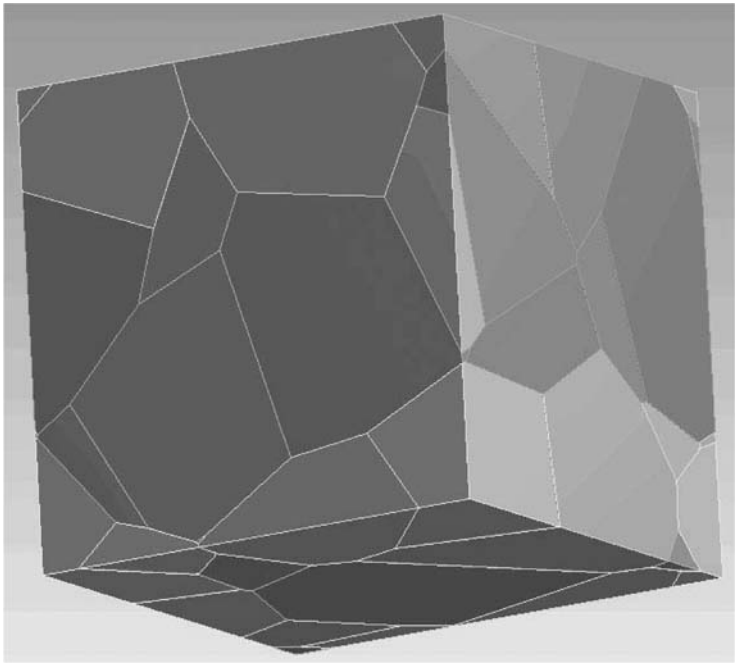
When N_{LCF} and N_{RF} are equal, the difference in the value of N_f obtained by Eqs (10.3) or (10.4) is a maximum, i.e. the life predicted by Eq. (10.4) is 50 % of that predicted by Eq. (10.3). When N_{LCF} and N_{RF} differ, the difference in life predictions is smaller.

10.3.4 Modelling microstructural evolution by ratcheting

Plastic ratcheting (directional accumulation of plastic strain) alters the microstructure of the rail and wheel material. Generally, the material is severely sheared in the direction of shear stress in the contact area. Very close to the surface, strains of the order of 1000 % are not uncommon; Fig. 10.2 shows how an unstrained matrix shears heavily. Plastic deformation strain hardens the material (Tyfour and Beynon, 1994b; Tyfour *et al.*, 1995). The material fails either by LCF or, more likely, by RF after accumulating a critical strain to failure (as described in the section above). Modelling crack initiation and wear requires representing the rail material microstructure realistically, subjecting it to passing wheel–rail contact stresses, determining plastic deformation and any strain-hardening and modification of the microstructure with each pass of the load. This process needs to be repeated over millions of wheel passes. As the material deforms, it deteriorates and fails, generating wear debris and initiating cracks.

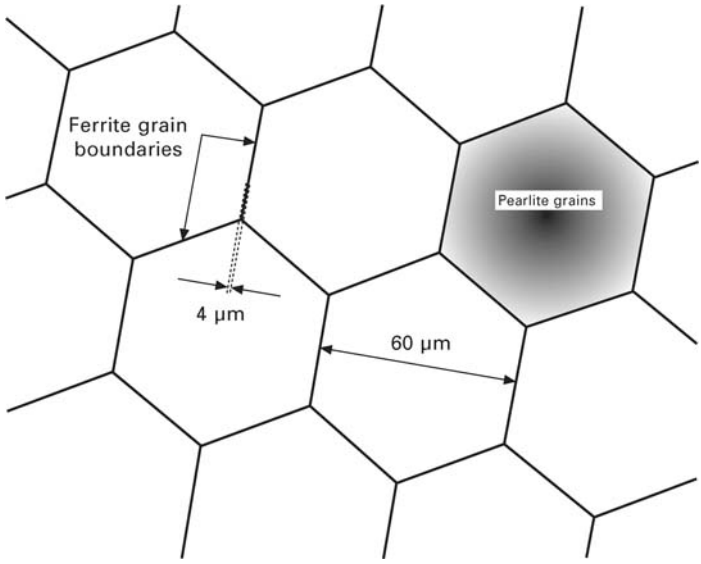
A popular technique for generating realistic random microstructures (polycrystals) in either 2D or 3D is to use the Voronoi method. A periodic 'unit cube' of Voronoi-generated microstructure, created using Qhull (Barber *et al.*, 1996), is shown in Fig. 10.12. Nygårds and Gudmundson (2002a, b) have studied pearlitic steel using this technique. The contact stresses determined using isotropic half-space assumption are not accurate in modelling the mesoscopic stress within the grain. An important consideration is the effect of grain anisotropy and the differing orientations of pearlite between, and even within, PA grains (Rodin, 1996; Li, 2002), and how crystallographic orientation affects crack tip displacement (Simonovski and Cizelj, 2007; Simonovski *et al.*, 2007). However, most of the research in this area has so far focused on undeformed materials in uniaxial loading; in contrast, the near-surface rail microstructure is heavily deformed and subject to a rapidly changing triaxial applied stress.

A simpler approach was taken by Kapoor and Franklin (2000) and Ringsberg *et al.* (2000b) where the isotropic half-space was divided up in layers and traversed by a line contact. The deformation of each layer was calculated using previous experimental results from a twin-disc machine. The effect of contact pressure, traction, etc. on wear rate and crack initiation time was



10.12 Periodic 'unit cube' of polycrystalline microstructure generated using the Voronoi method.

modelled. Over time this model has been progressively refined by incorporating microstructural detail, partial slip and thermal effects (Franklin *et al.*, 2001, 2003a, 2008; Alwahdi *et al.*, 2005; Franklin and Kapoor, 2007; Widiyarta *et al.*, 2008) and uses the data from SUROS twin-disc testing and metallurgical analysis described in the first half of this chapter (see also Garnham *et al.*, 2007; Garnham and Davis, 2008). A cross-section through the deforming material, parallel with the direction of traction and thus the direction of maximum strain, is modelled as a mesh of rectangular elements; material properties are assigned individually to the elements in a pattern representing the steel microstructure. Properties for PE ferrite and pearlite in BS11/220 grade rail steel were determined through nano-hardness measurements, and a hexagonal pattern of pearlite 'grains' (representing PA grains) with PE at the grain boundaries is simple to generate and configure (see Fig. 10.13). A series of orthogonal shear stress distributions can then be applied (either 2D, like twin-disc contact, or 3D, like wheel–rail contact) and the simulated microstructure accumulates plastic shear strain until the elements reach the limit of their ductility and start to fail. Failed elements at the surface may be removed as wear debris; subsurface, they are considered as sites of crack initiation, or as material through which crack propagation will be rapid.

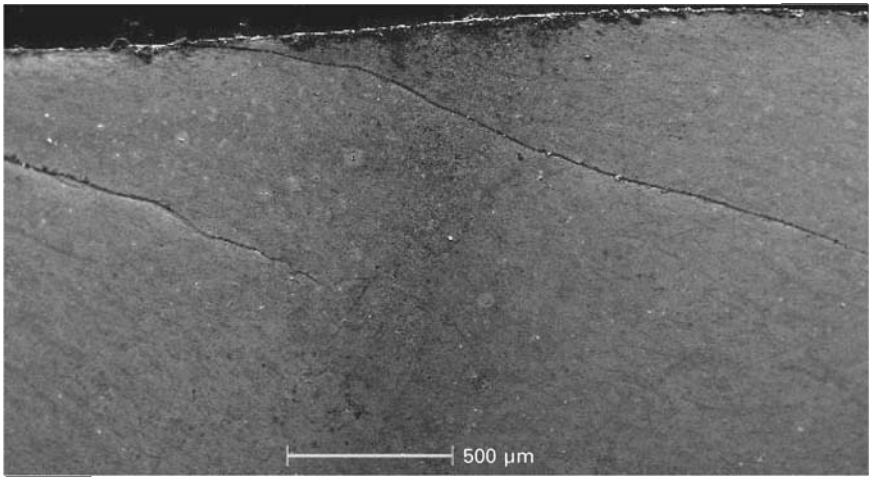


10.13 Representation of pearlitic microstructure as hexagonal prior-austenite grains of pearlite surrounded by pro-eutectoid ferrite at the grain boundaries. Grain size, grain boundary width, microstructure orientation and material properties of pearlite and ferrite can be configured easily.

10.3.5 Wear by ratcheting

Plastic ratcheting is the directional accumulation of plastic shear strain and manifests in two forms. In one, thin slivers of material are extruded from the edges of the contact region, and this can lead to 'lipping' in rails, for example. Technically this is wear, as the material has been removed from the contact region, even though the weight of the component has not changed. The material has just plastically flowed away from the contact area. Roughness of real surfaces produces very high contact pressures at asperities which drives plastic flow and lipping on a microscopic level (Kapoor and Johnson, 1992; Kapoor and Cocks, 1994; Kapoor and Johnson, 1995; Kapoor *et al.*, 1996; Kapoor, 1997, 1999). In a wheel-rail contact the nominal contact pressure itself may be large enough to drive ratcheting and cause lipping.

In the other manifestation of ratcheting, plastic shear deformation builds up beneath the surface. Figure 10.14 shows cracks in a rail cross-section which are aligned with the plastic shear deformation. The accumulation of plastic shear deformation leads to wear and crack initiation caused by ratcheting failure. This has been studied in controlled twin-disc tests of pearlitic rail steel specimens by Tyfour *et al.* (1995, 1996), by Su and Clayton (1997) and by Clayton and Su (1996), and more recently by Garnham *et al.* (2007) and Garnham and Davis (2008).



10.14 SEM image of a cross-section through a trafficked rail showing two 'head check' cracks following the angle of the plastically sheared rail steel.

In contrast to LCF, RF depends on the direction of strain accumulation, and twin-disc tests have confirmed that changing the rolling direction (and thus the direction of traction and therefore strain accumulation) reduces the wear rate for a period (Tyfour and Beynon, 1994b). The propagation of cracks can also be slowed by this reversal (Tyfour and Beynon, 1994a), which has the effect of 'unwinding' the accumulated plastic deformation. Driving wheels, for example, experience both braking and acceleration for proportions of the route, so there will be corresponding reversals of the direction of ratcheting strain accumulation. Particular track locations can sometimes be associated with either braking (approaching a station) or acceleration (leaving a station) so strain accumulation will be unidirectional. On curves there will also be a component of transverse strain accumulation in both wheels and rails. A model of wear by RF in a rail or a wheel requires an accurate determination of both reversing and accumulating components of plastic strains in all three directions over all cycles and is expected to be very involved. Modelling described in the previous section has involved simplifying the 3D strain distribution (as would be obtained in a wheel-rail contact) to 2D (as in a twin-disc machine).

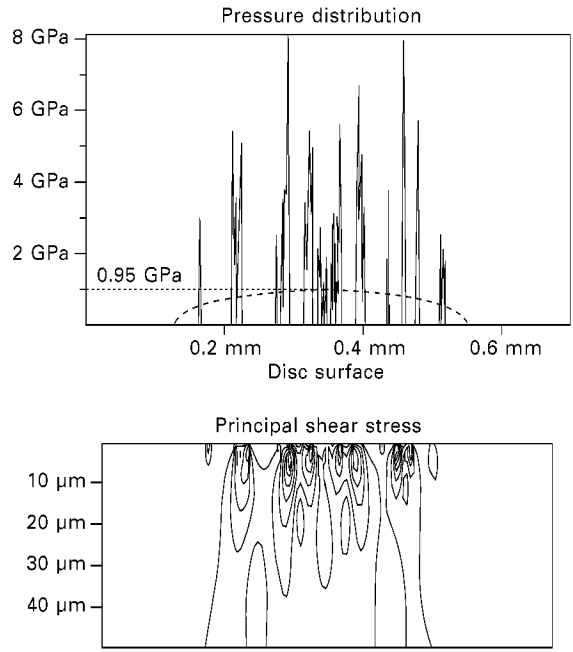
10.3.6 Rolling contact fatigue crack initiation

Surface-initiated cracks are associated with ratcheting in wheels (Ekberg and Sotkovszki, 2001) and rails (Eden *et al.*, 2005). RF has been used to predict tramway rail life by Beretta *et al.* (2005) and Desimone *et al.* (2005).

Both failure mechanisms (LCF and RF) were included in FE simulations of crack initiation in rails conducted by Ringsberg *et al.* (2000a), Ringsberg and Josefson (2001) and Ringsberg (2001).

Ratcheting simulations can be analysed to provide a plot of crack length against the number of wheel passes for the rail (Fletcher *et al.*, 2003); the number of cycles to crack initiation can be picked up from this plot for any desired crack length. In an earlier ratcheting model, where the rail steel was modelled as homogeneous, life until crack initiation was estimated by running the simulation until complete ratcheting failure of the material layer at the desired depth (Ringsberg *et al.*, 2000b).

Material within about 50 microns of the rail surface experiences very high stress caused by surface micro-roughness, i.e. asperities (not to be confused with corrugation) – see Fig. 10.15 (Kapoor *et al.*, 2002; Franklin *et al.*, 2003b). Corrugation also increases contact pressures and stresses,



10.15 Effect of surface micro-roughness on pressure distribution in a twin-disc contact. The dotted line represents the Hertz pressure distribution for smooth-surface contact, with peak pressure 0.95 GPa. The roughness causes highly localised contacts with pressures up to 8 GPa. (This is for a completely elastic calculation and, in practice, plasticity would reduce these pressure peaks.) The principal shear stress peaks at about 2 GPa within a few microns of the surface. (From Kapoor *et al.*, 2002)

increasing wear (Xie and Iwnicki, 2008). Although wheel and rail appear to make contact over a roughly elliptical area of 50–100 mm^2, the real area of contact is considerably less and supports a load of 5–20 tonnes, depending on the vehicle. The consequence of this is that some plasticity is inevitable, even for cases in which macro-contact is predicted to undergo shakedown, although in relatively lightly loaded cases, or for harder steels, this is confined to within about 15 microns of the surface.

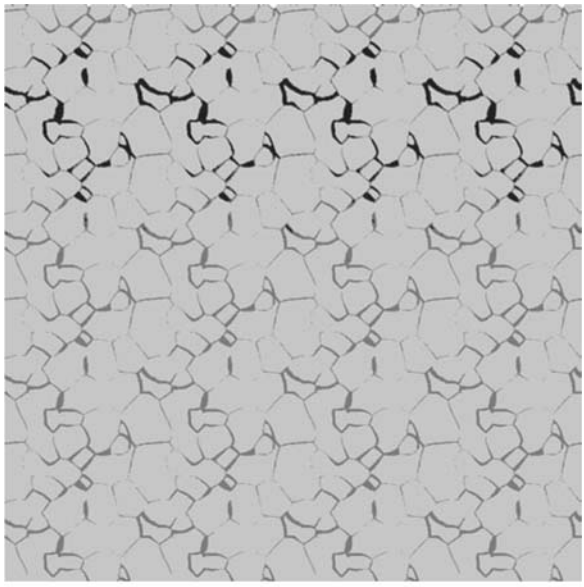
Repeated plastic deformation will lead eventually to RCF crack initiation and material loss through wear. How and when this happens depends critically on the microstructure. Most rails in use today are pearlitic steel which has a microstructure of lamellae (thin planes) of hard cementite (Fe_3C) sandwiched between ferrite. As the steel deforms plastically, the microstructure changes and the steel hardens. Eventually the regular planes of cementite bend or break up, and sometimes the material within 10–100 microns of the surface transforms into brittle 'white etching layer' (WEL, or sometimes WEC) in which almost none of the original pearlitic structure can be observed (Fig. 10.5). This forms as a result of plastic deformation and high temperatures (Baumann *et al.*, 1996), which can arise as a result of high friction and wheel slip (Fischer *et al.*, 1997; Ertz and Knothe, 2002). The buildup of dislocations in pearlite and the formation of WEL has been studied by Lojkowski *et al.* (2001a,b), and the effect of pearlite deformation on cracks in WEL has been studied using twin-disc testing by Carroll and Beynon (2007a,b).

10.3.7 Wear–fatigue interaction

The ratcheting model is useful for studying the balance between wear and crack initiation. Chapter 9 discusses wear–fatigue interaction further. To study the early stages of crack propagation, the hexagonal pattern of grains and grain boundaries is too regular, providing no obstacle to crack development. More realistic microstructures can be simulated, as in Fig. 10.16, but even here the 3D nature of the microstructure is not being taken into account, and cracks exist and grow in three dimensions.

10.4 Conclusions

Detailed metallurgical investigations of worn rails and (rail contact simulation) test discs have shown that RCF crack initiation and early propagation are strongly influenced by microstructure, for a given contact condition. The effect of different rail steel grades has been shown and modelled in two dimensions. Further RCF crack propagation becomes progressively more stress field driven. This work has shown that this early surface wear 'flaking' and RCF crack initiation and propagation cannot be fully modelled in two dimensions, given the granular nature of the formative and final microstructures.



10.16 Ratcheting wear simulation. (a) Internal representation, plastic shear deformation not shown. (b) Same simulation, but with deformation shown. Light grey indicates pearlite; dark grey indicates ferrite; black indicates material which has failed; white elements at the surface indicate where material has been removed as a result of wear.

It has also been shown that normally benign, ductile inclusions, when transversely strain-flattened near and at the rail surface, can also weaken the microstructure, and this effect within different rail steel grades requires further examination.

This has led to ongoing work examining three-dimensional aspects of this early crack growth. By integrating a fully three-dimensional model of microstructure into a computer simulation of plastic ratcheting, it is possible to study crack growth along grain boundaries in strained metallurgical structures. Crack initiation at inclusions can also be studied this way. These models can be used to study the transition from short crack growth accelerated by the severe plastic deformation near the rail surface to longer stress-driven crack growth.

10.5 Acknowledgements

EPSRC for financial support; from the Univ. Birmingham, Dept. Metallurgy & Materials, H. C. Eden for metallurgical analysis support; Dr Jiang Zia for nano-hardness test support; Mayorkinos Papalias for digital X-ray support and Ellies Muyapa for FIB support; from Corus Rail Technologies, John Hempshall and Jay Jaiswal for supply, machining & chemical analysis of rail materials; the University Loughborough for FIB test support; Mike Frolish (Univ. Sheffield, Dept. Engineering Materials) for twin-disc test support; Carillion & Bombardier for materials.

10.6 Nomenclature

BS11 British Standard specification for railway rails: Grade 220 ('normal grade')
CAT Computer-aided tomography
FE Finite element
FIB Focused ion beam
grade rail steel specification [the number (220/260/etc.) reflects the typical hardness]
HCF High cycle fatigue
HV1 Vickers hardness measurement with 1 kg load [HV0.1 for 0.1 kg load, etc.]
LCF Low cycle fatigue
PA Prior-austenite [usually referring austenite grain structure before cooling]
PE Pro-eutectoid [usually referring to ferrite at prior-austenite grain boundaries]
pearlite planar/lamellar sandwich (micro-)structure of ferrite and cementite (Fe_3C)

RCF Rolling contact fatigue
RF Ratcheting failure
SEM Scanning electron microscopy
SUROS Sheffield University ROlling Sliding twin-disc test machine
WEC white etching constituent (alternatively WEL for white etching layer)

10.7 References

Allery, M. B. P. (1986) British Rail Research, personal communication.

Alwahdi, F., Franklin, F. J. and Kapoor, A. (2005), The effect of partial slip on the wear rate of rails, *Wear*, **258**(7–8), 1031–37.

Barber, C. B., Dobkin, D. P. and Huhdanpaa, H. T. (1996), The Quickhull algorithm for convex hulls, *ACM Transactions on Mathematical Software*, **22**(4), 469–83, http://www.qhull.org.

Baumann, G., Fecht, H. J. and Liebelt, S. (1996), Formation of white-etching layers on rail treads, *Wear*, **191**, 133–40.

Beretta, S., Braghin, F., Bucca, G. and Desimone, H. (2005), Structural integrity analysis of a tram-way: load spectra and material damage, *Wear*, **258**, 1255–64.

Bower, A. F. and Johnson, K. L. (1991), Plastic flow and shakedown of the rail surface in repeated wheel–rail contact, *Wear*, **144**, 1–18.

Carroll R. I. and Beynon J. H. (2007a), Rolling contact fatigue of white etching layer: Part 1: Crack morphology, *Wear*, **262**(9–10), 1253–66.

Carroll R. I. and Beynon J. H. (2007b), Rolling contact fatigue of white etching layer: Part 2. Numerical results, *Wear*, **262**(9–10), 1267–73.

Clayton, P. and Su, X. (1996), Surface initiated fatigue of pearlitic and bainitic steels under water lubricated rolling/sliding contact, *Wear*, **200**, 63–73.

Desimone, H., Beretta, S. and Kapoor, A. (2005), Rail life prediction for tramcars under full slip regime, *Proceedings 11th International Conference on Fracture*, Turin, Italy, 20–25 March, available at: www.icf11.com/proceeding/EXTENDED/5422.pdf, accessed March 2009.

Dyson, I. N., Williams, J. A. and Kapoor, A. (1999), The effect of surface hardening on the elastic shakedown of elliptical contacts, *Proceedings of the IMechE: Part J, Journal of Engineering Tribology*, **213**, 287–98.

Eden, H. C., Garnham, J. E. and Davis, C. L. (2005), Influential microstructural changes on rolling contact fatigue crack initiation in pearlitic rail steels, *Materials Science and Technology*, **21**(6), 623–9.

Ekberg, A. and Sotkovszki, P. (2001), Anisotropy and rolling contact fatigue of railway wheels, *International Journal of Fatigue*, **23**, 29–43.

Ertz, M. and Knothe, K. (2002), A comparison of analytical and numerical methods for the calculation of temperatures in wheel/rail contact, *Wear*, **253**, 498–508.

Fischer, F. D., Werner, E. and Yan, W.-Y. (1997), Thermal stresses for frictional contact in wheel-rail systems, *Wear*, **211**, 156–63.

Fleck, N. A., Kang, K. J. and Ashby M. F. (1994), The cyclic properties of engineering materials, *Acta Metallurgy and Materials*, **42**, 365–81.

Fletcher, D. I. and Beynon, J. H. (1998), The influence of lubricant type on rolling contact fatigue of pearlitic rail steel, in Dowson, D., Priest M., Taylor, A., Ehret, P.,

Childs. T. H. C., Dalmaz, G., Berthier, Y., Flamand, L., Georges, J. M. and Lubrecht, A. A. (eds), *Lubrication at the Frontier, Proceedings 25th Leeds-Lyon Conference on Tribology, Paris 1998*, Elsevier, Amsterdam, the Netherlands, 299–310.

Fletcher, D. I. and Beynon, J. H. (2000a), Development of a machine for closely controlled rolling contact fatigue and wear testing, *Journal of Testing and Evaluation*, **28**(4), 267–75.

Fletcher, D. I. and Beynon, J. H. (2000b), Equilibrium of crack growth and wear rates during unlubricated, rolling-sliding contact of pearlitic rail steel, *Proceedings of the IMechE, Part F: Journal of Rail and Rapid Transit*, **214**(2), 93–105.

Fletcher, D. I. and Beynon, J. H. (2000c), The effect of intermittent lubrication on the fatigue life of a pearlitic rail steel in rolling-sliding contact, *Proceedings of the IMechE, Part F: Journal of Rail and Rapid Transit*, **214**(F3), 145–58.

Fletcher, D. I., Franklin, F. J. and Kapoor, A. (2003), Image analysis to reveal crack development using a computer simulation of wear and rolling contact fatigue, *Fatigue and Fracture of Engineering Materials and Structures*, **26**(10), 957–67.

Fletcher D. I., Hyde P. and Kapoor, A. (2008), Modelling and full scale trials to investigate fluid pressurisation of rolling contact fatigue cracks, *Wear*, **265**, 1317–24.

Franklin, F. J. and Kapoor, A. (2007), Modelling wear and crack initiation in rails, *Proceedings of the IMechE, Part F: Journal of Rail and Rapid Transit*, **221**, 23–33.

Franklin, F. J., Widiyarta, I. and Kapoor, A. (2001), Computer simulation of wear and rolling contact fatigue, *Wear*, **250–51**, 949–55.

Franklin, F. J., Chung, T. and Kapoor, A. (2003a), Ratcheting and fatigue-led wear in rail-wheel contact, *Fatigue and Fracture of Engineering Materials and Structures*, **26**(10), 949–55.

Franklin, F. J., Fletcher, D. I. and Kapoor, A. (2003b), Computer models for studying crack initiation, *Proceedings 6th World Congress on Railway Research*, Edinburgh, UK, 28–30 September, on CD.

Franklin, F. J., Garnham, J. E., Fletcher, D. I., Davis, C. L. and Kapoor, A. (2008), Modelling rail steel microstructure and its effect on wear and crack initiation, *Wear*, **265**, 1332–41.

Garnham, J. E. (1995), *The Wear of Bainitic and Pearlitic Steels*, PhD Thesis, University of Leicester, Leicester, UK.

Garnham, J. E. and Beynon, J. H. (1991), The early detection of rolling-sliding contact fatigue cracks, *Wear*, **144**, 103–16.

Garnham J. E. and Beynon, J. H. (1992), Dry rolling-sliding wear of bainitic and pearlitic steels, *Wear*, **157**(1), 81–109.

Garnham, J. E. and Davis, C. L. (2008), The role of deformed rail microstructure on rolling contact fatigue initiation, *Wear*, **265**, 1363–72.

Garnham, J. E., Franklin, F. J., Fletcher, D. I., Kapoor, A. and Davis, C. L. (2007), Predicting the life of steel rails, *Proceedings of the IMechE, Part F: Journal of Rail and Rapid Transit*, **221**, 45–58.

Ham, G., Rubin, C. A., Hahn, G. T. and Bhargava, V. (1988), Elasto-plastic finite element analysis of repeated two-dimensional rolling sliding contacts, *Journal of Tribology*, **110**, 44–9.

Johnson, K. L. (1962), A shakedown limit in rolling contact, *Proceedings, 4th US National Congress of Applied Mechanics*, Berkeley, CA, USA ASME, New York, USA, 971–5.

Johnson, K. L. (1992), The application of shakedown principles in rolling and sliding contact, *European Journal of Mechanics A/Solids*, **11**, 155–72.

Jones, C. P., Tyfour, W. R., Beynon, J. H. and Kapoor, A. (1997), The effect of strain hardening on shakedown limits of a pearlitic rail steel, *Proceedings of the IMechE, Part F: Journal of Rail and Rapid Transit*, **211**(2), 131–9.

Kapoor, A. (1994), A re-evaluation of the life to rupture of ductile metals by cyclic plastic strain, *Fatigue and Fracture of Engineering Materials and Structures*, **17**(2), 201–19.

Kapoor, A. (1997), Wear by plastic ratchetting, *Wear*, **212**, 119–30.

Kapoor, A. (1999), A mechanics model of erosion of ductile materials by backward extrusion, *Wear*, **233–235**, 182–90.

Kapoor, A. and Cocks, A. C. F. (1994), Wear through the plastic interaction of cylindrical asperities in sliding, *International Journal of Mechanical Sciences*, **36**(11), 1045–59.

Kapoor, A. and Franklin, F. J. (2000), Tribological layers and the wear of ductile materials, *Wear*, **245**, 204–15.

Kapoor, A. and Johnson, K. L. (1992), Effect of changes in contact geometry on shakedown of surfaces in rolling/sliding contact, *International Journal of Mechanical Sciences*, **34**(3), 223–39.

Kapoor, A. and Johnson, K. L. (1995), Plastic ratchetting as a mechanism of erosive wear, *Wear*, **186–187**(1), 86–91.

Kapoor, A. and Williams, J. A. (1994), Shakedown limits in sliding contacts on a surface hardened half-space, *Wear*, **172**, 197–206.

Kapoor, A. and Williams, J. A. (1996), Shakedown limits in rolling-sliding point contacts on an anisotropic half-space, *Wear*, **191**(1–2), 256–60.

Kapoor, A., Johnson, K. L. and Williams, J. A. (1996), A model for the mild ratchetting wear of metals, *Wear*, **200**(1–2), 38–44.

Kapoor, A., Franklin, F. J., Wong, S. K. and Ishida, M. (2002), Surface roughness and plastic flow in rail wheel contact, *Wear*, **253**, 257–64.

Kulkarni, S., Hahn, G. T., Rubin, C. A. and Bhargava, V. (1991), Elasto-plastic finite element analysis of three-dimensional pure rolling contact above the shakedown limit, Trans ASME, *Journal of Applied Mechanics*, **58**, 347–53.

Li, X.-D. (2002), Numerical correlation of material structure weaknesses in anisotropic polycrystalline materials, *Acta Mechanica*, **155**, 137–55.

Lojkowski, W., Djahanbakhsh, M., Bürkle, G., Gierlotka, S., Zielinski, W. and Fecht, H.-J. (2001a), Nanostructure formation on the surface of railway tracks, *Materials Science Engineering*, **A303**, 197–208.

Lojkowski, W., Millman, Y., Chugunova, S. I., Goncharova, I. V., Djahanbakhsh, M., Bürkle, G. and Fecht, H.-J. (2001b), The mechanical properties of the nanocrystalline layer on the surface of railway tracks, *Materials Science Engineering*, **A303**, 209–15.

Morales-Espejel, G. E., Kapoor, A. and Rodriguez-Sanchez, S. (1999), Shakedown in dry and lubricated surfaces, in Dowson, D., Priest, M., Taylor, A., Ehret, P., Childs, T. H. C., Dalmaz, G., Berthier, Y., Flamand, L., Georges, J. M. and Lubrecht, A. A. (eds), *Lubrication at the Frontier, Proceedings 25th Leeds-Lyon Conference on Tribology, Paris 1998*, Elsevier, Amsterdam, the Netherlands, 255–65.

Nygårds, M. and Gudmundson, P. (2002a), Micromechanical modeling of ferritic/pearlitic steels, *Materials Science and Engineering*, **A325**, 435–43.

Nygårds, M. and Gudmundson, P. (2002b), Three-dimensional periodic Voronoi grain models and micromechanical FE-simulations of a two-phase steel, *Computational Materials Science*, **24**, 513–19.

Ponter, A. R. S., Hearle, A. D. and Johnson, K. L. (1985), Application of the kinematical

shakedown theorem to rolling and sliding contact point, *Journal of the Mechanics and Physics of Solids*, **33**, 339–62.

Ponter, A. R. S., Afferrante, L. and Ciavarella, M. (2004), A note on Merwin's measurements of forward flow in rolling contact, *Wear*, **256**, 321–8.

Ringsberg, J. W. (2001), Life prediction of rolling contact fatigue crack initiation, *International Journal of Fatigue*, **23**, 575–86.

Ringsberg, J. W. and Josefson, B. L. (2001), Finite element analyses of rolling contact fatigue crack initiation in railheads, *Proceedings of the IMechE, Part F: Journal of Rail and Rapid Transit*, **215**, 243–59.

Ringsberg, J. W., Bjarnehed, H., Johansson, A. and Josefson, B. L. (2000a), Rolling contact fatigue of rails – finite element modelling of residual stresses, strains and crack initiation, *Proceedings of the IMechE, Part F: Journal of Rail and Rapid Transit*, **214**, 7–19.

Ringsberg, J. W., Loo-Morrey, M., Josefson, B. L., Kapoor, A. and Beynon, J. H. (2000b), Prediction of fatigue crack initiation for rolling contact fatigue, *International Journal of Fatigue*, **22**(3), 205–15.

Ringsberg, J. W., Franklin, F. J., Josefson, B. L., Kapoor, A. and Nielsen, J. C. O. (2005), Fatigue evaluation of surface coated railway rails using shakedown theory, finite element calculations, and lab and field trials, *International Journal of Fatigue*, **27**, 680–94.

Rodin, G. J. (1996), Eshelby's inclusion problem for polygons and polyhedra, *Journal of the Mechanics and Physics of Solids*, **44**, 1977–95.

Simonovski, I. and Cizelj, L. (2007), The influence of grains' crystallographic orientations on advancing short crack, *International Journal of Fatigue*, **29**, 2005–14.

Simonovski, I., Nilsson, K.-F. and Cizelj, L. (2007), The influence of crystallographic orientation on crack tip displacements of microstructurally small, kinked crack crossing the grain boundary, *Computational Materials Science*, **39**, 817–28.

Su, X. and Clayton, P. (1997), Ratchetting strain experiments with a pearlitic steel under rolling/sliding contact, *Wear*, **205**, 137–43.

Tyfour, W. R. and Beynon, J. H. (1994a), The effect of rolling direction reversal on fatigue crack morphology and propagation, *Tribology International*, **27**(4), 273–82.

Tyfour, W. R. and Beynon, J. H. (1994b), The effect of rolling direction reversal on the wear rate and wear mechanism of pearlitic rail steel, *Tribology International*, **27**(6), 401–12.

Tyfour, W. R., Beynon, J. H. and Kapoor, A. (1995), The steady state wear behaviour of pearlitic rail steel under dry rolling–sliding contact conditions, *Wear*, **180**, 79–89.

Tyfour, W. R., Beynon, J. H. and Kapoor, A. (1996), Deterioration of rolling contact fatigue life of pearlitic rail steel due to dry-wet rolling–sliding line contact, *Wear*, **197**, 255–65.

Widiyarta, I. M., Franklin, F. J. and Kapoor, A. (2008), Modelling thermal effects in ratcheting-led wear and rolling contact fatigue, *Wear*, **265**, 1325–31.

Williams, J. A., Dyson, I. N. and Kapoor, A. (1999), Repeated loading, residual stresses, shakedown, and tribology, *Journal of Materials Research*, **14**(4), 1548–59.

Wong, S. K. and Kapoor, A. (1996), Effect of hard and stiff overlay coatings on the strength of surfaces in repeated sliding, *Tribology International*, **29**(8), 695–702.

Wong, S. K., Kapoor, A. and Williams, J. A. (1997a), Shakedown limits on coated and engineered surfaces, *Wear*, **203–204**, 162–70.

Wong, S. K., Kapoor, A. and Williams, J. A. (1997b), Shakedown limits on coated surfaces, *Thin Solid Films*, **292**(1–2), 156–63.

Xie, G. and Iwnicki, S.D. (2008), Calculation of wear on a corrugated rail using a three-dimensional contact model, *Wear*, **265**, 1238–48.

Copyrighted Materials
Copyright © 2009 Woodhead Publishing Retrieved from www.knovel.com

11
Rail corrugation

S. L. GRASSIE, Stuart Grassie Engineering Ltd, Germany

Abstract: The quasi-periodic irregularity that appears on the running surface of rails, which is known as rail corrugation, is reviewed here. The review concentrates on particular characteristics of different types of corrugation, understanding its root causes and proposing practical methods for avoiding or alleviating the phenomenon. Different types of corrugation can be identified by their 'wavelength-fixing' and 'damage' mechanisms. Six different types of corrugation are thus identified, for four of which wear is the 'damage mechanism'. The most common wavelength-fixing mechanism is the P2 resonance of a vehicle's unsprung mass on the track stiffness, which is significant for three of the six types of corrugation. An understanding of this phenomenon and development of efficacious treatments have advanced significantly in the last 15 years as a result of research having been applied and the effects accurately and comprehensively monitored.

Key words: rail corrugation, railways, corrugation, vehicle–track interaction, wheel–rail interface, friction, rail maintenance, rail re-profiling, rail corrugation measurement.

11.1 Introduction

Corrugation of rails occurs in one manifestation or another, and often in several, on nearly every railway system. Its removal by grinding provides work for several international grinding companies and was already costing the industry worldwide at least US $\$10^8$ per annum in the early 1980s.[1] The phenomenon is a more or less periodic irregularity of the running surface which is often visible to the naked eye. It gives rise to high dynamic loads between wheel and rail, degradation of the ballast and other track components, and noise. The last is particularly annoying with the shorter wavelength corrugations which excite vibration in the audible frequency range. Corrugation is therefore a phenomenon of great practical concern to the railway industry: substantial savings could in principle be made if it could be prevented rather than simply treated, and if there existed guidelines which could be followed to make corrugation less likely on new railway lines. The purpose of this chapter is to review the topic of rail corrugation in this light, and particularly

This chapter is based on the paper entitled 'Rail corrugation: characteristics, causes and treatments' by S L Grassie and J Kalousek, published in the *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* (ISSN 0954-4097), 1993, Vol. 207, No. F1, pp. 57–68, DOI 10.1243/PIME_PROC_1993_207_227_02

to consider what has been learned from work which has been done in this area since the 1970s. Six different types of corrugation with significantly different characteristics are identified, the mechanism giving rise to each type of corrugation is discussed, and recommendations are made as to what has been done or could be done to treat the malady most effectively, and where possible prevent it.

The chapter is a revised and updated version of a review that was published in 1993 by the Institution of Mechanical Engineers.[2] The intervening 15 years have enabled several holes to be filled in understanding and a few things to be simplified. It has also allowed some results to be reported of applying a variety of treatments including some suggested in reference 2. There are few areas in which the practical recommendations and conclusions of that review have been found wanting.

It is fundamentally important to appreciate that 'corrugation' is not a single phenomenon with a single cause and a single solution or treatment: considerable confusion can otherwise arise. This may be apparent from a bibliography on rail corrugation which was compiled by the Research Department of the then British Railways (BR) in 1961[3] which contains references from almost every year in the period 1904–1960. The flow of literature has not abated since then and shows no signs of doing so. Considerable contradiction apparently exists, particularly in the early literature, regarding possible causes or contributory factors: for example, some investigators believed that 'soft' rails were responsible whereas others thought the opposite; in some places wear was held to be the cause whereas elsewhere the cause was thought to be 'flow' of the rail. As early as 1922 a paper appeared which was optimistically entitled 'How to avoid or overcome rail corrugation':[4] had this indeed been possible, both a great deal of subsequent research and also the present chapter would be unnecessary. Since BR's bibliography appeared, the structured approach which has developed to understand the phenomenon has borne fruit in the realization that corrugation is a family of phenomena with superficially similar features: primarily the periodically irregular running surface at roughly similar wavelengths. Although some differentiation is often adopted between 'short' and 'long' wavelength corrugation (30–100 mm and 100–300 mm approximately), classification by wavelength alone is insufficient. Corrugations differ in the detailed appearance of the rail, the root cause and, consequently, also in the treatment which can most effectively be adopted. These factors are discussed here.

11.2 Classification of corrugation

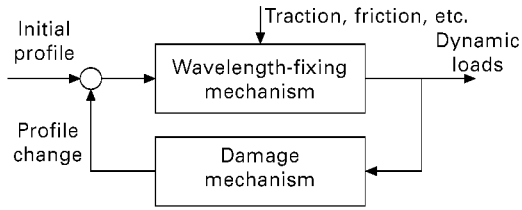
A general mechanism of corrugation formation, which is used in one form or another by most of those working in this area, can usefully be regarded as comprising two features: a so-called 'wavelength-fixing mechanism' and

a 'damage mechanism', as illustrated in the basic feedback loop of Fig. 11.1. The rail is initially uncorrugated, but the profile has components of roughness at all wavelengths, and inevitably some irregularities are larger than others. This initial roughness in combination with other factors such as traction, creep and the friction characteristic at the wheel–rail contact excites dynamic loads which cause damage of some type, thereby modifying the initial profile. Provided that sufficient trains pass over the site at a similar speed, the wavelength at which the dynamic load varies is similar from one train to another, and corresponds to the specific wavelength-fixing mechanism. The same irregularities excite each train, and the damage caused by one train tends to exacerbate vibration of subsequent trains, leading to further damage at a specific wavelength. The dynamic loads may be either normal to or in the plane of the wheel–rail contact.

A classification of corrugation is made in Table 11.1 by wavelength-fixing and damage mechanisms, giving rise to six types of corrugation:

- 'heavy-haul' corrugation;
- 'light-rail' corrugation;
- other P2 resonance corrugation;
- 'rutting';
- 'pinned-pinned resonance' corrugation;
- trackform-specific corrugation.

Wear is the most common damage mechanism. Other damage mechanisms are plastic flow and plastic bending. These are significant in the initial stages of corrugation formation only when the P2 resonance is the wavelength-fixing mechanism and gives rise to what are termed here 'heavy-haul' and 'light-rail' corrugation respectively. Plastic flow can become significant when a rail has become corrugated by another mechanism, as is shown in examples here. The nomenclature of Table 11.1 is intended to be useful while not substantially altering terms that have previously been used more loosely. The term 'roaring rails' appears at an early date in BRs bibliography[3] and is (with the benefit of hindsight) pertinent primarily to what is called here 'pinned-pinned resonance' corrugation for reasons that are clarified in Section 11.8. Although this is the most common type of 'short wavelength'



11.1 Corrugation mechanism.

Table 11.1 Characteristics, causes and treatments of rail corrugation

	Type	Wavelength-fixing mechanism	Where?	Typical frequency [Hz]	Damage mechanism	Relevant figures	References	Treatments demonstrably successful	Should be successful
1	Heavy haul	P2 resonance	Straight track or curves	50–100	Plastic flow in troughs	2,3	[35–37]	Hard rails	Reduce cant excess when corrugation on low rail
2	Light rail	P2 resonance	Straight track or curves	50–100	Plastic bending	4	[38]	Increase rail strength and EI	Reduce unsprung mass
3	Other P2 resonance	P2 resonance	Straight track or high rail in curves	50–100	Wear	5–7	[22,39,40]	Hard rails, highly resilient trackforms	Reduce unsprung mass
4	Rutting	2nd torsional resonance of driven axles	Low rail of curves	250–400	Wear	8–10	[4,5,22–34]	Friction modifier, hard rails, reduce cant excess, asymmetric profiling in curves	Reduce applied traction in curving, improve curving behaviour of vehicles dynamic vibration absorber
5	Pinned-pinned resonance (roaring rails)	Pinned-pinned resonance	Straight track, high rail of curves	400–1200	Wear	5,11	[6–21,47,51]	Hard rails, control friction	Increase pinned-pinned frequency so that corrugation would be < 20 mm wavelength
6	Trackform-specific	Trackform specific	Straight track or curves	–	Wear	12–14	[22,41]	Hard rails, friction control	Avoid 'peaky' resonances, improved steering

(30–100 mm) corrugation, it is nevertheless clear from Table 11.1 that other types of corrugation have similar wavelength. The following information is also tabulated:

- characteristic corrugation frequency or range of frequency; the corrugation wavelength is the quotient of predominant vehicle speed v and the characteristic frequency f, that is:

$$L = \frac{v}{f} \qquad [11.1]$$

- some references to the particular type of corrugation in the literature;
- reference to relevant figures in this paper.

Rolling contact fatigue was considered in reference 2 to be an independent damage mechanism giving rise to an independent type of corrugation. This distinction is not made here because there is more evidence that RCF is a secondary damage mechanism associated with heavy-haul corrugation than the primary mechanism that causes a change in profile of the rail as a result of dynamic loading, as in Fig. 11.1. To some extent this is an academic question, since both corrugation and RCF on heavy-haul railways have become very much better controlled since 1993 as a consequence of routine, preventative rail grinding.

A category of 'other P2 resonance' corrugation is proposed here to deal with a surprisingly common type of corrugation whose damage mechanism is wear and whose wavelength-fixing mechanism is the P2 resonance. This is discussed in Section 11.5.

Shortly before the review paper of 1993 appeared, work had appeared independently from groups in Europe and North America on the phenomenon of what was termed in ref [2] 'booted sleeper' corrugation. Subsequent work suggests that the booted sleeper trackform is not alone in having an occasional propensity to corrugate, so the wider category of 'trackform-specific' corrugation is proposed here. Characteristics that one should avoid in selection of a trackform and possible reasons why the booted sleeper trackform, for example, corrugates in some circumstances and not others are suggested.

11.3 Heavy-haul corrugation

11.3.1 Characteristics

This type of relatively long wavelength corrugation is associated particularly with heavy-haul railways, where there are high wheel loads (in excess of 15 tonnes), unit trains and low, consistent speeds. Examples are shown in Figs 11.2 and 11.3. Their cause (both the wavelength-fixing mechanism and

11.2 'Heavy haul' corrugation.

11.3 'Heavy haul' corrugation on low rail of curve on a mixed traffic line.

the damage mechanism) was probably understood earlier than that of any other type of corrugation. Research into their cause and possible treatments was undertaken primarily in Australia, see for example.[5–7] The corrugations propagate from welds, joints and other discrete railhead irregularities. The Australian research suggested that, although they are not restricted to curves,

they do occur preferentially on the 'high' rail in curves (as in Fig. 11.2) regardless of track super-elevation or operating speeds.[6] It appears that this type of corrugation also occurs on the low rail of mixed traffic lines (Fig. 11.3). The typical wavelength of 200–300 mm corresponds to a frequency of about 30 Hz for the low speeds of loaded trains. There is gross plastic flow in the corrugation troughs, 'mushrooming' of the railhead and, where this occurs on the low rail, significant plastic flow to the inside of the curve. At corrugation sites the ballast around the sleepers is much disturbed, appearing like white pebbles recently swept up on the dark beach of the remaining ballast: this is a common feature of all types of corrugation on ballasted track, or indeed at any site where there are high dynamic loads and excessive sleeper vibration.

11.3.2 Cause

The pertinent damage mechanism is clearly gross plastic flow as a result of excessive contact stresses. The wavelength-fixing mechanism is resonance of the unsprung mass (in this case usually the unsprung mass of heavily loaded wagons) on the track stiffness. This loaded track resonance gives rise to what are now commonly known as P2 forces, following British Rail's terminology, for example.[8] The periodically high dynamic loads excited by railhead irregularities are superposed on the high static loads, giving rise to periodic plastic flow. Consistent types of vehicle (and hence resonant frequency) and consistent speeds of loaded trains give rise to a fairly constant corrugation wavelength.

Whether or not corrugation of this type occurs preferentially on high or low rail in a curve depends on the tangential load. Plastic flow occurs more readily the higher the traction ratio and normal contact stress. For a vehicle curving at or above balance speed the traction ratio is greater at the high rail than on the low rail for almost all circumstances. It is accordingly reasonable to expect that plastic flow would occur more readily on the high rail. However, for a bogie curving with cant excess, the traction ratio on the low rail wheel increases significantly. Moreover, this occurs largely because the leading wheelset moves towards the high rail and the trailing wheelset moves in towards the low rail, thereby increasing the angle of attack. Contact on the low rail at the leading wheelset is towards the field side of the wheel, which is often slightly convex, whereas contact on the high rail is in the flange root of the wheel, which is concave. Contact stresses accordingly tend to be higher on the low rail, giving rise to plastic flow. The direction of plastic flow, towards the inside of the curve (Fig. 11.3), is consistent with lateral forces arising from the angle of attack of the leading wheelset. Where corrugation occurs by this mechanism on the low rail of mixed traffic lines, it is often exacerbated by the fact that the lower speed traffic is higher axle

load freight traffic, whereas the lines have been canted for a few high-speed passenger trains.

11.3.3 Treatment

Practical treatments for this type of corrugation are to select a sufficiently hard rail steel and to ensure that railhead irregularities are sufficiently small to reduce the contact force below the level which can be borne by the rail. Accordingly, the treatment that has been developed for existing corrugation is to grind the rail and straighten welds, thereby minimizing the initial roughness which excites vibration of the vehicle on the track stiffness and gives rise to high dynamic loads. Typically, rails are ground regularly to maintain low dynamic loads and allow high static loads to be carried. In principle the dynamic loads could also be reduced using primary suspension on wagons and bogie-hung traction motors on locomotives to reduce unsprung mass, but these are expensive modifications whose benefits would be difficult to justify by reducing corrugation alone.

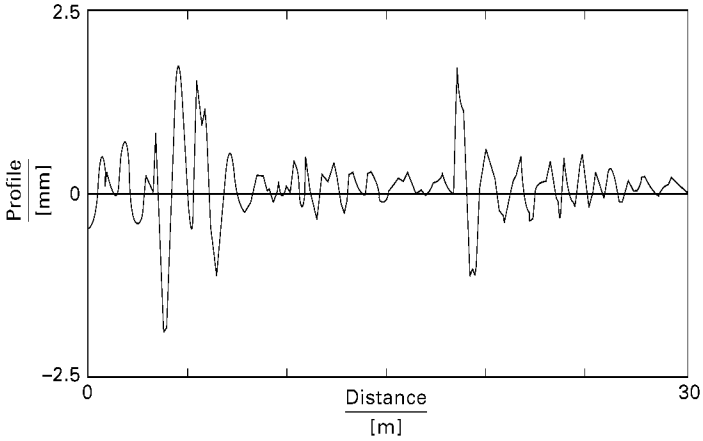
Initially the rail should be selected with adequate yield and tensile strengths to resist corrugation for typical railhead irregularities. Marich and Maass have tabulated the appropriate yield and tensile strengths to avoid corrugation by this mechanism,[7] and a design method has been proposed to select rail accordingly for particular railhead irregularities.[9]

Some difficulties can arise with aluminothermic welds and occasionally also with flash-butt welds, particularly on head-hardened rail, because the welds are softer than the parent material; these soft areas can dip under heavy wheel loads.[10] Improvements in welding are accordingly desirable to reduce this differential hardness. On mixed traffic lines where this type of corrugation is caused by low speed, loaded traffic, it would be desirable to cant curves for this traffic rather than for higher speed, usually passenger, traffic. Cant deficiency is a far more desirable condition than cant excess.

11.4 Light-rail corrugation

11.4.1 Characteristics

A corrugation which bears many similarities to that experienced on heavy-haul railways was identified in the 1980s on track operated by Australian National Railways (AN),[11,12] but is also present elsewhere. The corrugation propagates from welds and has a wavelength in the range 500–1500 mm: the profile of a badly corrugated rail is shown in Fig. 11.4. The typical wavelength of 700 mm again corresponds to an excitation frequency of about 30 Hz at predominant vehicle speeds (which are substantially higher than on heavy-haul railways). Amplitudes of 1 mm or more are often measurable



11.4 Measurement of 'light-rail' corrugation.

on both railhead and railfoot. The phenomenon is associated particularly with relatively light, 47 kg/m rail, although some is present also on older 53 kg/m rail; 60 kg/m rail has to date been unaffected. Gross plastic flow of the railhead is uncommon.

11.4.2 Cause

The wavelength-fixing mechanism for light-rail corrugation is the same as that for heavy-haul corrugation: resonance of the vehicle's unsprung mass on the track stiffness, excited primarily by irregular welds. In this case the critical vehicles were identified as locomotives, which had relatively high static wheel loads (about 11 tonnes) and high unsprung mass.[12] The damage mechanism was shown to be the yield of the rails in bending, which gave rise to full section deformation (or 'crippling') of the rail. The rail steel used at that time for the 47 kg/m and 53 kg/m rails had a relatively high tensile and low yield strength, so that the load required to cause plastic bending was lower than that required to cause plastic deformation of the work-hardened surface layer. For most types of rail steel, plastic bending occurs at much higher loads than plastic flow. The relatively wide range of wavelengths arose from the different speeds and types of vehicle: where these were consistent, wavelengths are also consistent.

11.4.3 Treatment

The treatment which has been adopted by AN for this type of corrugation has been successful to date in eliminating existing corrugation and preventing its recurrence. The principal components of the treatment are the same as

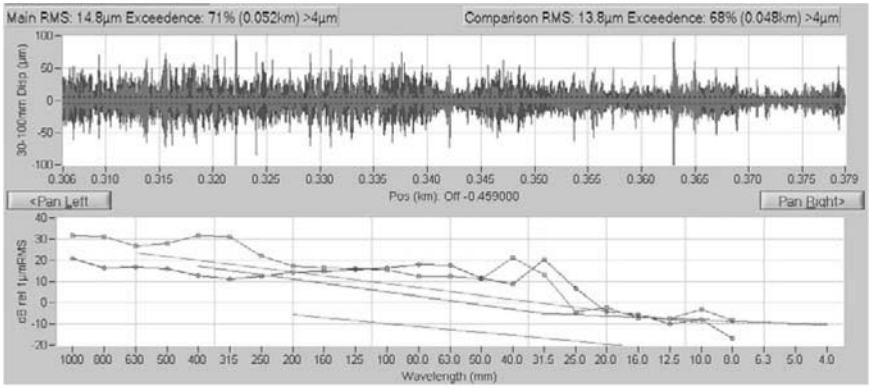
those for corrugation on heavy-haul track: reduction of railhead irregularities, particularly at welds, to a sufficiently small level that the sum of dynamic and static loads is insufficient to yield the rail in bending.[12] For the conditions on AN, it was sufficient to reduce the effective ramp irregularity to an amplitude of about 7 milliradians.[13] To allow higher speeds to be attained, locomotives with a lower unsprung mass have been acquired, thereby reducing the P2 force for a particular irregularity. The rail steel now used has higher yield strength and greater flexural rigidity, which gives greater resistance to this type of corrugation.

11.5 Other P2 resonance corrugation

11.5.1 Characteristics

Evidence from field measurements and observations undertaken since the 1990s indicates that the P2 resonance is significant not just in heavy-haul and light rail corrugation but also as one of the most prominent and frequent causes of corrugation on a wide variety of railways. For example, the longer wavelength corrugation in Fig. 11.5 has resulted from excitation of the P2 resonant frequency, and this is also the source of corrugation at 300–400 mm in the measurements (from the same railway) shown in Fig. 11.6. Also, while it was proposed in reference 2 that corrugation on tramways was the result primarily of excitation of the fundamental torsional resonance of axles, it appears that the most common source of corrugation on tram systems is in fact the P2 resonance, e.g. Fig. 11.7.

11.5 'P2 resonance' and 'pinned-pinned' resonance corrugation on a metro system



11.6 Measurements of short wavelength corrugation (30–100 mm) from two sites on a metro system.

11.7 'P2 resonance' corrugation on a tram system.

Corrugation can be particularly severe in those locations where the torsional resonance of wheelsets is also excited and occurs at substantially the same frequency. This is often the case in tight curves on metro railways, as in the case of the rapid transit system in Paris (RATP) cited by Tassilly and Vincent,[14] although this type of corrugation is also found in straight track. On

metros it can be a particular problem because the relatively low frequency 'rumble' is transmitted well into buildings and sometimes amplified by their own natural resonances that occur at similar frequencies.

11.5.2 Causes

The P2 resonance is, by definition, the wavelength-fixing mechanism for this type of corrugation and wear is the damage mechanism. If there are circumstances in which the fundamental torsional resonance can be excited and if the resonant frequencies coincide, this exacerbates the effects of the P2 resonance alone.

Empirical evidence from five transit systems cited in references 15 and 16 suggested that corrugation resulting from the P2 resonance on 'direct fixation' fastening systems was absent if the track support was more resilient than about 30 MN/m per fastening assembly. This has been supported by subsequent measurements, which demonstrate (for example) that corrugation can be quite severe on non-ballasted trackforms that use a railpad as the only resilient layer. Corrugation arising from the P2 resonance is shown in Fig. 11.5 for a trackform in which timber sleepers embedded in concrete provide the main resilient layer, while measurements on the same metro system (Fig. 11.6) show that longer wavelength corrugation associated with the P2 resonance is absent on a commercially available trackform on the same railway whose support stiffness is about 5 MN/m per assembly.[17] In this trackform the rail itself is held in rubber wedges, which provide the ideal combination of high vertical resilience and high lateral stiffness.

11.5.3 Treatments

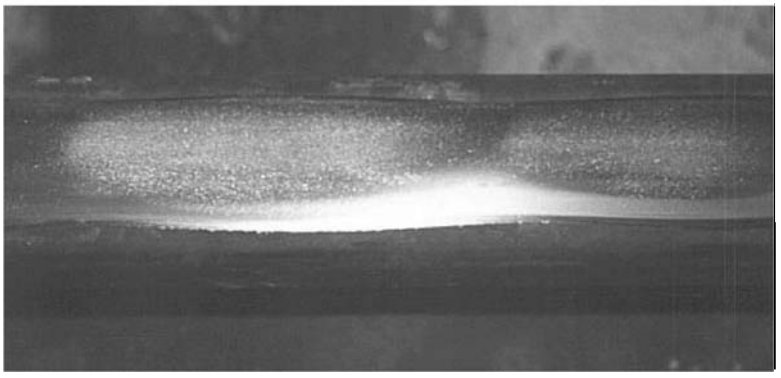
The most effective method of preventing this type of corrugation appears to be to use a support of sufficient resilience, as demonstrated by measurements of Fig. 11.6. If the track already exists the same measures that are effective in reducing other types of P2 resonance corrugation (Sections 11.3.3, 11.4.3) should also work here i.e. reducing the irregularities that trigger such corrugation and reducing unsprung mass. Since wear is the damage mechanism, hard rails should also help reduce the problem provided sufficient care is taken with profiling to ensure other problems are not introduced.

11.6 Rutting

11.6.1 Characteristics

'Rutting' is the name that has been given to the type of corrugation that occurs primarily on the inside rail in curves[2] e.g. Figs 11.8 and 11.9. Rutting

11.8 'Rutting' corrugation on a mainline railway.



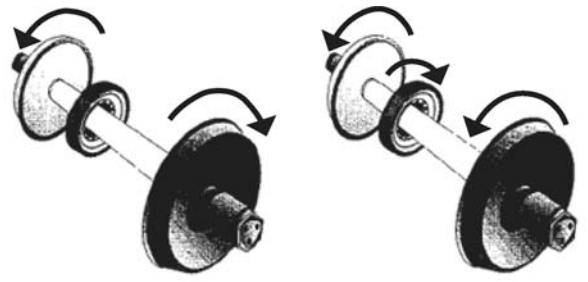
11.9 'Rutting' corrugation on a metro.

may also occur in straight track where traction or braking is severe. Discrete irregularities such as welds and joints (e.g. Fig. 11.8) are common 'triggers' for corrugation and often fix the position of corrugation along the rail. Rutting usually has an extremely uniform wavelength and appearance, with very little modulation. The corrugation can develop extremely quickly to a depth of tenths of a millimetre. Signs of wear debris are sometimes apparent and, for well-developed corrugation, also plastic flow. These characteristics are apparent in Fig. 11.9.

11.6.2 Cause

The wavelength-fixing mechanism for the most common type of corrugation that occurred on the inside rail in curves in metro systems was found in a study undertaken in North America to be the second torsional resonance of driven wheelsets.[15,16] The difference between the first and second torsional resonance is illustrated in Fig. 11.10. Since it has been found in subsequent work that the frequency associated with low rail corrugation in curves is commonly that of the second torsional resonance, it is proposed that the term 'rutting' be used to refer to this form of corrugation. Corrugation is also associated with the first torsional resonance, although less commonly than appeared from the literature in 1993 (see Section 11.5).

'Rutting' corrugation occurs where the traction ratio (the ratio of tangential to normal force, T/N) is close to the friction limit so that one wheel slips and drives the opposite wheel in the wheelset in a stick-slip oscillation: the 'traction, friction' input to the wavelength-fixing mechanism in Fig. 11.1 is accordingly critical. It is usually the wheel on the leading wheelset in a bogie on the high rail that slips, since the tangential force from curving is greatest on this wheel and the coefficient of friction is often less (because the gauge face of the high rail is lubricated to reduce wear, and there is usually some migration of lubricant). Applied traction increases the tangential force on the outer leading wheel and reduces it on the inner leading wheel, thereby increasing the difference in tangential force across the wheelset and exacerbating the stick-slip oscillation. The trailing wheelset 'steers' well for all cases, so there is a relatively low traction ratio at both outer and inner wheels. The damage mechanism for rutting is clearly wear, which is quasi-periodic with high peaks corresponding to the slip phase of a roll-slip oscillation. There has been considerable work on rutting corrugation since



11.10 Torsional modes of vibration of a wheelset (from reference[15]).

its identification in the mid-1990s as a significant problem and a separate phenomenon from 'roaring rails', e.g. references 18–24.

11.6.3 Treatment

A particularly successful treatment of this type of corrugation has proved to be so-called 'friction modifier', whose success has been well documented in tests on various metros, e.g. references 20–24. This substance is applied to wheel–rail contact, on either the wheel (usually with a replaceable 'stick') or the rail (usually in water-based fluid applied at the field side of the rail). It controls the coefficient of friction to a value of about 0.35, which is more than sufficient for traction and braking but limits several types of surface damage on both wheels and rails. The coefficient of friction also increases slightly with sliding speed, which in principle eliminates the conditions for stick-slip oscillation.

Grinding is necessary to remove this type of corrugation but does not by itself provide a means to prevent its recurrence. Harder rails should also limit wear and therefore the rate at which such corrugation develops. Low rail corrugation occurred in tests undertaken at the Facility for Accelerated Service Testing (FAST) in the USA, apparently as the result of a mechanism involving torsional resonance in combination with a peak in the vertical contact force associated with a sleeper resonance. In these tests lubrication of the gauge corner of the high rail was found to be the most successful treatment in reducing low rail corrugation, with some benefit from harder rail.[25]

Improving the ability of a bogie to steer around a curve, by whatever means, should reduce the propensity to form corrugation in curves. This can be done with more resilient yaw suspension or a steering bogie. Similarly, it would be desirable to use bogies in which wheelsets are not coupled through the traction motor. Asymmetric profiling of rails is also beneficial, essentially in order artificially to increase the rolling radius difference between wheels running on the outer and inner rails. However, this must be done in full awareness of the fact that it conflicts with the most widely used treatment of the rolling contact fatigue (RCF) defect known as 'gauge corner cracking' (GCC): re-profiling of rails to encourage wheels to run along the gauge corner of the high rail increases GCC, while reducing the tangential force required for curving reduces GCC.

Since the wavelength-fixing mechanism for rutting appears to be the second torsional resonance of driven wheelsets, the possibility exists in principle to develop a torsional dynamic vibration absorber to attenuate this resonance, as proposed in ref [15]. Although this remains a practical possibility, the benefits would be limited unless the treatment was adopted over a complete train fleet.

11.7 Roaring rails/'pinned-pinned resonance' corrugation

11.7.1 Characteristics

The term 'roaring rails' is used here to refer to the type of corrugation which is now commonly associated with high-speed, mainline track with relatively light axle loads (that is usually less than about 20 tonnes: a 10 tonne wheel load). It occurs predominantly on tangent track or in gentle curves in which there is no contact between the wheel flange and the gauge side of the rail. Examples from a metro system and from a mainline railway are shown in Figs 11.5 and 11.11, respectively. On the metro, this corrugation is superposed on a longer wave corrugation. A measurement of the shorter wavelength component (30–100 mm) of the corrugation is shown in Fig. 11.6 from two other sections of track on the same metro line.

At any one site roaring rail corrugation appears quite uniform to the naked eye, but measurements indicate a rather broad spectrum of corrugation wavelengths.[26] Well documented experimental evidence from field tests in the UK indicates that typically a martensic 'white phase' develops incrementally into one or two fairly continuous running bands on the railhead, and is then

11.11 'Pinned-pinned resonance' corrugation on mainline railway.

worn away periodically to leave the corrugations: these appear as relatively bright peaks and rather dull troughs.

It has been observed in the UK and elsewhere that Acid Bessemer rail steel corrugates more quickly than Open Hearth rail steel;[27,28] corrugations develop at an intermediate rate on the Oxygen rail steel which is now used in UK. Although it had appeared that there was no consistent variation in rates of corrugation formation on rails of different hardness,[29] some recent measurements suggest that harder rails corrugate significantly less quickly.[30] On main lines in the UK, where the speeds are typically 160–200 km/h, the severity of corrugation often varies visibly at sleeper pitch, being more severe on the approach to sleepers than in between.

11.7.2 Cause

Observations made in the field indicate that the damage mechanism for roaring rail corrugations is wear. This is consistent with the fact that the vertical wheel–rail force is least towards the corrugation troughs: indeed, a corrugation depth of 0.1 mm is sufficiently severe to cause loss of contact.

Although there was some uncertainty in 1993 about the wavelength-fixing mechanism for this type of corrugation, it is now clear that this is the rail's so-called 'pinned–pinned resonance'. This has been demonstrated primarily in the work undertaken at TU Berlin in Germany and Chalmers in Sweden.[31–34] At the pinned–pinned resonance the rail vibrates as if it were almost pinned at the sleepers, so that a wavelength is double the sleeper pitch. The frequency of this resonance is approximately

$$f = \frac{\pi}{2L^2}\sqrt{\frac{EI}{m}}\left(1 - \frac{1}{2}\left(\frac{\pi r_g}{L}\right)^2\left(1 + \frac{2(1+\upsilon)}{K}\right)\right) \qquad [11.2]$$

where m and EI are the rail's mass per unit length and bending stiffness, respectively, L is the sleeper spacing, r_g is the radius of gyration, υ is Poisson's ratio and K (≈ 0.34) is the shear constant of the cross-section. Clearly the wavelength is longer the greater the sleeper spacing and the lower the rail stiffness. A typical frequency is about 770 Hz for 56 kg/m rail and a 0.7 m sleeper spacing, e.g. Fig. 11.11, or 1200 Hz for 60 kg/m rail and 0.6 m spacing. The corrugation in Fig. 11.5 was on a metro railway with widely spaced supports and relatively light rail section, for which the pinned-pinned resonance frequency is 460 Hz, thereby giving a corrugation of similar appearance and wavelength to that on the mainline railway (Fig. 11.11) where trains are travelling close to 200 km/h. Corrugation occurs not because the resonance is excited, but rather because the supports, which are the nodes of the resonance, appear dynamically stiff (as noted in reference 35) and give a high dynamic contact force.

The pinned-pinned resonant frequency is significantly higher than those of other wavelength-fixing mechanisms (Table 11.1) and accordingly gives the shortest wavelength for a given train speed. An extremely practical question that arises is why corrugation is seldom measured with a wavelength of less than 30 mm and rarely with a wavelength of less than 20 mm. Irregularities with a wavelength of less than about 20 mm, which often result from rail grinding, are indeed usually worn out by traffic.[36] The answer to this question is found in the influence of contact mechanics between wheel and rail, from which it can be shown that non-steady-state contact mechanics (NSSCM) act essentially as an extremely powerful filter that attenuates wavelengths of less than 20 mm.[37]

Roaring rail corrugation is exacerbated by factors that increase the propensity of the wheel to slip and wear those areas of the rail which become corrugation troughs. Such factors include high vertical dynamic loads and high lateral and longitudinal creep, such as occurs when wheelsets are misaligned in a bogie.

The periodicity of corrugation at sleeper pitch is explained satisfactorily by the track's vertical dynamic behaviour alone: if the corrugation excites the pinned-pinned resonance, the variation in dynamic contact force is greater on the approach to sleepers than elsewhere.[35] With regard to the effect of rail metallurgy, an investigation by British Rail (BR) of different rail steels has suggested that Acid Bessemer steel may be more prone to corrugate because it wears relatively quickly and is more resistant to plastic deformation under cyclic loading above yield.[38]

11.7.3 Treatment

Grinding is the principal treatment for roaring rail corrugations. It has been shown in the UK[39,40] and in Germany[29] that grinding new rail also significantly delays the onset of new corrugation, which is consistent with exacerbation of the corrugation by vertical vibration excited by railhead roughness. Grinding of new rail is now undertaken routinely on many systems.

Whatever the detailed wavelength-fixing mechanism might be, it is desirable to reduce misalignment of wheelsets in bogies and thereby reduce the propensity of wheels to slip. Undoubtedly one of the reasons that roaring rail corrugation became more prevalent in the 1970s and 1980s is that an improved understanding of vehicle dynamics led to a deliberate increase in the in-plane stiffness of bogies to achieve high-speed stability: this increases the likelihood of slip even on unmotored wheels.

The limited measurements from test sites that indicate a reduced rate of corrugation formation on harder rails are consistent with wear as the damage mechanism.[30] More field tests of this type would be desirable to quantify

the effect of harder steels, but it nevertheless appears that these should offer one means of reducing roaring rail corrugation.

It was suggested in 1993 that resilient railpads should reduce the rate of corrugation formation by reducing vertical dynamic loads, particularly those caused by sleeper resonances and the pinned-pinned resonance of the rail. Although stiffness of track support appears to be critically important in formation and avoidance of P2 resonance corrugation (Section 11.5.3), it has subsequently become clear that railpads have relatively little influence on dynamic contact force associated with the pinned-pinned resonance. This is demonstrated by the measurements in Fig. 11.6, which shows corrugation measurements made on two sites within a few hundred metres of one another in similar curves on the same metro line at a similar time after grinding. (Immediate post-grind measurements are not available, but it can reasonably be assumed that residual irregularities were similar on the two sites, particularly in the 30–100 mm wavelength range of interest here.) The one-third-octave band spectra (the lower graphs in the figure) show that the amplitude of short wavelength corrugation is similar at these two sites, despite the very different support stiffness at the two locations.

Although it may appear superficially attractive to treat this type of corrugation using a dynamic vibration absorber to attenuate the pinned-pinned resonance, this would be unsuccessful because 'roaring rails' arise from the high dynamic loads caused by the high dynamic stiffness of the supports. If a dynamic vibration absorber were attached to the supports, they could move even less and 'roaring rails' might even be exacerbated. At mid-span between supports, the rail has a low dynamic stiffness, dynamic loads are low, and corrugation develops less quickly (Fig. 11.11).

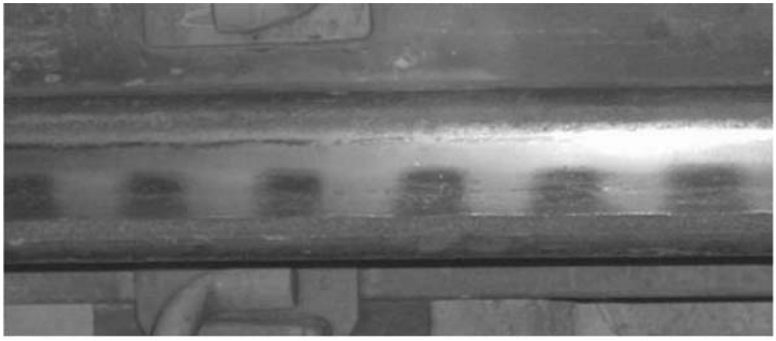
It was claimed in reference 2 that 'roaring rails' had been treated comprehensively and successfully on the Vancouver Skytrain system. If the pinned-pinned resonance is indeed the wavelength-fixing mechanism for roaring rails, evidence from Skytrain should be reassessed to determine whether this is consistent with corrugation having arisen from this mechanism. The success of the treatments adopted on Skytrain is nevertheless both unquestionable and also consistent with corrugation formation by several mechanisms involving slip and wear, particularly as a result of bogie misalignment and poor wheel and rail profiles. The treatment comprised the following three components:

1. reduction of the misalignment of wheelsets in the bogies;
2. grinding to remove existing corrugation and to re-profile the rails transversely: the latter is done both to reduce conformity between wheel and rail and to vary the effective gauge so as to change the contact point on the wheel, thereby giving more uniform wear;
3. use of a proprietary friction modifier with the characteristics described in Section 11.6.3.

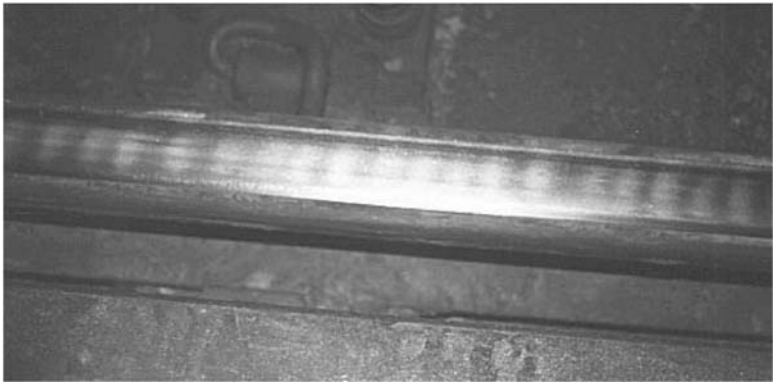
11.8 Trackform-specific corrugation

11.8.1 Characteristics

Some trackforms appear to be more sensitive to corrugation than others, and unusual types of corrugation appear to occur on some trackforms but not on others. Non-ballasted track systems are prone to these eccentricities. The review of 1993 referred to work that had been published regarding so-called 'booted sleepers' on metro systems in Paris, Baltimore and London.[14,41,42] These sleepers are used to introduce resilience in the support and thereby reduce ground-borne vibration from the track. However, corrugation occurred at some places on these metro systems on the low rail in curves with a radius of less than about 400 m on these particular trackforms but not on adjacent sections of track of different construction. Corrugation of measurable amplitude and about 50 mm wavelength developed within days of grinding to remove all previous traces of corrugation.[14,42] Examples of corrugation that appear only on particular trackforms are shown in Figs 11.12 and 11.13,



11.12 Corrugation on track with sleepers in resilient boots.

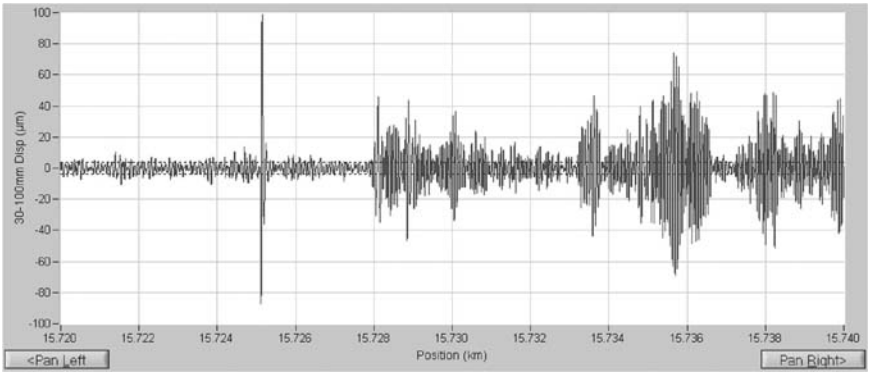


11.13 Corrugation on track with resiliently-mounted baseplate.

and a measurement that highlights the difference between corrugation on two trackforms is shown in Fig. 11.14.

11.8.2 Cause

The damage mechanism for these types of corrugation is differential wear of the corrugation troughs. A general statement can be made that the wavelength-fixing mechanism for this generic type of corrugation is a resonance that has relatively low damping and that is excited in some circumstances and not others. In some circumstances the resonance appears to be one that is excited primarily by lateral forces, so differential wear at a particular frequency occurs primarily because of periodic maxima in slip. This appears to be the case, for example, with the booted sleeper trackform. On the other hand, some trackforms have a maximum in the vertical dynamic contact force, e.g. the trackform responsible for the corrugation in Fig. 11.13 has a large baseplate mounted on a resilient railpad, giving rise to a 'peaky' and ill-damped resonance with a corresponding maximum in the vertical dynamic contact force. Curves are more prone to corrugate by a mechanism involving fluctuations in forces and slip in the plane of the track for the same reason as this occurs with 'rutting' corrugation (Section 11.6.2): tangential forces are high in curves and closer to the limit that can be carried by friction. If the wheel on the high rail on the leading wheelset slips, this forces the low rail wheel to slip. The frequency of the roll-slip oscillation that results from this is that of the resonance that is most readily excited. Where the significant dynamic behaviour is associated with vertical loading e.g., from resonance of a baseplate on a railpad, this is most likely to be excited by railhead irregularities.



11.14 Measurements showing difference in corrugation on different trackforms (to left and right of weld at 15.725 km).

11.8.3 Treatment

It is claimed in references 14–42 that where corrugation existed on the booted sleeper trackform it was reduced using resilient railpads to decouple vibration of the sleeper from that of the rail and also by lubricating the high rail. Grinding was used to remove existing corrugation.

Some improvement should be possible if bogies steered better (for example, with higher conicity wheel profiles), and if the rails were asymmetrically profiled to improve steering. A friction modifier (Section 11.6.3) would reduce the propensity for stick–slip vibration of the wheel and rail. Since wear is the damage mechanism, wear-resistant head-hardened or alloy rails should be beneficial.

Clearly corrugation does not appear on such trackforms everywhere they are used. There are many non-ballasted trackforms in existence, some of which corrugate as a result of unusual resonances while some apparently corrugate as a result of the P2 resonance (Section 11.5.3). In the absence of other information, the supplier's advice should be sought regarding experience with a particular trackform.

11.9 Conclusions and recommendations

A wide variety of types of rail corrugation exists in practice. Two critical characteristics of a corrugation have been considered:

1. the mechanism which gives rise to the periodicity (the wavelength-fixing mechanism);
2. the mechanism which causes the perceived damage (the damage mechanism).

By this means, six types of corrugation are distinguished and have been denoted by the terms heavy-haul, light-rail, other P2 resonance corrugation, rutting, roaring rails and trackform-specific corrugation. Both damage and wavelength-fixing mechanisms have been identified for all six types of corrugation. In all cases the wavelength-fixing mechanism is excitation of a characteristic frequency by trains passing at a relatively small range of speeds: the corresponding wavelength is simply:

$$L = \frac{v}{f} \qquad [11.3]$$

where v and f are the appropriate train speed and resonant frequency, respectively. Resonance of the unsprung mass of vehicles on the track stiffness (the P2 resonance, Table 11.1) is the most prevalent wavelength-fixing mechanism, and it is accordingly critically important to reduce the unsprung mass to as low a level as is practical in order to reduce the track's propensity to corrugate by any one of a number of mechanisms.

Whereas at the time of a previous review in 1993 it appeared that 'roaring rails' were the most perplexing member of the corrugation family, research that has been done in the meantime has conclusively demonstrated that this arises from excitation of the rail as if it were pinned at sleepers, or more correctly from the high dynamic loads that result from passing over the dynamically stiff supports when the rail is vibrating in this mode. Unfortunately, it appears that this type of corrugation is an inevitable consequence of periodically supported rails, and is accordingly likely to be with us for some time to come. However, because the corrugation tends to be quite short (30–100 mm wavelength) it usually develops to only a shallow depth and is readily removed by fairly light grinding.

Treatments which involve some measure of prevention, rather than simply grinding away the irregularity and allowing it to recur, have been identified for all six types of corrugation. All types of corrugation are at least alleviated by grinding. More detail of the treatments for individual types of corrugation can be found in the appropriate references (Table 11.1). Where excessive vertical dynamic loading is an essential feature of the corrugation mechanism (that is, heavy-haul and light-rail varieties), maintenance of a sufficiently smooth railhead profile is an essential means of preventing corrugation on existing track. For heavy-haul corrugation, improvements in welding procedure are required to minimize differential deformation of the weld and parent materials. Means are available to specify allowable irregularities for specified types of rail and traffic to prevent heavy-haul and light-rail corrugation. Grinding is important also because it can be used to profile rails transversely, thereby improving steering in curves and controlling conformity of wheel and rail.

Selection of an appropriate rail for new track helps to prevent corrugation. Sufficiently high flexural rigidity and yield strength are required to resist corrugation by plastic bending (light-rail corrugation), and adequate yield and tensile strengths are required to resist corrugation by plastic flow (heavy-haul corrugation): for all practical purposes, these are 'hard' rails. Hard rails also resist wear and, accordingly, resist formation of corrugation in those cases in which wear is the damage mechanism.

A resilient support is critically important in resisting corrugation formation by excitation of the P2 resonance. This is particularly important for non-ballasted trackforms, in which the track support is often relatively stiff. For slab track systems, the significant resilience in this respect is that between the rail and the slab rather than of the rail through to ground.

Control of friction between the wheel tread and the railhead is an important means of reducing all types of corrugation in curves. Control of roll-slip oscillations by controlling friction is a particularly effective method of reducing rutting.

All types of corrugation which occur in curves are exacerbated by poor steering of bogies and would accordingly be less likely to form if steering were improved. In many cases this is possible by asymmetric profiling to increase the rolling radius difference between inner and outer wheels. Rolling stock with more resilient yaw suspension should be less prone to corrugation. Steering is also improved, and both tangential loads on the low rail and the conditions for roll-slip oscillation reduced if bogies run close to or above 'balance speed', i.e. at cant deficiency rather than with excessive cant. Re-canting where possible could significantly reduce corrugation in curves, and would also reduce other types of surface damage of both wheels and rails.

Examples are given here of measurements that can now be obtained to document the effect of changes on corrugation formation, but which were not possible in 1993. The excellent measurements in Fig. 11.6 that document the success of a resilient trackform in eliminating long wavelength corrugation from the P2 resonance were obtained almost by accident, as a result of measurements that were undertaken routinely by the grinding contractor to show the extent of corrugation before and after the annual grinding programme. Since the long wave corrugation is both deep and difficult to remove by current re-profiling techniques, its avoidance represents a substantial saving in recurrent maintenance costs that would significantly offset the costs of installation of a more expensive trackform.

The scope for similarly interesting and almost zero-cost revelations is as yet rather limited since it is rare for grinding contractors routinely to obtain measurements of such accuracy. However, this is becoming more common as the equipment (Fig. 11.15 and references 43 and 44) becomes more widely available and as railway systems realize the value of insisting on low amplitudes of residual irregularity following re-profiling. Routine measurements of this form will contain a wealth of information in years to come on rates of corrugation growth in different conditions, effects of trackform, traffic and the like. Such measurements have been obtained to date only as part of research projects or monitoring programmes that had been established for a specific purpose, and the use of a variety of equipment (sometimes of questionable accuracy) has made comparison difficult.

11.10 Acknowledgements

Permission of the Institution of Mechanical Engineers to base this chapter on the paper that was originally published in the Institution's proceedings[2] is gratefully acknowledged. The author is grateful to Brian Whitney of Network Rail for the use of a photograph and to Schweerbau GmbH for the use of measurements made using the Corrugation Analysis Trolley (CAT). Although an attempt has been made to revise and update that review, this

11.15 Corrugation analysis trolley, as used for measurements in Figs 11.6 and 11.14[43–44]

has been far from comprehensive and the author apologizes to those whose work he has neglected.

11.11 References

1. Focus on corrugations, *Progressive Railroading*, **28**, 42–6.
2. Grassie, S.L. and Kalousek, J. (1993), Rail corrugation: characteristics, causes and treatments, *Proceedings of the IMechE, Part F: Journal of Rail and Rapid Transit*, **207**, 57–68.
3. British Railways Research Department, *Bibliography on Corrugation of Rails*, British Rail Research, Derby, UK.
4. Clark, C. H. (1922), How to avoid or overcome rail corrugation, *Engineering and Contracting*, **58**, 120 [cited in (**2**)].
5. Mair, R. I. (1977), Natural frequency of rail track and its relationship to rail corrugation, *Transactions of the Institution of Engineers, Civil Engineering*, **CE19**, 6–11.

6. Mair, R. I., Jupp, R. A. and Groenhout, R. (1984), Characteristics and control of long pitch rail corrugation at heavy axle loads, *Rail Research Papers,* Vol. 1, BHP Melbourne Research Laboratories, Melbourne, VIC, Australia.
7. Marich, S. and Maass, U. (1986), Higher axle loads – economics and technology agree, *Proceedings Third International Heavy Haul Railway Conference*, Vancouver, BC, Canada, 13–17. October, Paper 1A–1.
8. Frederick, C. O. and Round, D. J. (1985), Vertical track loading, in *Track Technology*, Thomas Telford, London, 135–49.
9. Grassie, S. L. (1991), A contribution to dynamic design of track, *Vehicle System Dynamics*, **20**, 195–209.
10. Dudley, N., Oswald, S. and Vines, M. J. (1987), Rail welding for heavy axle load unit train operations, *Rail Research Papers*, Vol. 1, BHP Melbourne Research Laboratories, Melbourne, VIC, Australia, 297–322.
11. Mullen, J. D. (1989), Bending and grinding smooths long-wave corrugations, *Railway Gazette International*, August 550–51.
12. Grassie, S. L. (1989), Corrugation on Australian National: cause, measurement and rectification, *Proceedings Fourth International Heavy Haul Conference*, Brisbane, QLD, Australia, 11–15 September, 188–92.
13. Shelley, S. J. and Williams, M. D. (1990), Future track maintenance strategies at Australian National, *Proceedings eighth RTAA Conference*, Sydney, NSW, Australia.
14. Tassilly, E. and Vincent, N. (1991), Rail corrugation: analytical model and field tests, *Wear*, **144**, 163–78.
15. Grassie, S. L. and Elkins, J. A. (1998), Corrugation on North American transit lines, *Vehicle System Dynamics*, **28**(Supplement), 5–17.
16. Brickle, B. J., Elkins, J. A., Grassie, S. L. and Handal, S. J. (1998), Rail corrugation mitigation in transit, *Research Results Digest*, number 26, Transit Cooperative Research Program, National Research Council, Washington, DC, USA.
17. Grassie, S. L. and Edwards, J. W. (2008), Development of corrugation as a result of varying normal load, *Wear*, **265**, 1150–55.
18. Daniel, W. J. T., Horwood, R. J., Meehan, P. A. and Wheatley, N. (2008), Analysis of rail corrugation in cornering, *Wear*, **265**, 1183–92.
19. Sun, Y. Q. and Simson, S. (2008), Wagon-track modelling and parametric study on rail corrugation initiation due to wheel stick-slip process on curved track, *Wear*, **265**, 1193–1201.
20. Wu, W. X., Smith, J. H., Brickle, B. V. and Luo, R. K. (1996), The effects of misaligned wheelsets and rolling surface conditions on the formation of rail corrugations, in Zobory, I. (ed.), *Proceedings of 2nd Mini Conference on Contact Mechanics and Wear of Rail/Wheel Systems*, Budapest Budapest University of Technology, Budapest Hungary, 333–40.
21. Matsumoto, A., Sato, Y., Tanimoto, M. and Kang, Q. (1996), Study on the formation mechanism of rail corrugation on curved track, *Vehicle System Dynamics*, **25** Supplement, 450–65.
22. Ishida, M., Moto, T. and Takikawa, M. (2002), The effect of lateral creepage force on rail corrugation on low rail at sharp curves, *Wear*, **253**, 172–77.
23. Egana, J. I., Vinolas, J. and Gil-Negrete, N. (2004), Effect of liquid high positive friction (HPF) modifier on wheel-rail contact and rail corrugation, *Tribology International*, **20**, 1–6.
24. Eadie, D. T., Santoro, M., Oldknow, K. and Oka, Y. (2008), Field studies of the

effect of friction modifiers on short pitch corrugation generation in curves, *Wear*, **265**, 1212–21.
25. Devine, T. J., Daniels, L. E. and Blume, N. (1982), Rail corrugation investigations at FAST: December 1979 through August 1981, Proceedings Conference on *FAST Engineering*, Report no FRA/TTC-82/01, 165–74.
26. Harrison, D. (1979), *The Corrugation of Railway Rails*, PhD thesis, University of Cambridge, Cambridge, UK.
27. Pearce, T. G. (1976), *The development of corrugations in rails of Acid Bessemer and Open Hearth steels*, Internal memo IMDA 257, British Railways Board, London, UK.
28. Hiensch, M., Nielsen, J. C. O. and Verheijen, E. (2002), Rail corrugation in the Netherlands – measurements and simulations, *Wear*, **253**, 140–9.
29. Kaess, G. (1983), Results of DB test sections installed for the study of rail corrugation, *Seminar on Rail Corrugation*, Institution of Mechanical Engineers, London, UK 27 September.
30. Heyder, R. and Girsch, G. (2005), Testing of HSH rails in high-speed tracks to minimise rail damage, *Wear*, **258**, 1014–21.
31. Hempelman, K. and Knothe, K. (1996), An extended linear model for the prediction of short pitch corrugation, *Wear*, **191**, 161–9.
32. Igeland, A. and Ilias, H. (1996), Rail head wear calculations based on high frequency wheelset/track interaction – a comparison between different models, in Zobory I (ed.), Proceedings of 2nd Mini Conference on Contact Mechanics and Wear of Rail/Wheel Systems, Budapest, Budapest University of Technology, Budapest, Hungary, 304–14.
33. Mueller, S. (1999), A linear wheel-track model to predict instability and short pitch corrugation', in ref [2] op cit, pp 899–913.
34. Ilias, H. (1999), The influence of railpad stiffness on wheelset/track interaction and corrugation growth, *Journal of Sound Vibration*, **227**, 935–48.
35. Grassie, S. L., Gregory, R. W., Harrison, D. and Johnson, K. L. (1982), The dynamic response of railway track to high frequency vertical excitation, *Journal Mechanical Engineering Science*, 1982, **24**, 77–90.
36. Grassie, S. L. (1996), Short wavelength rail corrugation: field trials and measuring technology, *Wear*, **191**, 149–60.
37. Knothe, K. and Gross-Thebing, A. (2008), Short wavelength rail corrugation and non-steady-state contact mechanics, *Vehicle System Dynamics*, **46**, 49–66 (also comments on paper and response to comments, *Vehicle System Dynamics*, **46**, 67–70).
38. Frederick, C. O., Clark, R. A., Bugden, W. G. and Allery, M. (1984), The surface damage of rails, Proceedings AAR/BR Conference on *Vehicle track interaction*, Princeton, NJ, USA, April, 5.
39. Bugden, W. G. (1983), Characteristics of short wavelength corrugation on rails, *Seminar on Rail Corrugation*, Institution of Mechanical Engineers, London, 27 September.
40. Clayton, P., Allery, M. B. P. and Bolton, P. J. (1981), Surface damage of rails, in Frederick, C. O. and Round, D. J. (eds), *Rail technology: proceedings of a seminar, organised by British Rail Research & Development Division & American Association of Railroads, held at Nottingham University, 21–26 September*, British Rail Research, Derby, UK, 179–92.
41. Tassilly, E. and Vincent, N. (1991), A linear model for the corrugation of rails, *Journal of Sound and Vibration*, **150**, 25–45.

42. Ahlbeck, D. R. and Daniels, L. E. (1991), Investigation of rail corrugations on the Baltimore metro, *Wear*, **144**, 197–210.
43. Grassie, S. L., Saxon, M. J. and Smith, J. D. (1999), Measurement of longitudinal rail irregularities and criteria for acceptable grinding, *Journal of Sound and Vibration*, **227**, 949–64.
44. See www.railmeasurement.com

Copyrighted Materials
Copyright © 2009 Woodhead Publishing Retrieved from www.knovel.com

12
Rail welds

M. J. M. M. STEENBERGEN, Delft University of Technology, The Netherlands; R.W. van BEZOOIJEN, Id^2 Consultancy, The Netherlands

Abstract: Welding is the standard for joining rails in modern railways. This chapter discusses several aspects of rail welding. The basic welding technologies are presented, including metallurgical aspects and post-welding treatments. Different damage mechanisms initiated at or by rail welds are addressed. The dynamic wheel–rail interaction at rail welding irregularities is investigated, leading to quantitative relationships between key parameters describing this interaction. A method for force-based weld assessment is elaborated as an alternative for the conventional method based on tolerances. The chapter concludes with energy considerations in relation to welding irregularities and track deterioration.

Key words: rail welds, weld geometry, dynamic wheel–rail interaction, weld assessment, track deterioration.

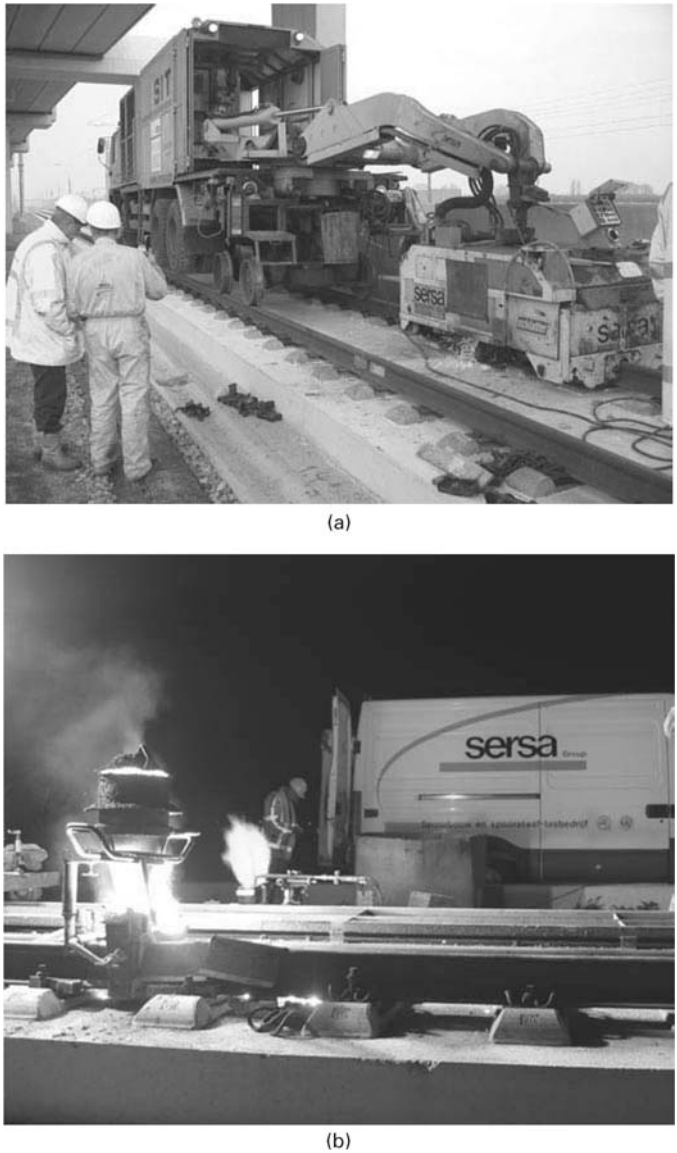
12.1 Introduction

Continuously welded rail (CWR) was introduced on the railways in the 1930s, and nowadays is the standard in modern railway tracks. In many cases, the traditional bolted or fish-plated rail connections are present in the track only in the form of insulated rail joints for detection and signalling purposes. The deflection of a bolted rail joint under static train axle loading leads to a high dynamic impact component, which is a source of rapid track deterioration and high noise levels (Kabo *et al.*, 2006; Steenbergen, 2006; Cai *et al.*, 2007). Therefore, the welded continuous connection was a significant improvement.

In the present chapter, the main different rail welding technologies will be discussed briefly in Section 12.2. After that, damage mechanisms that may be initiated at or by rail welds are addressed in Section 12.3. An important issue is the dynamic wheel–rail interaction at rail welds; this is the subject of Section 12.4. In Section 12.5 the Dutch rail welding regulations will be explained as a method for force-based weld assessment, as an alternative to the conventional method purely based on geometry. Section 12.6 will conclude with energy considerations in relation to welding irregularities and track deterioration.

12.2 Rail welding processes

The most common types of rail welding are electrical flash-butt welding and aluminothermic or thermite welding (Fig. 12.1). Enclosed arc welding is applied only in specific situations such as making closures. The application of the



(a)

(b)

12.1 Flash butt (a) and aluminothermic (b) rail welding on the Dutch HSL-South.

first type of welding is almost fully automated, including the positioning of the rail ends to be joined, and it is therefore often applied in the construction of new tracks or large renewal projects. It can be applied in stationary plants as well as mobile machines. The second welding process requires handicraft and is therefore often applied in repair and small maintenance works, or installation of insulated rail joints, switches and crossings. Both rail welding procedures are explained below.

The mobile flash-butt rail welding process (Wöhnart and Wenty, 2002) begins with the clamping of the rail ends by jaws in the welding unit. They act as water-cooled electrodes that ensure the transfer of the welding current to the rails. The rail ends are aligned and brought together by a hydraulic movement of the two aggregate halves of the welding unit, thus producing arcing. Now follows the welding process in three stages: pre-heating, a stage with constant pressing together of both rail ends and a progressive stage of accelerated burn-off. During this process, current intensity and flash-burn speed depend on the stage and the rail grade. The welding process ends with a final compressing stroke, with the aim to avoid welding nuggets, oxide inclusion and impurities from flashing. After welding, shearing knives, following the rail cross-sectional geometry, encompass the rail and trim the welding seam while still red hot.

Special alloy and head-hardened rails require special processing during the cooling stage. Special alloy rails should be prevented from cooling down too quickly and follow a post-heating program in order to prevent the development of brittle martensite parts. On the other hand, head-hardened rails need to be cooled down quickly in order to keep the hardness level of the parent material intact. This can be achieved by 'air quenching', yielding hardening of the railhead.

The aluminothermic welding method is a fusion welding process and makes use of a mixture of aluminium powder and iron oxide which is converted into alumina and steel at a high temperature (about 2500 °C). The process is as follows: both rail ends are aligned; a mould is installed around the gap in the joint (about 20–25 mm) with a crucible on top; the rails are pre-heated to about 1000 °C using propane burners; the mixture in the crucible is ignited, yielding an exothermic chemical reaction and filling the gap within the mould; then both crucible and mould are removed, and the weld collar is stripped and ground.

The alignment (a small overheight of about 1.2–1.4 mm is given) of the rail ends for both welding methods is done in order to account for the unequal contraction of the head and the foot of the rail profile after welding during cooling down; in this way 'dipped' welds are avoided. Cooling down after welding requires special attention in order to maintain the pearlitic crystalline rail structure. Special so-called TTT (time–temperature transformation) diagrams provide relationships between cooling rates and steel phases for

different rail steel grades and alloys, such that austenite formed during rail heating is re-transformed into pearlite and the formation of brittle martensite is avoided. The manual thermite welding process in particular requires craftsmanship, in order to control and optimize the steel phases and their specific properties with pre-heating, controlled cooling down and post-weld treatment under all circumstances and conditions.

Generally, the fatigue resistance of flash-butt welds is superior to that of aluminothermic welds; estimations for heavy-haul traffic (Union Pacific Railroad) are 365 MGT for thermite welds against 877 MGT for flash-butt welds (Wimmer *et al.*, 2002). In the case of severe welding geometry irregularities, the fatigue life of welds may be significantly affected by the effect of the dynamic wheel load and the occurrence of loose sleepers (Ishida *et al.*, 1999; Steenbergen, 2008b).

In order to obtain a good geometry, the final grinding of the rail surface is in most cases done manually for both types of welding. In some cases, however, the welds are treated using a grinding train, on newly built tracks, yielding a high surface quality (Fig. 12.2).

12.3 Rail welds and damage formation

The rail weld is, in comparison to plain rail, particularly susceptible to damage formation. This is mainly due to three reasons. First, the presence of the weld yields a geometrical irregularity along the wheel–rail interface, leading to dynamic axle load variations and dynamic contact stress amplifications. This concerns both normal stresses and tangential stresses and their distribution in the rolling contact. A second reason is the presence of material inhomogeneities along the wheel or rail surface. This includes a non-constant hardness distribution along the rail and microstructural disturbances, non-metallic inclusions or steel phase differences along the surface or subsurface. A third reason is the presence of internal material inhomogeneities or clusters of them in the weld; the microcleanliness of the weld metal in cast condition is generally inferior to that of the parent material. In general, none of the previous mechanisms occurs independently when leading to damage: they interact, significantly increasing the propagation speed of existing damage.

It is obvious from the above that the aluminothermic weld in particular is susceptible to these phenomena. The geometrical irregularity at the rail weld which is present in the majority of the cases can be due to many reasons. These include the effects of unstraightened rail ends, inaccurate rail end positioning, the manual grinding of the rail surface and the influence of inhomogeneous material shrinkage after cooling down in combination with early grinding. The length of the irregularity is generally less than 1 m. Therefore, in general the rail welds yield a significant contribution to the

(a)

(b)

12.2 Manual rail weld grinding (a) and grinding train (Speno) (b) on the Dutch HSL-South.

short-wave contribution of the track irregularity spectrum, which determines the rate of the rail damage progress and track deterioration: dynamic forces may damage the rail or railhead, the concrete sleepers and rail fastenings

and influence, for example, wheel seat fretting fatigue; furthermore, they are responsible for non-uniform ballast settlement or local decompaction. It is important to note that, especially on newly built tracks, the short-wave contribution is almost entirely determined by the rail weld geometry.

Normally, new welds are being tested ultrasonically for internal damage, cracks, voids or inclusions before acceptance. In general this is combined with magnetic particle or eddy current inspection for the lop layer (5–7 mm). When a not-detected serious internal material defect, for example a lack of fusion or porosity, is present in a weld, it often breaks at an early stage (Shitara *et al.*, 2003; Terashita and Tatsumi, 2003), especially in cold winters with high tensile forces in the rail. Therefore, internal material or fusion defects are in general not responsible for long-term deterioration at the wheel–rail interface.

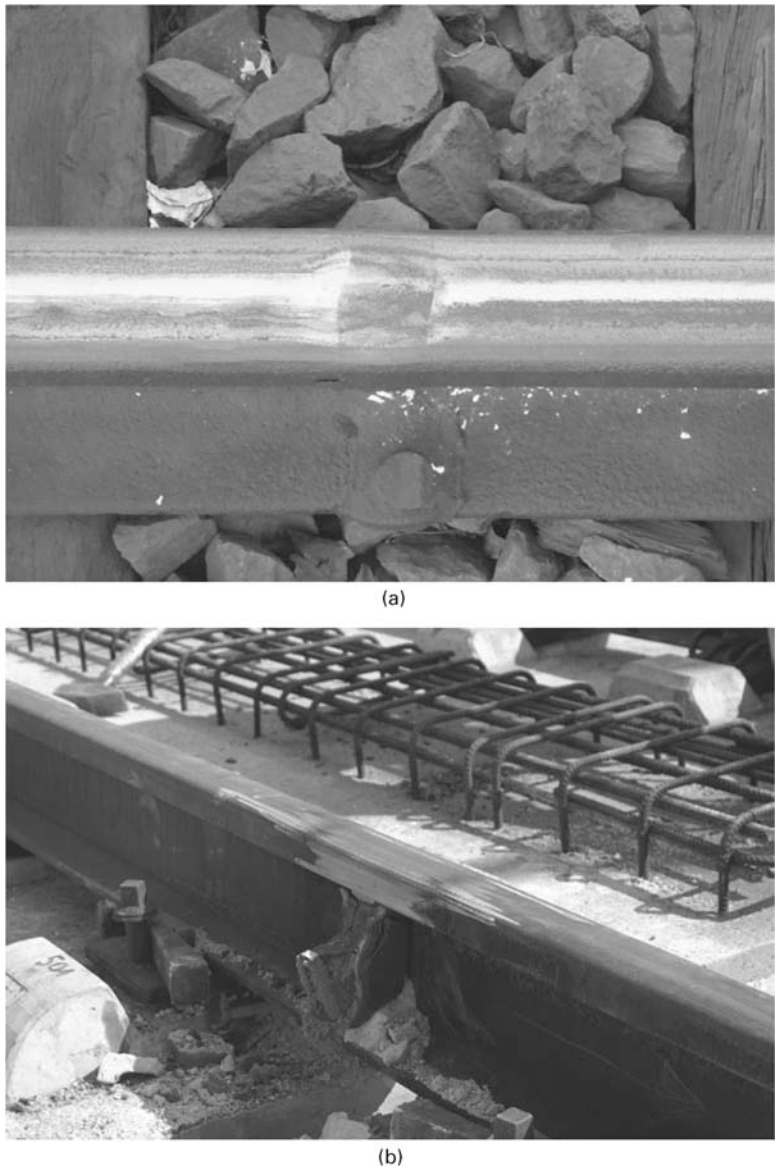
The effect of locally varying metallurgical properties along the rail surface however is particularly relevant for rail welds. As has been mentioned, often short indentations along the rail surface occur at the centre of rail welds, at the transitions from the weld material to the parent material, the so-called heat-affected zone (HAZ) (Fig. 12.3). The phenomenon is due to shrinkage after grinding and further cooling down, especially when the steel temperature was not low enough at the moment of final grinding. It can be observed particularly for aluminothermic welds in rail repair works, executed under time pressure and non-optimal conditions.

These geometrical indentations typically coincide with the local hardness minima along the rail surface in the HAZ at both sides of the weld metal (Mutton, 2000). This is shown qualitatively in the diagram in Fig. 12.4. Vickers hardness levels (at 5 mm below the surface) typically fluctuate between 250 and 400 HV. For comparison, the hardness of the parent rail generally starts around 260 HV for new common rail grades (increasing to 300 HV after grinding and initial cold work-hardening due to train operation) up to 400 HV for new head-hardened rails.

The unevenness in the rail surface, resulting in a local fluctuation of contact stresses in the rolling contact, may induce differences in shakedown, work-hardening and plastic deformation of the rail surface along the rail (Böhmer and Klimpel, 2002), leading to non-homogeneous wear, surface cracking and a progressively deteriorating geometry. This process is also called weld batter (Mutton and Alvarez, 2004). Examples of this phenomenon are shown in Fig. 12.5. In *in-situ* welding it is almost impossible to avoid these indentations, except for new track construction works, where the grinding can be done separately from the welding.

Due to the effects that have been discussed previously, rail welds are often considered, and also observed in practice, as corrugation or squat initiators (Hiensch *et al.*, 2002; Li *et al.*, 2006, 2007). In Figs 12.6 and 12.7 examples are shown of squats on aluminothermic and flash-butt welds, respectively.

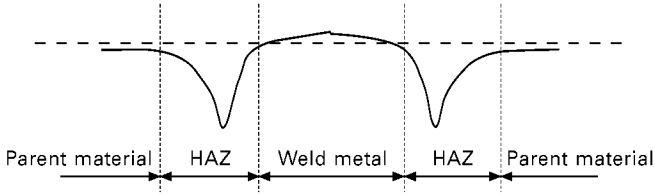
The residual stress field which is induced over the cross-section of the rail as a result of welding and heat transfer processes may be a further cause of internal damage propagation or failure (Chen *et al.*, 2006a,b). Webster *et al.* (1997) showed that in aluminothermic welds the longitudinal residual stresses are compressive at the top and the bottom of the rail, which is



(a)

(b)

12.3 (a) Weld and parent material; (b) the HAZ after thermite welding and grinding.



12.4 Qualitative behaviour of the steel hardness distribution along the rail surface.

beneficial for the fatigue lifetime. In the web region, however, residual stresses are strongly tensile, which makes that region very susceptible to damage initiation and growth. Tawfik *et al.* (2006) performed a similar analysis for flash-butt welds, and showed again a strong residual tensile stress field in the web of the rail, leading to specific failure modes originating from the rail web. Since the rail web mostly accounts for the shear stresses in the cross-section, these are failure modes induced by shearing under high axle loads, such as the horizontal split web mode. Tawfik *et al.* (2008) showed that the tensile residual stress field may be alleviated by localised post-weld heat treatment at the base of the foot directly after welding. Without heat treatment, stress values in longitudinal direction may be as high as 300 MPa. Skyttebol *et al.* (Skyttebol, 2004; Skyttebol *et al.*, 2005) showed that the residual stress field in combination with stresses resulting from ambient temperature loading and train axle loading may lead to very rapid crack or damage growth and a significantly reduced fatigue life. Mutton and Alvarez (2004) observed that the fatigue process leading to the horizontal split web failure mode on heavy-haul lines (especially on curves and tangent track) is in many cases low-cycle fatigue with few high axle overloads and rapid fracture growing from weld collar slag inclusions. Ilic *et al.* (1999) showed experimentally the effect of post-weld heat treatment also on the crystalline structure of thermite welds (grain size reduction and homogenizing) leading to higher ductility.

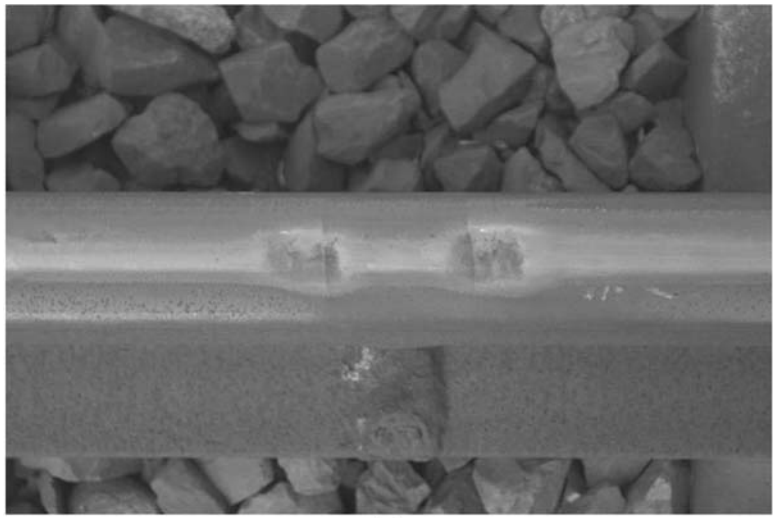
12.4 Rail welding irregularities and dynamic effects in the wheel–rail interface

As has been discussed previously, the rail surface is in general not perfectly smooth in the presence of rail welds, giving rise to dynamic wheel–rail interaction in the contact patch. In Fig. 12.8 (Steenbergen and Esveld, 2006b), finite element (FE) simulation results are shown for a train vehicle, passing an artificial irregularity in the track at a speed of 140 km/h. The irregularity is typical for a rail welding irregularity: it comprises both a smooth irregularity with a longer length-scale (a harmonic wave with a length of 1 m and a top

(a)

(b)

12.5 Deteriorating geometry resulting from non-uniform shakedown and work hardening along a rail weld in the HAZ; (a) flash-butt weld; (b) thermite weld.

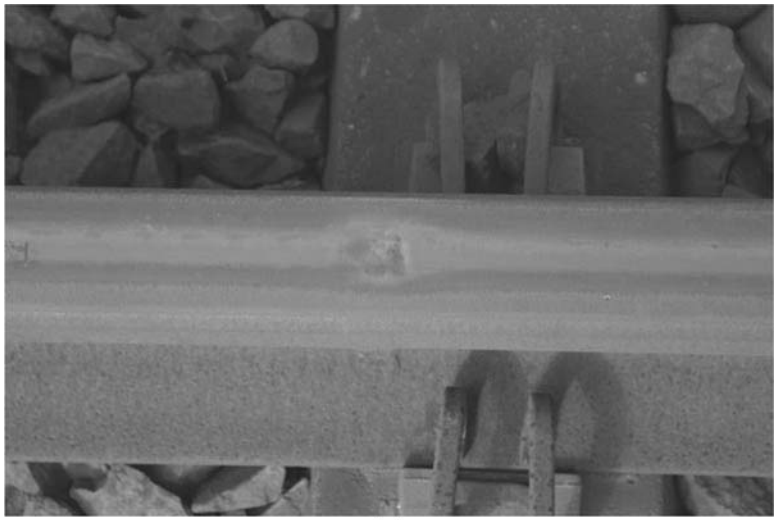


(a)

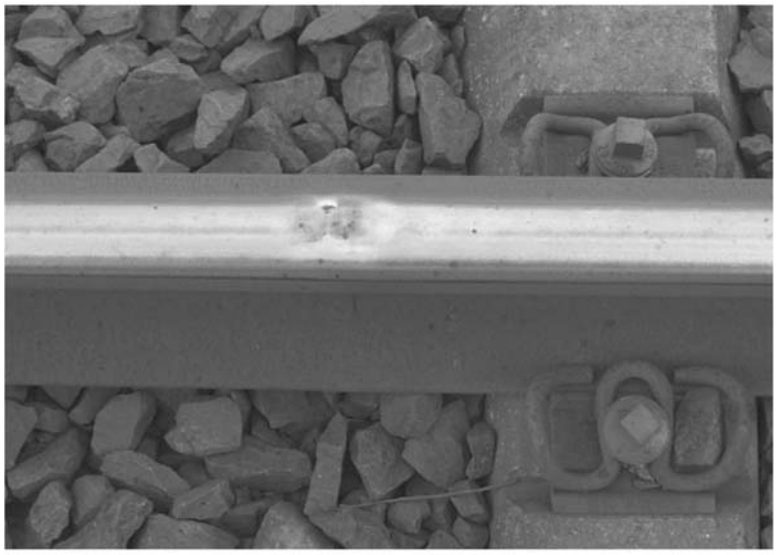
(b)

12.6 Squat initiation on thermite rail welds or in the HAZ (a); accompanied by crack initiation (b).

value of 1 mm), typically due to unstraightened rail ends or the upset of rail ends before joining, and a second non-smooth irregularity with a short length-scale (a triangular peak with a basis of 100 mm and a top of 0.2 mm), typically due to welding or shrinkage.



(a)



(b)

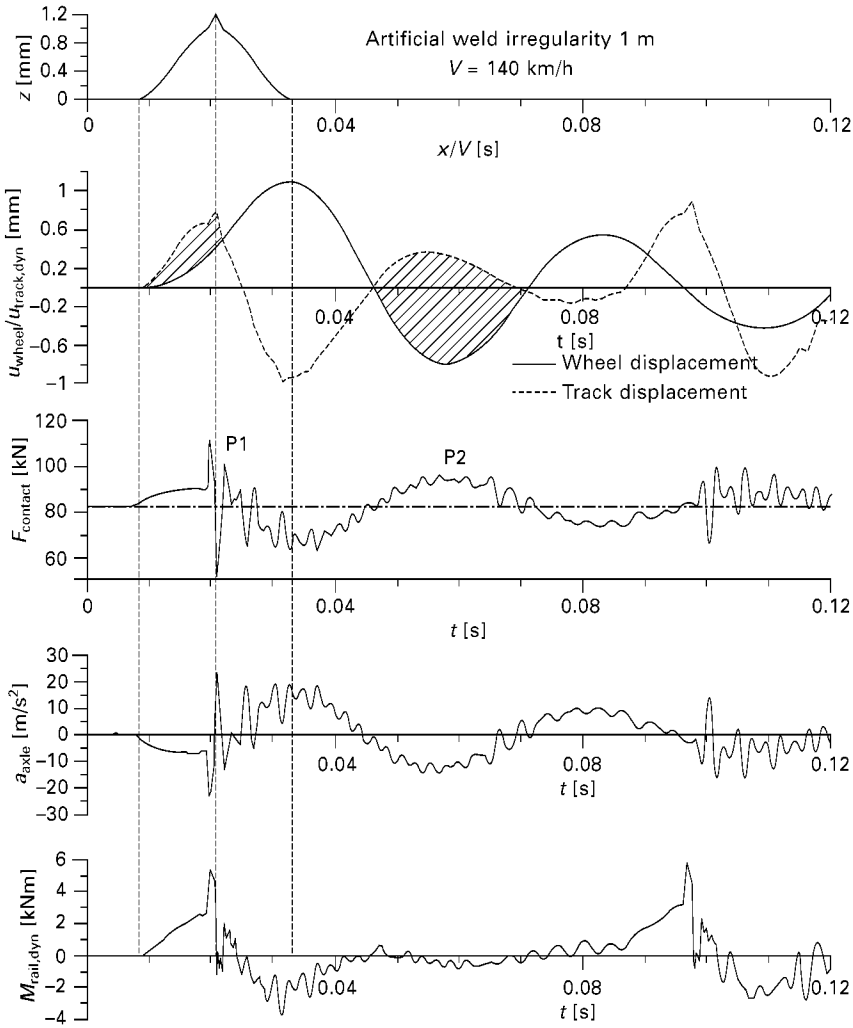
12.7 Squat initiations on flash-butt rail welds or in the HAZ.

The results in Fig. 12.8 have been obtained from FE simulation with the DARTS-NL package. Parameter values were adopted as follows: a wheel mass (half unsprung mass) of 970 kg; a sleeper mass of 300 kg; the 54E1 rail profile is used; the primary suspension stiffness equals $1.8.10^6$ N/m (per wheel); the railpad stiffness $1.2.10^9$ N/m and the ballast stiffness 30.10^6 N/m per sleeper. The latter value is rather small, to account for the generally bad compaction of the ballast underneath the sleepers close to the irregularity;

due to 'ballast pumping' loose sleepers often occur (Ishida *et al.*, 1999). The package uses a non-linear Hertz contact model.

In Fig. 12.8 the following quantities are shown as a function of time:

- the geometry of the irregularity z;
- the dynamic component of the vertical wheel displacement of the first wheel of a passing bogie, defined positive in upward direction and calculated in a convective reference frame moving along with the wheel;

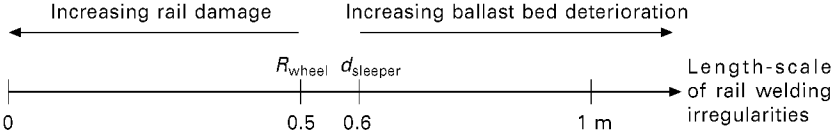


12.8 Dynamic response of the wheel–rail system to an artificial rail weld with both a long and short length-scale irregularity.

- the dynamic component of the displacement u of the rail/track, calculated in a static co-ordinate system with its origin at the centre of the irregularity and defined positive in downward direction;
- the wheel–rail contact force E for the wheel concerned (the dynamic force is superimposed on the static value), in a convective reference frame;
- the axle box acceleration a, in a convective reference frame;
- the dynamic component of the bending moment m in the rail, in a fixed coordinate system at the centre of the irregularity.

From Fig. 12.8, a number of general observations can be made. The rail (and this can to a large extent be generalised to the track) follows the vertical irregularity quasi-instantaneously; this is clearly visible especially for the short-length irregularity. The wheel displacement shows a delay in response relative to the track. It does not show any influence of the short peak in the irregularity, whereas this peak is reproduced almost exactly in the rail displacements. The maximum dynamic contact force is determined by the irregularity with the shortest length-scale. However, the corresponding force peaks (P1 forces) are rather narrow, and the corresponding amount of energy is rather small. These high-frequency peak forces damage mainly the rail itself, as can be observed also from the dynamic rail bending moments. The corresponding energy input will vanish by wave propagation in the rail and dissipation in – mainly – the railpads. The highest energy of the dynamic contact force is contained in the 'carrier frequency', which can be clearly observed. This carrier frequency is related to the delayed reaction of the wheel mass on the overall track stiffness to the irregularity (the P2 force). The relatively large energy contained in the relatively low carrier frequency is mainly responsible for ballast bed deterioration, as it is not efficiently dissipated in the rail and the railpads. Generally it can be stated that components of irregularities having a longer length-scale than the wheel radius or the sleeper span (0.5–0.6 m) damage the ballast-bed, whereas components with shorter length-scales damage the railhead and rail (Fig. 12.9).

As has been mentioned, short components in a welding irregularity lead to P1 forces, whereas longer components lead to P2 forces. This is mainly due to the difference in inertia between the unsprung wheel mass and the



12.9 Types of deterioration of the track system depending on the length-scale of rail welding surface irregularities.

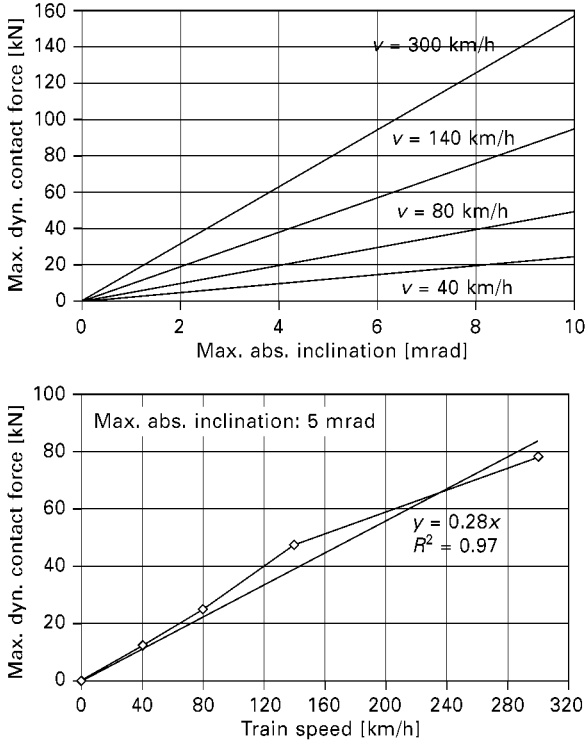
equivalent track mass: assuming an equivalent track mass comprising two fully effective concrete sleepers each of 300 kg and 4 m of effective rail mass (60 kg/m; 60E1 profile) the equivalent track mass can be estimated at 850 kg. This is less than 50 % of the unsprung axle mass, which is approximately 1800 kg for conventional trains. The response of the wheelset to the short excitation therefore shows a delay relative to the response of the track (rail and sleepers). The P1 force, which is a quasi-instantaneous amplification of the wheel-rail contact force, originates from the reaction of the rail and is mainly determined by the rail properties (bending stiffness and inertia). The P2 force results from the reaction of the wheelset on the track stiffness to the excitation. Its magnitude is determined mainly by the unsprung wheel mass and the equivalent track stiffness; the timescale of the peak is larger due to the relatively low natural frequency of the dominating mass–spring system.

Steenbergen and Esveld (2006b) investigated the quantitative relationship between the actual rail weld geometry and the corresponding dynamic wheel–rail response with the help of a FE model. Simulations were performed for a sample of 239 rail weld measurements from the Dutch conventional network. Model parameters were used as defined previously, except for the ballast stiffness, which was taken as $78 \cdot 10^6$ N/m per sleeper, which is a conventional value. The simulations were performed for different train velocities: 40, 80, 140 (conventional lines) and 300 km/h (high-speed lines) respectively. For conventional lines, dynamic wheel–rail forces were generally below 70 kN. For a static nominal wheelload of 112.5 kN, this yields a dynamic amplification factor of 1.6. This value is confirmed experimentally by Mutton and Alvarez (2004). Figure 12.10 summarizes computational results from the FE simulations. At the left, the relationship between the maximum dynamic wheel-rail contact force and the maximum absolute first derivative (on 25 mm basis of a running five-point-average through the measurement signal, which has a resolution of 5 mm) is shown for the different velocities for which simulations were performed. For each velocity, a best linear fit through the origin is displayed, disregarding any scatter.

Figure 12.10b shows the relationship between the contact force and the train speed for a predefined maximum absolute first derivative, which is taken as 5 mrad (again on 25 mm basis). In general, the relationship between the force and the speed can be well approximated as linear. In terms of the train speed V [m/s] and the maximum absolute rail inclination θ_{max} [mrad] (on 25 mm basis) the following expression is obtained (in [kN]):

$$F_{dyn,\,max} = \xi \cdot V \cdot \theta_{max}: \xi = 0.22 \qquad [12.1]$$

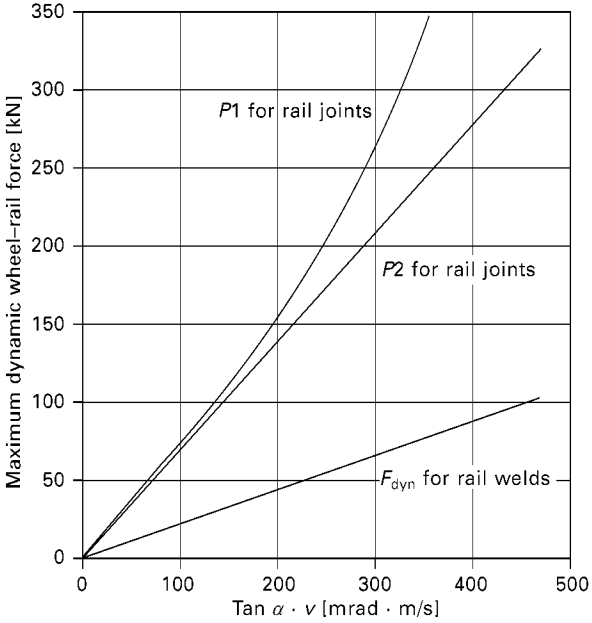
It is observed that ξ is a dimension-bearing coefficient. The issue has been addressed by Steenbergen (2008a), showing that the validation coefficient includes an equivalent track mass, a length dimension and a dynamic stiffness,



12.10 Dependence of the dynamic wheel–rail contact forces at rail welds on the train velocity, found from FEM simulations.

which is responsible for the scatter in the relationship. The linear relationship between the maximum dynamic contact force (or the dynamic amplification factor) and the train speed is confirmed experimentally by the results obtained by Mutton and Alvarez (2004). A very similar result was found, both from simulations and experiments, by Jenkins *et al.* (1974) at British Rail for the peak forces occurring at dipped rail joints. This result was reflected in the well-known Jenkins formula for the calculation of P1 and P2 forces. They found these peak forces to be governed by the dip angle and the train speed; furthermore the relationship between maximum dynamic contact forces and both variables proved to be linear in a good approximation.

Jenkins' results for rail joints, together with the results obtained for rail welds, are shown in the graph in Fig. 12.11. In this figure, the dynamic peak forces for rail joints and welds are depicted as a function of the product of the maximum absolute inclination and the train speed; $a \approx \tan a$ is defined as the angle of a dipped rail end with the horizontal for rail joints, and as the maximum absolute inclination (on 25 mm basis) of the geometry for rail welds. It is clear that the force level for rail welds is significantly lower than

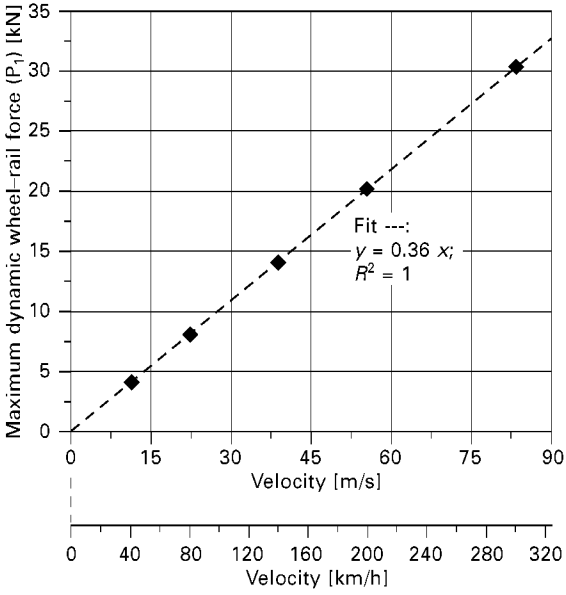


12.11 Maximum dynamic wheel–rail forces (F_{dyn}) for rail welds and rail joints, as a function of the product of the maximum absolute inclination (on 25 mm basis) or the dip angle and the speed.

for rail joints; the reduction factor is approximately three. This is partly due to the quasi-static rail deflection under axle loading, which is much larger for a jointed rail (behaving as a hinged beam) than for a welded rail. This quasi-static deflection increases the dip angle at the moment of wheel passage (Steenbergen, 2006).

Steenbergen (2008a,b) also investigated the dynamic wheel–rail interaction at rail welds with a linear numerical–analytical frequency–domain model, comprising a rail on elastic foundation and a train axle up to the primary suspension level. This study was performed for a better understanding of the backgrounds of wheel–rail interaction at short irregularities and trend behaviour. Analytical results show a perfectly linear relationship between the slope of a ramp in the rail geometry and the magnitude of the dynamic contact force in the time domain. The linear trend of this contact force as a function also of the train velocity is shown in Fig. 12.12, for the case of a ramp of 2 mrad and a length of 25 mm. Thus, for the elementary case of a ramp in the rail surface, the first peak force (P1) may be considered as directly proportional to the speed and the ramp angle, according to the expression:

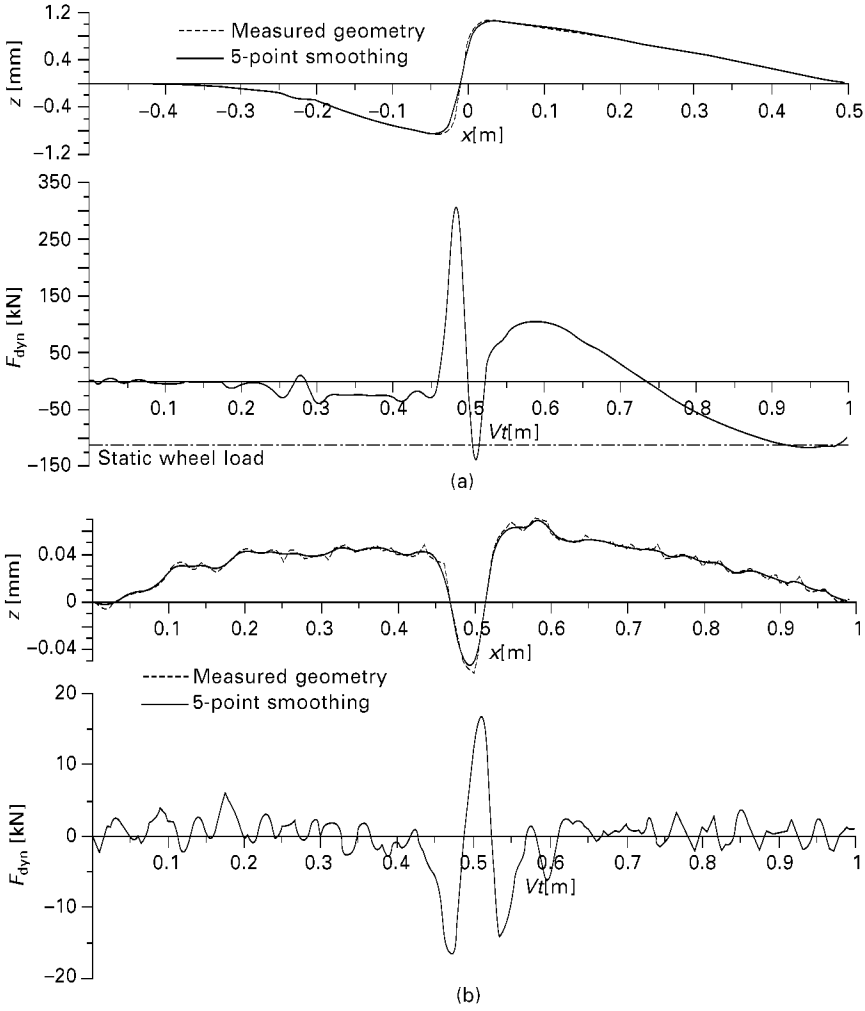
$$F_{dyn,max} = \xi \cdot V \cdot \theta_{max}: \xi = 0.18 \qquad [12.2]$$



12.12 Relationship between the maximum dynamic contact force (*P*1) and the velocity, for a given ramp inclination (2 mrad) and basis length (25 mm).

In the analytical simulations, a running five-point average of the measurement signal of a weld has been used as calculation input, with a discretization basis of 5 mm. The graphs in Fig. 12.13 show examples of the computed time histories of the contact force for two measured welds, for a conventional passenger train velocity of 140 km/h. The first weld is rather well aligned, but shows the effect of non-homogeneous shrinkage of the weld and parent material after the grinding process, which has obviously been performed at too high a temperature. The second weld is smooth on a micro-scale, but shows a severe misalignment before the rail ends have been welded together, resulting in a pronounced step of about 2 mm in rail height over 25 mm. For this particular weld, contact loss will occur as soon as the dynamic force component exceeds the static pre-load.

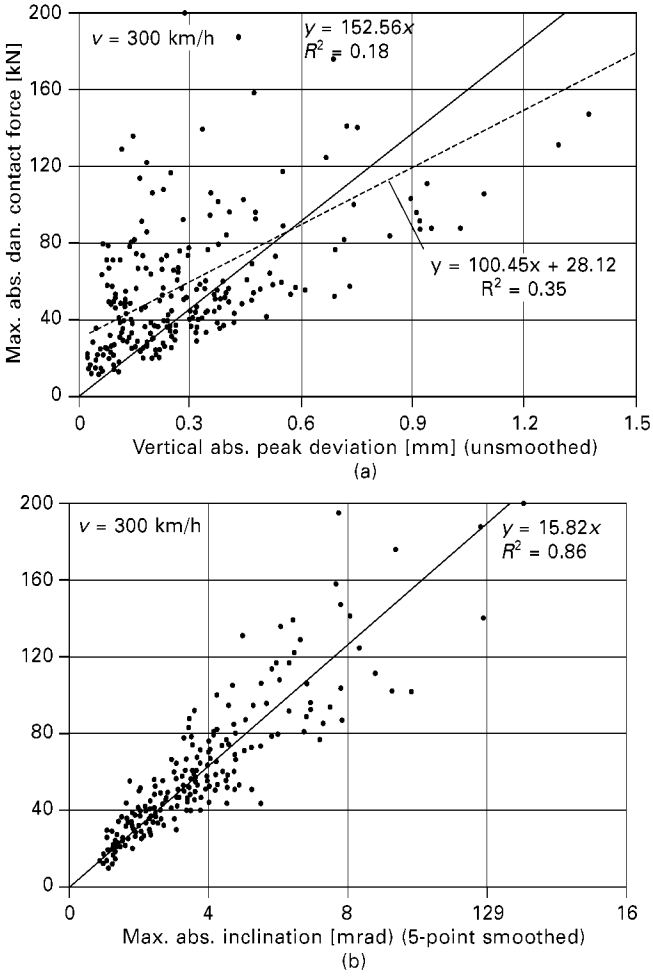
In Fig. 12.14 an example is shown of the correlation between the calculated maximum absolute dynamic wheel–rail contact forces (during the time interval of overpassing the weld) and the vertical absolute peak deviations (Fig. 12.14a) or the maximum absolute inclinations of the welds (Fig. 12.14b), for a train velocity of 300 km/h. Similar to the results obtained earlier with the FE model, the correlation between forces and maximum inclinations is found to increase significantly with the velocity. This can be explained by a more flat transfer of the system as the excitation frequency band shifts towards higher frequencies with increasing speed.



12.13 Examples of measured welds, their geometry after 5-point smoothing and calculated time histories of the dynamic wheel–rail contact force at 140 km/h.

The linear relationships, obtained from the numerical–analytical computations, between the dynamic force and the maximum absolute inclination are summarized in Fig. 12.15a for all train velocities. Figure 12.15b shows the dependence of the dynamic force on the train speed for a given maximum inclination (5 mrad, on 5 mm basis of the five-point averaged measurement signal). Again, the relationship between the dynamic force and the train speed can be very well described as directly proportional.

Figure 12.15 yields the value $\xi = 0.19$ for the coefficient in the general relationship (Eq. 12.1). From the FE analysis $\xi = 0.22$ was found. This

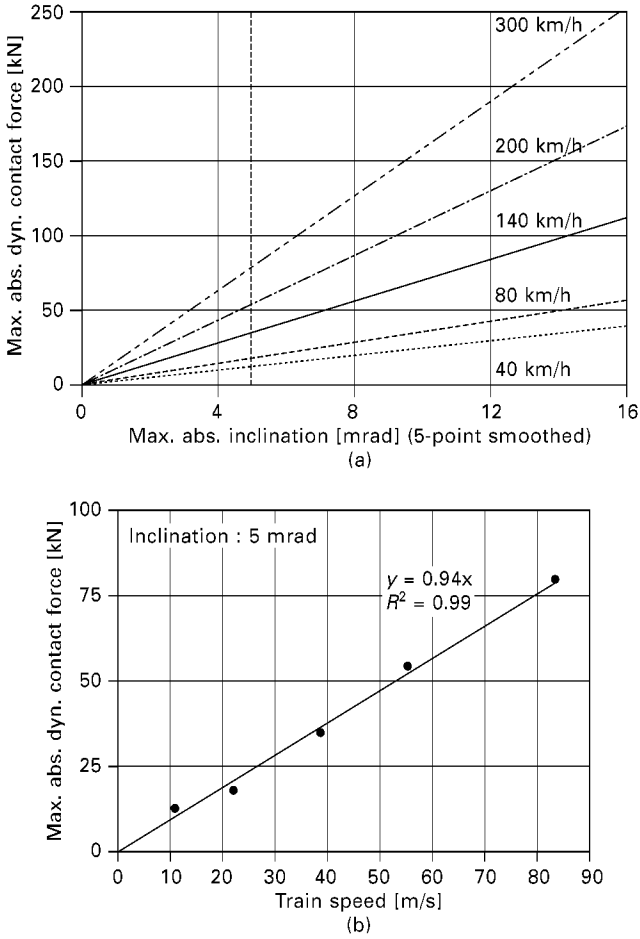


12.14 Peak value of the dynamic wheel–rail force as a function of the vertical peak deviation (a) and the maximum absolute inclination (b) respectively, at 300 km/h.

slightly higher value can be explained from the fact that inclinations on a 25 mm basis have been used in FEM analysis. The inclination decreases when the sample interval is extended from 5 to 25 mm. It can be concluded that $\xi \approx 0.2$ in practice (for 54 E1 rail).

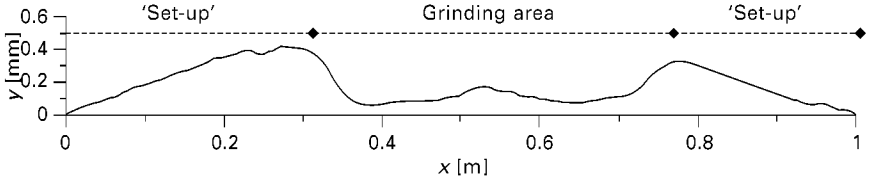
12.5 Rail weld geometry assessment; the Dutch rail welding regulations (2005)

Traditionally, the vertical geometry of rail welds is assessed manually with a steel straightedge in combination with a feeler gauge, using the principle



12.15 Dependence of the dynamic wheel–rail contact force at welds on the maximum absolute inclination (for different train speeds) (a) and on the train speed (for 5 mrad, 5 mm basis) (b).

of tolerances. These tolerances are defined for a given basis length, which is commonly 1 m. A common value for the vertical tolerance is 0–0.3 mm (Esveld, 2001). Two drawbacks of this traditional method are that no restrictions are being posed to the 'smoothness' of the rail surface, whereas this shape has a direct relation to the magnitude and the spectrum of the dynamic wheel–rail interaction force, and the fact that the train speed has no influence, whereas its effect on the contact forces is far from negligible. The spectrum of the wheel–rail forces in its turn is related to the rate of track degradation. A typical result of the current rail welding assessment methods (especially aluminothermic *in-situ* welding) is shown in Fig. 12.16. After the



12.16 Typical example of the vertical geometry along an *in-situ* made rail weld, resulting from traditional assessment methods.

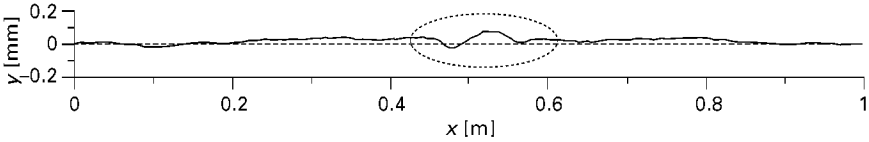
welding process, the 'top' resulting from the aligning of both rail ends, as far as it does not fit within the tolerance, is ground off. The result is often far from smooth, and introduces large dynamic variations in contact stresses.

The development of digital straightedges that sample the vertical rail geometry enables more advanced assessment methods, that do account for the 'smoothness' of the rail geometry. It is not feasible to compute the wheel–rail contact forces in each step of weld grinding, in order to limit the dynamic contact stresses. A compromise between theoretical efficiency and practical feasibility, at the cost of exactitude, can be looked for by correlating a parameter describing the geometry of the weld and the maximum contact force that would occur for a predefined reference track and wheelset, and the train speed. The relationship derived earlier (Eq. 12.1) could be used for this purpose, accounting for both the 'smoothness' of the rail surface through its maximum first derivative and for the train speed. For assessment purposes, the dimensionless rail weld quality index (QI) is proposed. It is defined as the actually measured maximum absolute first derivative (on 25 mm basis) normalized with a norm value, which can be defined differently for different values of the line section speed:

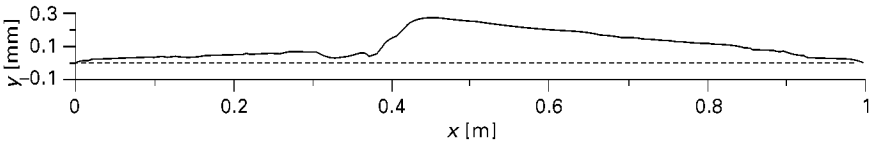
$$QI = \frac{\left|\frac{dz_i(x)}{dx}\right|_{max,\,actual}}{\frac{dz}{dx_{norm}}} \quad (\leq 1\text{: acceptance; } >1\text{: rejection}) \qquad [12.3]$$

The norm values for the inclinations dz/dx_{norm} for different line section speeds must then be established such that the dynamic force, according to Eq. (12.1), is constant. Line section speed intervals are chosen according to typical values: 40, 80, 140, 200 and 300 km/h. Typical examples of welds for which a more 'dedicated' weld assessment would make sense are shown in Figs 12.17–12.19 (Steenbergen and Esveld, 2006a).

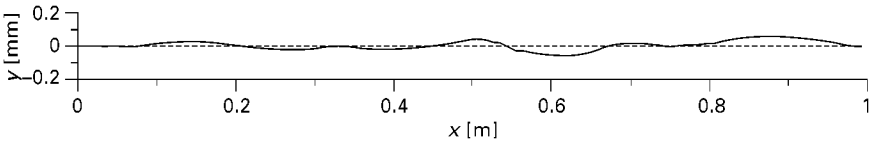
Figure 12.17 shows an almost perfectly straight weld; however, showing small indentations due to non-uniform shrinkage after welding and hot grinding. The weld in Fig. 12.18 shows an aggressive step due to a bad alignment of the rail ends before welding. This is accepted according to



12.17 Example of a weld with indentations due to non-uniform shrinkage after welding and grinding.



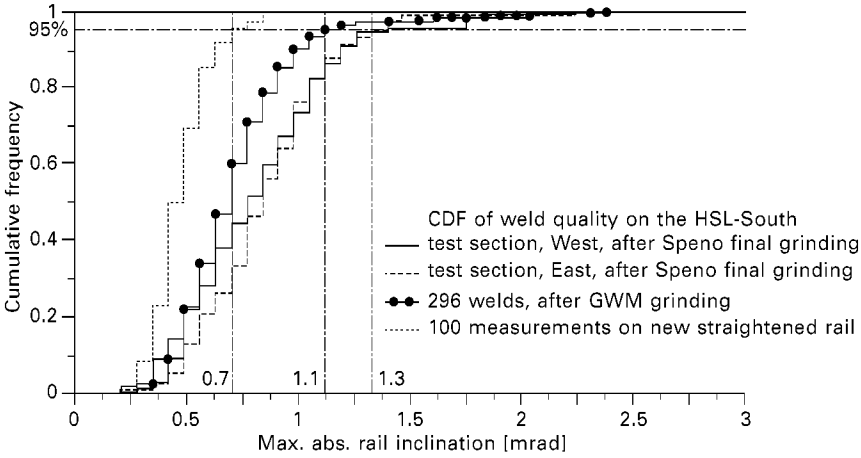
12.18 Example of a weld with a pronounced step due to bad rail end alignment.



12.19 Example of a weld with a 'smooth' surface but with negative height coordinates.

many standards allowing a vertical tolerance of +0.3 mm. However, owing to the large magnitude of the first derivative at the transition between both rail ends the weld would be rejected. In Fig. 12.19, an example is shown of an almost perfect weld. However, because it has some negative height co-ordinates, it is rejected according to traditional standards, which do not allow negative values. The derivatives do not show pronounced peaks and therefore the weld would be accepted.

The norm value for high-speed lines (300 km/h), would be the strictest one in the speed range. Its value can be based on the geometrical quality of new straightened rail: there is no reason why rail welds should have a better geometrical quality than the rail itself. This quality has been determined by measurements on new straightened longrails (120 m); a sample population of 100 segments of 1 m length of newly laid rails on the Dutch high-speed line HSL-South have been measured (Esveld, 2005). The distribution function of the absolute maximum first derivatives (25 mm basis) for these segments is shown in Fig. 12.20. The 95 percentile value of the first derivative in this figure is 0.7 mrad. Therefore, 0.7 mrad could be taken as an appropriate value for the intervention level of first derivative of the weld geometry for high-speed lines. This accuracy is very close to the maximum obtainable accuracy in welding and grinding with grinding trains. Tests in this regard have been



12.20 Cumulative distributions of maximum absolute first derivatives (25 mm basis) on the Dutch HSL-South; comparison of measurements on new straightened rail segments and measurements of rail welds after grinding with grinding trains.

carried out during the installation of the Dutch HSL-South (Winter *et al.*, 2007). In Fig. 12.20, the cumulative distribution of the maximum absolute first derivatives (25 mm basis) of 296 test welds (a mix of flash-butt and thermite welds) is also shown before and after grinding with a GWM 550 grinding train. Before grinding with the GWM the 95 percentile value of the maximum first derivative is 5.3; this reduces to 1.1 afterwards. Finally, Fig. 12.20 displays the cumulative distributions of the maximum absolute first derivative of the rail welds on a double track test section (East and West) of 6 km on the HSL-South. These welds were manually pre-ground, and finally ground with a Speno grinding train (types RR 24 and 48). Based on the results in Fig. 12.20, the 0.7 mrad limit seems too strict a value, and 1 mrad is proposed as a limit value in the QI determination for welds in new high-speed tracks (with norm speed 300 km/h).

In order to consistently determine speed-dependent norm values for the maximum inclination, a pre-defined dynamic force level should be adopted. This force level should be related to the damage level but, because little is known of this relationship and furthermore this is strongly frequency-dependent, any choice of this level is rather arbitrary. Two facts may be taken as a point of departure:

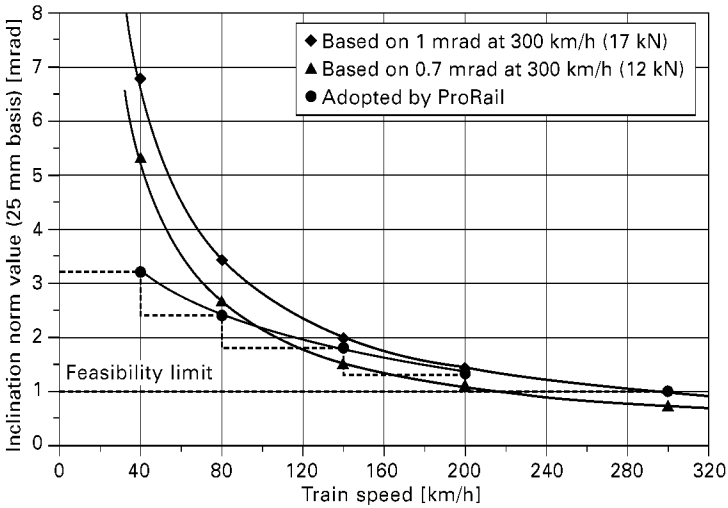
1. the empirical vertical tolerance of 0.3 mm has been used worldwide for several decades and has not led to systematic problems; the maximum peak deviation resulting from the standardized inclination should, for 140 km/h (which is a conventional passenger train speed), not have a different order of magnitude;

2. the norm value for high-speed lines (300 km/h) has been established at the feasibility limit of 1 mrad.

The norm value of 1 mrad at 300 km/h leads, according to Eq. (12.1) (with $\xi = 0.2$), to a dynamic force level of 17 kN. The value of 0.7 mrad, obtained for new straightened longrail, leads to a force level of 12 kN. The curves according to Eq. (12.1) corresponding with both force levels are depicted in Fig. 12.21. The following speed intervals have been used: 0–40 km/h, 40–80 km/h, 80–140 km/h, 140–200 km/h (conventional lines) and 200–300 km/h (high-speed lines).

The norm values should ideally be situated in the hatched area between the two curves in Fig. 12.21, with a cut-off by the feasibility limit. However, due to the hyperbolical behaviour, the values for low speeds grow to infinity, whereas it is easy to grind, even manually, to a much better quality. Therefore, the smooth function depicted in Fig. 12.21, leaving the hatched area at 40 and 80 km/h, has been adopted by the Dutch Rail Infra Manager ProRail. In 2005, the geometrical standards as shown in Table 12.1 were introduced for metallurgical rail welds in the Netherlands.

Table 12.2 gives the amplitudes of a sinusoidally shaped weld with a half wavelength of 1 m, and corresponding to the inclinations in Table 12.1. These amplitudes give an impression of the maximum vertical peak deviations corresponding to the inclination values. Using the 1.8 mrad criterion for 140 km/h, the maximum amplitude equals 0.57 mm, which is larger than – but in the same order of magnitude as – the 0.3 mm tolerance (1).



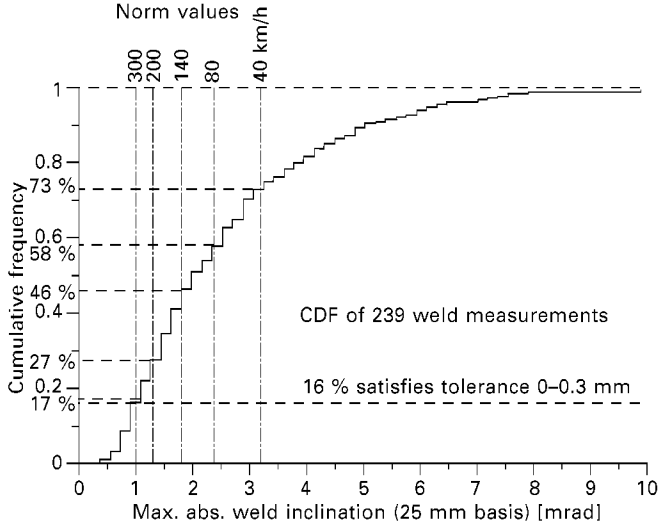
12.21 Determination of speed-dependent norm values in the QI definition.

Table 12.1 Inclination norm values adopted by ProRail for the QI determination (ProRail, 2007)

Train speed [km/h]	Max. absolute inclination (25 mm basis) [mrad]
0–40	3.2
40–80	2.4
80–140	1.8
140–200	1.3
200–300	1
NB: 60, 100 (special tracks)	*2.8, 2.2*

Table 12.2 Inclination norm values (25 mm) and the resulting 2 m wave amplitude

Train speed [km/h]	Inclination values [mrad]	Resulting 2 m wave amplitude [mm]
40	3.2	1.02
80	2.4	0.76
140	1.8	0.57
200	1.3	0.41
300	1	0.32



12.22 Comparison of acceptance levels of rail welds according to the traditional method based on tolerances and the QI-method.

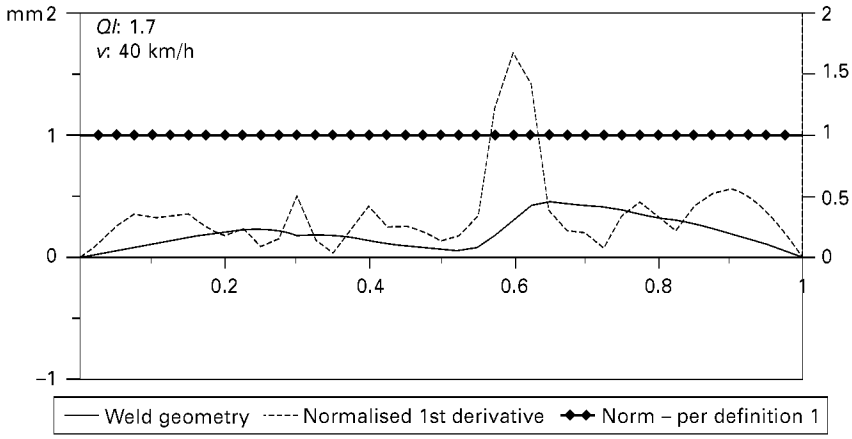
In Fig. 12.22, a comparison is made of the acceptation level of welds, according to the conventional standards based on vertical tolerances (0–0.3 mm, independent of the speed), and the method based on first derivatives. The analysis has been performed for the sample of 239 welds which has

been used for the computations earlier in this chapter. It is observed that the acceptance level increases drastically (except for 300 km/h). At the conventional 140 km/h the acceptance increases by 30 %. This relaxation, which, according to the investigation in this chapter should not lead to an increasing rate of deterioration, is due to the fact that peak deviations larger than 0.3 mm are accepted, provided that the contact geometry is smooth.

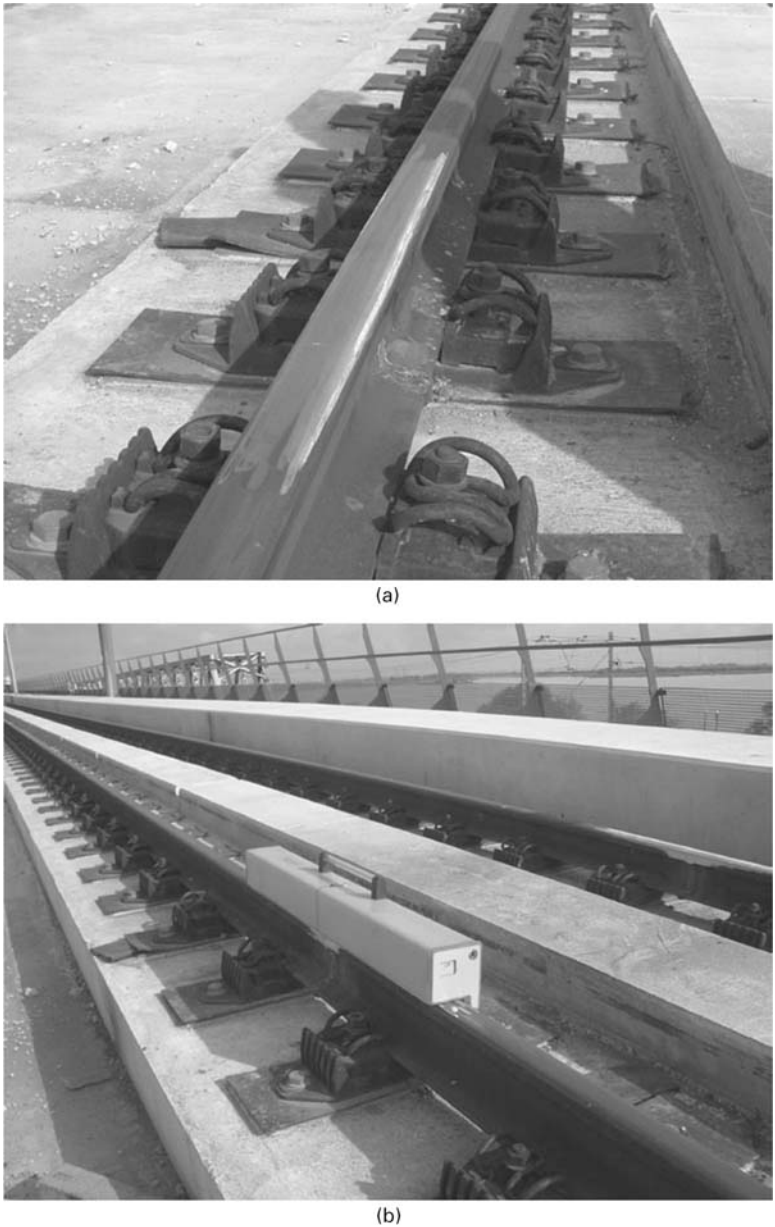
The weld geometry assessment method elaborated in this section can be simply implemented in practice. This is briefly illustrated in the following. After welding and cooling down of the rails, the quality of the geometry can be measured using an electronic digital straightedge. In the processor the routines for the calculation of the quality index can be programmed. The device samples the rail geometry and plots the normalized first derivative (dependent on the train speed) and the QI on a screen. With the help of this output, the geometry can be optimised such that the standards are met. An example of such a screen output is given in Fig. 12.23. At the location where the longitudinal rail geometry shows a relatively large inclination, the requirement is not met (the QI is 1.7), and the weld should be ground before acceptance. In Fig. 12.24, an example is shown of rail weld geometry assessment using an electronical straightedge.

12.6 Welding irregularities, energy considerations and deterioration

Previously, an assessment method of welding irregularities based on a limitation of dynamic forces in the wheel–rail interface was described. However, in many cases deterioration of a track component is directly related to energy



12.23 Example of the screen output of a digital measuring device for the assessment of the geometry of rail welds.



(a)

(b)

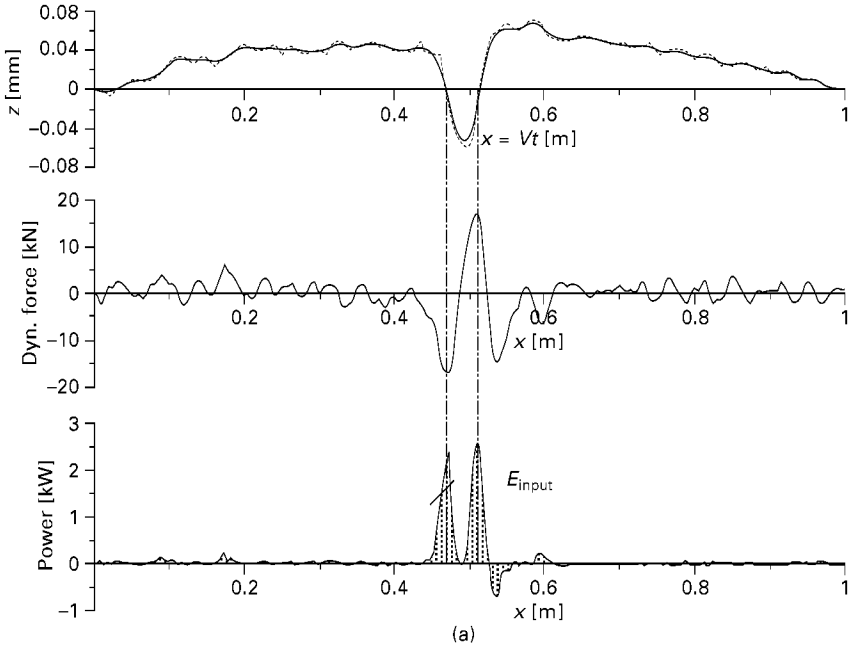
12.24 Ground rail weld (a) and geometry assessment (b) using a digital straightedge.

dissipation. Damage occurrence can then be translated into reaching a saturation limit in energy dissipation capacity in time. Therefore, in future regulations it would seem more appropriate to limit the power spectrum corresponding

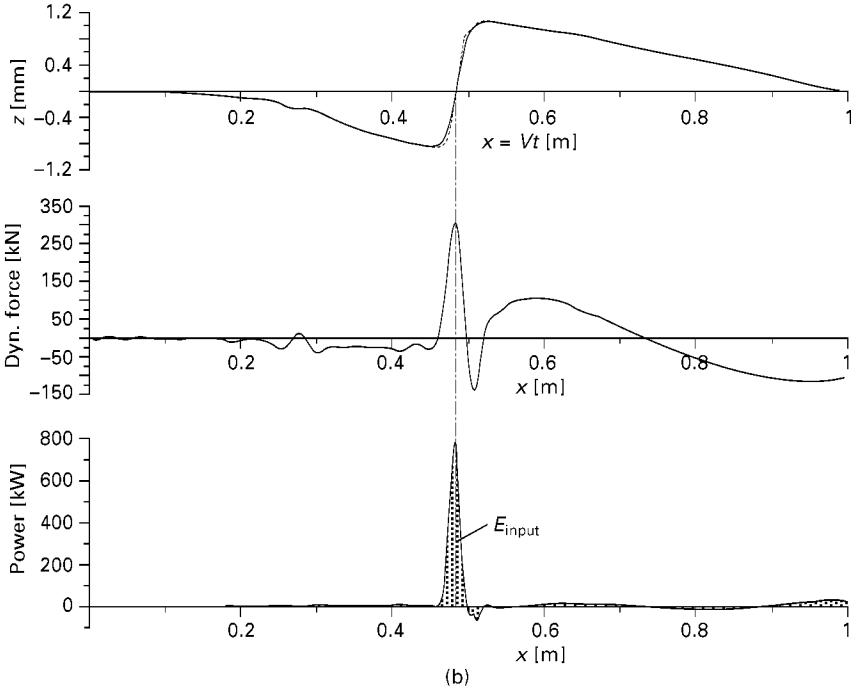
to a given track geometry in the whole wavelength or frequency domain. Although the step from purely geometrical to force-based track assessment is being made, the step from force-based to input-power-based assessment is not yet feasible. However, some examples focusing on energy-based assessment are given in this section for illustration.

Basically, a correlation between the maximum first derivative (or another geometrical parameter describing the weld geometry) and the energy or maximum power input into the track can be sought. This correlation is different from that between the maximum first derivatives and the maximum contact forces, due to the fact that the dynamic stiffness spectrum of the track (including effects from both elasticity and inertia) is a non-uniform spectrum. Depending on the geometry of the excitation, different frequency regions of the dynamic track stiffness spectrum play a role.

Figure 12.25 shows the contact force, the vertical rail velocity and the power input along the track (or energy input into the track section) for both welds displayed in Fig. 12.13, when a train axle passes at 140 km/h. The first weld, which is rather smooth but suffers from some shrinkage of the weld material, yields a maximum power input only between 2 and 3 kW. The second weld, containing a severe step, yields a maximum power input



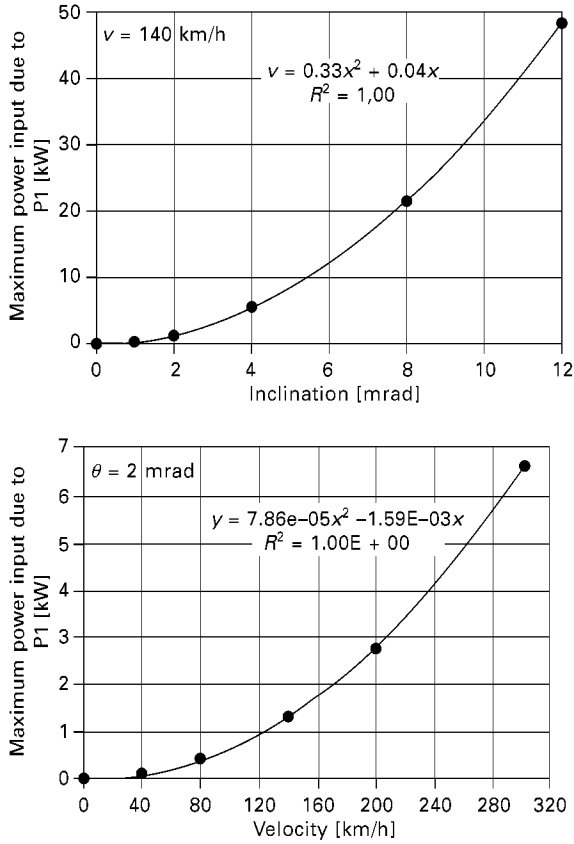
12.25 Power input into the track per train wheel for the welds depicted in Fig. 12.13 at 140 km/h.



12.25 Cont'd

between 700 and 800 kW, for each passing wheel (using a linear calculation). Given that modern locomotives generally have a traction power between 2000 and 8000 kW (the latter for high-speed trains), a power loss of 800 kW per wheel into the track due to a bad weld is very severe.

Figure 12.26 shows computational results for the elementary case of a small ramp in the rail surface, which was considered in Fig. 12.12. In Fig. 12.26a, maxima of the power input into the track due to P1 are shown as a function of the ramp inclination, for a given train speed of 140 km/h. In Fig. 12.26b this is repeated as a function of the velocity, for a given inclination of 2 mrad. The basis length of the ramp equals 25 mm in both cases. In both cases, the relationship is exactly second-order polynomial. Both relationships with the maximum dynamic force were linear. This indicates the severity of rail surface irregularities for long-term track behaviour, which is governed by energy dissipation mechanisms. These mechanisms need future investigation.



12.26 Maxima of the power input into the track due to P1 for a ramp with a basis of 25 mm. (a) Values as a function of the ramp inclination, for 140 km/h; (b) values as a function of the velocity, for an inclination of 2 mrad.

12.7 References

Böhmer A and Klimpel T (2002), Plastic deformation of corrugated rails – a numerical approach using material data of rail steel, *Wear*, **253**, 150–61.

Cai W, Wen Z, Jin X and Zhai W (2007), Dynamic stress analysis of rail joint with height difference defect using finite element method, *Engineering Failure Analysis*, **14**, 1488–499.

Chen Y, Lawrence F V, Barkan C P L and Dantzig J A (2006a), Heat transfer modelling of rail thermite welding, *Proc IMechE, Part F: Journal of Rail and Rapid Transit*, **220**, 207–17.

Chen Y, Lawrence F V, Barkan C P L and Dantzig J A (2006b), Weld defect formation in rail thermite welds, *Proc IMechE, Part F: Journal of Rail and Rapid Transit*, **220**, 373–84.

Esveld C (2001), *Modern Railway Track* (2nd edn), Zaltbommel, the Netherlands, MRT-productions.

Esveld C (2005), *Measurements of Rail and Weld Geometry on HSL-South*, ECS report, Zaltbommel, the Netherlands.

Hiensch M, Nielsen J C O and Verheijen E (2002), Rail corrugation in the Netherlands – measurements and simulations, *Wear*, **253**, 140–49.

Ilic N, Jovanovic M T, Todorovic M, Trtanj M and Saponjic P (1999), Microstructural and mechanical characterization of postweld heat-treated thermite weld in rails, *Materials Characterisation*, **43**, 243–50.

Ishida M, Moto T, Kono A and Jin Y (1999), Influence of loose sleeper on track dynamics and bending fatigue of rail welds, *Quarterly Report of RTRI*, **40**, 80–85.

Jenkins H H, Stephenson J, Clayton G A, Morland G W and Lyon D (1974), The effect of track and vehicle parameters on wheel/rail vertical dynamic forces, *Railway Engineering Journal*, January, 2–16.

Kabo E, Nielsen J C O and Ekberg A (2006), Prediction of dynamic train-track interaction and subsequent material deterioration in the presence of insulated rail joints, *Vehicle System Dynamics*, **44**, 718–29.

Li Z, Zhao X, Esveld C and Dollevoet R (2006), Causes of squats: correlation analysis and numerical modeling, *Proceedings 7th International Conference on Contact Mechanics and Wear of Rail/Wheel Systems*, Brisbane, Qld Australia, 24–27 September, 439–46.

Li Z, Zhao X, Esveld C and Dollevoet R (2007), Rail stresses, strain and fatigue under dynamic wheel-rail interaction, *Proceedings International Heavy Haul Association Specialist Technical Session*, Kiruna, Sweden, 11–13 June, 389–96.

Mutton P J (2000), Material aspects of weld behaviour in wheel-rail contact, *Proceedings Fifth International Conference on Contact Mechanics and Wear of Rail/Wheel Systems*, Tokyo, Japan, 25–27 July, 131–5.

Mutton P J and Alvarez E F (2004), Failure modes in aluminothermic rail welds under high axle load conditions, *Engineering Failure Analysis*, **11**, 151–66.

ProRail (2007), Directives RLN00127-1&2, Part 1 – Operationele eisen voor metallurgische lassen in bovenbouwconstructies; Part 2 – Kadereisen voor metallurgische lassen in bovenbouwconstructies, ProRail, Utrecht, the Netherlands.

Shitara H, Terashita Y, Tatsumi M and Fukada Y (2003), Nondestructive testing and evaluation methods for rail welds in Japan, *Quarterly Report of RTRI*, **44**, 53–8.

Skyttebol A (2004), *Continuous Welded Railway Rails: Residual Stress Analyses, Fatigue Assessments and Experiments*, PhD Thesis, Chalmers University of Technology, Gothemburg, Sweden.

Skyttebol A, Josefson B L and Ringsberg J W (2005), Fatigue crack growth in a welded rail under the influence of residual stresses, *Engineering Fracture Mechanics*, **72**, 271–85.

Steenbergen M J M M (2006), Modelling of wheels and rail discontinuities in dynamic wheel-rail contact analysis, *Vehicle System Dynamics*, **44**, 763–87.

Steenbergen M J M M (2008a), Quantification of dynamic wheel-rail contact forces at short rail irregularities and application to measured rail welds, *Journal of Sound and Vibration*, **312**, 606–29.

Steenbergen M J M M (2008b), *Wheel–Rail Interaction at Short-Wave Irregularities*, PhD Thesis, Delft University of Technology, Delft, the Netherlands.

Steenbergen M J M M and Esveld C (2006a), 'Rail weld geometry and assessment concepts', *Proc. IMechE, Part F: Journal of Rail and Rapid Transit*, **220**, 257–71.

Steenbergen M J M M and Esveld C (2006b), Relation between the geometry of rail welds and the dynamic wheel–rail response; numerical simulations for measured welds, *Proc IMechE, Part F: Journal of Rail and Rapid Transit*, **220**, 409–24.

Tawfik D, Kirstein O, Mutton P J and Chiu W K (2006), Verification of residual stresses in flash-butt-weld rails using neutron diffraction, *Physica B*, 385–6, 894–6.

Tawfik D, Mutton P J and Chiu W K (2008), Experimental and numerical investigations: alleviating tensile residual stresses in flash-butt welds by localised rapid post-weld heat treatment, *Journal of Materials Processing Technology*, **196**, 279–91.

Terashita Y and Tatsumi M (2003), Analysis of damaged rail weld, *Quarterly Report of Railway Technical Research Institute*, **44**, 59–64.

Webster P J, Mills G, Wang X D, Kang W P and Holden T M (1997), Residual stresses in alumino-thermic welded rails, *Proc. IMechE, Journal of Strain Analysis*, **32**, 389–400.

Wimmer W E, Connell D A and Boos M J (2002), Joint elimination through mobile flash-butt welding on Union Pacific Railroad, *Proc. of Innotrans 2002*, Berlin, Germany, 24–27 September.

Winter T, Meijvis P A J, Paans W J M, Steenbergen M J M M and Esveld C (2007), Track quality achieved on HSL-South – reduction of short-wave irregularities cuts life cycle cost, *European Railway Review*, **13**(3), 48–53.

Wöhnhart A and Wenty R (2002), Mobile flash-butt welding: three decades of experience, *Rail Engineering International*, **3**, 11–16.

Copyrighted Materials
Copyright © 2009 Woodhead Publishing Retrieved from www.knovel.com

13
Squats on railway rails

Z. LI, Delft University of Technology, The Netherlands

Abstract: Squats are becoming an increasing rail top damage problem for many railways. In this chapter the definition, classification and characteristics of squats are introduced and past research is reviewed. Correlation of squat occurrence with track parameters is discussed. Recent work on stress and strain analyses of three-dimensional frictional rolling contact is then presented; this concerns mainly the inclusion of plasticity and continuum dynamics, in view of the continuous accumulation of plastic deformation at squats under repeated high-frequency dynamic contact force. What follows explains, with the help of numerical analyses, how squats initiate due to differential wear and differential plastic deformation and how they grow. Subsequently detection methods and counter measures are presented, and further research directions are discussed.

Key words: squats, rolling contact fatigue (RCF), dynamic rolling contact, friction, plasticity, track short defect, detection, differential wear, differential plastic deformation.

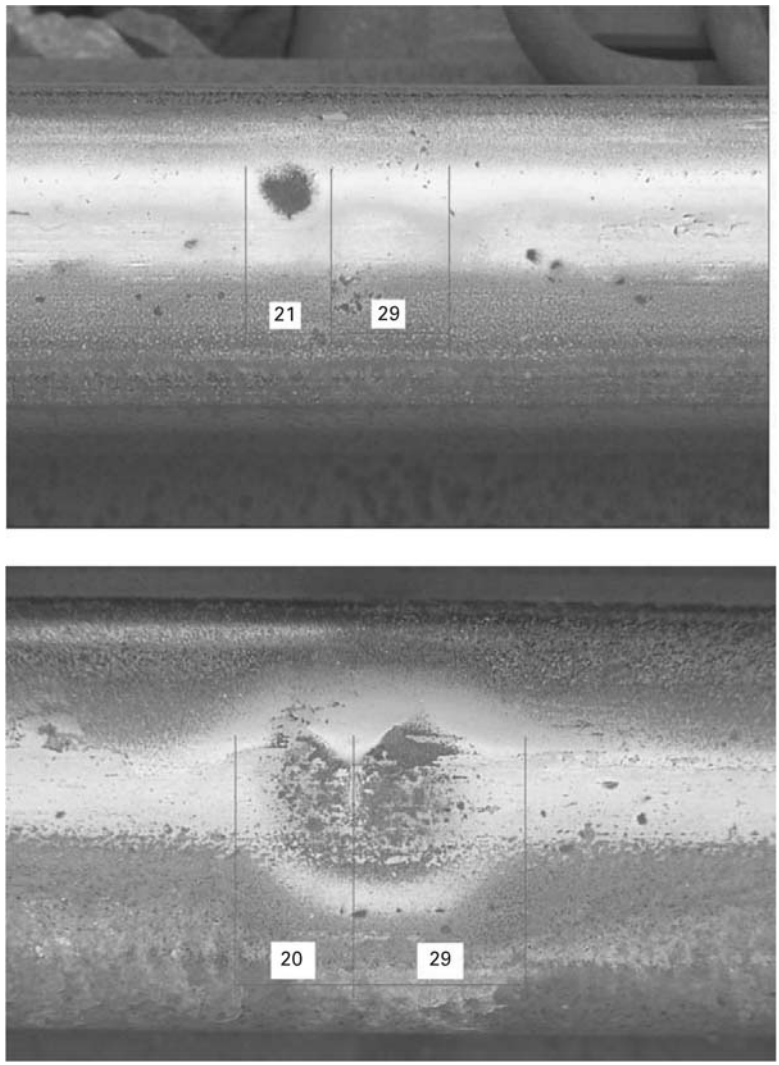
13.1 Introduction

Squatting is a type of rolling contact fatigue (RCF) which was hardly known about 30 years ago in Europe (Clayton *et al.*, 1983; Cannon and Pradier, 1996). It was first identified as a distinct failure in the 1970s (the Blue Book of RailTrack, 2001). Before that it was probably categorized as other more commonly occurring and similar-looking failure types, such as taches ovales. Now it has become one of the major sources of RCF damages on many railways (the Blue Book of RailTrack, 2001; Smulders, 2003).

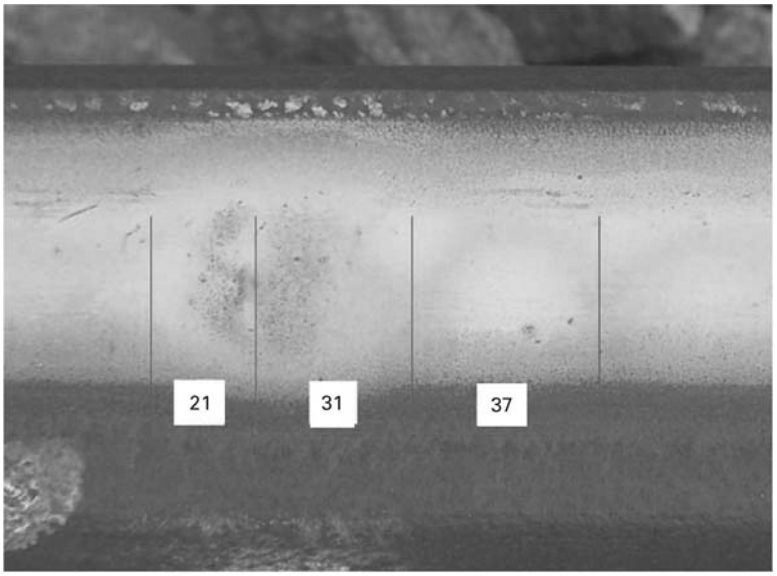
Squats usually occur on top of rails in the running band in straight track and large curves (Fig. 13.1). They are found on all types of track: ballast and slab track, with wooden or concrete sleepers, lines with passenger, freight or mixed traffic, and high-speed, conventional and metro lines. A fundamental characteristic of squats is large local plastic deformation – no shakedown takes place. Some of them are related to rail top surface vertical irregularities such as at indentations and welds of continuously welded rails (CWR). Excessive dynamic wheel–rail contact force is the direct cause for their growth; changes in the traffic–track interactive systems since the 1970s lie behind its increasing occurrence. These changes include, among other things, higher axle load and traffic density, different maintenance policy and deterioration in track quality, new wheel and rail materials, new rolling stock

with traction control and Anti-lock Braking Systems, and stiffer concrete sleepers, etc.

In The Netherlands, squats are visually classified into three categories – light, moderate, and severe (Smulders, 2003), or classes A, B, and C, correspondingly, see Fig. 13.1. Cracks can usually be seen at class B and C squats. The high wheel–rail dynamic force at squats causes accelerated local track deterioration, which spreads rapidly out. The large impact force,



13.1 Squats at different stages of development. Numbers indicate length in mm. (Li *et al.*, 2008a)



13.1 Cont'd

together with the crack accompanying the squats, may result in catastrophic rail break.

13.2 Review of past research

So far research on squats has been limited, falling mainly into the following three categories.

13.2.1 Metallurgical research

Clayton *et al.* (1983) reviewed a metallurgical research program at British Rail on surface-initiated rail problems; squats were found to be a problem of great concern on some routes in the UK. Longitudinal/vertical section of rail specimens showed surface-initiated cracks which can branch downwards; when they reach a critical size, brittle fracture can result in a broken rail. It was found that fatigue life is reduced as contact pressure is increased and that creepage plays an important role. A factor of ten differences has been observed in the RCF resistance in two tested materials. More of this squats-related work can also be found in Clayton and Allery (1982) and Clayton and Hill (1987).

Marich (2006) presented the Australian experience with squats. He linked squats with the hard and brittle 'white etching' layer which is most commonly found on infrequently ground rail. Such a layer can form due to adiabatic

shear between the wheel and rail surfaces caused by the micro-slip under traction. He concluded, therefore, that the development of squats is very similar in nature, but not in degree, to the development of wheel burns. Nevertheless, he found that this explanation alone could not explain why squats usually develop only on one rail, and what particular operating/track conditions enhance their development.

13.2.2 Stress and crack growth research

Bold *et al.* (1991) studied the growth rate of squat type shallow-angled surface-initiated RCF cracks. In the 1990s the European Rail Research Institute (ERRI) carried out a RCF research program (Cannon and Pradier, 1996). Squats crack was modeled by Bogdański, leading to a series of publications, e.g. Bogdański *et al.* (1998), Bogdański and Brown (2002), Bogdański and Lewicki (2008). Bogdański *et al.* (1998) modeled a squat as a plane oblique semi-elliptical crack. The state of stress in the vicinity of the crack front was determined, and the values and ranges of the stress intensity factors at the crack front were calculated. By combining crack front loading histories with mixed-mode fatigue crack growth rate data Bogdański and Brown (2002) further analyzed the growth of squat-type cracks. In Bogdański and Lewicki (2008), the effect of entrapped liquid was modeled.

Dang Van and Maitournam (2002) presented, for the case of squats, calculations of stresses and strains in the rail subjected to repeated moving contacts. Stationary methods were employed. Busquet *et al.* (2005) computed plastic flows in the near-surface layer as a function of traction coefficient; surface contact load distribution was based on the solution of Kalker (Kalker, 1990; Chollet, 1999).

13.2.3 Initiation and growth mechanisms

If the initiation and growth mechanisms of squats are known, the problem may then be tackled at the root causes. This may lead finally to preventive maintenance actions based on detection, measurement, and prediction.

Since 2006, some research works have concentrated on the mechanism of squat initiation and growth, and on their root causes (Li, 2006; Zhao *et al.* 2007; Li *et al.*, 2008a, b). Their approach consists of correlation analyses, numerical analyses and monitoring. The correlation analyses relate squat occurrence to certain parameters in the vehicle–track interaction system, and to observations of phenomena around squats in the tracks. It helps eliminate many of the less influential parameters and helps significantly simplify the subsequent numerical model. The numerical analyses are employed to quantify the relation between the influential parameters identified in the correlation analysis and the dynamic rolling contact forces, stress and strain. A few of

the most influential parameters can thereby be further isolated. Monitoring in the tracks provides realistic inputs to the analyses, as well as data for validation. This chapter will follow this line of research.

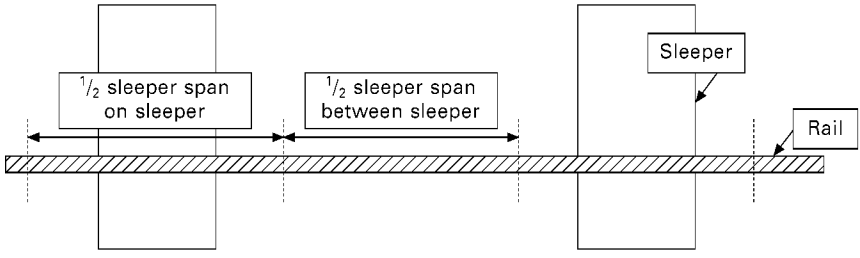
13.3 Correlation of squats with track parameters

Correlations have been established based on analyses of measured track geometry data and field observations of the Dutch Railways (Li, 2006).

13.3.1 Correlation with track structure parameters

Correlation has been identified with short-wave irregularities. It was found that about 74 % of the squats were on the ½ rails centered on sleepers and the rest were on the other ½ rails centered between two sleepers (see Fig. 13.2). This suggests that the stiffness and damping characteristics of the track, particularly those of the rail and the railpad, and their technical status may have played a role.

It was also found that short pitch corrugation with various severities could be seen in the neighborhood of about 72 % of the squats, with a wavelength ranging between 2 and 6 cm. Later a numerical analysis (Li, 2008a) showed that the growth of a squat is accompanied and promoted by a dynamic wheel–rail interaction force, which is excited by the squat itself and which has the wavelength of short pitch corrugation. This suggested that for the corrugation seen in the neighborhood of the above-mentioned 72 % of squats, some might not really be corrugation, but could be wave patterns of the non-uniform plastic deformation or wear caused by the dynamic contact force. The available data were, however, not suitable for proving such a proposition, because the traffic directions were not known. Another field survey was therefore subsequently conducted; it confirmed the proposition: in total 74 % of the investigated squats had short pitch corrugation-like rail surface waves around them, among which 33 % was indeed short pitch corrugation, which means that the waves are simultaneously before and

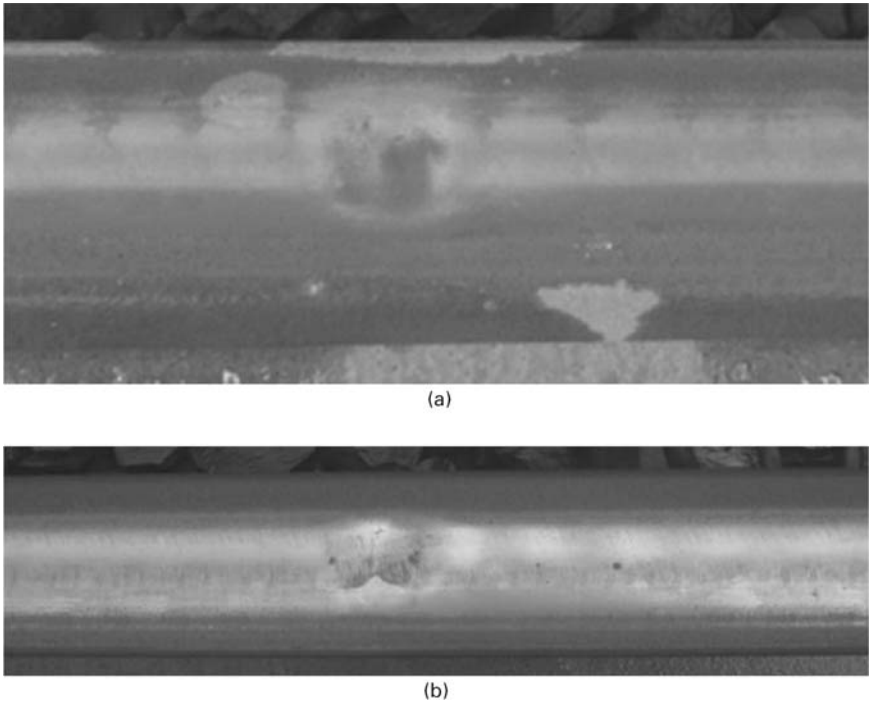


13.2 Definition of the half sleeper span centered on sleepers and the other half between sleepers.

after the squats. For the other 41 %, the wave pattern was found only after the squats (see Fig. 13.3 for illustration). For all the investigated squats the wavelength of the wave pattern was again in the range of that of the short pitch corrugation.

The similarity of squats and short pitch corrugation in their wave pattern has some significance. The squats found at corrugation should have initiated directly from the rail surface irregularity of the corrugation. The other 41 % should have initiated from other sources, the wave pattern following them being the consequence of the dynamics force excited by them. In both cases the wavelength should have been determined by the eigen characteristics of the local wheel–track system. In view of the similarity in the wavelengths of the corrugation and the wave pattern, the eigen systems for the occurrence of short pitch corrugation and of the squats wave pattern should have some characteristics in common.

These statistics are largely in line with an investigation in the UK: according to historic research 75 % of squats are associated with one of the following features: corrugation, welds and periodic indentations in the rail



13.3 Short pitch corrugation and corrugation-like wave pattern after a squat. Their wavelength is usually in the range of 2–6 cm. Traffic from left to right.

running surface caused by hard objects brought forward by the wheels (the Blue Book of RailTrack, 2001).

The relationship between squat occurrence and track irregularities such as alignment, cross-level, gauge and vertical profile was analyzed and the effects of rail vertical wear and rail foot inclination examined. There was no clear correlation found, as was expected, because these measurements are of the long-wave type, with a resolution of 25 cm.

13.3.2 Correlation with welds and materials

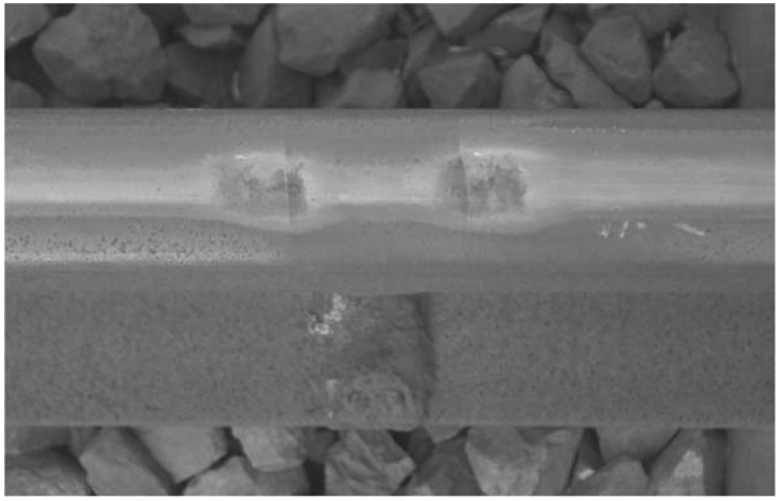
Out of different field surveys it was found that about 10–15 % of squats were at welds. Figure 13.4 shows two typical cases. The vulnerability of welds, both thermite and flash-butt, to squats, may be explained by two reasons: material inhomogeneity in the heat-affected zone and geometry deviation.

Figure 13.5a illustrates the hardness distributions of thermite and flash-butt welds. Part, if not all, of the weld and the heat-affected zone have lower hardness than the parent material. Comparing Fig. 13.4 with Fig. 13.5, resemblance can clearly be seen between the damage, the hardness variation and the rail top longitudinal profiles. Obviously the hardness difference in the heat-affected zone will lead to variations in plastic deformation and wear behavior. This local non-uniform deformation and wear, when accumulated after repeated wheel passages, may cause significant increase in local dynamic contact force which, in turn-results in further differential deformation and wear, like a positive feedback. If there is already geometrical deviation due to imperfect grinding after welding, the problem will be exacerbated.

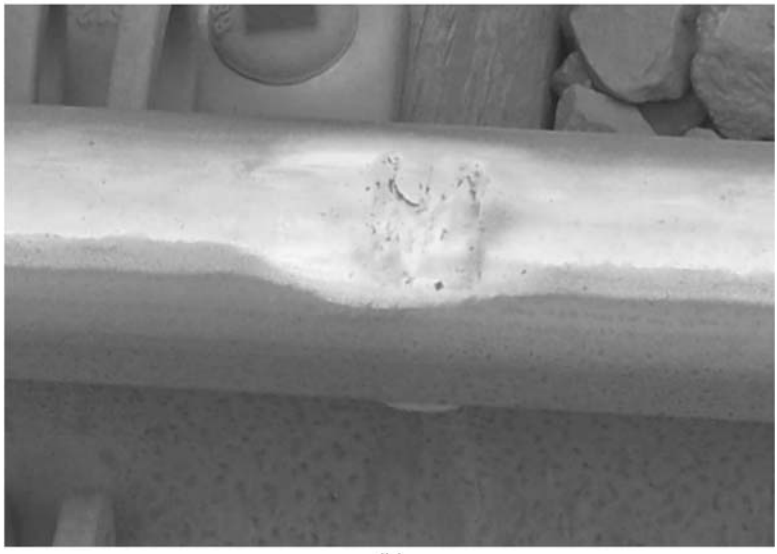
13.3.3 Correlation with rail surface irregularities

In nature, squats are visually rail top geometry deviation due to large plastic deformation. They must therefore grow from some small rail top irregularities. These can be:

- Indentations by hard alien objects between wheels and rail. These could be, for instance, a hard ball from an aerosol paint can, or a ball or roller from a bearing. They can indent into the wheel and be brought forward. Such indentations can be recognized by their periodicity of the wheel circumference – usually about 3 m.
- Vertical misalignments of the rails, e.g. at switches and crossings.
- Short pitch corrugation.
- Skidding and sliding damages by wheels during traction and braking. This has to do with the low adhesion due to contamination on railheads, with the high traction power of modern motorized passenger cars and locomotives, and with the characteristics and performance of their



(a)

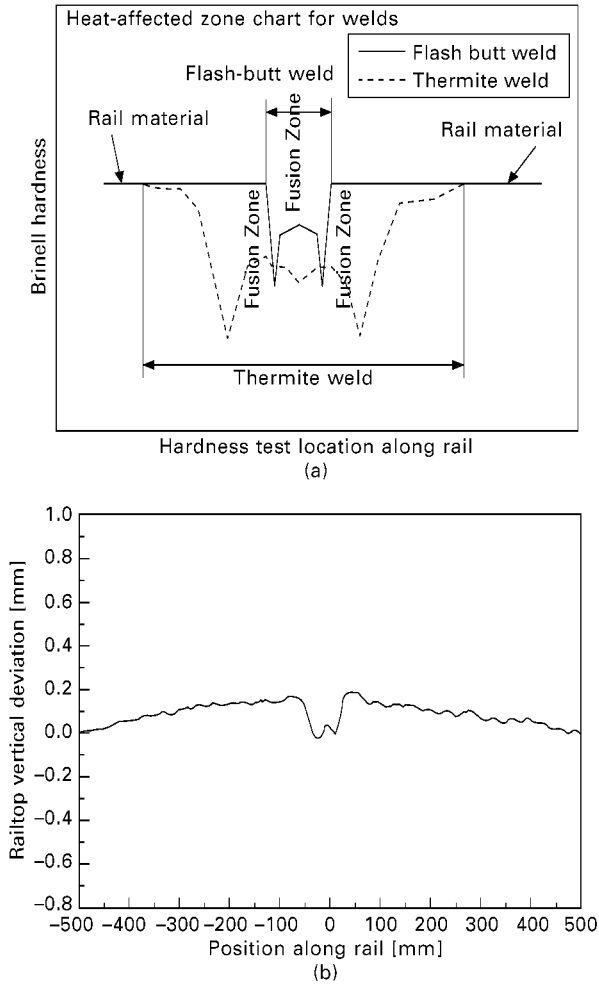


(b)

13.4 Squats at welds.

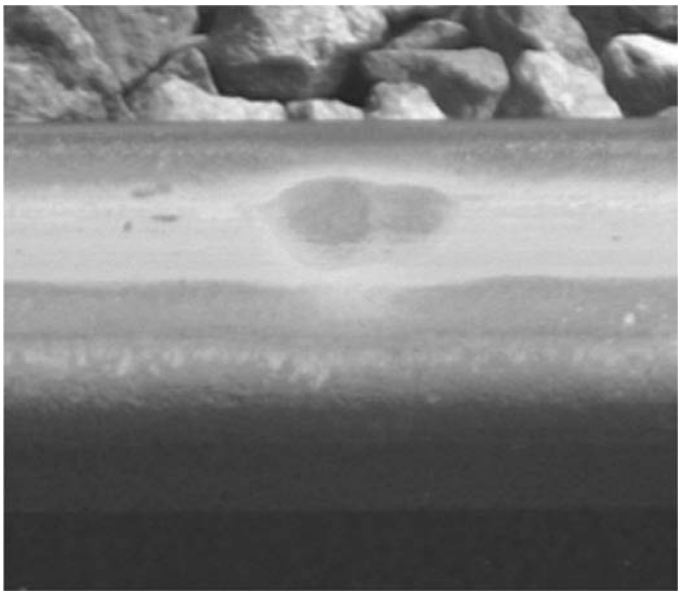
traction control systems and Anti-lock Brake Systems, which allow large slip. Wheel burns can often be observed at and near stations, but also elsewhere.

Some authors make a distinction between wheel burns and squats. In this chapter they are considered as an initiation source of squats because defects developed from wheel burns in the later stages bear all the characteristics of squats. Figure 13.6 shows such an example.



13.5 Hardness (a) and geometry (b) variations at welds.

- Differential wear and differential plastic deformation. Rails inevitably experience wear, and often also plastic deformation. This wear and deformation is usually uniform along the rails. If, for some reasons, non-uniform local wear or plastic deformation takes place repeatedly at fixed places and accumulates, differential wear and deformation occur. Welds, due to their heat-affected zone and poor finish geometry, can often have such wear and deformation, as is discussed above. Short pitch corrugation, with its damage mechanism being plastic deformation and wear (Kalousek and Grassie, 2000), may also be of the type of differential wear and differential deformation which repeats itself over a long length of rail. The wave pattern following a squat discussed above is also due



13.6 A squat developed from a wheel burn.

to differential wear and plastic deformation caused by the dynamic force excited by the squat itself. Poorly laid or maintained tracks have short defects in the rail, railpad, fastening, sleepers, and sleeper supports, etc. They may excite dynamic contact force of the necessary wavelength repeatedly at the same location so that differential wear and plastic deformation occur. Such wear and deformation, when accumulated to a certain degree, become sources of squat initiation. One such case is presented in Section 13.6 below.

13.3.4 Correlation with friction

There have been observations which show that squat occurrence is much higher at locations of high traction and braking efforts. European railway statistics show significant increase in squats occurrence at ascending gradients higher than 1 %. Marich (2006) pointed out the relation between squat occurrence and the micro-slip between wheel and rail.

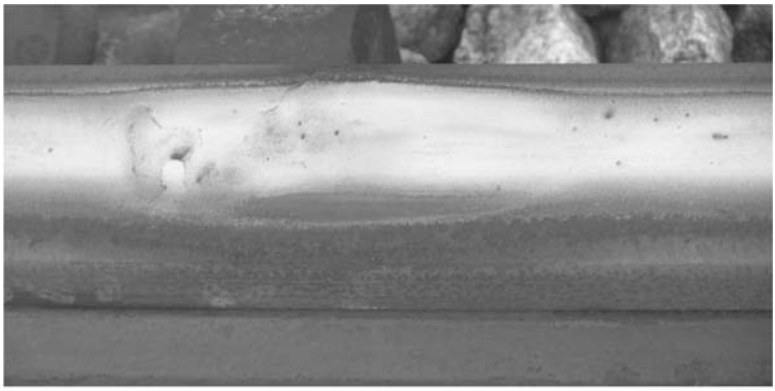
13.4 Characteristics of squats

The characteristics of squats are summarized as follows:

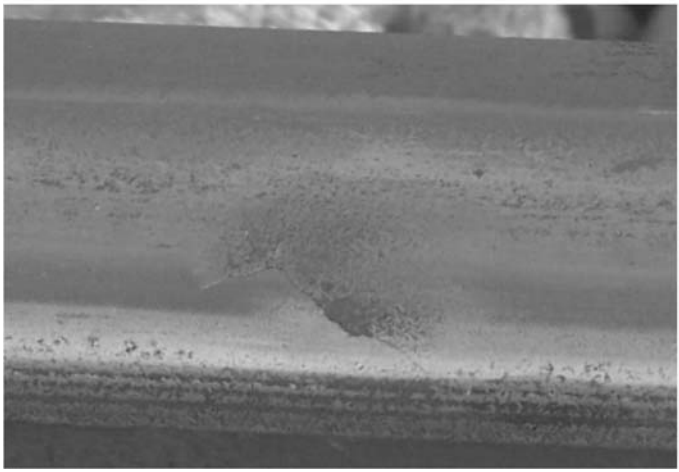
- An individual squat is local visible surface deformation on the rail top in the running band.

- A squat begins from a small rail top vertical irregularity. It excites dynamic wheel–rail interaction force, which promotes further continuous growth in three dimensions of the defect in an accelerated way. No shakedown will therefore take place.
- Squats are usually found on straight tracks and large curves.
- Opinion is still divided about whether squats are surface- or subsurface-initiated. The subsurface-initiation opinion bases its arguments on the fact that with advanced squats cracks are usually found in subsurface and, with some indentations, cracks were also observed by sectioning.

  There seems to be a tendency to shift to the surface-initiation opinion. This opinion is supported by the correlation discussed above: squat occurrence is mainly associated with rail top surface irregularities, weld and high traction and braking efforts. There should usually be no crack there in the beginning. The high tangential force shifts the maximum shear stress from subsurface to surface. They are therefore distinct from some subsurface-initiated defects which may appear similar to squats in one way or the other. Such defects may initiate from subsurface defects like inclusion. With visual inspection and eddy current measurement of small growing squats, usually no crack can be found. Cracks initiate and grow in the growing process of a squat when ductility of the material is exhausted under the repeated and ever-increasing dynamic force.
- Squats occur often in isolation, in a seemingly random way. Here it is called 'seemingly random', because it is believed that there must be some deterministic mechanisms behind them which should be detectable or traceable.
- If they occur in multiplication, they may often be associated with corrugation.
- Squats in their advanced stages (classes B and C) usually have visible accompanying V- or U-shaped cracks (see Figs 13.1b, 13.4b and 13.7 for examples). In the last stage, when the rail has to be replaced, the network of cracks is more complicated (see Fig. 13.7). In the Blue Book of RailTrack (2001), there is a more detailed discussion about the cracks.
- Sometimes severe head checks may also have a squat-like appearance due to the impact force caused by the subsurface cracks of the head checks (see Fig. 13.8). They can be distinguished from squats because when head checks occur they usually occur with many of them next to each other continuously on curves on the gauge shoulder or gauge corner.
- The identification of class A squats by visual inspection is rather debatable. For instance, very often indentations of larger than a certain size are recorded as class A squats, but strictly speaking only those indentions which will grow into class B, and therefore further into class C, can



13.7 A severe squat with a network of surface and subsurface cracks, the deepest of which is 16 mm (ultrasonic measurement). The length has exceeded greatly the usual size of squats of 2–6 cm. (Li *et al.*, 2008a)



13.8 Squat-like head check (compare with Fig. 13.1(b) for difference).

be called squats. The problem is how, based on visual information, to assess the potential of an indentation to grow. Therefore the accuracy of statistics of class A squats obtained by visual inspection depends greatly on the experience of inspectors, and may vary considerably.

- For many squats the middle part is often rusty in contrast to the surrounding area. For class A squats this may simply be due to the fact that the indentation is too deep to have wheel–rail contact there. For more advanced ones, a network of cracks may be found beneath the dark surface (Clayton *et al.*, 1983).

- A squat occurring in isolation may be less dangerous than head checks. Multiple squats within the length of two sleeper spans should be treated with more care.
- Another outstanding feature is the wave pattern and corrugation associated with squats. They indicate the important role of the high-frequency dynamic wheel–rail interaction on the initiation and growth of squats (Li, 2008a).

13.5 Three-dimensional dynamic rolling contact solutions in elastoplasticity

Kalker's (1990) theory, particularly his program CONTACT is usually employed for solutions of frictional wheel–rail rolling contact when detailed surface stresses and micro-slip are sought. It is, however, based on linear elasticity and half-space approximation in statics. The solution of CONTACT is basically of a boundary element method nature.

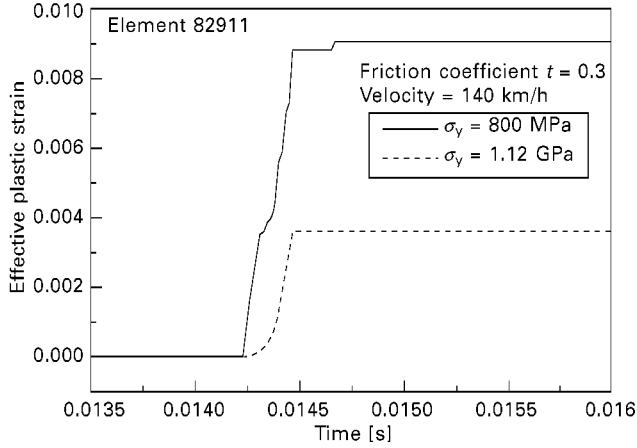
As has been discussed, the development of squats is related to plastic deformation under high-frequency dynamic interaction. The geometry in the vicinity of the contact area may not be flat. This means that a finite element (FE) solution should be sought. To such end, a method for solution of 3D dynamic contact in elastoplasticity with friction was developed by Zhao *et al.* (2007) and Li *et al.* (2007), and applied to the initiation and growth of squats and their root causes in Li *et al.* (2008a, b). In view of the fact that some of the track and vehicle parameters have strong influence on the initiation and growth of squats, the rolling contact in these works was treated in the vehicle–track interaction systems, with all the relevant factors taken into account. A model of such a system is presented in Section 13.6.

13.5.1 Effect of material strength

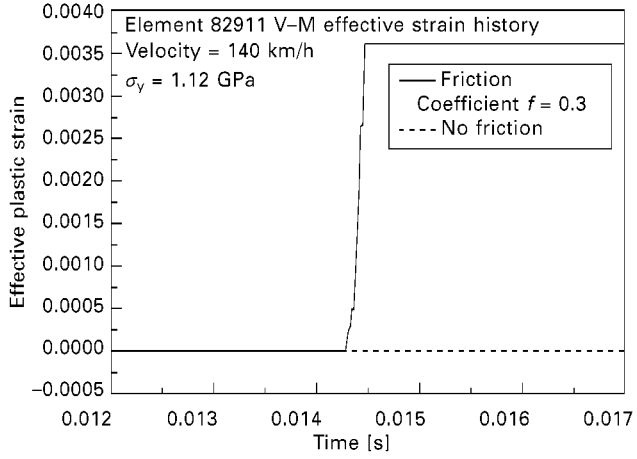
Material with high yield strength can withstand high stress without deformation. Figure 13.9 shows the time histories of the von-Mises stresses and plastic strains on the rail surface when a wheel rolls over it (Li, 2008a). The friction coefficient was 0.3, and it was assumed that the rail was free of residual stress. With all the other conditions being the same, it can be seen that with high yield stress, the plastic deformation is smaller.

13.5.2 Effect of friction

High traction and braking efforts bring the maximum shear stress from subsurface to surface. High wheel load combined with high friction will result in plastic deformation. Figure 13.10 shows the effect of friction on the plastic deformation of a rail surface layer run over by a wheel.



13.9 Time histories of effective strain σ_Y is yield stress. (Li *et al.*, 2008a)



13.10 Comparison of plastic strain under different friction conditions. (Li *et al.*, 2008a). (V.M = von Mises)

13.5.3 Effect of unsprung mass

For high-frequency dynamic contact like at squats, the effect of unsprung mass on the dynamic force may be different, depending on its distance from the contact point. Based on an FE calculation (Li *et al.*, 2007), the dynamic factor of a lumped mass on wheel axle is 2–2.5, while the dynamic factor of distributed mass of wheel tire can be up to 12. Here the dynamic factor is defined as increase in dynamic force over increase in weight. This means

that for high-frequency wheel–rail interaction, the error may be large if the unsprung mass is still considered as one single mass on a Hertz spring.

13.5.4 Forces and surface stresses distribution of dynamic contact

When a wheel of smooth surface rolls on a smooth rail, vibrations of the system will be excited. Waves of different frequencies are present. Some of their wavelengths may be related to the development of squats and corrugation.

Figure 13.11 shows the time histories of the vertical contact forces and maximum von-Mises stress of rolling contact between smooth wheel and rail surfaces. The speed is 140 km/h (Zhao *et al.*, 2007). It can be seen that the level of tangential force has only small influence on the vertical force, but it has large effect on the von-Mises stress. Apart from a major frequency component, there are also higher frequency components in the waves. This is especially noticeable in Fig. 13.11b. Such high-frequency components can also be observed in the surface stress distributions (see Fig. 13.12).

Figure 13.13a shows the tangential traction distribution. In general, it matches the pattern for solution for static rolling contact (Fig. 13.13b), with the stick and slip areas distinguishable. Nevertheless the disturbance of the high-frequency components is visible.

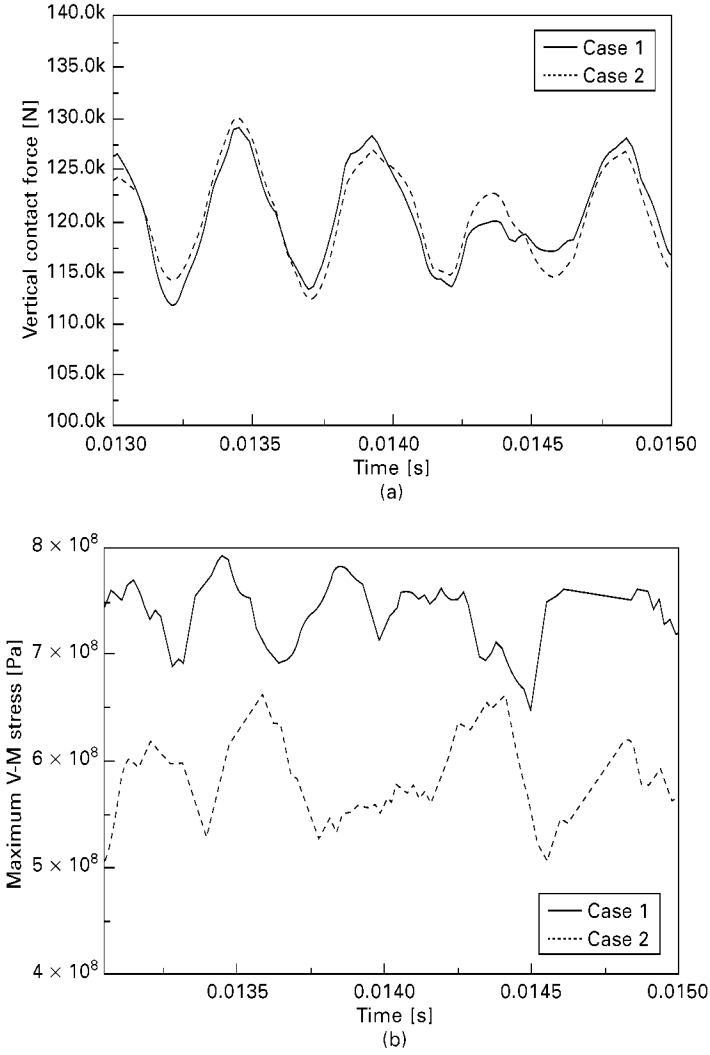
13.6 Squats initiation due to differential wear and differential plastic deformation

13.6.1 How differential wear and deformation occur

Differential wear and differential plastic deformation have been mentioned as one of the major initiation sources of squats, but how can such wear and deformation occur?

It has been observed frequently that squats can be found at locations of stiffness change in the tracks, such as at the end of fish-plates and in switches and crossings (S&C). Figure 13.14 shows such an example. Obviously stiffness change should not be the only cause, otherwise there will be squats in almost every S&C and at each fish-plated joint. There should be other influential factors behind.

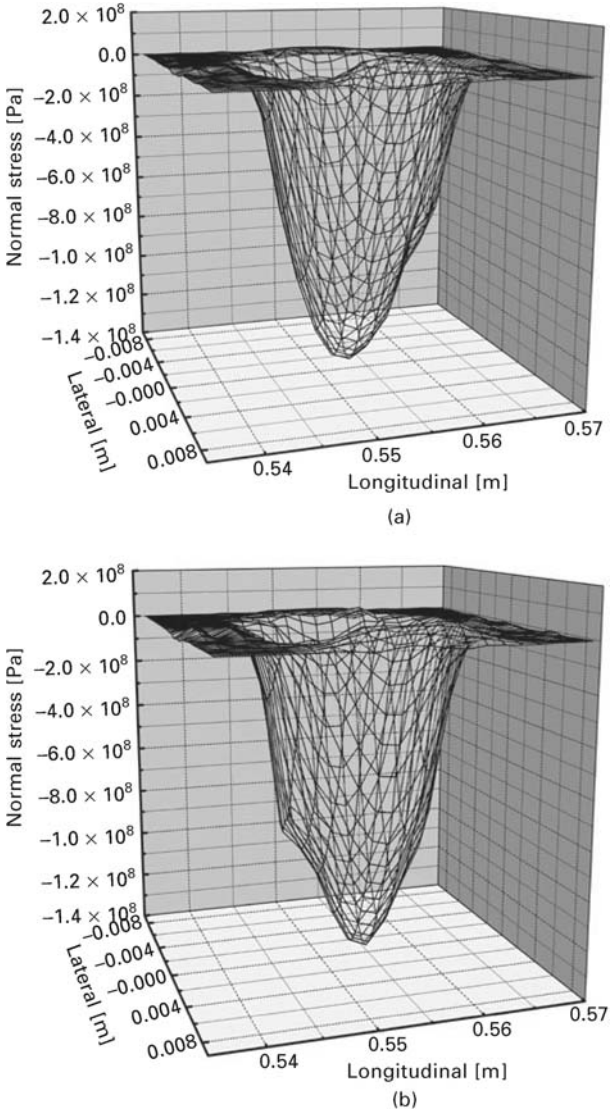
The squat in Fig. 13.14 provides a good case for study. A transient FE model was employed (Li *et al.*, 2008b), as shown in Fig. 13.15. The fish-plates were pressed against the rails by the pre-load of the four bolts numbered 1 to 4 in Fig. 13.15a. Contact and friction were defined at the interface between the fish-plates and the rails. A wheel rolled on the rail at a speed of 140 km/h, a typical line speed on the Dutch railway network. Parameter



13.11 Time histories of the vertical contact forces and maximum von-Mises stress of rolling contact between smooth wheel and rail surfaces. Friction coefficient is 0.3. The tangential force is 30 % of vertical force for case 1 (a) and 10 % for case (b). (Zhao *et al.*, 2007)

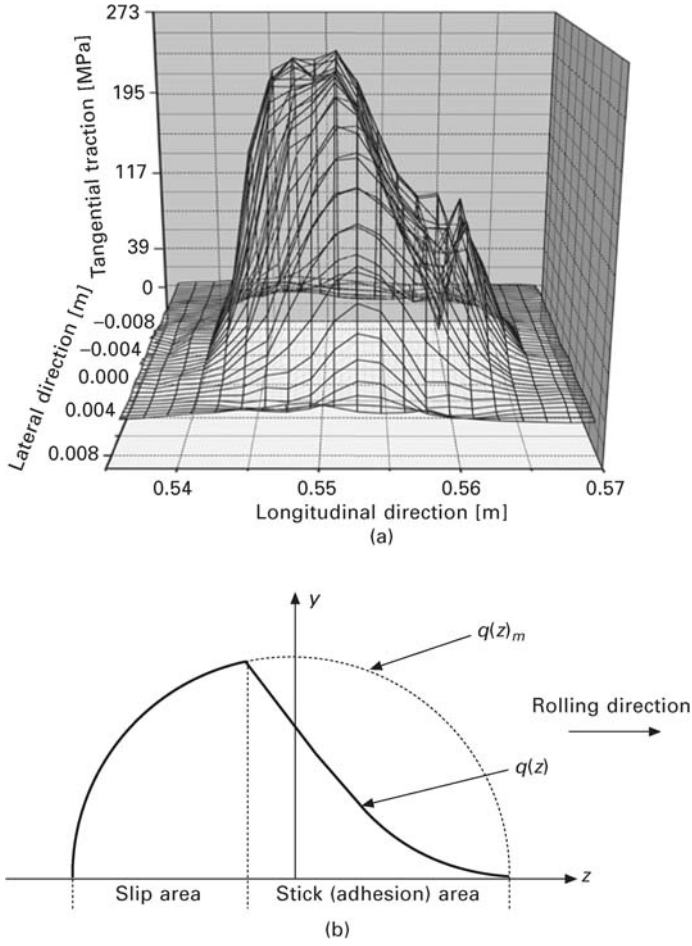
variation analysis was performed with different ballast and fastening stiffness and damping, and different fish-plate bolt pre-load. It was found that it was the fish-plate pre-load condition which played the most important role.

Figure 13.16 shows the vertical and longitudinal contact forces under different fish-plate pre-load. They have the following characteristics: at the fish-plate end the vertical forces are all broader and flatter at the peak than



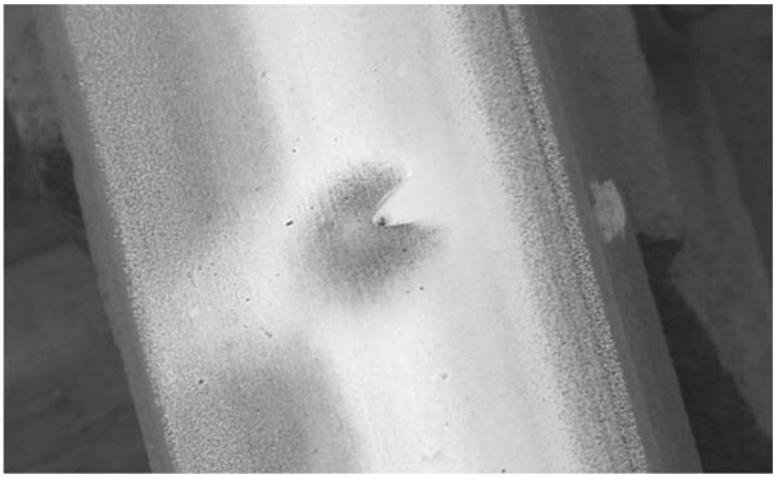
13.12 Normal pressure distributions with effect of high-frequency waves visible. Friction coefficient is 0.3. The tangential force is 30 % of vertical force for case 1 (a) and 10 % for case (b). (Zhao *et al.*, 2007)

elsewhere. When the 4th bolt is loose, the peaks, of the longitudinal force in particular, are the highest at the fish-plate end. This means that the stiffness change at the fish-plate end, together with the loose bolt being closest to the fish-plate end, causes at each wheel passage at the same location a larger



13.13 Tangential surface stress distribution. (a) 3D distribution of dynamic rolling contact solution. (b) Static rolling contact solution. (Zhao *et al.*, 2007). ($q(z)$ is the tangential surface stress and $q(z)_m$ is the limit of the tangential surface stress)

contact force with a peak of the necessary width. This force causes more wear at this location than in its neighborhood. Differential wear arises as a consequence. If the force is large enough, differential plastic deformation will also occur. The width of the peak is of importance because if it is too narrow with respect to the size of the contact area, differential wear and deformation may actually not be able to take place. The fish-plate and the loose bolt fix the wear and deformation at the same location. Without the location fixing mechanism, the wear and deformation will not accumulate because each



13.14 A squat at the end of a fish-plate. The traffic is downwards. (Li *et al.*, 2008b)

passing wheel is somewhat different from another in wavelength, phase and magnitude, so that statistically everywhere along the rail the chance is equal for wear and deformation, and the wear and deformation will be uniform.

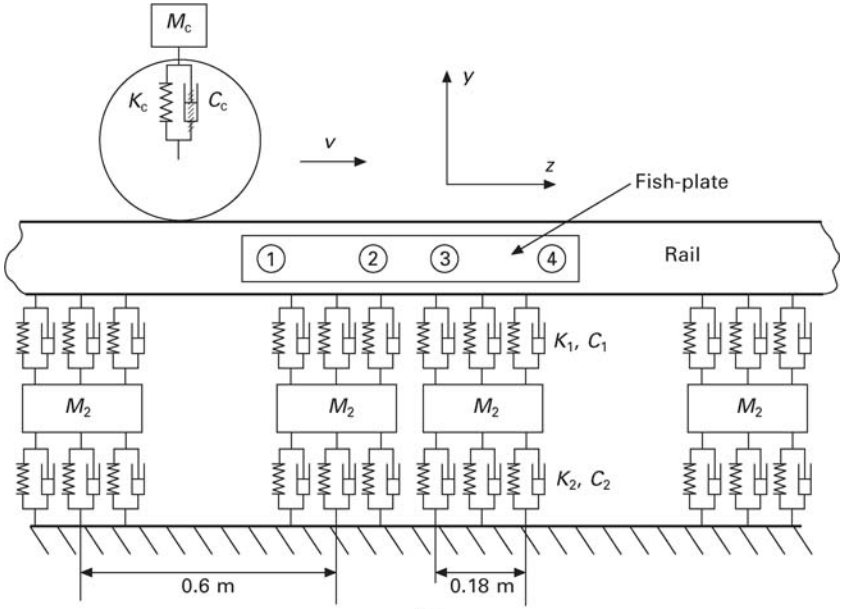
The analyzed case shows that one single parameter may not be sufficient to cause a squat. Squats may often be a consequence of the interplay of multiple parameters. This is also confirmed by correlation analyses.

13.6.2 Differential wear and deformation cause increased dynamic force

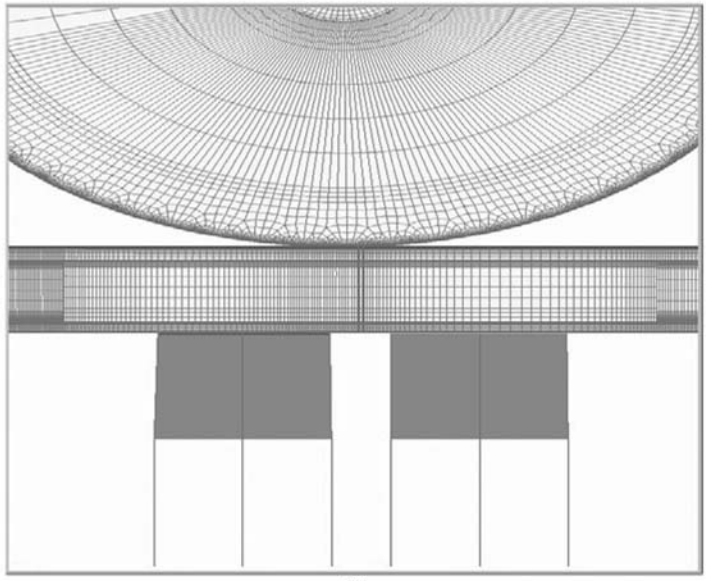
With the accumulation of differential wear and deformation, a small local geometry deviation occurs on the rail surface. Such a defect, as assumed in Fig. 13.17a, will cause an obvious increase in the dynamic force at the same location, see Fig. 13.17b, which will further promote wear and deformation.

13.7 Squats growth process

The process of squats growth from a small surface geometrical deviation to the typical shapes of Figs 13.1b and 13.1c was first postulated based on numerical simulation (Li *et al.*, 2008a). It has now been proved true by monitoring of squats in the field. In short, the process is such that a class A squat causes a series of dynamic force peaks at its own location and thereafter. The first two peaks turn the squat A itself, and the part of the rail

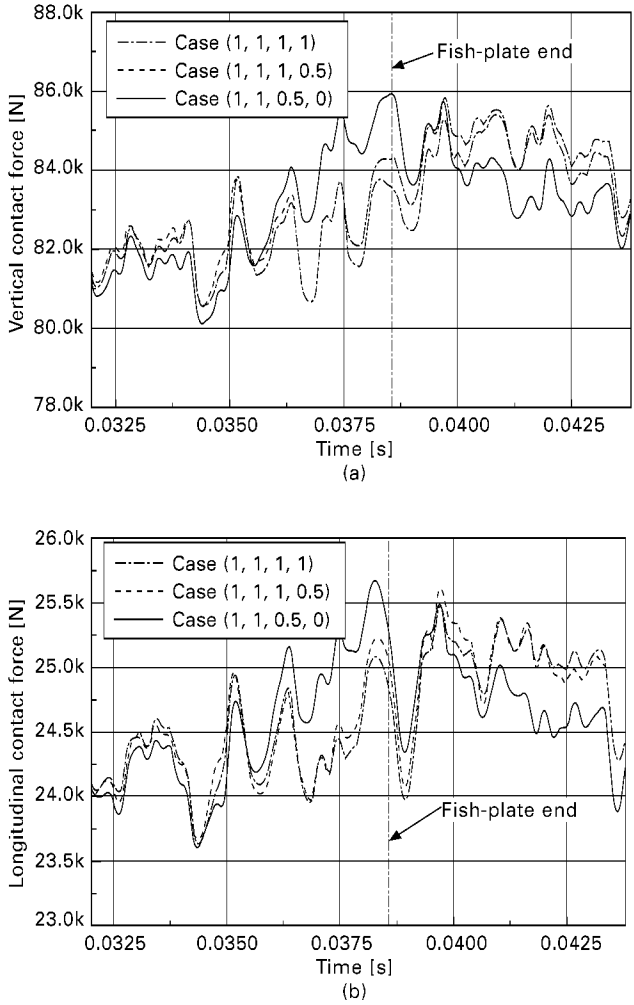


(a)



(b)

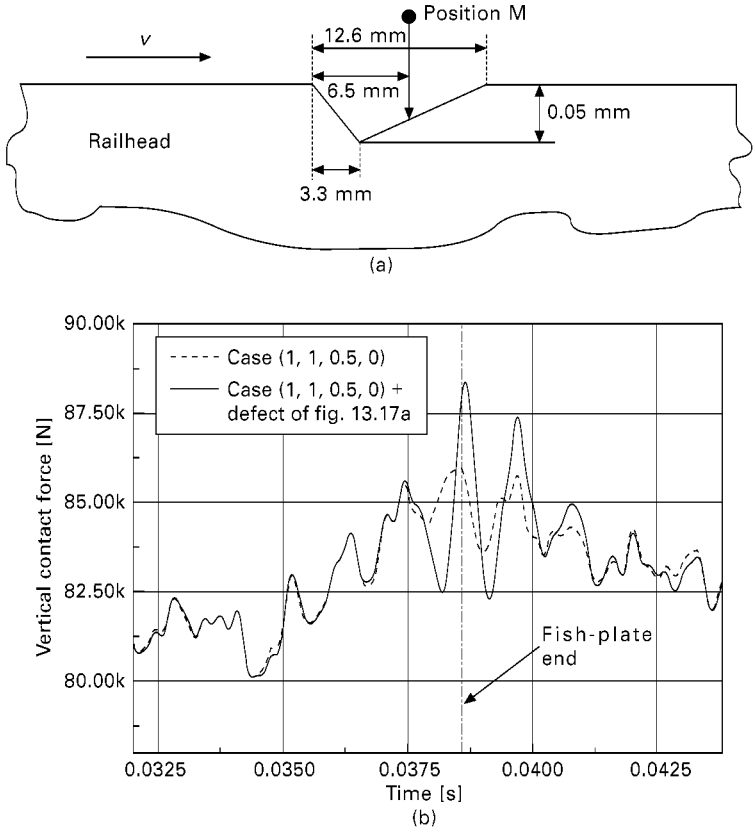
13.15 FE modeling of a wheel rolling over fish-plated rails. The wheel, rail, fish plates and sleepers were modeled as continuum. Other elements in the system were modeled as massed, springs and damper (Li *et al.*, 2008b) [*v* is the rolling velocity; *m*, *C* and *K* signify mass, damping and stiffness respectively; subscripts *C*, 1 and 2 signify car, fastening with railpad and sleeper with ballast, respectively]



13.16 Vertical and longitudinal contact forces under different fish-plate preloads. (a) Influence of the 3rd and 4th bolt preloads on vertical force and (b) the corresponding longitudinal force. Case (1, 1, 0.5, 0) means that at bolts 1 and 2 the bolts were fully fastened, while bolt 3 is half-fastened and bolt 4 is completely loose. The friction coefficient is 0.3. (Li *et al.*, 2008b)

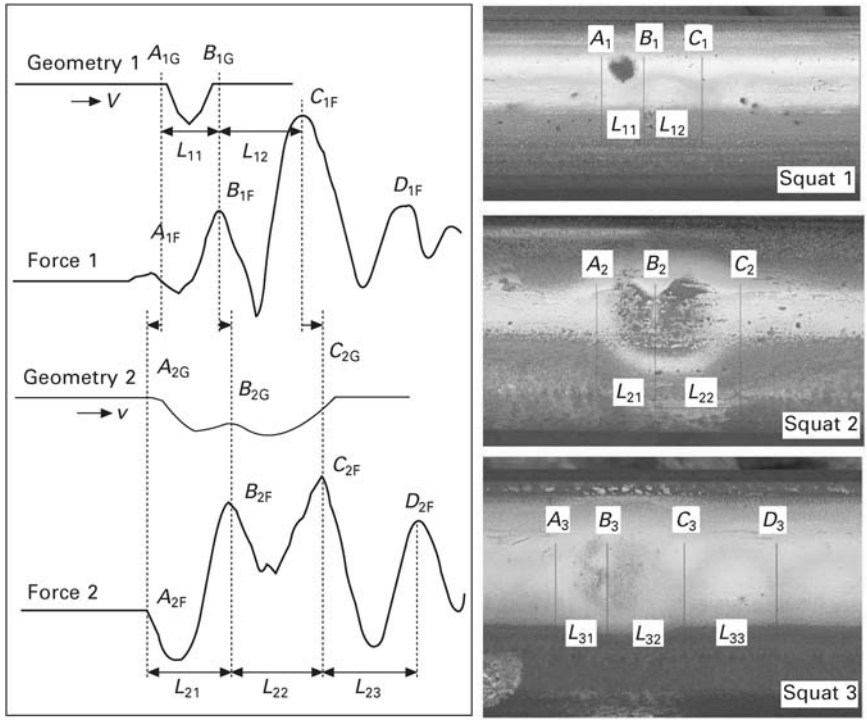
immediately after it, into a class B squat. A class B squat will cause larger forces, and the squat continues to grow into class C. A detailed description is as follows.

(1) For a class A squat growing into class B, as shown in Fig. 13.18, with proper size and track and traffic conditions:



13.17 Dynamic contact force increases due to differential wear and deformation: The dynamic contact force increase in (b) is due to the geometry deviation of (a). (Li *et al.*, 2008b)

- A typical class A squat will has a first impact with a passing wheel at B_1 in the photo of squat 1 (i.e. at B_{1G} in geometry 1), causing a peak contact force B_{1F}. This peak force after many wheel passages turns point B_1 into B_2, and B_{1G} into B_{2G}. Note that B_2 and B_{2G} usually shift forward with respect to B_1 and B_{1G}, and A_2 and A_{2G} shift backward with respect to A_1 and A_{1G}, so that $L_{11} < L_{21}$.
- A second peak force C_{1F} follows B_{1F}. In the early stage of the class A squat development peak C_{1F} generates the wave pattern around the position C_1. It is a result of the differential wear and plastic deformation of the dynamic force excited by geometry 1.
- As wear and deformation at C_1 accumulate, the increased dynamic force gradually turns the wave pattern around C_1 into the large plastic deformation between B_2 and C_2. That is to say that in such



13.18 Squats growth process.

a process the part of rail top surface around C_1 deforms gradually into part of the squat (the part between B_2 and C_2), and C_1 shifts gradually forward to C_2. The geometry 1, which is typical of squat 1, deforms into geometry 2, which is typical of squat 2.

- In the process of turning wave pattern C_1 into part of the squat, there may be for a certain period of time no visible wave pattern after (i.e. to the right of) C_1 and C_2, because the force peak D_{1F} and D_{2F} may not yet be large enough to make a new wave of visible differential wear and plastic deformation. This may partly explain why in the correlation analyses, it is about 72–75 %, but not 100 % of squats which have corrugation or wave pattern with them.

(2) For a class B squat growing into class C:

- The geometry 2 of class B squats, together with other track and traffic conditions, causes large impact force so that they grow further into class C. Wave patterns D_3 and so forth develop after C_3, which corresponds to C_2.
- When a class C squat develops further, the wave patterns marked

with A_3, B_3, C_3 and D_3 may be wiped out by the large impact force. The squat may then take a shape like that shown in Fig. 13.7.

It is noticed that B_{1F}, B_{2F} and C_{2F} are results of forced vibration, while C_{1F}, D_{1F} and D_{2F} are due to free vibration. The former excites the latter and they both cause a squat to grow. Under certain conditions the latter may also causes a squat to propagate forward, so that it is no longer an isolated one but a cluster of squats that will form.

It is observed that the wavelengths of the forces of the forced vibration and the free vibration are similar. They are usually between 20 and 40 mm, similar to that of short pitch corrugation. Keeping in mind the strong resemblance in appearance between the wave pattern following squats and short pitch corrugation, it appears that squats and short pitch corrugation have something in common in their development mechanism related to the eigen characteristics of the vehicle–track interaction system.

With the growth of a squat, its wave length increases simultaneously. In the most severe cases it can be up to 60 mm or higher. During the development of a squat the following relations hold:

$$A_{1G}B_{1G} < A_{2G}B_{2G},\ L_{11} < L_{21} < L_{31} \text{ and } L_{12} < L_{22} < L_{32}$$

In Fig. 13.18, force peak B_{1F} is smaller than C_{1F}. This may partly be attributed to the fact that in geometry 1 the edges A_{1G} and B_{1G} are sharp, which results in a smaller contact area, and hence lower contact stiffness. Generally speaking the relative magnitudes of the force peaks change constantly due to the continuous change of the contact geometry at the squats.

It should be pointed out that the contact forces shown in Fig. 13.18, particularly their wavelength, were obtained with a certain set of parameters corresponding to a track short defect which is believed to be one of the most common situations for squat to grow. Under normal track conditions the dominant frequency components of the dynamic force and the associated wavelength are different. This signifies the importance of track quality for high-frequency dynamics, and helps to explain why squats take place in a seemingly random way and why some squats grow faster than others. Further investigations into the effects of different short defects, means of detection, the appropriate maintenance policy and possible optimal design to avoid them are needed.

From the above it is clear that there are two major factors which play the most important roles in the growth of a squat, namely the magnitude of the force and its wave length. They are particularly important for the process of a class A squat growing into B because, based on the current visual classification, only a small fraction of class A squats grow into B, while it is certain that a class B will grow further into C. In the quest for the root causes of squats initiation and growth, focus should be on the factors in the

vehicle–track system which have the most critical influence on the magnitude and wavelength of dynamic contact force.

13.8 Detection of squats

From a maintenance point of view, squats and their root causes should be detected as early as possible so that predictive and preventive actions can be taken in time. Currently, the most widely employed automatic inspection for squats is ultrasonic detection. This method detects cracks, and it is reliable only when the cracks are more than 5–7 mm deep. When a crack reaches such a depth, it is often too late for grinding. Eddy current can detect surface cracks of a depth of about 0.1–2.5 mm. Surface defects which do not have cracks cannot be found. This encompasses a large number of class A squats which do not have detectable cracks yet while they are growing.

Because of the impact contact between wheel and rail at squats, squats should be detectable by instrumented wheel or axle box acceleration. The detection may be based on the magnitude of the impact force and the frequency components. In general, track short defects may cause high dynamic wheel–rail interaction force of certain frequency contents, and they should be detectable. These methods have the advantage of being able to measure or give indication of the magnitude of the force. The tendency of the squats to grow may therefore be assessed based on measurement. This may lead to much more accurate detection and classification of squats than visual inspection. Those small rail surface defects which tend not to grow can then be excluded from being counted as squats. The disadvantages of these methods are the high costs of the instrumented wheel and the difficulties in the instrumentation and in signal processing.

13.9 Counter measures

Class A squats and some class B squats with shallow cracks can be removed effectively by grinding. With severe squats, rail replacement is often inevitable. With each replacement there come two new welds, which are disadvantageous. As an alternative a squat can also be repaired by first removing the damaged part of the railhead and then filling the cavity by welding. This is often applied to parts of switches and crossings. It is very important to guarantee the quality of the welding process and the ensuing grinding to reduce the chance for new squats to occur at the welds.

The large dynamic force at squats causes damage to railpads, fastening, sleepers and ballast. They should also be repaired when squats are removed, otherwise such track short defects may cause squats to reappear. It should be pointed out that when grinding the non-uniform subsurface plastic deformation at the squats due to the local high dynamic force should also be taken into

account. Otherwise the remaining inhomogeneity in the surface layer of rail material may promote re-occurrence of squats.

Because squats grow always from small rail top surface defects, preventive and cyclic grinding will greatly reduce their occurrence. Preventive grinding should be applied shortly after new rail is installed. Interval and depth of cyclic grinding should be determined optimally from a life-cycle costs point of view, with loading conditions taken into account. Reducing the width of the heat-affected zone of welds to below a critical size should help reduce squats at welds. Such technology is being developed and tested. The best counter measures are always predictive and preventive and based on fundamental understanding of the problems and optimal design of the system. To this end, further research is needed.

13.10 Further research

The need for more researches on squats has been rising in recent years due to the increase in the problem and the associated costs. The following fields are identified:

- Determination of optimal interval and depth for cyclic grinding. This should be based on the properties of the rail materials, track structure and operation conditions. Damage models of the materials have to be developed. This may have two directions: crack initiation models for preventive actions and crack growth models for residual life prediction.
- Development of squatting resistant materials and welding technology. One such possibility is reduced heat-affected zone of welds with width smaller than the critical value for squats to develop. Observations indicate that rails of different materials and manufacturing processes behave differently with respect to squat occurrence under the same operation conditions. It means that development of more squatting-resistant material is possible, although much effort may be required for in-depth mechanical and metallurgical studies.
- Detection of squats and some of the root causes at their early stages, and their tendency to grow, based on measurement of dynamic response. To that end problems in the following aspects need to be solved: hardware instrumentation, such as strain-gauge instrumented wheel and axle box acceleration measurement, signal processing and numerical modeling for the determination of quantitative relation between the signals and the defects.

A prerequisite condition for all these research and development areas is advanced modeling techniques which can determine with sufficient accuracy the dynamic force and the associated stresses and strains under realistic operating conditions. The numerical approach presented in this chapter can

meet many of these needs. It requires further refinement and the computing speed should be improved. Particularly for material development purpose and damage study, the material behavior under high hydro-static pressure and high strain rate may have to be investigated and included in the modeling.

In view of the similarities between corrugation and squats, researches on these two problems may combine some of their efforts. By way of state-of-the-art numerical analyses, together with measurement of dynamic responses for input and validation, some of the controlling parameters for squats development may be identified and made detectable. Squats prevention or reduction may then be achieved by predictive actions or optimal design of the total vehicle–track system.

13.11 References

Bogdański B and Brown MW (2002), Modeling the three-dimensional behavior of shallow rolling contact fatigue cracks in rails, *Wear*, **253**, 17–25.

Bogdański S and Lewicki R (2008), 3D model of entrapped liquid in rolling contact fatigue cracks in rails, *Wear*, **265**, 1356–62.

Bogdański S, Olzak M and Stupnicki J (1998), Numerical modelling of 3D rail RCF 'Squat' – type crack under operating load, *Fatigue and Fracture of Engineering Materials and Structure*, **21**, 923–35.

Bold PE, Brown MW and Allen RJ (1991), Shear mode crack growth and rolling contact fatigue, *Wear*, **144**, 307–17.

Busquet M, Baillet L, Bordreuil C and Berthier Y (2005), 3D finite element investigation on the plastic flows of rolling contacts—correlation with railhead microstructure observations, *Wear*, **258**, 1071–80.

Cannon DF and Pradier H (1996), Rail rolling contact fatigue – research by the European Rail Research Institute, *Wear*, **191**, 1–13.

Chollet H (1999), *Contact roue–rail*, Internal Report, INRETS, Arcueilcedex, France.

Clayton P and Allery MBP (1982), Metallurgical aspect of surface damage problems in rails, *Canadian Metallurgical Quarterly*, **21**(1), 31–46.

Clayton P and Hill DN (1987), Rolling contact fatigue of a rail steel, Gladwell GML, Ghonem H and Kalousek J. (eds), *Proceedings of the Conference on Contact Mechanics and Wear of Rail/Wheel Systems II*, University of Rhode Island, Kingston, RI, USA 8–11 July, University of Waterloo Press, Waterloo, ON, Canada, 361–78.

Clayton P, Allery MBP, and Bolton PJ (1983), Surface damage phenomena in rails, in Kalousek J, Dukkipati RV and Gladwell GML (eds), *Proceedings of the Conference on Contact Mechanics and Wear of Rail/Wheel Systems*, Vancouver, BC, Canada, 6–9 July, University of Waterloo Press, Waterloo, ON, Canada, 419–43.

Dang Van K and Maitournam MH (2002), On some recent trends in modelling of contact fatigue and wear in rail, *Wear*, **253**, 219–27.

Kalker JJ (1990), *Three Dimensional Bodies in Rolling Contact*, Kluwer, Dordrecht/Boston/London.

Kalousek J and Grassie SL (2000), Rail corrugation: causes and cures, *International Railway Journal*, July, 24–6.

Li Z (2006), *Correlation Analysis of Squats*, Internal project report, 5 March, Railway Engineering group, Technical University of Delft, Delft, the Netherlands.

Li Z, Zhao X, Esveld C and Dollevoet R (2007), Rail Stresses, Strain and Fatigue under Dynamic Wheel-Rail Interaction. *Proceedings International Heavy Haul Association Specialist Technical Session*, Kiruna, Sweden, 11–13 June, 389–96.

Li Z, Zhao X, Esveld C, Dollevoet R and Molodova M (2008a), An investigation into the causes of squats – correlation analysis and numerical modeling, *Wear*, **265**, 1349–55.

Li Z, Zhao X, Dollevoet R and Molodova M (2008b), Differential wear and plastic deformation as a cause of squat at track local stiffness change combined with other track short defects, *Vehicle System Dynamics*, **46**, (Supplement), 237–46.

Marich S (2006), Practical/realistic implementation of wheel/rail contact technologies – the Australian experience, *Proceedings 7th International Conference on Contact Mechanics and Wear of Rail/Wheel Systems*, Brisbane, Qld, Australia, 24–26 September, 3–21.

Rail Damages, the Blue Book of RailTrack, UK (1 February, 2001).

Smulders J (2003), Management and research tackle rolling contact fatigue, *Railway Gazette International*, June 439–42.

Zhao X, Z Li, Esveld C and Dollevoet R (2007), The dynamic stress state of the wheel-rail contact, *WSEAS Transactions on Applied and Theoretical Mechanics*, **2**(2), 52–59.

Copyrighted Materials
Copyright © 2009 Woodhead Publishing Retrieved from www.knovel.com

14

Effect of contaminants on wear, fatigue and traction

S. R. LEWIS and R. S. DWYER-JOYCE,
University of Sheffield, UK

Abstract: The wheel–rail contact inevitably operates in a contaminated environment. Rain water, leaves, and artificial materials may all be present on the railway track. Some materials are deliberately applied to the rail to increase traction, reduce wear, or postpone the onset of fatigue. These act as lubricants, to minimise friction, or friction modifiers, to maintain traction within controlled limits. Solid particles can become entrained into the contact and crushed by the wheel passage. This causes damage to the wheel and rail surface and can accelerate wear. Liquids, especially high-viscosity oils, can form a partially separating film between the wheel and rail and so reduce traction. This chapter describes the processes of damage by solid materials and film formation by liquids and shows how traction is affected by both. The processes by which these mechanisms occur are complex. Modelling is therefore complex. Twin-disc machine type test-rigs have been used to illustrate the mechanisms, but field data availability is limited. The important case of the effect of leaves on the railway line is also discussed and shown to have some unusual features distinct from either solid or liquid contaminants.

Key words: contamination, leaves, friction modifier, particulate, sanding friction.

14.1 Introduction

In the context of the railway track, contamination refers to any material that is present on the rail and becomes entrained into the wheel–rail contact. This material may be solid, such as sand, ballast dust, leaves, or debris, or liquid such as water, grease, or oil. Depending on the circumstances, the contaminant may simply be blown or swept away by the passage of a railway vehicle. However, particles or liquid may enter into the contact. Then they are subject to extremely high pressure and relative sliding motion. The behaviour of the contaminant, its effect in the transmission of power between the wheel and rail, and the nature of damage to either surface is the subject of this chapter.

It is perhaps important at this stage to distinguish between contaminants, flange lubricants, and friction modifiers. Here we define *contaminants* as materials, either liquid or solid, unintentionally present on the rail that have undesired effects on the running performance of the rail network. These

negative attributes may be poor adhesion, accelerated wear, surface damage, or reduced fatigue life.

A *flange lubricant* is a substance, usually an oil or grease, that is deliberately spread along the rail gauge corner (counteracting with the wheel flange) on low-radius curves. They are applied to prevent the high wear rates seen under curving due to a shift of the contact patch.

Friction modifiers are deliberately placed on the rail in order to enhance network performance. These can be liquids (which once spread on the rail subsequently dry), or solids such as sand. The main purpose of a friction modifier is to create a target traction level. However, they are designed to tackle other rail-related issues, for example to reduce wear, mitigate noise, and even reduce ground-borne vibrations.

The distinction between a lubricant and a friction modifier may be somewhat blurred in many cases. Flange lubricants are principally designed to be used only on the rail gauge and minimise friction and reduce wear in tight bends. Friction modifiers, on the other hand, are used on the railhead and intended to achieve a targeted friction range, usually in areas of low traction caused by prior contamination. Conversely, lubricants are not intended to be used on the railhead, since in this region optimum friction levels are required for acceleration and braking.

14.2 Contaminants

14.2.1 Leaves

Every autumn, rail networks are plagued by the adhesion problems caused by leaves on the line. The cost of removal or service delays across the world's networks is very high. It is also an issue with a poor public perception where leaves are not usually regarded as a problem that should interrupt transportation services.

During the age of steam powered railways, leaves were not such a concern. Line side trees and vegetation were regularly cut back to prevent fire. Rail steels during this period were also not of today's quality and, with large diameter driving wheels (and hence a larger area of contact), wear rates were higher and track replacement more frequent. Thus any leaf layer which did form on the rail would be quickly worn away.

With modern day rail vehicles, leaves present a greater problem. Not only do modern rail steels have greater durability, wheels are smaller and vegetation felling is not as common. The increased speed of modern railway vehicles is also a contributory factor. Dead leaves which have fallen from trackside trees during the autumn are picked up by the turbulence of a passing train, and laid straight onto the railhead.[1] The leaf will then be crushed as it is rolled over by each consecutive wheel pass. The leaves are entrained

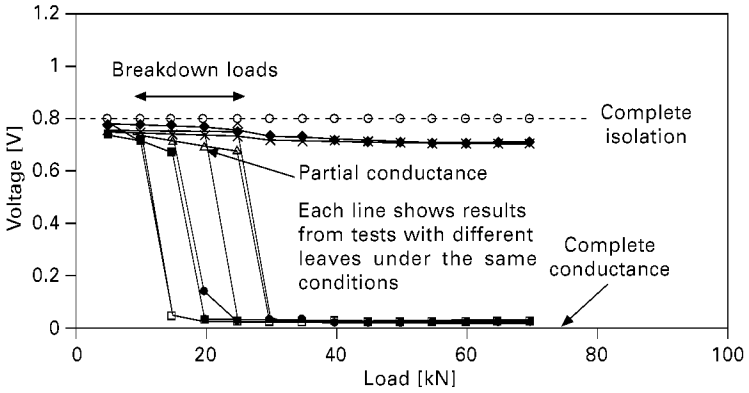
into the contact and form a wet bonded layer/film on the rail surface with a dark brown/black coloration. This film has been described as 'Teflon-like'[2] and is very strongly adhered to the surface of the rail (due to chemical bonding)[3] and thus is very difficult and expensive to remove. This dark patch was originally thought to be charring of the leaf material as it is repeatedly crushed. Cann[2] re-created the crushing of leaves in the laboratory using a ball on disc device. The leaves formed into a thin black layer on the disc very quickly-suggesting that there may be a chemical reaction between the rail steel and the leaf constituents.

The principal problems associated with leaves are traction and isolation. Traction coefficients as low as 0.03 have been reported in laboratory testing[2,4] and 0.01 from field tests.[5] These levels are far below the dry rail range of 0.3–0.6 and present a major lack of traction, presenting braking issues and delays due to wheel spin during maximum acceleration, i.e. when pulling out of a station.

The other problem caused by leaves is that of track circuit isolation. The position of trains on the rail network is detected electronically. Current is passed through sections of rail and received by a detector. If there is no train present on that section of track then all the input voltage will be received by the detector, indicating that the track section is free. However, if a train is present in a section of track, its wheels will short the circuit and no or minimal voltage will be received by the detector. An insulating material which becomes trapped between the wheel and rail may cause a line to appear free when it is in fact occupied. Isolation caused by an inhomogeneous leaf layer is clearly a complex problem. It is likely to be affected by many parameters including the leaf structure, quantity, water content, contact loading, and relative slip.

The problem is very difficult to appraise through field trials, and simulation in the laboratory can only ever be an approximation of the real conditions. However, laboratory tests have been used to investigate some of the important mechanisms. Static tests[6] showed that fresh leaves conduct current well when they are initially crushed by the wheel. Figure 14.1 shows the loads (18–30 N) at which various types of leaves are crushed/broken down and full conductance is achieved. This is because, as they are broken down, metal to metal contact is achieved. However, tests with dead leaves showed signs of insulation owing to a lack of moisture. Tests also tried to simulate sand and leaves as sanding is standard practice for combating adhesion loss (see below). The presence of sand and leaves in the contact can reduce conduction even further.

Surprisingly, tests have also shown evidence that dry leaves actually cause damage to wheel and rail. They reported deep scratches on the disc surfaces after testing, probably caused by leaf stalks.[7] The damage was not quantified, however, and it is assumed that once the leaf is crushed and the black layer



14.1 Isolation caused by leaves and the relationship between contact load and voltage across the contact.[6]

is created on the rail there would be negligible wear due to much reduced traction coefficient.

14.2.2 Solid particulates

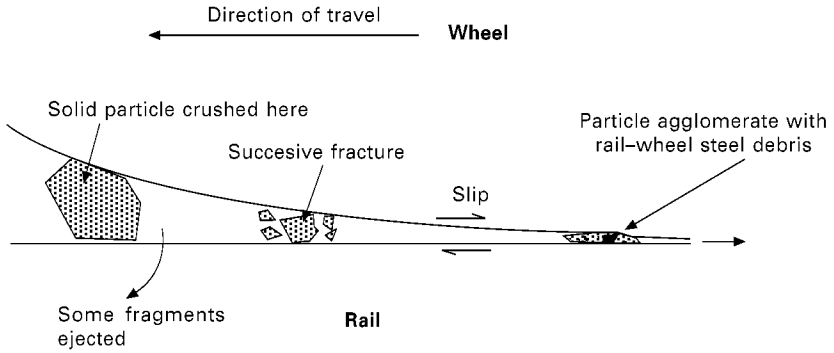
The commonest and most damaging contaminant class is that of solid particles. In the field there are likely to be particles of minerals such as ballast material, soil particles, concrete dust, or sand. The latter has been used since the early development of railways as a solution to low traction. The sand is usually air blasted directly into the wheel–rail contact of the driving wheels from a storage hopper located somewhere on board the vehicle. It can either be used at the driver's discretion or it is automatically dispensed when emergency braking is applied.

Particle entrainment into the wheel–rail contact

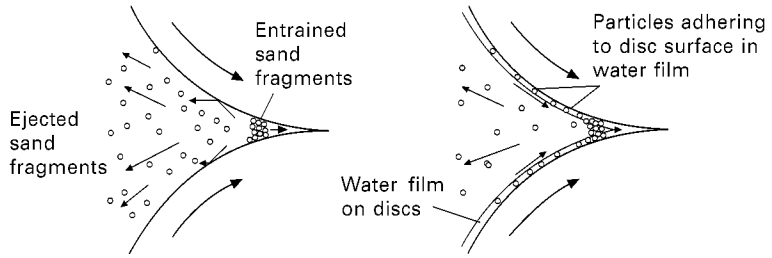
As the train approaches a stretch of solid particulate contaminated rail it is likely that many of the particles will be blown away. Anything remaining on the line will be crushed in the entry zone to the contact (shown schematically in Fig. 14.2).

The load is so high that successive fracture of a brittle material will occur until they are crushed down to tiny fragments. These fragments will be entrained into the contact. When sand particles enter into the wheel–rail contact they are crushed by the immense contact pressure. This breaking up of the sand particles causes some of the sand to be ejected from the contact as observed in twin-disc tests (see Fig. 14.3).[9]

Sand is not the only solid particulate that could potentially be found on the railhead. Crushed granite, the material used as ballast, could potentially also



14.2 Schematic of entrainment and succession fracture of brittle particles.[8]

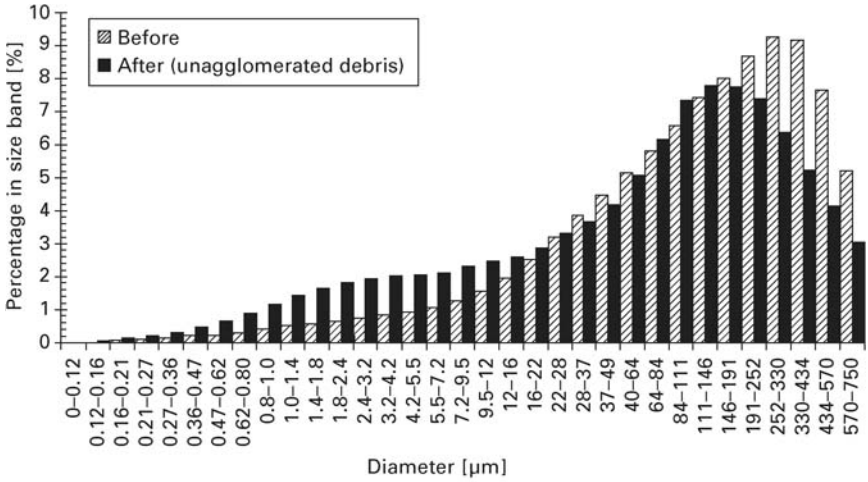


14.3 Ejection of sand particles from a twin-disc contact.[9]

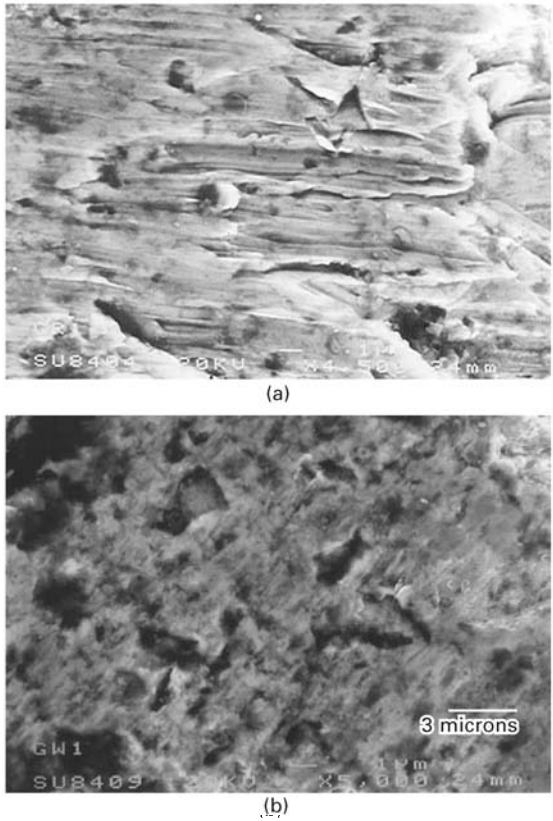
be present. Static tests were performed in reference 8, assessing the damage caused by sand and ballast. Ballast showed relatively small indentations in the wheel material and no damage on the rail material. Sand, on the other hand-showed much greater damage to the wheel and even a small amount of indentation in the rail. Figure 14.4 shows some measurements of size of crushed granite particles before and after entrainment through a twin-disc contact. The contact was between discs of wheel and rail steel used to simulate wheel–rail contact conditions.

Clearly, the particles are crushed to a smaller size (note not all particles passed through the contact) and so only a trace in size reduction is absorbed. Because the final crushed fragments are small, the level of surface change is small. Figure 14.5 shows some scanning electron microscope (SEM) images from the twin-disc surface. The contaminant particles will be subject to the micro-slip that occurs in the contact. This causes a groove to form (Fig. 14.5a) on the surface. Typically, the particle imbeds and sticks on the softer surface and causes a ploughed groove on the harder one. This is why there are no grooves on the softer of the twin-disc pair (Fig. 14.5b).

The particles tend to agglomerate in the contact under these conditions



14.4 Distribution of granite particles before and after testing.[8]

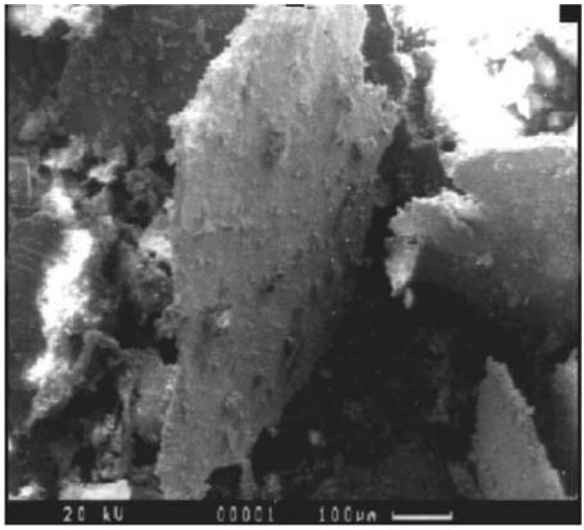


14.5 Images of rail (a) and wheel (b) after 6000 cycles of operation granite particle contamination.[8]

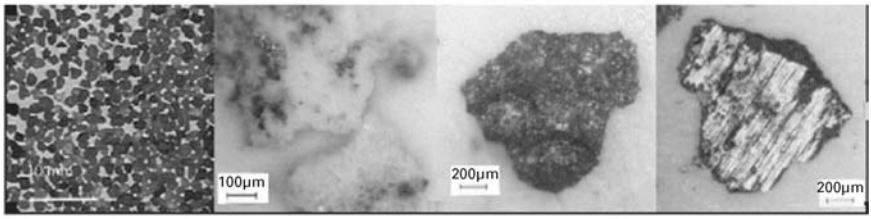
of high pressure and slip. The particles extracted from the contact appear to be mixtures of the fragments of the original material and steel wear debris. Figure 14.6 shows the extracted particles from a test with crushed granite, and Fig. 14.7 shows particles from a similar test with sand.

Traction sand and the effect on friction

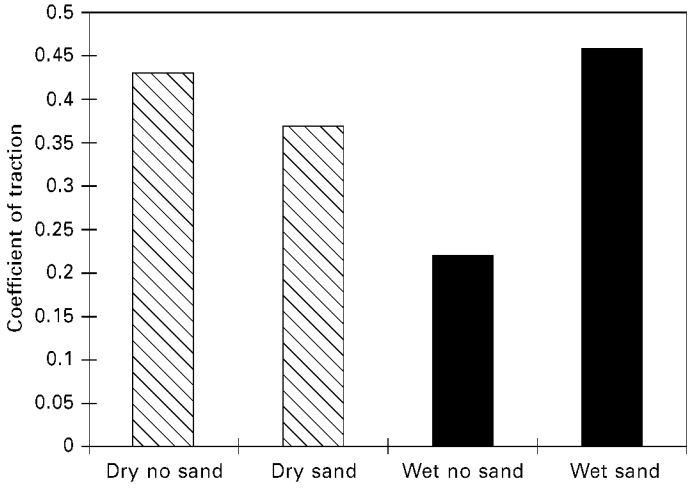
The effect of sand application on traction levels depends on what the sand is being used to counteract. For example testing[10] has shown that when sand is used in dry conditions the traction coefficient is actually reduced slightly (Fig. 14.8). This is a result of the broken down sand particles forming a low shear strength layer between the wheel and rail, thus acting as a somewhat poor lubricant. However, when sand is used under wet conditions, these crushed sand particles tend to agglomerate together and become entrained



14.6 SEM image of crushed granite particle.[8]



14.7 Series of images showing crushed sand, (a) before testing, (b) agglomerated sand particles with metal debris, (c) steel debris bottom view, (d) top view.[10]



14.8 The effect of sand on traction coefficients in a twin-disc contact.[10]

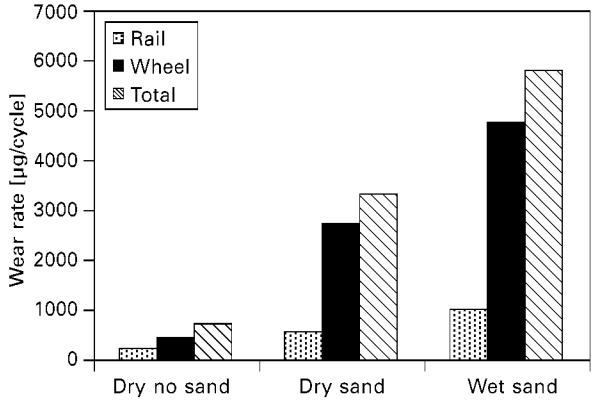
into the contact. In addition to this, the sand particles become adhered to the wheel and rail surface. Both of these actions will restore traction to dry levels or higher. When sand is added under leaf-contaminated conditions traction is restored to higher levels as the sand particles cut through the thin leaf layer.[7]

Abrasive wear by particles

Sand in the contact also has a significant effect on the wear of the wheel and rail. Laboratory tests[11,12] show that sand entrained into the contact can vastly increase wear rates by 10–100 times. This is known as a 'three-body abrasion' process.[13] As the sand particles are crushed in the contact they are subject to micro-slip and scratch the opposing surfaces (similar to those scratches observed in Fig. 14.5). Repeated micro-ploughing action leads to loss of material from the surfaces.

Figure 14.9 shows some data from twin-disc tests under various conditions of sand contamination. The difference in wear rates between the rail and wheel can be as much as 2.5 times.[9] This wear rate is even higher if water is added to the contact as shown below. The mechanisms for this are not clear, but it is probably due to the fact that it is easier to entrain wet agglomerated masses of sand into the contact where it causes more abrasive damage. This is a matter for concern, especially as sand is most likely to be found on the rail in low-adhesion situations, i.e. wet rails.[9]

However, in reference 9, even though three-body abrasion was observed and scratches formed mostly on the rail surface, results showed that it was actually the softer wheel that wore faster. In these tests it was apparent



14.9 Wheel and rail wear rates under conditions of sand contamination from twin-disc testing.[9]

that the action of the sand particles sticking into the wheel had initiated a low-cycle fatigue process, causing large chunks of wheel material to be lost as wear debris. Either way it is clear that the presence of sand in the wheel–rail contact is detrimental to the life of both wheels and rails. Grieve *et al.*[8] devised a model to predict the three-body abrasion process in the wheel–rail contact. The model calculates material removal rate assuming that an embedded particle removes a volume of material equivalent to its cross-sectional area. The damage to the whole rail infrastructure reaches beyond the wheel–rail contact; it can also cause damage to bearings, drainage systems and contamination of track ballast.[12]

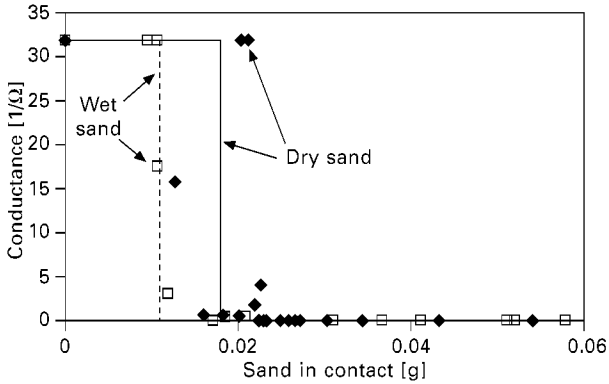
Isolation by sand particles

Sand, like leaves, can also cause a major isolation problem. In the right quantity sand can form an insulating layer blocking the 'track occupied' signal. Static testing[6] showed a critical mass of sand within the contact above which complete isolation was observed. For a dry rail this was 0.018 g/cm^2 and for a wet rail 0.011 g/cm^2 (see Fig. 14.10).[6] These critical masses translate to higher sand flow rates than are currently used by British Rail (this was also confirmed in dynamic testing.[14] However, this may not be the case for other rail operators around the world and could become an issue at lower train speeds.

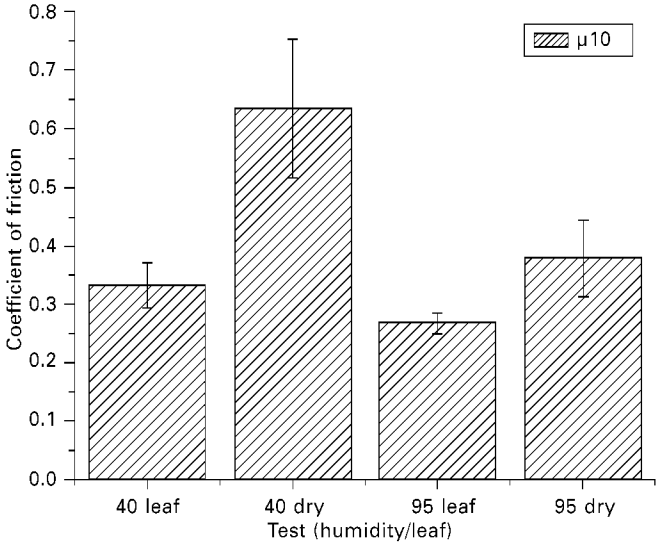
14.2.3 Water

Reduced traction

In most temperate climates, water will be present on the rail as a result of rain fall or morning dew. The main issue is that a thin film of moisture can



14.10 Wheel–rail conductance variation with concentration of sand in contact.[6]



14.11 Chart showing influence of humidity on a dry and leaf contaminated rail.[3]

cause a reduction in traction coefficient level typically down to around 0.3.[15] Normally, a traction coefficient of 0.3 is not low enough to seriously affect traction and braking.

Environmentally controlled lab tests carried out by Olofssen and Sundval[3] showed that atmospheric humidity will also affect the traction coefficient of a dry rail (see Fig. 14.11). Humidity, however, has little effect on leaf-contaminated contacts. This phenomenon was also reported in reference 7. From this evidence it is most likely that results reported in the field are

also affected by humidity but, due to lack of controllability, this cannot be confirmed.

Liquids and rolling contact fatigue

The presence of water has been linked to an increased probability of rolling contact fatigue (RCF). It does this by a mechanism known as 'fluid crack pressurisation'. Standing water on the railhead will be forced into the cracks in the rail steel as the wheel rolls over. The water then enters the crack and is forced toward the tip as the mouth of the crack closes due to compression in the contact.[16,17] This has two effects: firstly the water can lubricate the crack faces so that a shear stress is seen at the tip (mode II stress intensity), and second that the forcing of the water toward the crack tip causes a tearing/widening (mode I stress intensity) of the crack. These have the combined effect of increasing the RCF rate in the rail.

14.2.4 Oil or grease

Lubricating oil or grease is another commonly found contaminant on the rail. It can be dripped from leaking trains or deposited at level crossings on vehicle tyres or spilt goods.[5] Oil is the main cause of traction loss on rail networks, and it has been shown that only small deposits are required to decrease traction.[7] It has also been shown[7] that oil is quickly removed from the railhead by the rolling and squeezing action of the wheel. However, as only small amounts of oil are necessary to cause traction loss the effects of oil contamination can be long lasting.

Lubricating film formation

The process by which an oil or grease reduces the traction in the wheel rail contact is one of elastohydrodynamic lubrication. The oil becomes entrained into the contact where it is subjected to very high pressure. This causes the viscosity to rise considerably. This phenomenon, known as piezoviscosity, enhances the thickness of the oil film that forms from the hydrodynamic entrainment of the liquid. If there is sufficient oil on the rail surface then an oil film will form between the wheel and rail. If this oil film is thicker than the surface roughness then a very low traction coefficient will result. There are several theoretical methods for calculating the thickness of the oil film from the wheel and rail material properties, lubricant properties, and the geometry. For the wheel rail case, the most applicable is from Dowson and Hamrock.[18] The oil film thickness is defined in terms of four non-dimensional parameters:

$$H = 2.69U^{0.07}G^{0.53}W^{-0.067}(1 - 0.61e^{-0.73k})$$

The above dimensionless parameters are given by the following:

$$W = \frac{P}{2E^*R_x^2}, U = \frac{\eta_o u}{2E^*R_x}, k = 1.03\left(\frac{R_y}{R_x}\right)^{0.64}, G = 2\alpha E^*, H = \frac{h}{R_x}$$

where P = total normal load, E^* = reduced modulus of combined wheel and rail, R_x = reduced radius in the direction of motion, R_y = reduced radius perpendicular to direction of motion, u = mean surface speed, η_0 = viscosity of the lubricating oil as it enters the contact, α = pressure viscosity coefficient and, h = film thickness.

Typically, fluid films of the order of 0.1–1 μm are formed. The combined roughness of the rail and wheel is likely to be higher than this so the surfaces are not completely separated. Part of the load and traction will be supported by the fluid film and part by asperity contact regions. This is known as boundary lubrication.

Effects on rolling contact fatigue

Oil can also have the same effect on RCF as water, see above. Although the mechanism may be the same, oil can affect RCF in a different way.[19] Research has shown that whether or not oil affects RCF depends upon the oil application rate.[20,21] If lubrication is applied intermittently then it has a dramatic effect on RCF crack formation and growth rate. It was also shown that higher frequency lubrication will lower or prevent RCF compared to that seen under dry running conditions,[21] whereas lower frequency lubrication enhances RCF. This seems to suggest that the higher viscosity of oil, compared to water, means that it takes longer for it to seep into a crack. Thus if smaller amounts of lubricant are present on the rail with more wheels passing over, it will not have time to enter the crack before it is squeezed out of the contact by the wheel. If lubricant is left on the rail for longer, then there may be sufficient time for the oil to seep into the crack before it is rolled over by the wheel.

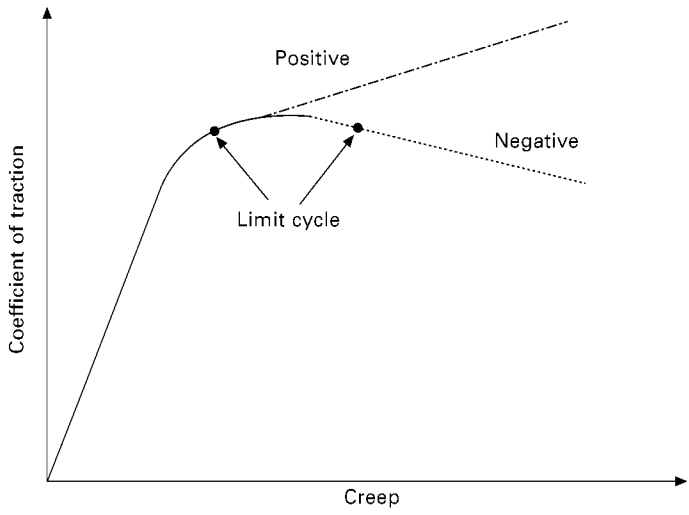
14.3 Friction modifiers

Clearly the traction between the wheel and rail contact is an important issue. If it is too low the wheel slides rather than rolls; if it is too high then the wheel and rail are subject to excessive shear stress that leads to increased wear rates. Modern rail networks are seeing increasing traffic levels, tighter running schedules, and increasing line speeds. Delays caused by low traction and wheel and rail life are increasingly important. This has seen the introduction

and increasing use of friction modifiers. Here we define the term friction modifier as a man-made product that is applied to the railhead to achieve a target friction level. The objective is to maintain friction levels within a defined range, in contrast to a lubricant which is usually applied simply to minimise friction. Ideally, they should be able to increase traction in cases where it might otherwise be low, but not cause surface damage like sand. Further they should maintain traction below an upper band so that rolling resistance is reduced.

There are a wide range of friction modifiers currently available on the market. The detailed constituents are not generally made public. Some use long-chain polymers, dispersed in a suspending liquid that evaporates to leave a solid film on the rail surface. There are others that consist of solid particles suspended in a gel. Friction modifiers in liquid form can be applied to the railhead via automatic trackside applicators. Recent developments, however, have seen the introduction of solid (and liquid)[22,23] products which can be applied from the vehicle; a move prompted by the network operators to shift costs to the passenger and freight companies.

A key objective of a friction modifier is to provide a positive friction characteristic to the wheel–rail contact.[24] A typical creep curve is shown in Fig. 14.12. Above a creep limit (typically 3 % for dry rail) there is a negative slope. This means that as saturation is reached traction levels will start to decrease with increasing creep. This presents a problem if creep exceeds saturation; then a limit cycle is created whereby creep will alternate between two points where the friction is the same. This creates a resonant frequency



14.12 Illustration of friction modifier effects on creep curve.[26]

in the wheel web creating a high-pitched noise,[25] but it can also affect wear of the rail in the form of corrugation. By providing positive friction this limit cycle is eliminated, hence noise and rail damage are reduced.[21]

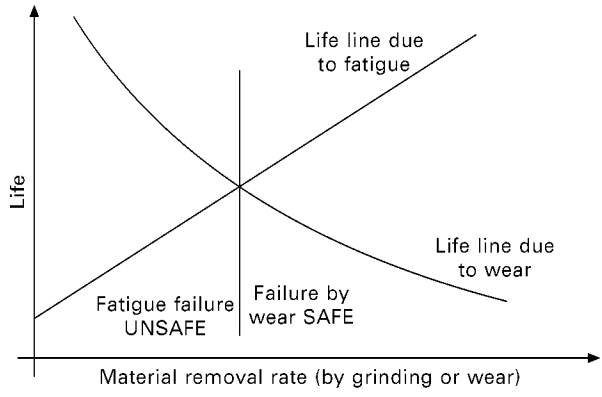
In reference 23 it has been shown that the correct application of friction modifier can virtually eliminate wear and corrugation. However, improper application can vastly increase RCF. Noise can also be vastly reduced by the application of friction modifiers. In badly affected areas sound pressure levels can be reduced by as much as 21 dB with and average reduction of 10–15 dB.[26]

When it comes to the design of friction modifiers, it is important to take into account the interaction between wear and RCF. Removal of material at the railhead will shorten or even eliminate RCF cracks. This process is known as 'crack truncation'.[27] Thus it can be seen that if a friction modifier retards the wear rate too much then RCF cracks will dominate. Figure 14.13 schematically shows how wear rates need to be managed as well as friction levels so that rail life is determined by wear and not fatigue.

14.4 Discussion

A contaminant in the wheel–rail conjunction affects the contact performance by causing surface damage, changing the friction coefficient, and altering the wear and/or RCF behaviour. The mechanisms by which the contaminating material causes these phenomena are complex and there are no general theories available.

For viscous liquid contaminants (i.e. lubricants) then the principles of elastohydrodynamic lubrication can be applied to understand the film-forming behaviour. However, even then, the tractive performance is difficult to predict, since this depends on how the fluid behaves under extreme pressure in the



14.13 Illustration of effects of crack truncation.[27]

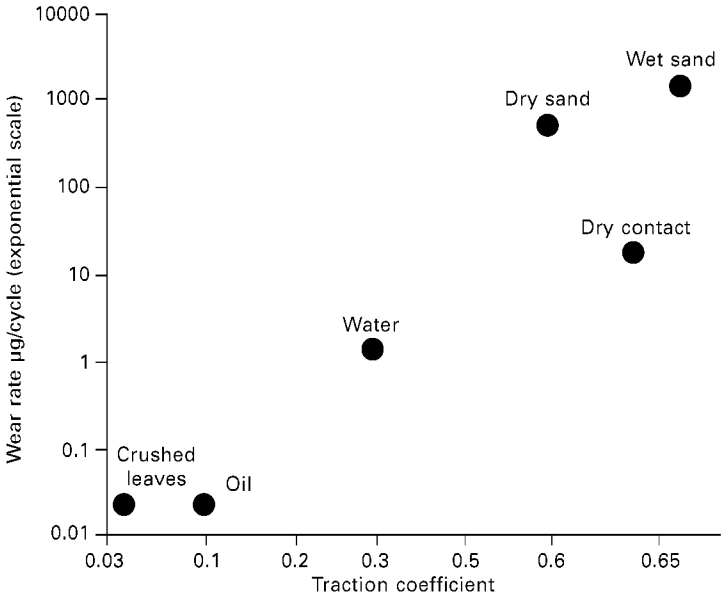
contact. The interaction with fatigue cracks is also complex and usually modelled using a combination of fracture mechanics and numerical modelling of the contact mechanics.[28,29]

Solid particulate contaminants are even more complex to model. The traction, damage, or wear performance depends on how the particles break down or deform when they enter into the contact. This will depend on the fracture of brittle materials in a complex stress field. The quantification of surface damage, traction, wear, and fatigue is usually on the basis of empirical observation.

There are three levels of experimentation where such empirical data can be obtained. Field data are the most useful but also the most costly and hardest to interpret. Typical field trials for surface contaminants include those carried out in references 22 and 26. Then there are a number of wheel–rail component test apparatus.[17,21] Either full size wheels or scaled wheel specimens are run against sections of rail. This provides a good controlled test environment, but specimens tend to be costly. The simplified method for studying contamination is through twin-disc testing. It is possible to simulate materials, contact pressure, and slip; but importantly the geometry is completely different. For this reason, the contaminant entrainment process cannot be simulated exactly since this is controlled by the geometry of the wheel and rail. Nevertheless mechanisms, effects on surface damage wear and traction can be explored.

Figure 14.14 and the accompanying Table 14.1 show the relationship between rail wear rates and associated traction coefficients for all the types of contaminant that have been discussed in this chapter. There is a clear trend of wear rates increasing with higher traction coefficients. The wear rates and traction coefficients have been gathered from laboratory-based twin-disc tests.[10] The chart is intended as an illustration of the effects of different contaminant types. There are no published data for wear rates seen under leaf contamination. Even though damage by entrainment of leaf stalks into the contact has been reported, it was not quantified. Hence the vertical position of crushed leaves on the above chart has been approximated to be close to that of oil. It is important to notice that the wear axis in the figure has a logarithmic scale and that the wear rate caused by oil is virtually negligible compared to that of other contaminants.

The aim of contamination control is to maintain a position in the centre of the plot shown as Fig. 14.14, where friction is held at a controlled level, just high enough to achieve the required traction, but low enough to avoid excessive shear stress and wear. Wear can never be eliminated since there will always be metal-to-metal surface contact (traction levels could not be maintained without this). However, the wear rate should be at an acceptable economic level and, under certain circumstances, balanced to achieve a wear rate that removes fatigue cracks as they propagate.



14.14 Chart illustrating relationship between wear and traction coefficient for various contaminants.

Table 14.1 Wear data and test conditions

	Wear rate (µg/cycle)	Test conditions
Dry	21	400 RPM 1 % slip, 1500 MPa
Water	1.23	400 RPM, 1 % slip, 1500 MPa
Dry sand	614.5	400 RPM, 20 % slip, 1500 MPa
Wet sand	1082.8	400 RPM, 20 % slip, 1500 MPa
Oil	0.02	400 RPM, 1 % slip, 1500 MPa
Crushed leaves	n/a	400 RPM, 0.5–5 % slip, 1500 MPa

14.5 Conclusions

Contamination of the wheel–rail contact may be in the form of solid particulates, water, grease or oil, or deliberately applied friction modifiers. Solid particles can become entrained into the wheel–rail contact. They cannot support the high loads and so crush or deform to smaller fragments. These pass into the contact and damage the running surfaces which can lead to excessive wear. However, the presence of solid materials can enhance the friction coefficient. For this reason sand is used in areas of low traction.

Liquids can form a thin separating layer between the wheel and rail. The thickness of this layer and hence the reduction in traction depends on the viscosity of the liquid. So whilst water causes reduction in traction it is

generally not severe enough to affect operation, the presence of lubricating grease or oil causes an unacceptably large drop in traction. The presence of leaves on the track is a special case of contamination, where there is an apparent chemical reaction between the rail steel and leaf constituents that leaves a highly durable very low friction solid film on the surface.

The quantitative prediction of the effects of these contaminants, in terms of traction, wear, and fatigue life is difficult. Most of the available knowledge has been obtained from field experience or laboratory-based twin-disc type testing.

The emergence of friction modifiers as a means to control railhead contamination provides a new approach to the wheel and rail interface management. The target of maintaining friction, and then wear and fatigue, at a controlled level is challenging; especially when the controlling mechanisms are complex and subject to a great many influencing parameters.

14.6 References

1. Johnson, T., 2006. *Understanding Aerodynamic Influences of Vehicle Design on Wheel/Rail Leaf Contamination*, Research Program report, Rail Safety and Standards Board, London, UK.
2. Cann, P M., 2006. The 'leaves on the line problem' – a study of leaf residue film formation and lubricity under laboratory test conditions, *Tribology Letters*, **24**(2), 151–8.
3. Olofsson, U., Sundvall, K., 2004. Influence of leaf, humidity and applied lubrication on friction in the wheel–rail contact: pin-on-disk experiments, *Proceedings of the IMechE, Part F: Journal of Rail and Rapid Transit*, **218**, 235–42.
4. Broster, M., Pritchard, C., Smith D A., 1974. Wheel/rail adhesion – its relation to contamination on British railways, *Wear*, **29**, 309–21.
5. Nagase, N., 1989. A study of adhesion between the rails and running wheels on main lines: results of investigations by slipping adhesion test bogie, *Proceedings of the IMechE, Part F: Journal of Rail and Rapid Transit*, **213**, 33–43.
6. Lewis, R., Massing, J., 2106. Static wheel/rail contact isolation due to track contamination, *Proceedings of the IMechE Part F: Journal of Rail and Rapid Transit*, **221**, 43–53.
7. Gallardo-Hernandez, E A., Lewis, R., 2008. Twin disk assessment of wheel/rail adhesion, *Wear*, **265**(9–10), 1309–16.
8. Grieve, D G., Dwyer-Joyce, R S., Beynon, J H., 2001. Abrasive wear of railway track by solid contaminants, *Proceedings of the IMechE, Part F: J. Rail and Rapid Transit*, **215**, 193–15.
9. Lewis, R., Dwyer-Joyce, R S., 2005, Wear at the wheel/rail interface when sanding is used to increase adhesion, *Proceedings of the IMechE, Part F: J. Rail and Rapid Transit*, **221**, 29–41.
10. Lewis, R., Dwyer-Joyce, R S., 2005. Foreign particles at the wheel/rail interface, *Proceedings Conference on Excellence in Railway Systems Engineering and Integration*, Derby, UK, November, 25–6.
11. Kumar, S., Krishnamoorthy, P K., Prasanna Rao, D L., 1986. Wheel–rail wear and

adhesion with and without sand, *Trans ASME, Journal of Engineering for Industry*, **108**, 141–47.
12. Jenks C W., 1997. *Improved Methods for Increasing Wheel Rail Adhesion in the Presence of Natural Contaminants*, Transit Co-operative Research Program, Research Results Digest, No. 17, Transportation Research Board, National Research Council, Washington DC, USA.
13. Rabinowicz, E., Dunn, L A., Russell, P G., 1961. Study of abrasive wear under three body conditions, *Wear*, **4**, 345–55.
14. Lewis, R., Dwyer-Joyce, R S., Lewis, J., 2003. Disk machine study of contact isolation during railway track sanding, *Proceedings of the IMechE, Part F: Journal of Rail and Rapid Transit*, **217**, 11–24.
15. Beagley, T M., Pritchard, C., 1975. Wheel/rail adhesion – the overriding influence of water, *Wear*, **35**, 299–313.
16. Bower, A F., 1988. The influence of crack face friction and trapped fluid on surface initiated rolling contact fatigue cracks, *Trans ASME, Journal of Tribology*, **110**, 704–11.
17. Fletcher, D I., Hyde, P., Kapoor, A., 2006. Investigating fluid penetration of rolling contact fatigue cracks in rails using a newly developed full-scale test facility, *Proceedings of the IMechE, Part F: Journal of Rail and Rapid Transit*, **221**, 35–44.
18. Hamrock, B. J., Dowson, D., 1979. Minimum film thickness in elliptical contacts for different regimes of fluid-film lubrication. in Dowson D (ed.), *Elastohydrodynamics and Related Topics, Proceedings 5th Leeds-Lyon Symposium on Tribology held in the Institute of Tribology, Department of Mechanical Engineering, the University of Leeds, England, Sept 1978*, Mechanical Engineering Publications for the Institute of Tribology, Leeds University, Leeds, UK and the Institut national des sciences appliquees de Lyon, Lyon, France, 22–7.
19. Kaneta, M., Yatsuzuka, H., Murakami, Yl., 1985. Mechanism of crack growth in lubricated rolling/sliding contact, *ASLE Transactions*, **28**(3), 407–14.
20. Fletcher, D I., Beynon, J H., 2000. The effect of intermittent lubrication on the fatigue life of pearlitic rail steel, *Proceedings of the IMechE, Part F: Journal of Rail and Rapid Transit*, **214**, 145–58.
21. Eadie, D., Elvidge, D., Oldknow, K., Stock, R., Pointner, P., Kalousek, J., Klauser, P. (2006), The effects of top of rail friction modifier on wear and rolling contact fatigue: full scale rail-wheel test rig evaluation, analysis and modelling, *Proceedings 7th International Conference on Contact Mechanics and Wear of Rail/Wheel Systems*, Brisbane, QLD, Australia, 24–27 September, 411–19.
22. Tomeoka, M., Kabe, N., Tanimoto, M., Miyauchi, E., Nakata, M., 2002. Friction control between wheel and rail by means of on-board lubrication, *Wear*, **253**, 124–29.
23. Suda, Y., Iwasa, T., Komine, H., Tomeoka, M., Nakazawa, H., Matsumoto, K., Nakai, T., Tanimoto, M., Kishimoto, Y., 2005. Development of on-board friction control, *Wear*, **258**, 1109–14.
24. Egana, J I., Vinolas, J., Gil-Negrete, N., 2005. Effect of liquid high positive friction (HPF) modifier on wheel–rail contact and rail corrugation, *Tribology International*, **38**, 769–74.
25. Eadie, D T., Kalousek, J., Chiddick, K C., 2002. The role of high positive friction (HPF) modifier in the control of short pitch corrugation and related phenomena, *Wear*, **253**, 185–92.

26. Eadie, D T., Santoro, M., Powell, W., 2003. Local control of noise and vibration with KELTRACK™ friction modifier and Protector® trackside application: an integrated solution, *Journal of Sound and Vibration*, **267**, 761–72.
27. Kapoor, A., Fletcher, D I., Franklin, F J., 2003. The role of wear in enhancing rail life, in Dowson D., Priest M., Dalmaz G. and Lubrecht A. A. (eds) (2003), *Tribological Research and Design for Engineering Systems, Proceedings 29th Leeds-Lyon Conference on Tribology, Leeds 2002*, Elsevier, Amsterdam, the Netherlands, 331–40.
28. Kaneta, M., Yatsuzuka, H., Murakami, Y., 1985. Mechanism of crack growth in lubricated rolling/sliding contact, *ASLE Transactions* **28**(3), 407–14.
29. Bogdanski, S., Olzak, M., Stupnicki, J., 1998. Numerical modelling of 3D rail RCF squat type crack under operating load, *International Journal of Fatigue and Fracture of Engineering Materials and Structures*, **21**, 923–35.

Copyrighted Materials
Copyright © 2009 Woodhead Publishing Retrieved from www.knovel.com

15

Effect of damage on vehicle dynamics

S. BRUNI and F. BRAGHIN, Politecnico di Milano, Italy

Abstract: This chapter discusses the influence of different types of wheel–rail surface damage on rail vehicle dynamics. Changes in wheel–rail transversal profiles produced by regular wear mainly affect the critical speed and curving behaviour of the vehicle, whereas damage-induced alterations of the longitudinal rail profile (corrugation) and of the wheel circumferential profile (wheel out-of-roundness) are responsible for increased train–track interaction effects in the 20–5000 Hz frequency range, and cause increased emission of noise and vibration. Finally, the effect of localised defects on wheel and rail surfaces, caused e.g. by rolling contact fatigue, is examined.

Key words: wear of railway wheels, stability of railway vehicles, curving of railway vehicles, equivalent conicity, rail corrugation.

15.1 Classification of damage that affects vehicle dynamics

The dynamics of a railway vehicle travelling along the track is dominated by wheel–rail contact forces, introducing non-conservative and non-linear effects that depend upon the complex wheel and rail geometry. Every change in the profile shapes, which may be induced by wear, or by other kinds of surface damage, will hence profoundly affect vehicle dynamics, since they will result in an alteration of the relationship between vehicle motion and the contact forces. Furthermore, the stiff steel-on-steel contact between railway wheels and rails is highly sensitive to even small geometric imperfections, so that any damage affecting the smoothness of the contacting surfaces is likely to result in higher train–track interaction effects, causing increased emission of noise and vibration, reduction of ride comfort and accelerated degradation of the track.

This chapter is aimed at describing the influence of different types of wheel–rail surface damage on vehicle dynamics. To this end, the following categorisation of damage occurring at the wheel–rail interface was chosen:

1. effect of variations in the cross-sectional geometry of wheel and rail profiles produced by wear;
2. effect of railhead corrugation, introducing modifications of rail geometry in the longitudinal direction;
3. effect of wheel out-of-roundness, producing modifications in the

circumferential geometry of the wheel rolling surface with irregularity patterns that are periodic with the revolution of the wheel;

4. effect of localised defects on wheel and rail surfaces, that are again associated with profile modifications in the longitudinal/circumferential direction, but in this case are restricted to small portions of the rolling surfaces.

Transversal profile wear is mainly related to vehicle motion in the horizontal plane, involving guidance, stability and steering behaviour of the vehicle, whereas the other three types of damage mainly affect the vertical vibration of the vehicle, although vibration in the horizontal plane cannot be neglected to explain phenomena like rail corrugation, rolling and squeal noise, etc. A distinction can also be made with respect to the range of frequency involved, since modifications of the transversal wheel–rail profiles induced by wear will mainly affect vehicle dynamics in the low-frequency range (typically below 20 Hz), whereas train–track interaction effects excited by wheel and rail corrugation and by local defects span over a much wider frequency range, up to several kHz.

Although rail corrugation, wheel out-of-roundness and localised defects all influence the same train–track interaction phenomena, a separate treatment will be proposed in this chapter because of the different type of excitation which is produced on the train–track system: periodic in the case of rail corrugation and wheel-out-of-roundness (but with periodicity related to the length of the sleeper bay in the first case and with the length of the wheel rolling circumference in the second one), impulsive in the case of localised defects.

15.2 Effects of transversal profile wear

The most important geometrical features of the wheel–rail contact in relation to vehicle dynamics are:

- the variation of rolling radius with lateral wheelset displacement, as this governs the conicity effect;
- the variation of the contact angle (inclination of the tangent to wheel–rail contact), as this affects the resistance of the vehicle against derailment.

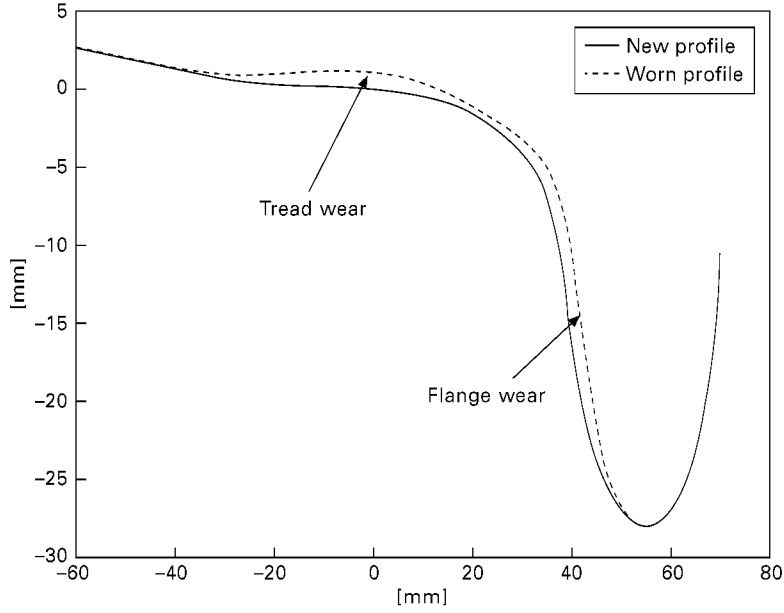
In this section, the effect of wheel and rail wear on the two above parameters and the main implications on vehicle dynamics are described.

In the most part of the text below, the case of a vehicle equipped with solid wheelsets is considered. In a solid wheelset the two wheels are mounted on a common rigid axle and hence assume the same angular speed, even in the presence of a difference in the rolling radius, a circumstance having

implications for vehicle guidance and curving performances. For a vehicle equipped with independently rotating wheels (IRW), profile modifications normally have less critical implications on vehicle dynamics, although some specific issues deserve attention and are mentioned in the following sections.

15.2.1 Wear regions of wheel and rail profiles

On the wheel surface, wear takes place in two almost distinct regions of the transversal profile: on the tread and on the flange, see Fig. 15.1. The main effect of material removal from the tread is to modify the original coned shape, affecting the equivalent conicity, which will in turn affect the guidance, stability and curving performances of the vehicle (see Sections 15.2.2 and 15.2.3). In general, the equivalent conicity of a wheelset becomes higher with the progress of wear. Although the increase of conicity enables the wheelset to perform better in steering, it will reduce vehicle stability and, above a certain level of wear, wheel re-profiling will be required to restore the original profile and avoid vehicle instability at service speeds. In the case of severe wear, the tread assumes a hollow form, which produces a conformal contact with the railhead. Conformal contact alters the distribution of normal pressures and tangential tractions exchanged by the wheel and the

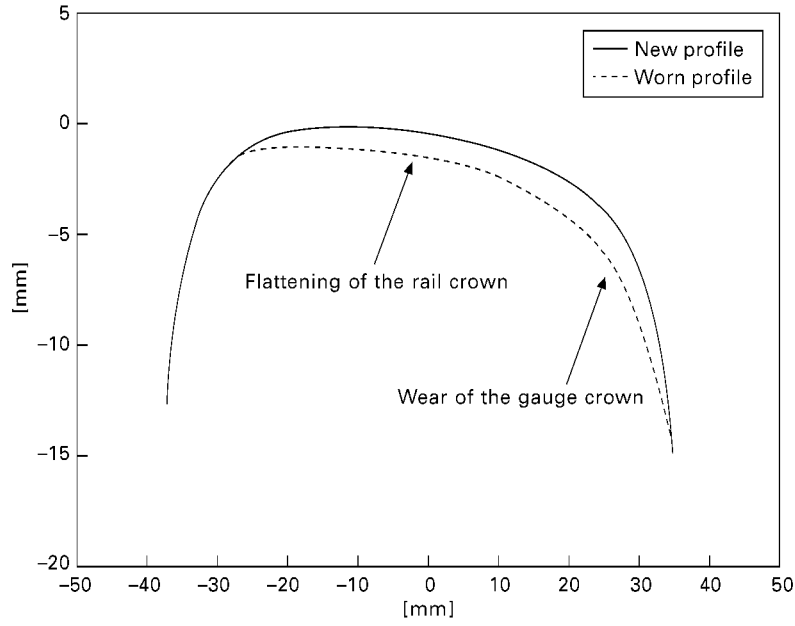


15.1 Regions of wear on the wheel and on the rail.

rail in the contact patch and is responsible for sudden jumps in the position of the contact point as the wheelset moves transversally across the flangeway clearance. Some specific effects produced by hollow worn wheel profiles on running dynamics are described in Section 15.2.5.

Wear of the wheel flange is mostly produced by curving, since the outer wheel of the leading wheelset in a bogie enters into flange contact with the gauge corner of the high rail. Flange wear affects the maximum value of the contact angle (the slope of the common tangent to the wheel and rail profiles in the contact point), which is an important parameter in assessing vehicle resistance to flange climb derailment. Furthermore, flange wear affects the possible formation of two-point contact between the flanging wheel and the high rail, which may substantially alter the steering behaviour of the vehicle, especially in sharp curves. The effects of flange wear are discussed in Sections 15.2.3 and 15.2.4.

As far as wear of the rail transversal profile is concerned, two main effects have to be mentioned: flattening of the rail crown and wear of the gauge corner (see Fig. 15.2), the latter occurring especially on curved paths on the high rail. Flattening of the rail crown is reported to increase the equivalent conicity (Pearce, 1996), although this effect is less pronounced than that of tread wear on the wheel. On the other hand, wear of the gauge corner has an effect on the location of the contact point(s) during curve negotiation and



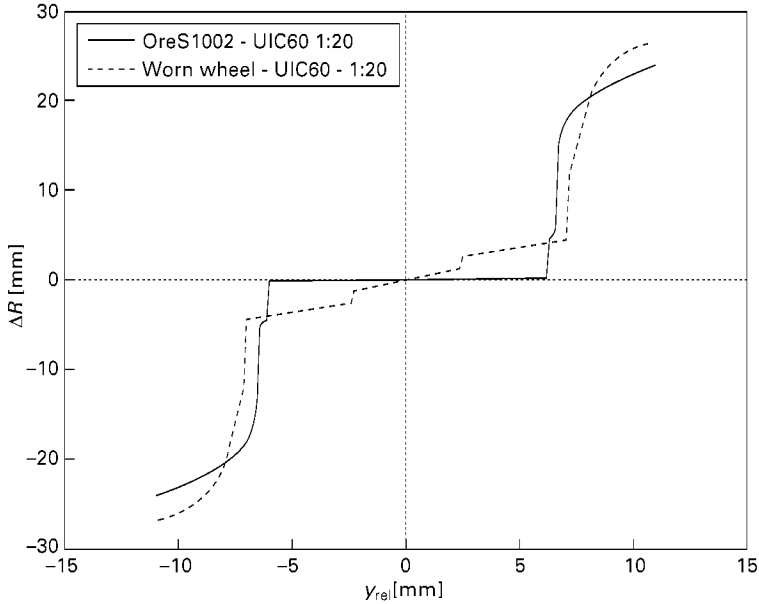
15.2 Regions of wear on the rail.

hence affects the attitude of the vehicle along the curve and its resistance to derailment, as better detailed in Section 15.2.4.

15.2.2 Effect of profile wear on guidance and stability

Guidance is an inherent property of a solid wheelset with coned wheels: in the absence of lateral forces, a wheelset with symmetric wheel profiles running on a pair of symmetric rails will naturally roll along the central position across the flangeway clearance. If the wheelset is displaced towards one side, the wheel approaching flange contact rolls on an increasing radius and the wheel moving away from flange contact rolls on a decreasing radius. Because the wheels have the same angular speed, the difference in the rolling radii steers the wheelset back towards the centre of the track, providing the guidance effect.

Therefore, guidance is governed by the variation of the rolling radius difference produced by a lateral displacement of the wheelset across the track. This is normally represented for a given wheel–rail pair by the graph plotted in Fig. 15.3, where the case of a wheelset with theoretical ORE S1002 profiles running on UIC60 rails laid at 1435 mm gauge and 1:20 cant

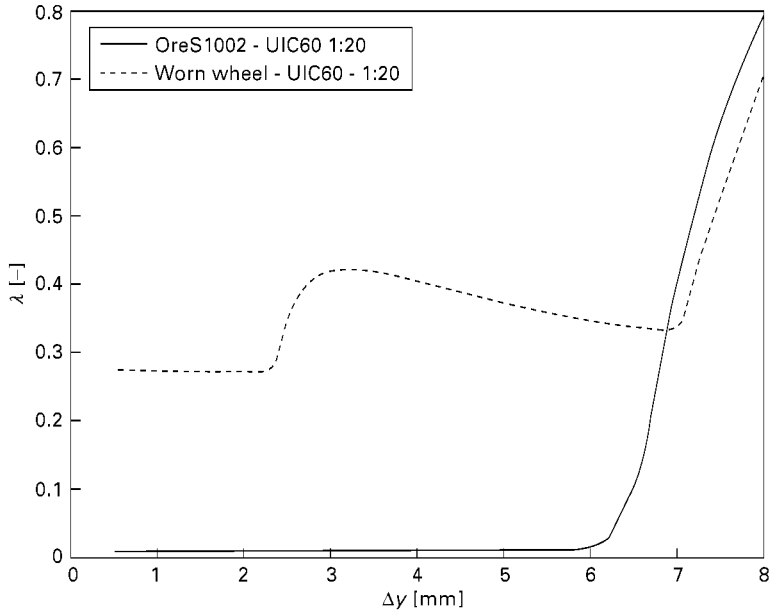


15.3 Rolling radius variation ΔR plotted vs wheel set lateral displacement y_{rel} for UIC60 1:20 rail coupled with new and worn ORES1002 wheel profiles.

(solid line) is compared to the case of a wheelset with symmetric worn ORE S1002 profiles running on the same track (dashed line).

As can be seen in Fig. 15.3, the rolling radius difference graph often shows a linear central portion, whose steepness depends on the shape of the contacting profiles and varies with time on account of wear progress. Since the slope of the graph in the central portion strongly affects the guidance, the stability and (to a certain extent) the steering behaviour of the vehicle, different measures have been proposed to quantify this parameter, resulting in different definitions proposed for the 'equivalent conicity', λ (Wickens, 2003). The procedure to compute the equivalent conicity is now standardised in Europe by the UIC leaflet 519 and is based on the calculation of the semi-cone angle of an equivalent coned wheel providing the same wavelength of kinematical oscillation as the actual wheel–rail profile considered (UIC, 2004). The equivalent conicity is expressed as a function of the amplitude of wheelset lateral oscillation across the track. Figure 15.4 shows the equivalent conicity graph for the same wheel–rail couples considered in Fig. 15.3.

It is observed in Figs 15.3 and 15.4 that the worn wheel profile has a higher equivalent conicity than the new one. This can be generalised to state that normally the equivalent conicity of the wheel–rail couple will increase with increasing wear of the wheel profiles, although Pearce (1996) mentions the possibility for a reduction of the equivalent conicity produced



15.4 Conicity graph for UIC60 1:20 rail coupled with new and worn ORES1002 wheel profiles. [λ = conicity, Δy = Lateral wheel set shift]

by wheel wear in the case of a vehicle with stiff suspensions travelling on a very curvaceous route. When coming to consider the rail profile, equivalent conicity is mostly affected by the flattening of the rail crown (see Fig. 15.2) which will also result in an increase of the equivalent conicity. It may be concluded then that in most cases an increase of wheel and rail wear will produce higher values of equivalent conicity and hence increased guidance for the solid wheelset.

Unfortunately, the variation of rolling radius difference on the solid wheelset is also responsible for the dynamic instability phenomenon known as 'hunting'. Indeed, pure rolling of the wheels over the rail rarely takes place, and a more complicated contact condition occurs where the effect of local elastic deformation in the wheel and rail surfaces cannot be neglected (Kalker, 1990). Based on the analysis of the distribution of normal pressures and tangential tractions acting in the contact patch, a relationship can be established between the kinematical deviation from pure rolling motion (referred to as creepage) and the tangential forces (or creep forces) arising in the contact patch. On account of the non-conservative nature of creep forces, a dynamic instability takes place at high speed, caused by a positive energy balance of the combined lateral and yaw vibration of the wheelset, despite the presence of dampers in the vehicle suspensions. The result is a self-excited vibration, whose amplitude is limited only by the non-linear effects associated with flanging of the wheels against the rails. The minimum forward speed of the vehicle corresponding to the onset of hunting vibration is called the 'critical speed' of the vehicle (Knothe and Böhm, 1999; Wickens, 2003).

Considering the equations of motion of one single wheelset elastically restrained by primary suspensions, or of the entire bogie/vehicle assembly, and linearising the creep force–creepage relationship, it is possible to show that an increase in the equivalent conicity produces a decrease of the vehicle critical speed. Therefore, an excessive increase of the equivalent conicity λ produced by wear of the wheel–rail profiles has to be kept under control since excessive wear may decrease the vehicle critical speed below the maximum service speed. Indeed, in most cases re-profiling of railway wheels is triggered by the equivalent conicity exceeding some threshold level which is assumed (or detected) to be incompatible with vehicle stability at service speeds.

In recent years, non-linear approaches have been developed to assess vehicle stability (Goodall and Iwnicki, 2004; Polach, 2006). These studies show that, depending on the trend of the equivalent conicity with the amplitude of the lateral wheelset motion, hunting instability may appear either as a gradual increase of the self-sustained lateral vibration with increasing vehicle speed (a behaviour known as 'subcritical Hopf bifurcation') or, more dangerously, as the sudden initiation of a large amplitude self-sustained vibration produced by a small increase of vehicle speed (leading to a 'supercritical Hopf

bifurcation'). Therefore, when considering the effect of wheel–rail wear on vehicle stability, not only the absolute value of the equivalent conicity, but rather the entire shape of the equivalent conicity diagram (Fig. 15.4) has to be considered.

Before closing this section, it shall be recalled that the considerations reported above refer only to vehicles equipped with solid wheelsets. Since the two wheels of an IRW axle are not forced to assume the same angular speed, an IRW vehicle will not experience hunting instability, regardless of the cross-sectional shape of the wheels and rails, but also will not provide any guidance effect. Therefore, an IRW vehicle running on a tangent track section in presence of irregularity will undergo high transversal displacements across the flangeway clearance, that are limited only by wheel flanging against the rails. This kind of behaviour is less affected by changes in the profile of the tread. However, it is possible that the sudden impacts occurring between wheel flanges and the gauge corners of the rails lead to a risk of derailment being triggered by large alignment defects in the track, as reported, for example, by Cheli *et al.* (2006).

15.2.3 Effect of profile wear on vehicle steering

When examining the steering attitude of a railway vehicle, it is necessary to consider that the wheelsets in the vehicle are constrained by the suspension. Therefore, the discussion of vehicle steering behaviour shall be based on considering the specific architecture of the vehicle. For the sake of conciseness, in this section the case of a two-axle bogie equipped with solid wheelsets will be considered, and the effects of interaction between the bogie and the carbody through the secondary suspension will be neglected.

There are basically two ways by which the bogie may run through a curve: if the curve radius is sufficiently large, flange contact can be avoided on all wheels, and the steering behaviour can be described with a sufficient level of accuracy by a linearised analysis, considering purely coned wheels with cone angle corresponding to the equivalent conicity λ of the wheel–rail pair. On sharper curves, however, flange contact between the outer wheel of the leading axle and the high rail cannot be avoided any longer and the entire geometry of wheel–rail contact has to be considered, including the possible formation of multiple points of contact between the same wheel–rail pair.

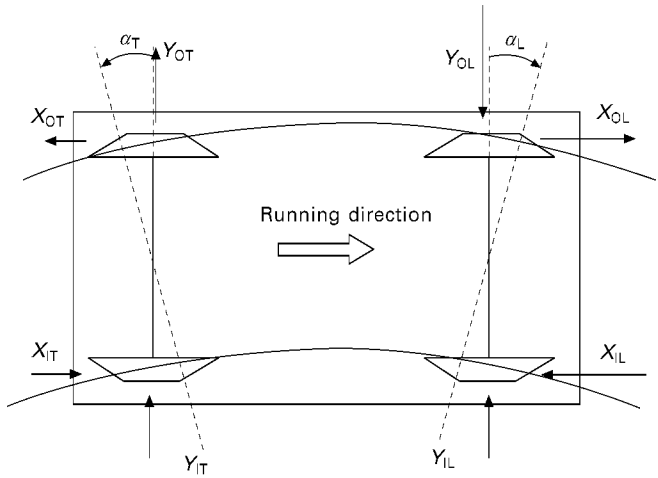
For large-radius curves, the results provided by the linearised approach (Wickens, 2003) show that an increase of the equivalent conicity improves the steering attitude of the vehicle, favouring the leading and trailing wheelsets to adopt the same angle of attack (rotation of the axle with respect to the radial direction) and thereby reducing to a minimum the lateral forces exchanged with the rails. The linear theory also shows that the minimum curve radius required to avoid wheel flanging (for a given wheelbase and for a given value

of primary suspension stiffness) is inversely proportional to the equivalent conicity, which means that wheel–rail pairs having a higher conicity allow the bogie to negotiate curves with smaller radii without entering into flange contact.

Since the equivalent conicity is normally increased by wear progress, there are cases where a reasonable amount of wear may improve the vehicle steering behaviour but, as described above, will negatively affect vehicle stability, an example of the well-known conflict between steering and stability in railway vehicles. The statement that wear improves steering should, however, be considered with care, because rail vehicle steering is a complicated process, which needs to be considered in several respects, including (at least): ride safety; forces applied on the track; and damage produced on the wheelsets and on the rails. Therefore a modification of the wheel–rail profiles being beneficial in some respects might turn out to worsen the vehicle's curving behaviour in other respects. An example of the trade-off between different requirements, all associated with vehicle negotiation, is provided later in this section.

Coming to consider the negotiation of relatively sharp curves, the typical attitude of the bogie during steady-state curving implies that flange contact is established at least on the outer wheel of the leading axle, whereas the lateral shift of the trailing axle depends upon a number of parameters including bogie wheelbase, stiffness of the suspension and cant deficiency. On the flanging wheelset, the rolling radius difference becomes greater than the value required to negotiate the curve in pure rolling, and therefore a longitudinal force X_{OL} arises on the outer wheel, pointing in the forward direction, as shown by Fig. 15.5. If the wheelset is not powered or braked, a longitudinal force X_{IL} having opposite direction takes place on the inner wheel, to balance the torsional rotation of the wheelset, and hence a steering torque is produced on the bogie, favouring curve negotiation. At the same time, on account of track curvature and of the two axles being linked by the suspension, an angle of attack will appear on both axles. On the leading axle, the angle of attack α_L is negative and the corresponding creep force is directed outwards the curve and has to be balanced by the flange force. On the trailing wheelset, the angle of attack α_T is positive, and the creep force is directed inwards the curve and counteracts the effect of the centrifugal force produced by cant deficiency.

The above described situation can be largely affected by modifications in the shape of the wheel and rail profile caused by wear. Worn wheel profiles normally provide a higher steering effect than new ones. This will reduce the guiding force acting on the outer wheel in the leading wheelset Y_{OL}, reducing the risk of flange climb derailment, and also damage to the wheel and to the high rail. However, a higher steering torque will also produce a re-distribution among the leading and trailing axles of the total lateral force



15.5 Attitude of a railway bogie during curve negotiation.

acting on the bogie due to cant deficiency. In tilting-body trains that are run at very high values of cant deficiency, it may be that the value of the track shift force on the trailing axle $Y_{OT} + Y_{IT}$ rises above the Prud'homme limit (Elkins and Carter, 1993), resulting in a limitation of the maximum cant deficiency allowed for the vehicle in service (Roberti and Bruni, 2001).

The curving behaviour of the bogie in a sharp curve is also heavily affected by the possible formation of two distinct contact points between the flanging wheel and the high rail, one on the tread and one on the flange. If this condition occurs, the rolling radius is much larger for the contact point on the flange than for the one on the tread, with the consequence that two longitudinal forces are generated pointing forward on the flange and backward on the tread. These two forces tend to cancel each other out, resulting in a reduction of the steering effect on the bogie. Depending on the initial shape of the profiles and on service conditions, wear effects on the wheel may have a different effect on the formation of two-point contact. New profiles, having a pronounced concavity of the flange root, will tend to favour the formation of two-point contact. In this case, the progress of wear may reduce the curvature of the profile and hence eliminate two-point contact. However, there are cases where the two contact point condition occurs as the consequence of the formation of hollow wear on the wheel (Sawley and Wu, 2005).

It must finally be recognised that the curving attitude of a railway bogie is a complicated and highly non-linear phenomenon, which makes it difficult to formulate general conclusions valid across a large range of vehicle types and for different curve radii. For a more detailed treatment of this subject the reader is referred to Wickens (2003).

15.2.4 Effect of profile wear on flange climb derailment

When considering the mechanism of flange climb derailment, wear occurring on the wheel flange and on the gauge corner of the rail plays an important role. Wearing of the high rail in the gauge corner results in a reduction of the maximum contact angle and, in sharp curves, the rail profile will tend to become conformal to the wheel flange, a situation which can cause specific problems if a vehicle with a different wheel profile is introduced in the line.

According to Pearce (1996), the maximum angle on the wheel flange will normally tend to increase slightly with wear when the train is running on a conventional railway. However, on curvaceous routes with many unlubricated sharp curves (where wear occurs at high values of the angle of attack), the effect of wear is to reduce the flange angle. This happens in particular for urban rail vehicles with IRW, often travelling on tracks with sharp curves and at large angles of attack due to the lack of a steering action in this type of axle. Indeed, flange climb derailment of IRW vehicles is a known problem of urban rail transport systems in which the wear of wheel and rail profiles plays a key role.

15.2.5 Effect of hollow-worn and asymmetric wheel profiles

In this section, the effect on vehicle dynamics of two specific wear patterns, hollow wear and asymmetric wear, is examined. Hollow wear is reported by Pearce (1996) to appear most frequently on vehicles with soft suspensions running on relatively straight routes. The same type of wear is known to appear frequently also in freight trains (Sawley *et al.*, 2005). Asymmetric wear may arise on account of asymmetries in the rolling stock, such as an initial mismatch of the wheel radii or a misalignment of the wheelsets in the bogie, or may be caused by the train constantly running on a route with asymmetric geometry.

Severely hollow worn wheels show a negative slope of the rolling radius difference diagram in its central region. Thus, the wheelset loses its natural guidance properties and is forced to run in flange contact with one of the rails. According to Sawley and Wu (2005), hollow wear significantly increases lateral forces in straight track and large-radius curves, possibly causing accelerated track degradation and increasing the risk of derailment caused by rail rollover. The same authors also indicate that hollow wheels cause an increase in the specific rolling resistance per unit normal load.

In the case of asymmetric wear, the rolling radius difference diagram will be offset in lateral direction, so that the position corresponding to zero

rolling radius difference will not coincide with the wheelset being centred across the flangeway clearance. In the case of severe asymmetric wear, the wheelset will be forced to run against one rail, and additional contact forces will be generated. Furthermore, the high slope of the rolling radius difference diagram corresponding to one wheel approaching the flange contact condition might lead to hunting instability at low speed (Sawley *et al.*, 2005).

15.2.6 Effect of profile wear on turnout passing

To conclude the discussion of transversal profile wear effects on rail vehicle dynamics, specific effects related to turnout negotiation have to be mentioned. The varying profile of the switch blade and of the crossing nose are designed to allow a smooth transition of the wheels from the stock rail to the blade and across the gap in the crossing. However, the design of these profiles has to be based on considering the theoretical shape of the wheel profile (i.e. neglecting wheel wear). Furthermore, localised wear occurring on the switch blade and on the crossing nose affect the wheel–rail contact and may be the cause of serious dynamic effects.

When the wheel is transferred from the stock rail to the blade, excessive wear of the blade causes a mismatch between the wheel and the blade which, in serious cases, may lead to derailment with the wheel climbing the tip of the blade. This problem is not just associated with the amount of wear accumulated on the blade, but especially with the variation of wear along the blade: a worn blade tip followed by a less worn profile of the blades should be avoided because this will produce conformal contact of the wheel flange with the blade tip, whereas the following recovery of the theoretical shape of the blade will form a sort of ramp, sustaining the flange climb process (Pearce, 1996).

Considering the negotiation of the crossing panel, wear of the wheel profile and of the crossing nose often results in high impact loading: even in presence of a 'perfect' geometry of the crossing nose, hollow wear on the wheel tread hinders the smooth transfer of wheel–rail contact from the stock rail to the crossing nose (Alfi and Bruni, 2009), and causes an impact of the wheel on the crossing nose. This situation is worsened by localised wear and damage occurring on the crossing nose and results in accelerated degradation of the crossing panel and in increased noise and vibration disturbance, which is particularly harmful in urban rail networks.

15.3 Effects of rail corrugation

The abnormal cyclic wear patterns known as corrugations which develop on initially smooth railheads occur in many diverse situations on straight and curved track. It is commonly accepted that they are of two types according

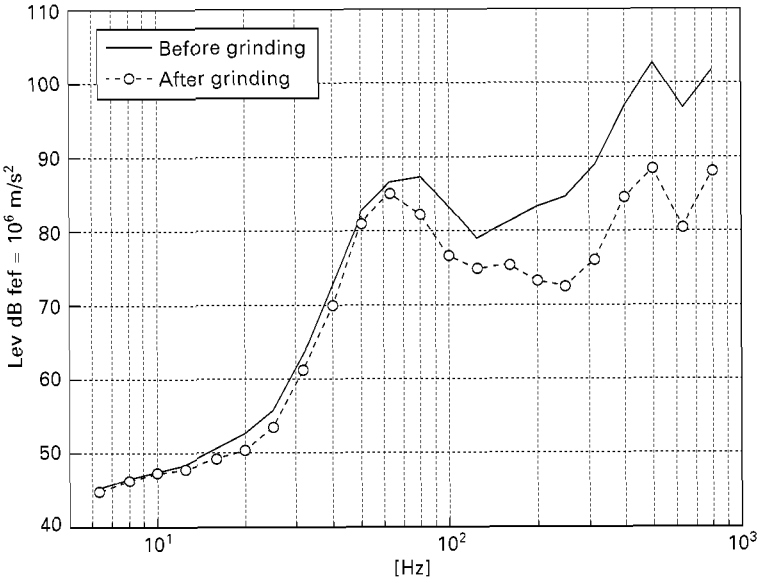
to their wavelength: long wavelength and short wavelength corrugation. When fully developed, long wavelength corrugations may display vertical undulations up to 5 mm deep with a wavelength in the range of 200–1500 mm. This length is related to the sleeper pitch (fractions or multiples). Short wavelength corrugations, instead, are generally less than 0.25 mm deep and have a wavelength between 40 mm and 100 mm, apparently irrespective of the train speed. In some situations, short-wave corrugations occur in groups spaced at sleeper pitch, and certain types of rail steel are more prone than others to short wave corrugation (Clark and Foster, 1983).

Corrugation gives rise to high dynamic loads between wheel and rail that lead to the degradation of the ballast and other track and train components. However, the most evident effects of rail corrugation are ground-borne vibrations and high-frequency noise emission. These two effects are of significant practical concern to the railway industry, especially for lines in urban areas at small distances from neighbouring houses or lines in shallow-depth tunnels under buildings.

The ground-borne vibrations can be perceived by the inhabitants via the floor vibrations, as well as via the noise re-radiated inside the houses by the vibrating building structures (secondary noise). Moreover, noise is particularly annoying with short wavelength corrugations which excite vibrations in the audible frequency range. Typical measures used to mitigate the effect of rail corrugation are frequent rail grinding and use of friction modifiers, both having implications on maintenance costs: corrugation is therefore the cause of a significant increase in maintenance costs.

Figure 15.6 shows tunnel wall acceleration levels measured before and after rail grinding on an underground metro line equipped with a direct fastening track system: it is shown that rail grinding reduces by 10 dB or more the vibration produced by train passage in the 100–800 Hz frequency range. Of course, this result depends on the considered track typology, and other track systems can be less sensitive to rail corrugation. Still, the result clearly shows the impact of rail corrugation on the vibro-acoustic annoyance produced by the railway line.

A large amount of research work has been carried out since the 1980s to find appropriate measures to prevent the formation of rail corrugation or at least to reduce the rate of corrugation growth: investigations have specifically addressed the design of track systems that do not present a wavelength fixing mechanism, the optimisation of track parameters to reduce corrugation phenomena and the design of the running gear to reduce train–track interaction and its effect on corrugation. A review of studies on rail corrugation can be found in Sato *et al.* (2002), whereas a classification of corrugation mechanisms, their causes and treatments can be found in Grassie and Kalousek (1993).



15.6 Measured tunnel wall acceleration level before and after rail grinding for direct fastening track.

15.4 Effects of wheel out-of-roundness

In the literature, wheels carrying any type of defect are referred to as 'out-of-round'. This can consist of either a discrete defect, such as a wheelflat, or a deviation from circularity occurring along the entire wheel circumference. In this section we will analyse this latter case, whereas the effect of concentrated defects will be treated in the next section.

Out-of-roundness defects can be classified into periodic and stochastic defects: in the first case, the deviation from circularity is represented by a single harmonic component or by the superimposition of a limited number of harmonics, whereas stochastic non-roundness consists of the random superimposition of several harmonic components defined along the wheel circumference (Nielsen and Johansson, 2000). Among periodic defects, we can further distinguish out-of-round wheels according to the wavelength of the irregularity: if this is lower than 0.5 m (more than five periods around the wheel circumference), we speak of wheel tread corrugation; otherwise of wheel polygonalisation (Meywerk, 1999).

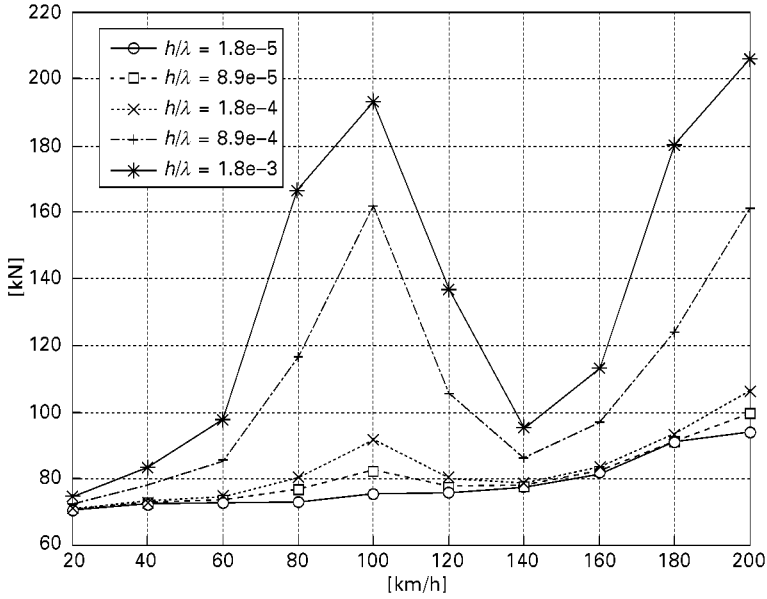
Polygonalised wheels have been detected only on disc-braked wheelsets of high-speed trains, and the defect amplitude is of the order of 1 mm. Several causes of wheel polygonalisation have been identified: according to some research, this type of defect is determined by inhomogeneous tread material properties (Müller *et al.*, 1995), while according to others the natural

vibrations of the wheelset excited by dynamic imbalance of the disc-braked wheelset are the true cause of this phenomenon (Vohla, 1996; Brommundt, 1997; Meinke and Meinke, 1999). In any case, it has been shown that the number of harmonics that dominate in polygonalised wheels is a function of the wheel stiffness, the track properties and the train speed (Pallgen, 1998).

Stochastic non-roundness of the wheel tread usually appears on wheels that are block braked, and the dominating circumferential wavelength is equal to 3–7 cm with an amplitude of less than 0.1 mm. The causes of such wheel tread defects are still subject of investigation: according to Vernersson (1999), during braking manoeuvres some regions of the wheel tread become warmer (hot spots) due to thermoelastic instability and suffer higher wear rates. However, the origin of hot spots is yet not clarified.

Considering the effect of wheel out-of-roundness on vehicle dynamics, stochastic non-roundness of the wheel is the cause of high-frequency impacts having small amplitude, whereas wheel polygonalisation determines low-frequency impacts having greater amplitude. Thus, wheel polygonalisation gives rise to high-cycle fatigue phenomena that reduce service life and cause discomfort to passengers owing to high vibration amplitudes. Stochastic non-roundness, instead, mainly increases noise and vibration emission in the frequency range between 60 and 2000 Hz. It is stated by Ahlbeck and Harrison (1988) that the amplitude of large impact loads is related to the ratio of depth to wavelength of wheel tread irregularity and that the most critical wheel defects are not always easily detected by visual inspection of the wheel.

Cai and Raymond (1992) show that one defective wheel can lead to impact loading on the wheel on the same side of the bogie and the opposite wheel of the same wheelset but of smaller magnitude, while Dong and Sankar (1994) conclude that impact loads increase with the train speed, reaching a maximum when the polygonalisation length, coupled with the vehicle speed, excites the first eigenfrequencies of the train–track system (usually around 70–110 Hz and 270–280 Hz for ballast track). This conclusion is supported by the results of calculations performed for the purposes of this chapter using a train–track interaction model. Figure 15.7 shows the peak impact load value corresponding to different values of the travelling speed and of the polygonalisation depth, for harmonic out-of-roundness having wavelength 550 mm (five periods along the wheel circumference): it is observed that the maximum peak load occurs at the speed of 100 km/h, corresponding to 51 Hz frequency, for which the resonance of the first vertical mode of vibration of the track coupled to the wheelset occurs.



15.7 Simulated peak vertical contact force component as a function of the ratio between defect depth and length for a high-speed vehicle running on a ballast straight track at speeds ranging from 50–200 km/h.

15.5 Effects of localised damage on wheel and rail profiles

Localised damages on wheels and rails are usually classified according to their length. Very small localised defects, having length in the range of 1 mm or less, are usually due to non-severe rolling contact fatigue (RCF) phenomena such as spalling and shelling. Although quite frequent, these small defects do not determine significant wheel–rail impact loads and will not be treated further in this section.

Larger localised defects (up to 50 mm length in longitudinal/circumferential direction) may be subclassified according to their origin: for wheel defects we consider wheelflats and other local defects, while on the rails defects caused by shelling and spalling may occur. Plastic deformation is usually common in conjunction with these type of wheel–rail defects.

Wheelflats are caused by poorly adjusted or faulty brakes, incorrect braking procedures or differences in friction coefficient on the rail, and also by wheelspin of driven wheels. Thus, two different mechanisms determine wheelflats: abrasive wear and spalling on an extended wheel tread area. Note that, when a wheelflat is formed, heat is generated in the wheel material, causing a phase transformation of the steel to the brittle martensite phase.

When loaded dynamically, some material pieces of the tread may fall off. This is followed by an elongation of the original wheelflat because of plastic yielding at its edges.

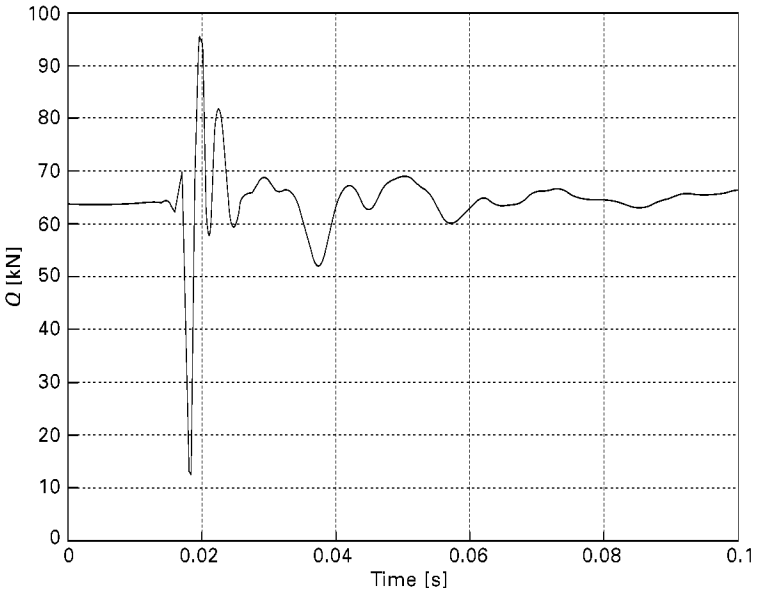
Other local wheel defects, instead, are due to differences in wheel tread material properties that determine different wear regimes and fatigue behaviour over the wheel rolling surface. Usually, local wheel defects occur in wheel steels with a low carbon content that have not been normalised in the production process, thus having a material structure with large grain sizes and reduced material strength. Wheels with such material, in combination with non-uniform material properties around the circumference, may be exposed to local surface- or subsurface-initiated RCF cracking, and a long local defect is formed.

Wheel and rail localised defects ranging from 1 mm to 50 mm determine non-negligible impact loads that may cause severe damages to both track and vehicle components. Moreover, impact loads determine increased wear and microstructural changes: in tempered martensitic steels, the hardening process prevents crack growth, whereas in a pearlitic steels, local cementite cracking occurs.

Small skid flats have been observed to roll out in service. This might be attributed to both a batter mechanism and to wear. Large skids, instead, rarely roll out but, by the previous mechanisms, may become larger. Another source of impact loads is represented by material transfer of wheel–rail debris or cast iron brake blocks debris to the wheel. As for wheelflats, severe cases of tread buildup are generally associated with faulty brake mechanisms or with a handbrake that is left on.

Wheelflats, as well as other local wheel–rail defects, determine wheel–rail contact forces that have a characteristic shape (see Fig. 15.8): the response begins with a sudden drop in wheel–rail contact force due to the increasing radial deviation from the nominal wheel radius. During this phase, the wheel moves downwards and the rail upwards to compensate for the missing wheel material (elastic return). Depending on the wheelflat depth and trainspeed, loss of contact between the wheel and the rail may occur in this phase. When the contact is displaced to the trailing region of the wheelflat the radial deviation of wheel surface from the nominal one decreases while the wheel keeps moving downwards on account of its large inertia, so that the rail is suddenly pushed downwards and a large peak is produced in the vertical wheel–rail force, followed by a damped transient response.

Experimental investigations (Johansson and Nielsen, 2003) as well as numerical simulations have shown that impact loads increase with increasing defect length, depth/height ratio and train speed. For wheel local defects other than wheelflats, Kalay *et al.* (1995) have shown that the length of the defect is most important: a wheel defect of 0.5 m length and 3.5 mm depth may cause peak impact loads greater than 540 kN at a train speed of



15.8 Simulated time history of the vertical wheel–rail contact force Q for a high-speed vehicle running at 100 km/h in presence of a wheelflat.

100 km/h and an axle load of 330 kN. Wheelflats and local wheel defects usually show an almost linear relationship between maximum wheel–rail contact force and train speed, whereas, as shown in Fig. 15.17, for polygonalised wheels there is usually a maximum of wheel–rail contact forces at a given train speed due to track resonance excitation.

High impact loads produced by this kind of defect can be particularly critical to the fatigue resistance of various track and vehicle components such as fasteners, sleepers, wheels, axleboxes, etc. as they may produce crack initiation and accelerate the subsequent crack propagation. Thus, the design of such components has to take impulsive loads into account. It should be observed that in the above list of track and vehicle components whose life is influenced by impact loads, axles were intentionally left out. In fact, results from a measuring campaign performed with a measuring wheelset on a test track (Alfi *et al.*, 2008) have shown that singularities in wheels and rails actually produce extremely high train–track interaction effects, but the stresses in the axle seem to be only marginally affected.

Besides causing high impact loads, concentrated wheel and rail defects increase fuel consumption (Barke and Chiu, 2005) and increase pass-by noise levels (Wu and Thompson, 2002). Rolling resistance of a wheel on a rail is low, provided the surfaces are smooth. A wheelflat immediately results in an increase in fuel consumption. Being the peak sound pressure level

proportional to the change in momentum, it is evident that wheel and rail defects may significantly increase the level of noise pollution. To minimise maintenance costs, decrease noise emission, reduce fuel consumption and avoid catastrophic events such as train derailments, it is therefore of great importance to detect and replace non-round wheels and to remove rail concentrated defects timely.

15.6 Conclusions

In this chapter, the various effects of wheel–rail surface damage on vehicle dynamics have been presented: changes in wheel–rail transversal profiles produced by regular wear mainly affect the critical speed and curving behaviour of the vehicle, whereas damage-induced alterations of the longitudinal rail profile (corrugation) and of the wheel circumferential profile (wheel out-of-roundness) are responsible for increased train–track interaction effects in the mid–high frequency range (20–5000 Hz) and cause increased emission of noise and vibration. All these effects have important implications for the behaviour of the train–track system, so that excessive damage may determine reduced performance of the system, increased vibro-acoustic annoyance and, in extreme cases, reduced levels of safety.

The accumulation of wear and damage effects at the wheel–rail interface has therefore to be kept under control either by grinding or by adopting appropriate design measures to reduce the rate of surface damage growth. In this last view, a number of means to reduce regular wear on the wheels are discussed in Chapter 6, whereas for details regarding causes and remedies for other defects the reader is referred to Chapters 8 and 11 on out-of-round wheels and rail corrugation, respectively.

15.7 References

Ahlbeck DR and Harrison HD (1988), The effects of wheel/rail impact loading due to wheel tread runout profiles, *Proceedings 9th International Wheelset Congress*, Montreal, QC, Canada, 12–15 September.

Alfi S and Bruni S (2009), Mathematical modelling of train-turnout interaction, *Vehicle System Dynamics*, **47**(5), 551–74.

Alfi S, Braghin F and Bruni S (2008), Numerical and experimental evaluation of extreme wheel-rail loads for improved wheelset design *Vehicle System Dynamics*, 46, (Supplement), 431–44.

Barke DW and Chiu WK (2005), A review of the effects of out-of-round wheels on track and vehicle components, *Proceedings of the IMechE Part F: Journal of Rail and Rapid Transit*, **219**, 151–75.

Brommundt E (1997), Simple mechanism for the polygonalization of railway wheels by wear, *Mechanics Research Communications*, **24**, 435–42.

Cai Z and Raymond GP (1992), Theoretical model for dynamic wheel/rail and track

interaction, *Proceedings 10th International Wheelset Congress*, Sydney, NSW, Australia, September 27–October 1.

Cheli F, Corradi R, Diana G, Facchinetti A and Gherardi F (2006), Effect of track geometrical defects on running safety of tramcar vehicles, *Vehicle System Dynamics*, **44**(Supplement), 302–12.

Clark RA and Foster P (1983), On the mechanics of rail corrugation formation, *Vehicle System Dynamics*, **12**(1), 35–9.

Dong RG and Sankar S (1994), The characteristics of impact loads due to wheel tread defects, *RTD Rail Transportation ASME*, **8**, 23–30.

Elkins JA and Carter A (1993), Testing and analysis techniques for safety assessment of rail vehicles: the state-of-the-art, *Vehicle System Dynamics*, **22**(3–4), 185–208.

Goodall RM and Iwnicki SD (2004), Non-linear dynamic techniques v equivalent conicity methods for rail vehicle stability assessment, *Vehicle System Dynamics*, **41**(Supplement), 791–99.

Grassie SL and Kalousek J (1993), Rail corrugation: characteristics, causes and treatments, *Proceedings of the IMechE, Part F: Journal of Rail and Rapid Transit*, **207**, 57–68.

Johansson A and Nielsen JCO (2003), Out-of-round railway wheels – wheel-rail contact forces and track response derived from field tests and numerical simulations, *Proceedings of the IMechE Part F: Journal of Rail and Rapid Transit*, **217**, 135–46.

Kalay S, Tajaddini A, Reinschmidt A and Guins A (1995), Development of a performance-based wheel-removal criteria for North American railroads, *Proceedings 11th International Wheelset Congress*, Paris, France.

Kalker J J (1990), *Three-Dimensional Elastic Bodies in Rolling Contact*, Dordrecht, the Netherlands, Kluwer.

Knothe K and Böhm F (1999), History of stability of railway and road vehicles, *Vehicle System Dynamics*, **31**(5–6), 283–323.

Meinke P and Meinke S (1999), Polygonalization of wheel treads caused by static and dynamic imbalances, *Journal of Sound and Vibration*, **227**, 979–86.

Meywerk M (1999), Polygonalization of railway wheels, *Archive of Applied Mechanics*, **69**, 105–20.

Müller R, Diener B and Diener M (1995), Verschleißerscheinungen an Radlaufflächen von Eisenbahnfahrzeugen, *ZEV DET Glasers Annalen (Zeitschrift für Eisenbahnwesen und Verkehrstechnik)*, **119**, 177–92.

Nielsen JCO and Johansson A (2000), Out-of-round railway wheels - a literature survey, *Proceedings of the IMechE, Part F: Journal of Rail and Rapid Transit*, **214**, 79–91.

Pallgen G (1998), Unrunde Räder an Eisenbahnfahrzeugen, *Eisenbahningenieur*, **49**(1), 56–60.

Pearce T G (1996), Wheelset guidance – conicity, wheel wear and safety, *Proceedings of the IMechE, Part F: Journal of Rail and Rapid Transit*, **210**, 1–9.

Polach O (2006), On non-linear methods of bogie stability assessment using computer simulations, *Proceedings of the IMechE, Part F: Journal of Rail and Rapid Transit*, **220**, 13–27.

Roberti R and Bruni S (2001), Development of operations of tilting train on Italian network, *Proceedings 5th World Congress on Railway Research*, Cologne, Germany, 25–29 November, on CD.

Sato Y, Matsumoto A and Knothe K (2002), Review on rail corrugation studies, *Wear*, **253**, 130–39.

Sawley K and Wu H (2005), The formation of hollow-worn wheels and their effect on wheel/rail interaction, *Wear*, **258**(7–8), 1179–86.

Sawley K, Urban C and Walker R (2005), The effect of hollow-worn wheels on vehicle stability in straight track, *Wear*, **258**(7–8), 1100–1108.

UIC (2004), *Method for Determining the Equivalent Conicity*, Leaflet 519, International Union of Railways, Paris, France.

Vernersson T (1999), Thermally induced roughness of tread braked railway wheels. Part 1: brake rig experiments, and Part 2: modelling and field measurements, *Wear*, **236**, 96–116.

Vohla G (1996), *Werkzeuge zur realitätsnahen Simulation der Laufdynamik von Schienenfahrzeugen*, PhD Thesis, MI Verlag, Düsseldorf, Germany.

Wickens A H (2003), *Fundamentals of Rail Vehicle Dynamics*, Lisse, the Netherlands, Swets & Zeitlinger.

Wu TX and Thompson DJ (2002), A hybrid model for the noise generation due to railway wheel flats, *Journal of Sound and Vibration*, **251**(1), 115–39.

Copyrighted Materials
Copyright © 2009 Woodhead Publishing Retrieved from www.knovel.com

16

Noise and vibration from the wheel–rail interface

D. THOMPSON and C. JONES, University of Southampton, UK

Abstract: This chapter discusses the various types of noise arising from the wheel–rail interface and means for controlling them. The main source is rolling noise which occurs on plain track and has a broad frequency content. Discrete features such as rail joints and wheel flats generate impact noise, whilst in curves very high squeal noise levels can be produced. In each of these cases, the noise has its origins at the wheel–rail interface, but it is not emitted from the interface region itself. In the case of both rolling and impact noise, vibration of the whole wheel structure and of a considerable length of track are involved in radiating the noise. For curve squeal, specific modes of vibration of the wheel radiate the noise. At low frequencies, vibration is transmitted through the ground and may be experienced either as feelable vibration or as a low frequency rumbling noise.

Key words: wheel–rail rolling noise, wheel–rail impact noise, curve squeal noise, ground-borne vibration.

16.1 Introduction

Noise is an important and unfortunate by-product of the process of the wheel rolling on the rail. It is remarkable that the noise energy produced by a rolling wheel is of the order of one millionth of the energy taken to drive the train. Yet the control of railway noise has become an important issue in recent years due to the introduction of legislation across Europe intended to control noise. This takes the form of noise limits for individual vehicles through the Technical Specifications for Interoperability (EC, 2001; EC, 2002a) and the requirement to draw up Action Plans for large conurbations and busy railway routes through the Environmental Noise Directive (EC, 2002b).

Noise from the wheel–rail interface takes a number of different forms. On straight track, noise with a broad frequency content occurs, known as rolling noise (Sections 16.3 and 16.4). Discrete features such as rail joints and wheelflats generate an impulsive noise (Section 16.5). In curves, very high noise levels can be produced which are often dominated by single frequencies. This phenomenon is known as curve squeal (Section 16.6). In each of these cases, the noise has its origins at the wheel–rail interface, but it would be wrong to assume that it is emitted from the interface region itself. For curve squeal, modes of vibration of the whole wheel are responsible for

the noise radiation. In the case of both rolling and impact noise, vibration of the whole wheel structure and of a considerable length of track is involved in radiating the noise.

At low frequencies, vibration is transmitted from the track through the ground and may be experienced either as feelable vibration (in the range 4–80 Hz) or as a low-frequency rumbling noise (30–250 Hz). These are described in Section 16.7. Other sources of railway noise, not covered in this chapter, include aerodynamic noise, which is important at high speeds, and traction noise, which is often dominated by the noise from fans. Nevertheless, the wheel–rail interface is the origin of the most important sources of noise from the railway system. Before discussing each of these sources of noise, their generation mechanisms and means for their control, a short introduction is given to the field of noise and vibration.

16.2 Basics of noise and vibration

This section gives only a very brief overview of some basic quantities in the field of acoustics. The interested reader is referred to textbooks on the subject for further details, e.g. Fahy and Walker (1998) and Kinsler *et al.* (1982).

Sound consists of audible fluctuations in air pressure which propagate through the air as waves with a wave speed, c_0, of about 340 m/s at 20 °C. To express the magnitude of a sound, the squared sound pressure is usually averaged over time to give the rms (root mean square) value:

$$p_{\text{rms}} = \left(\frac{1}{T} \int_{t_1}^{t_1+T} p^2(t)\, \mathrm{d}t \right)^{1/2} \qquad [16.1]$$

where $p(t)$ is the sound pressure at time t and T is the averaging time. Moreover, sound is usually analysed into different frequency components, e.g. using Fourier analysis. A sinusoidal sound pressure of circular frequency ω rad/s ($\omega = 2\pi f$ where f is the frequency in Hz) can be written as:

$$p(t) = p_0 \cos(\omega t - \phi) \qquad [16.2]$$

where p_0 is the amplitude and ϕ is a phase angle. Alternatively complex variable notation may be used.

The normal ear is sensitive to sound in the frequency range 20–20 000 Hz, although the upper limit reduces with age and with noise exposure. This defines the frequency range of interest in acoustics. The ear is also sensitive to a very large range of amplitudes (around six orders of magnitude). Logarithmic scales are therefore generally used to present acoustic data, not only because of these large ranges, but also to mimic the way the ear responds to sound. Amplitudes are expressed in decibels. For example, the *sound pressure level* is defined as:

$$L_p = 10 \log_{10}\left(\frac{p_{\text{rms}}^2}{p_{\text{ref}}^2}\right) = 20 \log_{10}\left(\frac{p_{\text{rms}}}{p_{\text{ref}}}\right) \quad [16.3]$$

where the reference pressure p_{ref} is usually taken as 2×10^{-5} Pa. This reference level is used because 0 dB then corresponds approximately to the threshold of hearing (at 1 kHz). In practice, the ambient noise level even in very quiet places is seldom below 30 dB. In contrast, sound pressure levels of over 100 dB are in the range where prolonged exposure would lead to a significant risk of hearing damage.

Frequencies are also generally plotted on logarithmic scales. The frequency range is then divided into bands that are of equal width on a logarithmic scale, with *one-third octave bands* being the most common form of presentation. The centre frequencies of each band are given by $10^{(N/10)}$ where N is the band number, although by convention they are rounded to particular standard values. Band number 13 thus corresponds to 20 Hz and band 43 to 20 kHz.

The total sound emitted by a source can be quantified by its power, W. In decibel form the *sound power level* is written as:

$$L_W = 10 \log_{10}\left(\frac{W}{W_{\text{ref}}}\right) \quad [16.4]$$

where the reference value W_{ref} is usually 10^{-12} W.

The sound pressure level at some distance r from a compact source emitting power W into free space can be written as:

$$L_p = L_W - 20 \log_{10}(r) - 11 - DI \quad [16.5]$$

where DI is the directivity index and depends on the direction of the receiver location. The square of the sound pressure generally increases in proportion to the power, 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, the sound pressure reduces as the receiver moves further from the source. For the compact point source of Eq. (16.5) this reduction is 6 dB per doubling of distance, while for a line source it is 3 dB per doubling.

Sound is generated by a variety of mechanisms, but they can mostly be grouped into two main categories:

- radiation by 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;
- sound produced by unsteady aerodynamic flow – thus wind, particularly turbulence and flow over solid objects, also produces sound, e.g. jet noise, turbulent boundary layer noise, exhaust noise and fan noise.

It is important to realise that sound generation is often a very inefficient process. A sound source of only 1 W corresponds to large sound pressure levels (109 dB at 1 m according to Eq. (16.5)), whereas this would typically be produced by a machine with an output of hundreds of kW. The *proportion* of the mechanical power of many machines that is converted into sound (often termed the *acoustic efficiency*) is thus often in the range 10^{-7} to 10^{-5} (Fahy and Walker, 1998).

Noise is actually *unwanted* sound. In practice, the acceptability of sound levels and signal content varies greatly between individuals. Various frequency weighting curves have been devised in an attempt to approximate the way that the ear perceives different sounds. The A-weighting is the most commonly used. As the ear is most sensitive around 1–5 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. To calculate this, the weighting curve is applied to the spectrum before determining the total level.

The vibration of a structure can be expressed in terms of its displacement, u, velocity, v or acceleration, a. For vibration at a frequency ω the corresponding amplitudes (ignoring their relative phases) are related by

$$a = \omega v = \omega^2 u \qquad [16.6]$$

The vibration can also be expressed in decibels in the same way as the sound pressure (see Eq. 16.3), but the corresponding reference values vary.

The sound radiation from a vibrating object in a particular frequency band can be written as

$$W = \rho_0 \, c_0 \, S \, \langle \overline{v^2} \rangle \, \sigma \qquad [16.7]$$

where ρ_0 is the density of air, c_0 is the speed of sound, S is the surface area of the structure and $\langle \overline{v^2} \rangle$ is the surface-averaged, mean-square (i.e. rms squared) velocity normal to the surface in the frequency band of interest. From Eq. (16.7), the radiated sound power is proportional to the square of the surface vibration amplitude. The factor σ is known as the radiation ratio or radiation efficiency. It is determined by the size and shape of the structure and by the distribution of its vibration (e.g. its structural wavelength), but it is independent of the amplitude of vibration. Generally σ is small at low frequency when the structure is small compared with the acoustic wavelength, and tends to unity at high frequency as the wavelength becomes small. Predictions of this factor can be obtained using numerical methods such as the boundary element method or, for simple cases, analytical models.

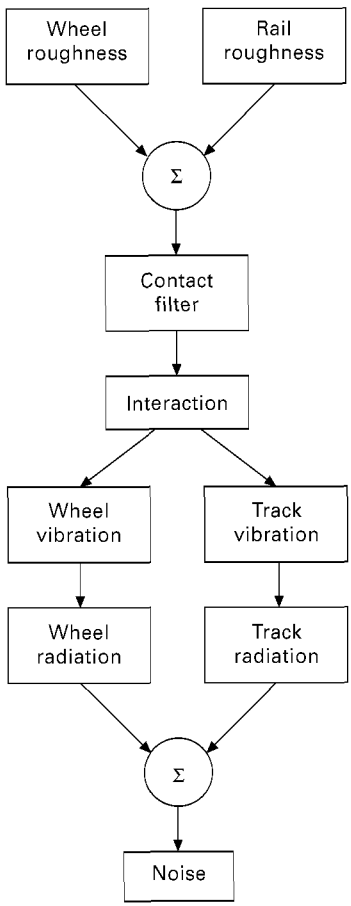
16.3 Rolling noise

16.3.1 Surface roughness

The most important form of noise generated at the wheel/rail interface is rolling noise. This is caused by the fact that the wheel and rail running surfaces are not perfectly smooth. The process by which rolling noise is generated is summarised in Fig. 16.1.

Undulations of the wheel and rail surfaces with wavelengths between around 5 and 250 mm produce vibrations in the audible frequency range when traversed at typical train speeds. This follows from the equation:

$$f = \frac{v}{\lambda} \qquad [16.8]$$



16.1 Model for rolling noise generation.

where, if v is the speed in m/s and λ is the wavelength in m, then f is the frequency generated in Hz. These undulations are generally referred to as *roughness*, but this should not be confused with micro-roughness which corresponds to a shorter range of wavelengths. At the wavelengths of interest to rolling noise, the roughness has amplitudes in the range from tens of microns at long wavelengths to less than a micron at short wavelengths. The amplitude is thus of the order of 10^{-4} times the wavelength and is not visible on what appears to be a smooth surface.

It may seem surprising that such small amplitudes can lead to high noise levels, but a simple estimate using the equations from the previous section shows that this is so. Taking an rms vibration amplitude of 1 μm at 2 kHz, and assuming that the whole wheel vibrates with this amplitude (in fact it may amplify the vibration at some frequencies), the corresponding velocity amplitude from Eq. (16.6) is 0.013 m/s. Substituting this into Eq. (16.7) along with the values $\rho_0 = 1.2$ kg/m^3, $c_0 = 340$ m/s, $S = 2\pi r^2 = 1.3$ m^2 (for a wheel radius $r = 0.46$ m) and assuming $\sigma = 1$, the sound power is estimated as 0.08 W, or 109 dB re 10^{-12} W. The corresponding sound pressure level at 7.5 m is 84 dB, based on radiation into a hemisphere above a rigid ground. This is clearly a significant source of sound.

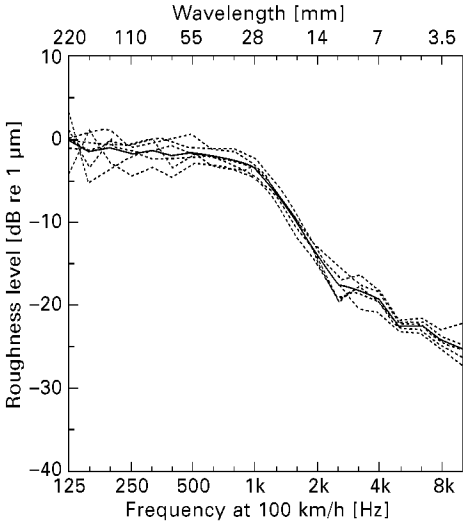
Since the wheel–rail contact occurs over an area (see Chapter 3), roughness wavelengths that are short compared with the contact patch length (typically 10–15 mm) are attenuated in their excitation of the system. This effect, known as the contact filter, tends to limit the effective sound radiation to frequencies below about 5 kHz. The attenuation is significant for frequencies from about 1–1.5 kHz for a speed of 160 km/h, and at even lower frequencies for lower speeds.

An example of the contact filter effect is shown in Fig. 16.2. This was determined using a numerical 'discrete point reacting springs' (DPRS) model (Remington and Webb, 1996) with roughness measurements obtained on multiple parallel lines. This shows results obtained using a series of such measurements in combination with the DPRS model, from Thompson (2003).

16.3.2 Wheel and track dynamics

The roughness excites the system by introducing a relative displacement between the wheel and rail. As shown in Fig. 16.1, both the wheel and the track vibrate as a result of the roughness excitation. These systems act as amplifiers to the noise, radiating sound over a much larger area than the contact itself.

The relative importance of the wheel and track components of sound radiation depends on details of their designs as well as on the train speed and the roughness spectrum but, in many cases, both the wheels and the



16.2 Contact filter effect due to contact patch of semi-axis length 5.7 mm. Numerical DPRS model for data from six wheels — mean of six wheels. (From Thompson, 2003, used with permission from Elsevier)

track are significant sources. As the noise radiation depends on the combined roughness of both wheel and track, it is possible that a rough wheel causes significant noise to be radiated mainly by the track vibration or vice versa. It is therefore difficult to assign noise contributions solely to the vehicle or infrastructure.

A railway wheel is a lightly damped, resonant structure. As with any structure, it can vibrate freely at a series of frequencies called its 'resonance' or 'natural' frequencies. The associated vibration patterns are called the mode shapes. Wheels are usually axisymmetric, and their normal modes of vibration can therefore be described in terms of the number of nodal diameters, n, and nodal circles, m. These node lines are lines at which the vibration pattern of the mode shape has a zero. In-plane radial modes with n nodal diameters and circumferential modes with n nodal diameters also occur.

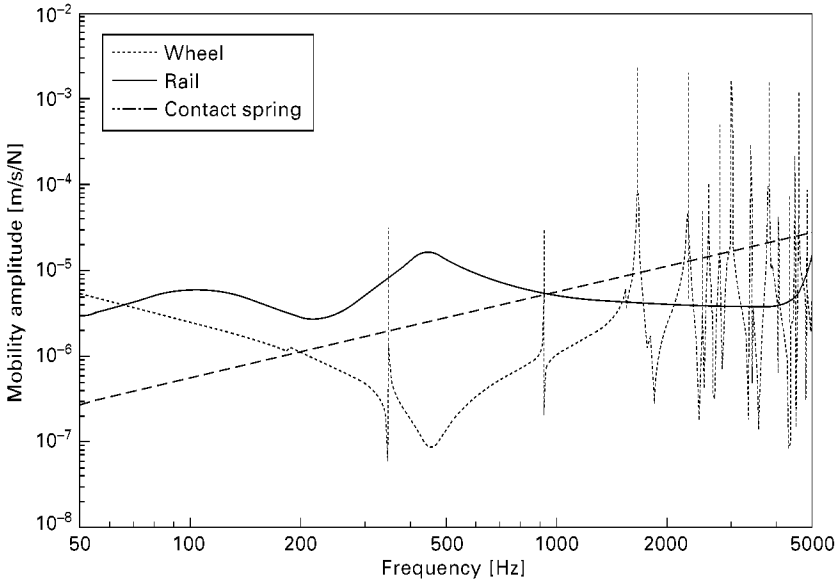
As rolling noise is caused by roughness excitation in the vertical direction, the wheel modes of most relevance have large radial motion at the tread region. As the wheel cross-section is asymmetrical, the radial and out-of-plane (axial) modes of a railway wheel are coupled, so that 1-nodal-circle axial modes are important as well as radial modes. The finite element method can be used quite effectively to calculate the natural frequencies and mode shapes of a railway wheel (Thompson, 1993a).

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 mobility, the vibration velocity due to a unit force as a function of frequency.

Figure 16.3 shows the radial point mobility of a wheelset calculated using the normal modes obtained from a finite element model. At low frequencies, the mobility is inversely proportional to frequency, corresponding to mass-like behaviour. Around 500 Hz, a trough appears and above this frequency the curve rises in stiffness-like behaviour until a series of sharp resonance peaks is reached at around 2 kHz. These peaks are the axial 1-nodal-circle and radial sets of modes.

The dynamic behaviour of track is characterised mainly by waves propagating away from the excitation point rather than resonant modes of vibration. A typical track mobility is also shown in Fig. 16.3. For simplicity, this is predicted using a model based on a continuously supported rail, which neglects the effects of the periodic support. Compared with the wheel, it varies much less as a function of frequency. A broad peak at around 100 Hz corresponds to the whole track vibrating on the stiffness of the ballast. At the second peak, at about 500 Hz, the rail vibrates on the railpad. The frequency of this peak depends on the railpad stiffness, which may vary considerably between tracks. The periodic nature of the support introduces



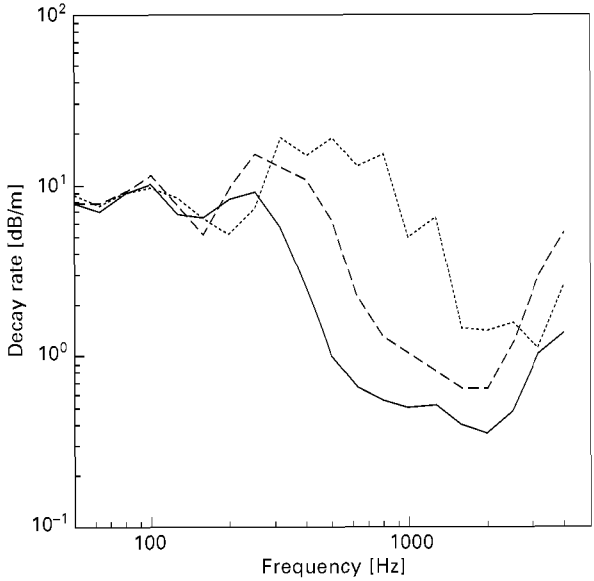
16.3 Vertical mobilities of the wheel–rail system. Radial mobility of UIC 920 mm freight wheel, vertical mobility of track with moderately soft pads and contact spring mobility. (From Thompson and Jones (2006), used with permission from Taylor and Francis Group)

a peak at around 1 kHz between sleepers and a corresponding dip above the sleepers (not present in Fig. 16.3). This is known as the pinned-pinned mode where a half wavelength in the rail matches the sleeper spacing. Although important for corrugation growth, it is of limited importance for noise radiation.

Above the frequency of the rail-on-pad 'resonance', 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, has a very significant effect on the noise radiation from the rail. The total sound power radiated by the rail is inversely proportional to the decay rate Δ (in dB/m):

$$W \sim \int_S v^2 \, \mathrm{d}S \sim \frac{v_R^2}{\Delta} \qquad [16.9]$$

where the integral is over the whole surface area of the (infinite) rail and v_R^2 is the velocity amplitude of the rail at its excitation point (the wheel–rail interface). Figure 16.4 shows measured decay rates of vertical vibration for three different railpads installed in the same track. The results for the



16.4 Measured decay rate of vertical vibration along the track for three different rail pads: —— 140, – – – – 300 and ·····1000 MN/m. (From Thompson *et al.*, 1999, used with permission from Professional Engineering Publishing)

middle value of pad stiffness correspond to the mobility in Fig. 16.3. A peak occurs in the decay rate between 300 and 800 Hz, depending on the pad stiffness. This peak corresponds to the anti-resonance region between the two resonance peaks in Fig. 16.3. 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.

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 non-linear (see Chapter 3), for small dynamic deflections it can be approximated by a linearised stiffness, k_H (Wu and Thompson, 2000). This is shown as a mobility (= $i\omega/k_H$) in Fig. 16.3.

The coupled wheel–rail system is excited by the roughness, which forms a relative displacement input. In this model, the forward motion of the wheel can be ignored and the system 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 velocity amplitude of the wheel (v_W) and rail (v_R) at a particular frequency can be written as:

$$v_W = \frac{Y_W i\omega r}{Y_W + Y_R + Y_C}; v_R = \frac{-Y_R i\omega r}{Y_W + Y_R + Y_C} \qquad [16.10]$$

where r is the roughness amplitude and Y_W, Y_R, Y_C are the vertical mobilities of the wheel, rail and contact spring, respectively. Where the rail mobility has a much larger magnitude than that of the wheel or contact spring, $v_R \approx -i\omega r$, that is the rail is pushed down at the amplitude of the roughness. From Fig. 16.3, this can be expected to occur over the frequency region between about 100 and 1000 Hz. Changing the rail mobility in this frequency region has little effect on the rail vibration at the contact point, although the same changes may affect the decay rates. In practice, coupling also exists in other directions as well as the vertical, notably the lateral direction. This modifies Eq. (16.10) to yield a matrix equation, but the principle remains the same (Thompson, 1993b).

16.3.3 Noise radiation

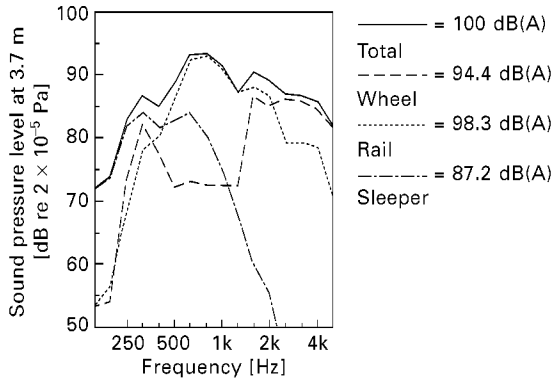
The vibrations of the wheel, rail and sleepers all produce noise. In general, the sound power radiated by a vibrating surface can be expressed using Eq. (16.10). Thus components radiate large amounts of noise if their vibration is large and/or their surface area is large and/or their radiation efficiency is high.

Figure 16.5 shows predictions of the noise from wheels, rails and sleepers (Thompson *et al.*, 1996a). 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) during the passage of a pair of similar bogies. The wheels are the most important source of noise at high frequencies, above about 2 kHz. From Fig. 16.3 it can be seen that this corresponds to the region in which many resonances are excited in the radial direction. Between about 400 and 1600 Hz, the rail is the dominant source of noise. Here the rail vibrates roughly with the amplitude of the roughness and has a relatively low decay rate. 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 the spectra in Fig. 16.5 differ for other wheel and track designs, train speeds and roughness spectra, 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 peak in the noise spectrum shifts towards higher frequencies, leading to a greater importance of the wheel in the total sound level. The shape of the roughness spectrum also affects the balance of wheel and track radiation in the total noise level.

16.4 Reduction of rolling noise

The flowchart of Fig. 16.1 indicates the main areas where mitigation measures may be used to reduce rolling noise. These can be summarised as:



16.5 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 railpads. (From Thompson and Jones, 2006, used with permission from Taylor and Francis Group)

- **reduction of excitation** – smooth wheels and rails;
- **reduction of vibration and sound radiation** – optimised wheel and track designs;
- **reduction of sound transmission** – acoustic shielding by barriers, shrouds, etc.

16.4.1 Smooth wheels

Rolling stock with cast iron brake blocks has long been known to generate higher levels of rolling noise than disc-braked stock, the difference being typically 10 dB (Hemsworth, 1979). The reason for this difference is the development of wheel corrugation with a wavelength between 4 and 8 cm on the running surface of tread braked wheels. This is formed by a combination of hot spot contraction, a thermoelastic instability phenomenon and a wear regime called 'galling' in which material is transferred from the brake blocks to the wheel running surface (Thompson and Gautier, 2006).

The widespread introduction of disc-braked passenger stock from the 1970s onwards has led to significant reductions in rolling noise levels. However, until recently freight vehicles in Europe continued to use exclusively cast iron brake blocks. Reasons for this include their lower cost, difficulties in running trains with a mixture of braking systems and, especially, the fact that rules for interoperability required the use of cast iron brake blocks. This has now changed with the introduction of the technical specifications for interoperability (TSIs) (EC, 2001) which implicitly prevent new vehicles from being fitted with cast iron blocks.

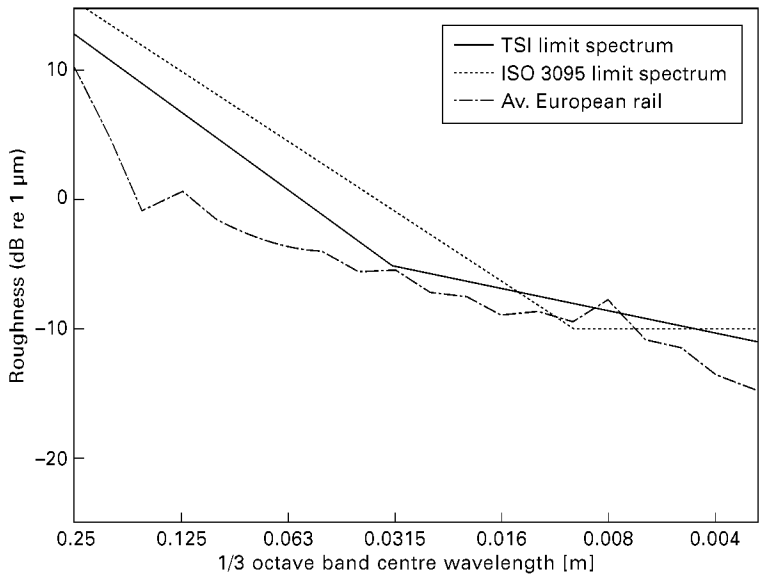
The UIC (International Union of Railways) has promoted the development and testing of alternative brake blocks made from composites (Hübner, 2001; de Vos *et al.*, 2006). These provide a similar wheel roughness and hence noise level as disc brakes, but for a lower cost. A number of so-called K-blocks have become available. However, their main disadvantage is the fact that the thermal loading of the wheel is altered compared with cast iron blocks and therefore simple retrofitting is not possible. Modified wheels and brake cylinders are required. So-called LL-blocks are intended for retrofitting, but their development is still at an earlier stage (Thompson and Gautier, 2006).

16.4.2 Smooth rails

As shown in Fig. 16.3, the wheel–rail system is excited by the *combined* wheel and rail roughness. Therefore it is only effective to reduce wheel roughness if the rail is also smooth. A corrugated rail is a particularly serious form of rail roughness, described in Chapter 11. A corrugated track can be 10 dB noisier than a smooth one for cast-iron block-braked wheels, while for disc-braked wheels the difference can be up to 20 dB (Hemsworth, 1979).

The main means of removing rail corrugations is to grind the railhead. Unfortunately, conventional grinding trains using rotating grinding stones may leave a residual roughness at a wavelength of 1–2 cm (Thompson and Gautier, 2006). Special acoustic grinding procedures have therefore been developed involving oscillating grinding stones which are used at least for the final pass. Grinding of the rail for acoustic purposes is carried out in Germany and The Netherlands to maintain special low-noise sections of track. These sections are allowed a reduced source term in the national noise calculation schemes. They are regularly monitored using an instrumented vehicle and, when their noise level reaches an intervention level, they are ground using acoustic grinding procedures.

The TSIs include limits for the rail roughness spectrum, as shown in Fig. 16.6. Although these roughness limits are intended specifically for qualifying a test track for pass-by measurements, they have also come to be seen more widely as a guide of best practice. In addition, limits for roughness (and also for track decay rates) have been included in the latest revision of the international standard for train exterior noise measurements (ISO, 2005), although there are detail differences between this and both the High-speed and Conventional Rail TSIs (EC, 2001; 2002a).



16.6 Limits for rail roughness applying during vehicle type testing in TSI (EC, 2001; EC, 2002a) and in ISO 3095 (ISO, 2005). Measured 'average European' rail roughness spectrum from Hardy (1997).

16.4.3 Optimised wheel design

Once the surface roughness has been reduced as much as possible, there remains considerable scope for reducing noise by careful design of both the wheels and the track. One means of reducing the amount of noise radiated by the wheels is to increase their damping. However, care is needed as, in effect, a wheel in rolling contact with the rail already experiences considerably more damping 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 and are commercially available to increase the damping of railway wheels and thereby reduce the noise produced (Thompson and Gautier, 2006). These include tuned absorbers, constrained layer damping treatments and friction dampers.

Reductions in the wheel component of radiated noise can also be achieved by careful attention to the wheel cross-sectional shape. In recent years, use has been made of theoretical models such as TWINS (Thompson *et al.*, 1996a,b) to assist in designing wheels for low noise. 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 also beneficial to reduce the wheel diameter. Smaller wheels have higher resonance frequencies, so it is possible to move most of the resonances out of the range of excitation (i.e. above about 5 kHz) (Thompson, 1993a; Thompson and Jones, 2002). The upper frequency limit 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 beneficial for noise. This also negates the increase in unsprung mass caused by increases in thickness. However, if the wheel size is reduced too much, the track noise will actually increase slightly due to the reduction in contact filter effect (Thompson and Jones, 2002).

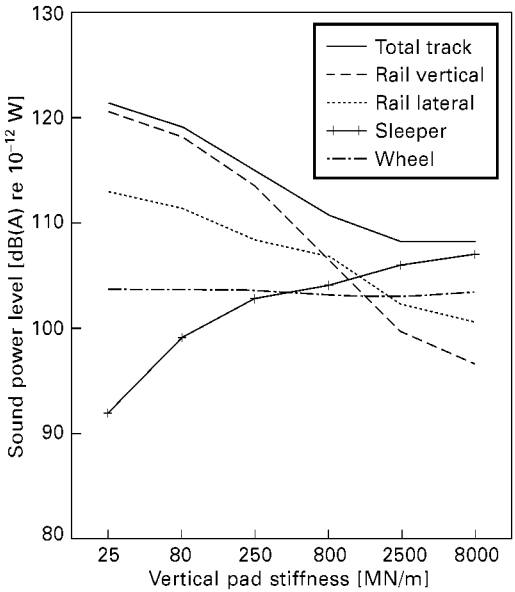
16.4.4 Optimised track design

To achieve significant reductions in overall noise, it is usually important to deal with the track component of noise as well as that produced by the wheel. Two parameters, in particular, affect the noise emission of the track and these are related to one another: the stiffness of the railpad and the decay rate of vibrations along the rail. A stiff railpad causes the rail and sleeper to be coupled together over a wide frequency range whereas a soft pad isolates the rail from the sleeper above a certain frequency. However, as the pad stiffness is reduced, waves can propagate more freely along

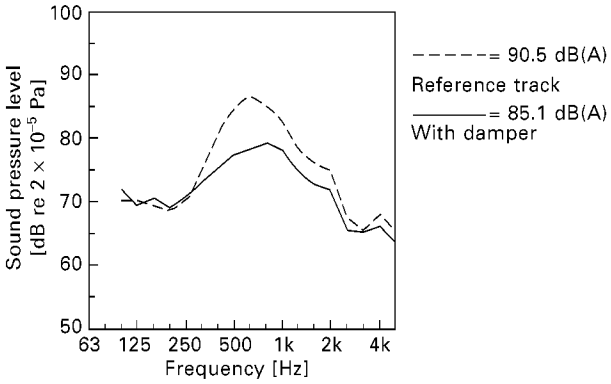
the rail. Figure 16.7 shows predicted noise levels from the rail, sleepers and wheel indicating the dependence on railpad stiffness (Vincent *et al.*, 1996). Although this does not take account of non-linear effects, which are known to be important (Thompson *et al.*, 1999), it does indicate the correct qualitative trends.

The railpad is designed, amongst other things, to protect the sleeper and ballast from large impact forces, for example due to wheelflats. Consequently, soft railpads have become increasingly commonplace in recent years, but this leads to an increase in the noise produced by the track. Although Fig. 16.7 indicates that stiffer pads would be beneficial for noise, and this has been demonstrated in field tests, the optimum for noise radiation is too stiff to be acceptable for other reasons. Stiff pads are also believed to lead to a higher incidence of corrugation growth which, in the long term, has a negative effect on the noise (see Chapter 11).

In order to increase the decay rate of vibrations in the rail without increasing the railpad stiffness, various types of rail damper have been developed. For example, in the Silent Track damper (Thompson *et al.*, 2007), 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. Figure 16.8 shows the noise reduction achieved in field tests (Thompson *et al.*, 2007). Here a low noise wheel was used to minimise the effect of the wheel on the total noise. The



16.7 Effect of railpad stiffness on predicted components of rolling noise. From Thompson *et al.*, 1999 used with permission from Professional Engineering Publishing)



16.8 Measured noise reduction from Silent Track rail damper during the passage of a low noise wheel at 100 km/h. (From Thompson *et al.*, 2007, used with permission from Elsevier)

overall reduction in the track component of noise was approximately 6 dB. In other situations, where the track in the starting point situation has a higher decay rate due to the use of stiffer pads, the effect will be less than here.

16.4.5 Shielding measures

From Fig. 16.1, the final area where noise reduction can be sought is in restricting the sound propagation. Local shielding measures may be applied in the region of the wheels and rails in the form of shrouds around the bogies and low barriers close to the rails. In Jones (1994) noise reductions of 8–10 dB were achieved by a well-designed enclosure around the bogie in combination with low trackside barriers. Such levels of attenuation rely on careful design involving the overlapping of the vehicle- and track-mounted devices to avoid any direct line-of-sight between the sources and the wayside receiver. The existence of different gauging requirements for both structures and vehicles on different railway systems makes the general adoption of such systems more difficult, although shrouds have been successfully used on tramway systems. Access for maintenance and the ventilation of braking systems are also issues requiring careful attention.

Conventional noise barriers built alongside the track, as also used for roads, are widely used in many countries. These can achieve noise reductions of 10–20 dB depending on their height and the terrain, but they are quite an expensive solution and are also visually intrusive. By using cost–benefit analysis it has been shown that reduction of noise at source, as discussed above, can be more cost-effective than the exclusive use of noise barriers (Oertli, 2003).

16.5 Impact noise

As well as random roughness on the railhead and wheel tread, larger discrete features on the wheel or rail running surface also cause noise. Examples include rail joints, dipped welds, wheelflats and gaps at switches and crossings. The mechanism of noise generation is similar to that due to roughness, where a vertical displacement is introduced between the wheel and rail. However, the features are larger and localised and, consequently, the noise is more impulsive in nature. In some cases, loss of contact can occur between the wheel and rail followed by large impact forces. The noise produced by such discrete features is often referred to as impact noise.

Wheelflats, areas of the wheel tread that have been worn flat, are usually attributable to locking of the brakes due to poor wheel–rail adhesion (see Chapter 17). Wheels with flats produce impact loading of the track which can lead to damage of track components. They also produce high levels of noise. Similarly, rail joints and dipped welds provide discrete inputs to the wheel–rail system that induce quite large contact force variations and noise.

In order to predict impact forces and noise it is not sufficient to use the linearised contact spring, as used for rolling noise. Instead, a full non-linear contact stiffness must be included, for example using a Hertzian deflection model (Chapter 3). Time-domain models incorporating the non-linearities in the contact zone have been used, for example, in Clark *et al.* (1982) and Nielsen and Igeland (1995). These models are large, yet they are limited to a maximum frequency of around 1500 Hz. In Wu and Thompson (2002) an alternative approach was adopted to model impact noise up to around 5 kHz, using simplified models of the wheel and rail in a time-stepping approach. This allowed the effects of the non-linearities to be determined. The results were then used to determine an equivalent roughness that could be used in the TWINS model to predict the noise radiation. Similar models have also been used for noise from rail joints (Wu and Thompson, 2003).

Mitigation measures for impact noise should clearly start with attempts to eliminate the troublesome discontinuities. While this is not always possible for operational reasons, wheelflats in particular should be avoided if possible using effective wheel-slide protection systems. Any flats occurring should be detected as quickly as possible using monitoring systems, allowing the vehicles to be removed from service for re-profiling.

Jointed track has been mostly replaced by continuously welded rail on main lines in the last 30 years. Inevitably some joints remain at expansion joints, track-circuit insulating joints and switches and crossings. Measures such as swing-nose crossings can be used to minimise impact forces and thus noise at crossings. It should also be ensured that welded rail joints are as level as possible by using rail straightening equipment, as dipped welds can produce as much noise as rail joints.

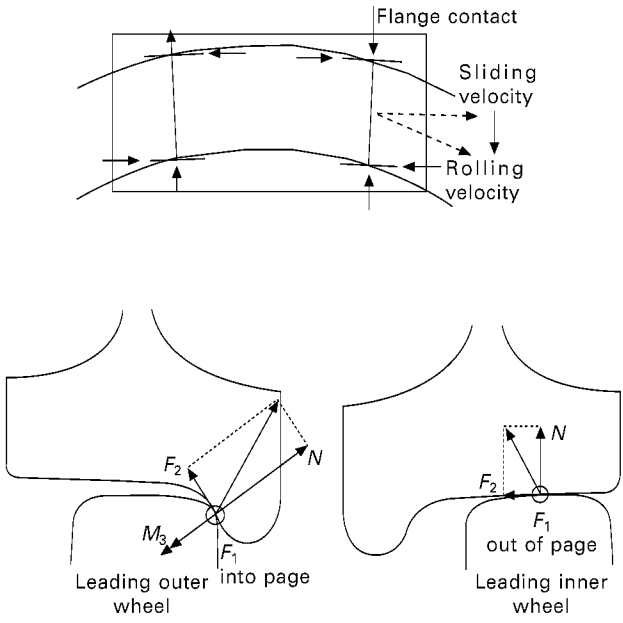
In addition, all counter-measures mentioned for rolling noise, such as wheel damping, wheel shape optimisation, rail damping and local shielding, can be expected to work equally well for impact noise.

16.6 Curve squeal

16.6.1 Mechanism of squeal noise generation

In tight curves, railway vehicles can produce an intense squealing noise. This is a particular problem in urban areas where curved tracks are commonplace. The mechanism of generation of curve squeal is quite different from rolling and impact noise, although both occur at the wheel–rail interface. Instead of a vertical excitation by a relative displacement, squeal is generated by unsteady lateral creep forces.

When traversing a curve, a railway wheelset in a bogie is misaligned relative to the tangential direction (see Fig. 16.9). This can lead to large creep forces, particularly on the leading wheelset of a bogie which can have a large yaw angle. The leading inner wheel in a bogie experiences a large lateral creep force; the contact with the rail is located towards the field side of the tread. The outer wheel has a similar yaw angle, but is usually in flange



16.9 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. (From Thompson and Jones, 2006, used with permission from Taylor and Francis Group)

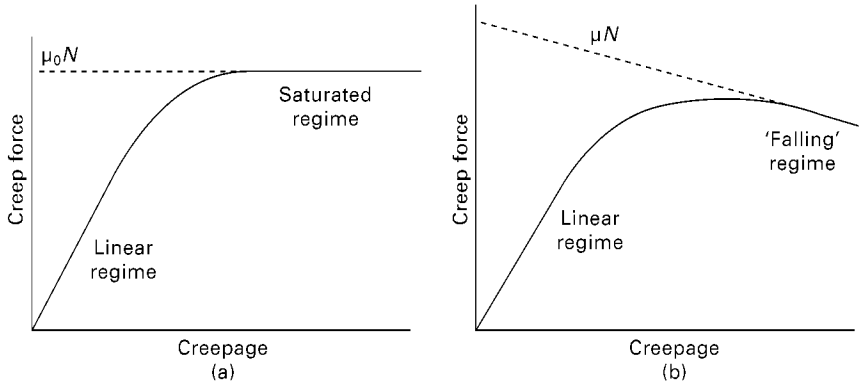
contact. Consequently, the normal force on the flange acts at a significant contact angle to steer the wheelset around the curve.

Creep forces increase as the relative velocity between wheel and rail increases. This relative velocity, normalised by the rolling velocity, is termed the creepage. Initially the dependence of creep force on creepage is linear, as shown in Fig. 16.10a. As creepage increases, more and more of the contact zone passes from adhesion to sliding behaviour. Above a certain value of creepage the creep force saturates, as the whole wheel–rail contact zone is in sliding. The maximum value of the creep force is therefore $\mu_0 N$, where μ_0 is the friction coefficient and N is the normal load.

As the sliding velocity increases, however, the friction coefficient μ tends to fall. Thus, as the creepage increases beyond the saturation point, the creep force once more reduces in amplitude (see Fig. 16.10b). It is this negative slope of creep force with increasing creepage that is believed to be the main reason for the unstable dynamic behaviour leading to squeal noise. The falling creep curve can be thought of as 'negative damping' – the force is proportional to velocity but, as the velocity increases, the force decreases. Since a railway wheel has a very low level of damping, it only takes a small amount of negative damping to destabilise the system, leading to a stick–slip motion and hence squeal noise.

As well as this effect there is also some evidence that squeal can occur even if the friction characteristic does not fall with increasing velocity. Similar phenomena occur in brake squeal (Kinkaid *et al.*, 2003).

Curve squeal excited by unstable lateral creep forces is usually associated with the leading inner wheel of a bogie or two-axle vehicle. The fundamental frequency of this squeal noise corresponds to one of the natural modes of



16.10 A typical creep force–creepage relationship (a) for constant friction coefficient, (b) for velocity-dependent friction coefficient. (From Thompson and Jones, 2006, used with permission from Taylor and Francis Group)

the wheel and is often in the range 200–2000 Hz. The corresponding wheel modes have a mode shape with maximum lateral amplitude at the wheel tread. In addition, a more intermittent higher frequency squeal noise is generated at the leading outer wheel due to contact between the wheel flange and the rail gauge face. This is known as flanging noise. Compared with squeal due to lateral slip, this has received much less attention.

Theoretical models for curve squeal have been developed by various authors, mostly based on the assumption that instability of the lateral friction force is the cause. Rudd (1976) gave a simple model which has been extended by Fingberg (1990) and Périard (1998) by including better models of the wheel dynamics, the friction characteristic and the sound radiation from the wheel in a time-domain approach. Heckl (2000) has also studied squeal using a simplified model and provided experimental validation using a small-scale model wheel. De Beer *et al.* (2003) extended these models, based on excitation by unstable lateral creepage, to include feedback through the vertical force as well as through the lateral velocity. A frequency-domain approach was used to determine instability and to predict which mode is most likely to be excited.

The frequency domain and time domain approaches have been compared by Huang (2007) and by Huang *et al.* (2007). They are found to give consistent results for large creepages but, close to the turning point of the creep force curve, the frequency-domain model may over-estimate the instability compared with the time-domain calculation. The curving behaviour of the vehicle is also included in a unified model in Huang (2007). This model allows for an arbitrary contact angle and includes longitudinal and spin creepage as well as lateral creepage. This means that it can also be used for flange contact.

16.6.2 Reducing squeal noise

Unlike rolling noise, where it is useful and important to achieve noise reductions of a few dB, for curve squeal it is necessary to eliminate the source. Quantified noise reductions therefore tend to say more about the severity of the starting point situation than about the effectiveness of the treatment. It is also important to realise that curve squeal is an extremely unreproducible phenomenon due to its high sensitivity on parameters such as temperature, humidity, train speed, track geometry and wheel and rail profiles. This makes consistent testing prone to considerable difficulties.

Friction management (Chapter 17) is important for the control of squeal noise, whereas it has little or no effect on rolling noise. Lubrication using either grease or water sprays is effective. However, it is important that the friction levels on the wheel and rail running surfaces are not adversely affected as this would lead to loss of adhesion and could compromise safety. Grease is therefore normally applied only to the rail gauge corner or wheel flange.

Water sprays have also been used effectively in a number of locations, e.g. Kootwijk-Damman (1996).

Friction modifiers that act by reducing or eliminating the negative slope of the creep force curve without adversely affecting the level of friction can also be used. These can be applied either to the track or on the vehicle. They have been shown to be very effective in eliminating squeal and can even be applied to the top of the railhead without compromising traction or braking (Eadie *et al.*, 2003).

Squeal noise can also be controlled by adding damping treatments to wheels. Unlike rolling noise where quite large increases in damping are necessary, relatively small increases in damping can be sufficient to eliminate squeal. As well as constrained layer treatments and tuned absorbers, ring dampers have also been used as a simple means of increasing the damping of a wheel (Kirschner *et al.*, 1983; Wetta and Demilly, 1995). Finally, 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.

16.7 Ground vibration and ground-borne noise

16.7.1 Overview of vibration phenomena

As well as air-borne sound, the wheel–rail interface produces vibration that is transmitted through the track into the ground. This leads to several different but related phenomena that are important in different situations.

Low-frequency ground-borne vibration is generated particularly by heavy axle load freight traffic, which may be travelling at relatively low speeds. Passenger trains may also cause significant levels of vibration, especially electric multiple unit trains with a high unsprung mass. This type of vibration may have large components at frequencies as low as 10 Hz or below. Unlike air-borne sound, the vibration is strongly dependent on the site and is especially associated with soft soil conditions. The vibration may propagate over distances of up to around 100 m. It interacts with buildings in the vicinity of the track inducing bouncing or rocking motion. Such low-frequency vibration is perceived as 'whole body' feelable vibration, which can be a source of annoyance or sleep disturbance. It is assessed under the principles of ISO 2631 (ISO, 1997). For the levels of vibration produced by trains there is unlikely to be any damage to buildings.

A second phenomenon concerns high-speed passenger trains running on soft ground. For particularly soft soils, it is possible that the train speed exceeds the wavespeed in the ground, in which case very large displacements of the ground can occur. This can affect the track structure as well as leading to high vibration levels propagating in the ground. Although this is rather rare

it has attracted attention in the research community in recent years (Krylov, 1995; Sheng, *et al.*, 1999b; Madshus and Kaynia, 2000).

Thirdly, where trains run in tunnels, vibration transmitted through the ground can produce noise in the buildings above. In this case, the frequency range of interest is somewhat higher than for feelable vibration, covering the low part of the audible range from about 30–200 Hz. Vibration at these frequencies excites bending modes of the floors and walls of buildings which radiate a rumbling noise into the rooms. This 'ground-borne noise' can be dealt with effectively by using vibration isolating track designs. While new railway projects are usually designed with the intention of keeping noise levels below 40 dB(A), there are many properties that are subject to such noise levels and above. For example in London it has been estimated that around 56 000 households are subject to at least this level (Edwards, 1996). These various phenomena are discussed in more detail in the following sections.

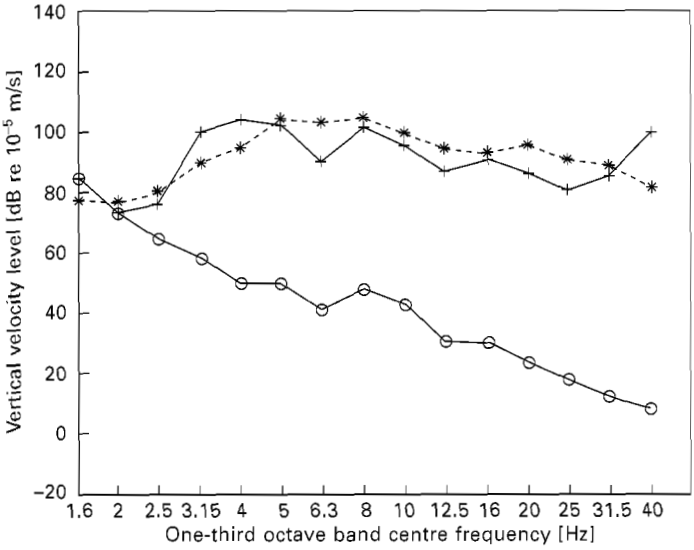
16.7.2 Ground-borne vibration from surface railways

The ground structure can often be defined in terms of a number of layers of homogeneous material. In such a layered ground, vibration propagates parallel to the ground surface via a number of wave types or 'modes'.

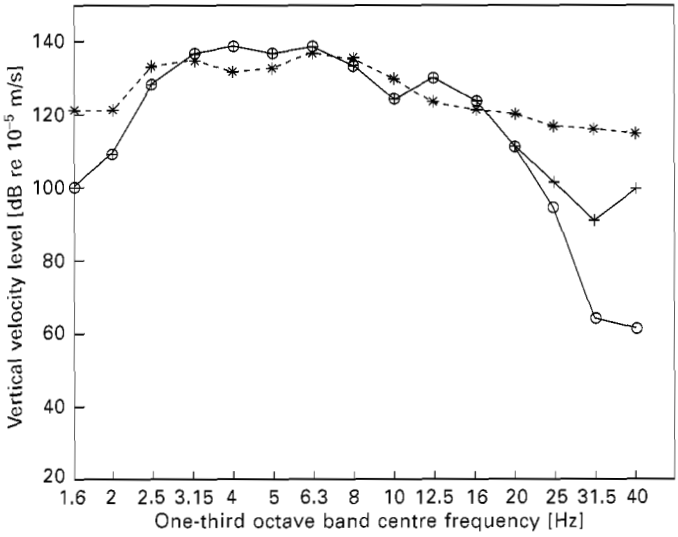
For a moving train, vibration is generated by two main mechanisms: (i) the movement of the quasi-static load along the track on the ground surface; (ii) excitation due to 'roughness' of the wheel and rail surfaces, in the same way as for rolling noise. If the train is moving at a speed that is lower than that of any waves in the ground, the first mechanism produces a large displacement at the track, but this does not propagate significantly in the ground. It is therefore only important very close to the track. However, if the train speed reaches, or exceeds, the wavespeed in the ground, vibration can be produced which does propagate through the ground. A detailed discussion of the critical speed, and the influence of the track parameters on this, is given in Sheng *et al.* (2004b). On the other hand, excitation by wheel or rail irregularities can be important at lower speeds. Models for these phenomena are found in (Sheng *et al.*, 1999a; Jones *et al.*, 2000).

Coupling a simple model for the dynamic behaviour of each vehicle in a train to these ground models allows the ground vibration to be predicted (Sheng *et al.*, 2004a). Comparisons have been made with measurements of vibration from a number of sites (Sheng *et al.*, 2003). Figures 16.11 and 16.12 show examples for a Swedish high-speed train at a site with very soft soil.

Figure 16.11 shows the measured and predicted one-third octave vibration spectra for a train speed of 70 km/h at a point 7.5 m from the track. Both the vibration due to the quasi-static moving loads and that due to the roughness



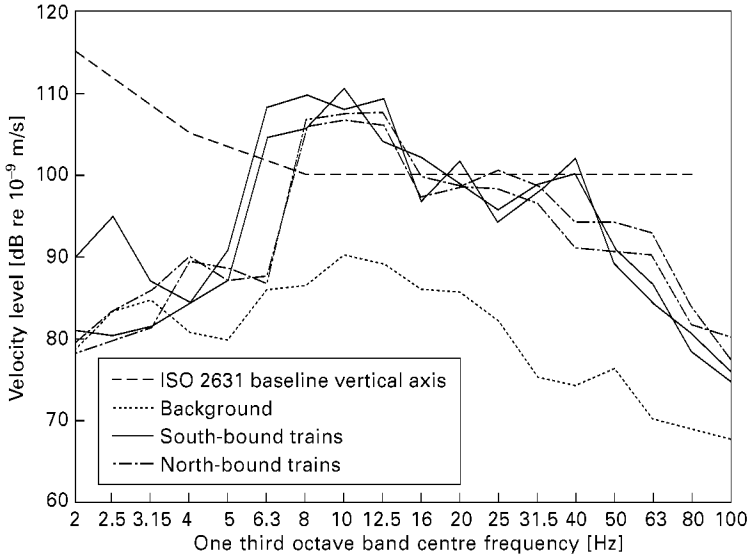
16.11 Vertical velocity level at Ledsgård for train speed of 70 km/h: o, predicted level due to quasi-static loads; + predicted total level; * measured level at 7.5 m. (From Sheng *et al.*, 2003, used with permission from Elsevier)



16.12 Vertical velocity level at Ledsgård for train speed of 200 km/h: o, predicted level due to quasi-static loads; + predicted total level; * measured level at 7.5 m. (From Sheng *et al.*, 2003, used with permission from Elsevier)

excitation are shown. Despite limitations in the input data, the results show the clear dominance of the roughness excitation for this speed. At the higher speed of 200 km/h, the results are as shown in Fig. 16.12. Here, the train speed is close to the 'critical' speed of the ground and the quasi-static excitation mechanism produces large vibration. This is about 35 dB higher than for the lower speed and covers frequencies up to about 20 Hz. This site is unusual in having such a low ground wave speed, and the phenomenon seen here is therefore not common. Even so, where high-speed lines, designed for speeds of 300 km/h and above, pass over soft ground the critical speed must be taken into account.

More usually, ground-borne vibration is caused by the irregular vertical profile of the track along with out-of-round wheels. Figure 16.13 shows more typical measured spectra of vibration for a particular site compared with the background vibration at that site and the baseline for perception of vibration in the spinal axis for humans from ISO 2631 (ISO, 1997). Spectra vary a lot between sites with different soil conditions and between different types of rolling stock. However, the example is typical in that vibration components between about 4 and 50 Hz are often the most significant. Levels going only a few dB above the nominal threshold of perception (corresponding to a velocity of about 0.1 mm/s above 8 Hz) are likely to give rise to complaints. Vibration at such low frequencies cannot easily be treated except by significant ground engineering works. However, since its source is the irregular vertical profile of the track, tamping maintenance can often help.

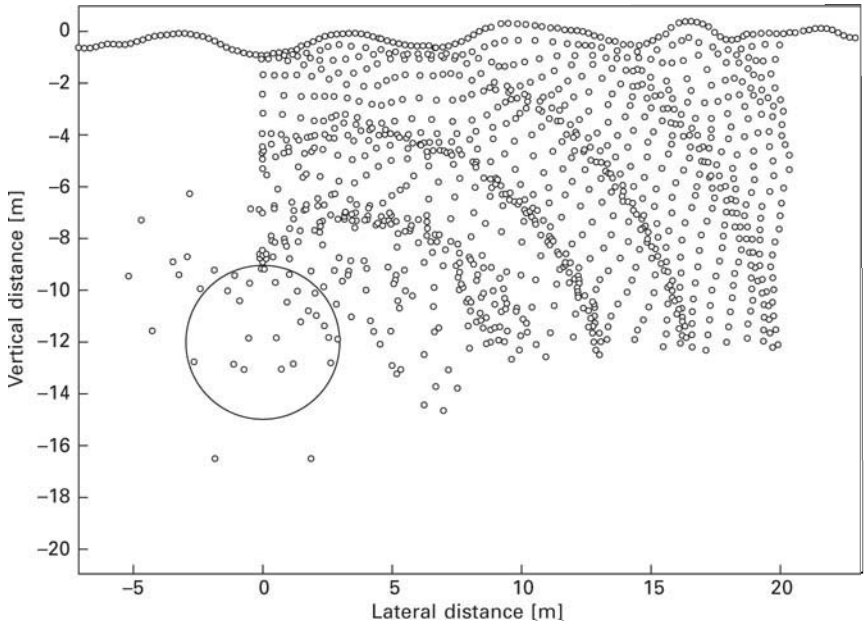


16.13 Spectra of trains measured at a house 50 m from the track compared to ISO 2631 baseline for vertical axis perception.

16.7.3 Ground-borne noise from trains in tunnels

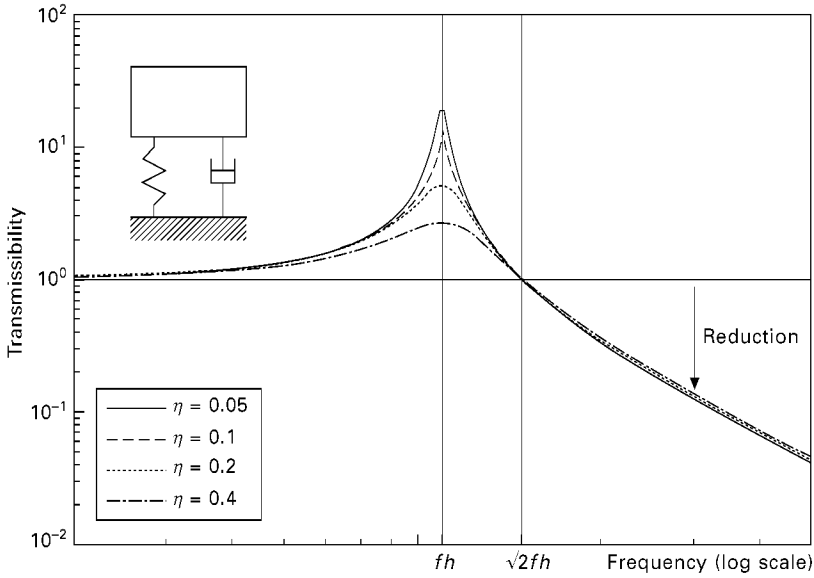
In order to predict the ground-borne noise from trains in tunnels, it is important to include the dynamic effects of the tunnel structure. Coupled boundary element and finite element models have been developed (Jones *et al.*, 2002) which allow these effects to be studied. Figure 16.14 shows some example results for a circular bore tunnel in a clay soil. Boundary elements are used for the ground surface and the ground–tunnel interface, and finite elements are used for the tunnel structure itself. Exaggerated instantaneous particle displacements are shown for a point force at 100 Hz applied at the base of the track. These results demonstrate that a wave pattern radiates away from the tunnel. At this frequency, a 'shadow zone' occurs above the tunnel so that the greatest amplitudes of vibration on the ground surface occur at a distance of 15–20 m away from the tunnel alignment.

The effect of the tunnel structure is illustrated in Fig. 16.15. This shows the instantaneous displacement of the tunnel itself for cases of an unlined tunnel and one with a concrete lining. Again the frequency is 100 Hz. Compared with the unlined tunnel, the amplitude of vibration at the crown of the lined tunnel is much higher. As a result the vibration level directly above the tunnel is strongly affected by the tunnel structure.



16.14 The vibration field around the unlined tunnel at 100 Hz. (From Jones *et al.*, 2002, used with permission from Thomas Telford Publishing)

16.15 Amplified representation of the response showing waves round the unlined (left) and lined (right) tunnel rings at 100 Hz (From Jones *et al.*, 2002, used with permission from Thomas Telford Publishing)



16.16 The force transmissibility of a hysteretically damped single degree of freedom system. (From Thompson and Jones 2006, used with permission from Taylor and Francis Group)

Ground-borne noise from underground railways can be controlled by using resilient elements in the track structure to provide vibration isolation. Figure 16.16 demonstrates the principle of isolation using the example of a single-degree-of-freedom (mass-spring) system. If a harmonic force is applied to the mass, the ratio of the force transmitted to the foundation to that applied to the mass is termed the transmissibility. The transmissibility has a peak at the resonance frequency, f_n, of the mass-spring system at which the force is amplified. Above $\sqrt{2}\, f_n$, the transmissibility drops below unity and continues to fall with increasing frequency. Figure 16.16 also shows the effect of the

damping in the support. Here a hysteretic damping model is used, defined by a constant damping loss factor η. The amplitude at the resonance can be controlled by increasing the damping, but the degree of vibration isolation at higher frequencies is not affected (whereas for the more commonly used viscous damping model it is).

The principle shown in Fig. 16.16 illustrates the reason why vibration isolating trackforms can reduce ground-borne noise. To be effective the resonance frequency of the track system must be sufficiently low that the resonance peak does not occur in the audible frequency range. This means relatively soft fastening systems, if possible in combination with a large mass. However, there are practical constraints which prevent the support stiffness from being too small, in particular the allowable vertical and lateral static displacements under the weight of the train. Consequently, vibration isolating tracks are unlikely to be useful in reducing low-frequency ground vibration problems as described in the previous section.

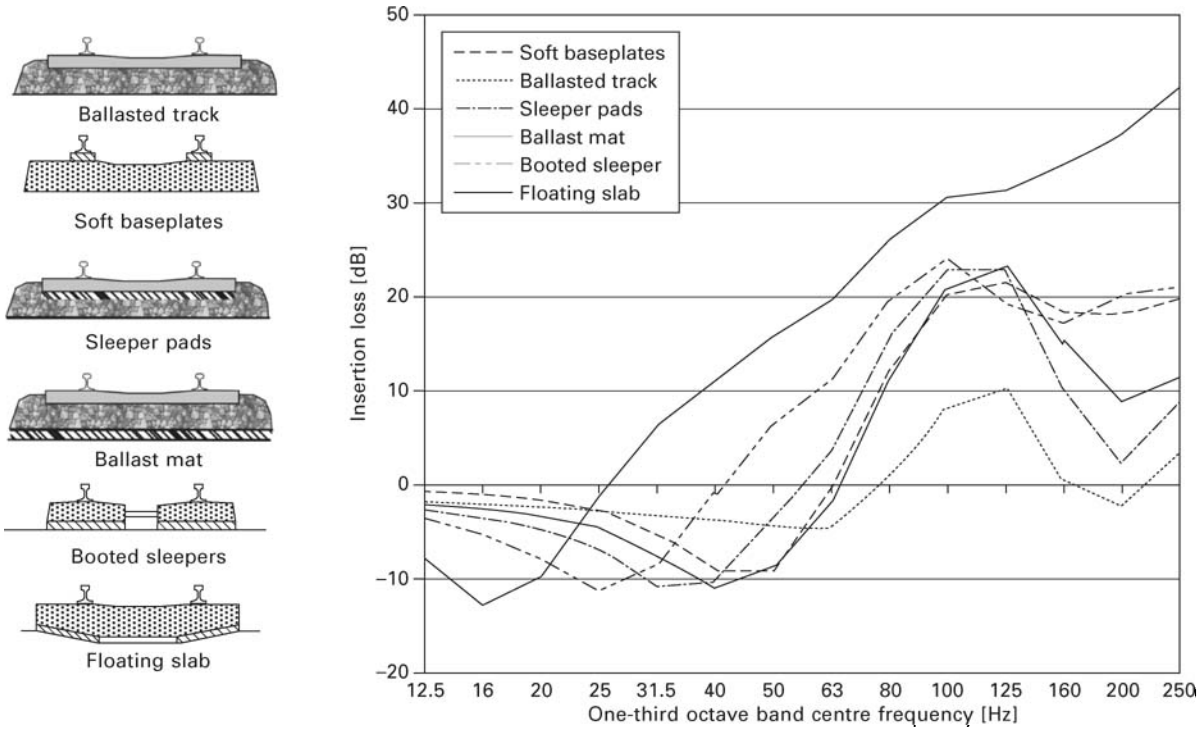
A number of basic design concepts for vibration isolating track are shown in Fig. 16.17. The railpad is not shown but may have some influence at the higher end of the frequency range of interest. The main resilient component is usually the ballast layer. Conventional slab tracks therefore lead to increased vibration compared with ballasted track. This can be rectified by using soft baseplates, sleeper soffit pads, ballast mats or a floating slab design. A limiting factor with soft baseplates is the need to prevent excessive rotation or lateral deflection of the rail.

In order to assess the relative performance of different track designs, it is often sufficient to calculate a vibration level difference between a proposed track and a reference situation. A relatively simple calculation model can be used for this, for example one that represents the ground simply as an elastic half-space (Jones, 1996). Care should be taken in that the vibration level difference at the tunnel invert is often greater than at some distance away from the track.

Example results are shown in Fig. 16.17 for various track forms relative to a slab track with a fairly stiff support. The degree of vibration isolation achieved (positive insertion loss) increases, generally in the order in which the track types are shown in Fig. 16.17 from top to bottom. The cost of such tracks also increases in an approximately corresponding order. However, note that in each case there are frequencies where the insertion loss is negative implying a worsening of the situation compared with the reference case.

16.8 Conclusions and future trends

It has been shown that a number of techniques exist that allow rolling noise from trains to be reduced substantially. The elimination of cast iron brake



16.17 Illustration of the insertion loss at 25 m for different resilient track types as estimated by a model of the track on an infinite homogeneous soil.

blocks from freight trains offers the greatest single benefit, with potential reductions of up to 10 dB, and would bring freight vehicles into line with modern passenger stock. Local acoustic shielding can also be effective, although there are considerable practical constraints. Apart from this, to achieve more than about 5 dB reduction in overall noise, combinations of measures would be required. An appropriate combination of measures applied to the wheel and track design could deliver reductions in excess of 10 dB, depending on the initial situation (Thompson and Gautier, 2006). Such measures, applied at source, are generally more cost-effective than secondary measures, such as noise barriers. However, their introduction requires coordinated action which may prove difficult where responsibility for vehicles and track is segregated.

Although various research projects have demonstrated both the noise reduction and the potential for implementation of prototype designs, the railway industry is cautious about introducing such new technology. Clearly, it is unrealistic to expect large new construction projects to adopt new technology, but a way has to be found to develop such technology from the prototype stage to the point where it is accepted as proven technology and no longer perceived as a source of unacceptable risk.

The development of legislation, particularly the TSIs (EC, 2002a; EC, 2001), will no doubt provide an increasing stimulus to consider noise-reducing measures as part of the procurement of new vehicles and track. Many of the techniques described here are therefore likely to become more commonplace in the future.

Impact noise is an extreme but localised form of rolling noise that can usually be dealt with by elimination of the defects. Curve squeal is a quite different type of noise caused by unsteady lateral creepage at the wheel–rail interface. Mitigation measures include friction modification as well as wheel damping treatments. The aim is to eliminate the noise rather than reduce its level. Ground-borne vibration can lead to low-frequency feelable vibration or higher frequency 'ground-borne noise' in buildings close to railways. In most cases roughness excitation dominates, but in some cases the movement of the axle loads may lead to high levels of vibration at sites where the train speed approaches the wavespeed in the ground.

'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 depends strongly on the tunnel structure. To reduce ground-borne noise, various vibration isolating track forms can be used depending on the vehicles and the noise reduction required at particular locations. To reduce ground vibration at lower frequencies, it is not sufficient to use isolating track forms. Instead, soil improvement, ground trenches or walls may be considered.

16.9 Sources of further information and advice

For further information it is recommended that the reader studies the recent book on railway noise and vibration (Thompson, 2008). Some additional material is also contained in (Thompson, 2001; Jones and Thompson, 2001; Thompson and Jones, 2006; Krüger *et al.*, 2006).

16.10 References

Clark, R.A., Dean, P.A., Elkins, J.A. and Newton, S.G. (1982), An investigation into the dynamic effects of railway vehicles running on corrugated rails, *Journal of Mechanical Engineering Science*, **24**, 65–76.

De Beer, F.G., Janssens, M.H.A. and Kooijman, P.P. (2003), Squeal of rail bound vehicles influenced by lateral contact position, *Journal of Sound and Vibration*, **267**, 497–507.

De Vos, P.H., Bergendorff, M., Brennan, M. and van der Zijpp, F. (2006), Implementing the retrofitting plan for the European rail freight fleet, *Journal of Sound and Vibration*, **293**, 1051–57.

Eadie, D.T., Santoro, M. and Powell, W. (2003), Local control of noise and vibration with Keltrack friction modifier and Protector trackside application: an integrated solution, *Journal of Sound and Vibration*, **267**, 761–72.

EC (2001), Directive 2001/16/EC of the European Parliament and of the Council of 19 March 2001 on the interoperability of the trans-European conventional rail system, *OJ*, **L110**, 2.4.2001, 1–27.

EC (2002a), 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, *Official Journal of the European Communities*, **L245**, 12.9.2002, 402–506.

EC (2002b), Directive 2002/49/EC of the European Parliament and of the Council of 25 June 2002 relating to the assessment and management of environmental noise, *Official Journal of the European Communities*, **L189**, 18.7.2002, 12–25.

Edwards, J.W. (1996), Survey of environmental noise and vibration from London Underground trains, *Proceedings Internoise '96*, Liverpool, UK, 30 July–2 August, Institute of Acoustics, St Albans, UK, 2029–2032.

Fahy, F. and Walker, J. (1998), *Fundamentals of Noise and Vibration*, E&FN Spon, London, UK.

Fingberg, U. (1990), A model of wheel-rail squealing noise, *Journal of Sound and Vibration*, **143**, 365–77.

Hardy, A.E.J. (1997), *Draft proposal for noise measurement standard for ERRI committee C163*, Report RR-SPS-97-012 Issue 1, BR Research – published through ERRI, Utrecht, the Netherlands.

Heckl, M.A. (2000), Curve squeal of train wheels, part 2: which wheel modes are prone to squeal? *Journal of Sound and Vibration*, **229**, 695–707.

Hemsworth, B. (1979), Recent developments in wheel/rail noise research, *Journal of Sound and Vibration*, **66**, 297–310.

Huang, Z. (2007), *Theoretical Modelling of Railway Curve Squeal*, PhD Thesis, University of Southampton, Southampton, UK.

Huang, Z.Y., Thompson, D.J. and Jones, C.J.C. (2007), Squeal prediction for a bogied

vehicle in a curve, in Schulte-Werning, B., Thompson, D., Gautier, P., Hanson, C., Hemsworth, B., Nelson, J., Maeda, T. and de Vos Pl. (eds) (2008), *Noise and Vibration Mitigation for Rail Transportation Systems: Proceedings of the 9th International Workshop on Railway Noise, Munich, Germany, 4–8 September 2007*, Berlin, Germany, Springer, 313–19.

Hübner, P. (2001), The Action Programme of UIC, UIP and CER "Abatement of Railway Noise Emissions on Goods Trains", *Proceedings Internoise 2001*, The Hague, Netherlands, 27–30 August, 107–112.

ISO (1997), ISO2631: 1997 *Evaluation of human exposure to whole body vibration, Part 1 General requirements*, International Standards Organisation, Geneva, Switzerland.

ISO (2005), ISO3095: 2005 *Railway applications – Acoustics – Measurements of noise emitted by railbound vehicles*, International Standards Organisation, Geneva, Switzerland.

Jones, R.R.K. (1994), Bogie shrouds and low barriers could significantly reduce wheel/rail noise, *Railway Gazette International*, July, 459–62.

Jones, C.J.C. (1996), Groundborne noise from new railway tunnels, *Proceedings Internoise '96*, Liverpool, UK, 30 July–2 August, Institute of Acoustics, St Albans, UK 421–426.

Jones, C.J.C. and Thompson, D.J. (2001), Means of controlling rolling noise at source, in Krylov, V.V. (ed.), *Noise and Vibration from High-Speed Trains*, Thomas Telford, London, 163–83.

Jones, C.J.C., Sheng, X. and Petyt, M. (2000), Simulations of ground vibration from a moving harmonic load on a railway track, *Journal of Sound and Vibration*, **231**, 739–51.

Jones, C.J.C., Thompson, D.J. and Petyt, M. (2002), A model for ground vibration from railway tunnels, *Proceedings of the Institution of Civil Engineers. Transport*, **153**, 121–29.

Kinkaid, N.M., O'Reilly, O.M. and Papadopoulos, P. (2003), Automotive disc brake squeal, *Journal of Sound and Vibration*, **267**, 105–66.

Kinsler, L.E., Frey, A.R., Coppens, A. and Sanders, J.V. (1982), *Fundamentals of Acoustics*, Wiley, New York.

Kirschner, F., Koch, J.E. and Cohen, H.C. (1983), Light weight vibration damping treatments for railroad wheels, *Proceedings 11th International Congress Acoustics*, Paris, France, 19–27 July, **5**, 129–32.

Kootwijk-Damman, C.M. (1996), Industrial noise in railway yards; the problem of squeal noise, *Journal of Sound and Vibration*, **193**, 451–2.

Krüger, F., Hecht, M., Hölzl, G., Jäcker, M., Jaquet, T. and Repczuk, A. (2006), *Schall- und Erschütterungsschutz im Schienenverkehr*, Expert Verlag, Renningen, Gemany.

Krylov, V.V. (1995), Generation of ground vibration by superfast trains, *Applied Acoustics*, **44**, 149–64.

Madshus, C. and Kaynia, A.M. (2000), High speed railway lines on soft ground: dynamic behaviour at critical train speed, *Journal of Sound and Vibration*, **231**, 689–701.

Nielsen, J.C.O. and Igeland, A. (1995), Vertical dynamic interaction between train and track – influence of wheel and track imperfections, *Journal of Sound and Vibration*, **187**, 825–39.

Oertli, J. (2003), The STAIRRS project, work package 1: a cost-effectiveness analysis of railway noise reduction on a European scale, *Journal of Sound and Vibration*, **267**, 431–37.

Périard, F. (1998), *Wheel–Rail Noise Generation: Curve Squealing by Trams*, PhD Thesis, Technical University of Delft, Delft, the Netherlands.

Remington, P.J. and Webb, J. (1996), Estimation of wheel/rail interaction forces in the contact area due to roughness, *Journal of Sound and Vibration*, **193**, 83–102.

Rudd, M.J. (1976), Wheel/rail noise – part II: wheel squeal, *Journal of Sound and Vibration*, **46**, 381–94.

Sheng, X., Jones, C.J.C. and Petyt, M. (1999a), Ground vibration generated by a harmonic load acting on a railway track, *Journal of Sound and Vibration*, **225**, 3–28.

Sheng, X., Jones, C.J.C. and Petyt, M. (1999b), Ground vibration generated by a load moving along a railway track, *Journal of Sound and Vibration*, **228**, 129–56.

Sheng, X., Jones, C.J.C. and Thompson, D.J. (2003), A comparison of a theoretical model for vibration from trains with measurements, *Journal of Sound and Vibration*, **267**, 621–35.

Sheng, X., Jones, C.J.C. and Thompson, D.J. (2004a), A theoretical model for ground vibration from trains generated by vertical track irregularities, *Journal of Sound and Vibration*, **272**, 937–65.

Sheng, X., Jones, C.J.C. and Thompson, D.J. (2004b), A theoretical study on the influence of the track on train-induced ground vibration, *Journal of Sound and Vibration*, **272**, 909–36.

Thompson, D.J. (1993a), Wheel/rail noise generation, part II: wheel vibration, *Journal of Sound and Vibration*, **161**, 401–19.

Thompson, D.J. (1993b), Wheel/rail noise generation, part IV: contact zone and results, *Journal of Sound and Vibration*, **161**, 447–66.

Thompson, D.J. (2001), Theory of generation of wheel/rail rolling noise, in Krylov, V.V. (ed.), *Noise and Vibration from High-Speed Trains*, Thomas Telford, London, UK, 3–26.

Thompson, D.J. (2003), The influence of the contact zone on the excitation of wheel/rail noise, *Journal of Sound and Vibration*, **267**, 523–35.

Thompson, D.J. and Gautier, P-E. (2006), A review of research into wheel/rail rolling noise reduction, *Proceedings of the IMechE, Part F: Journal of Rail and Rapid Transit*, **220**, (F4), 385–408.

Thompson, D.J. and Jones, C.J.C. (2002), A study of the use of vehicles with small wheels for determining the component of noise from the track, *Proceedings IOA Spring Conference*, Salford, UK, 25–27 March.

Thompson, D.J. and Jones, C.J.C. (2006), Noise and vibration from railway vehicles, in Iwnicki, S. (ed.), *Handbook of Railway Vehicle Dynamics*, CRC, Boca Raton, FL, USA, 279–325.

Thompson, D.J., Fodiman, P. and Mahé, H. (1996a), Experimental validation of the TWINS prediction program for rolling noise, part 2: results, *Journal of Sound and Vibration*, **193**, 137–47.

Thompson, D.J., Hemsworth, B. and Vincent, N. (1996b), Experimental validation of the TWINS prediction program for rolling noise, part 1: description of the model and method, *Journal of Sound and Vibration*, **193**, 123–35.

Thompson, D.J., Jones, C.J.C., Wu, T.X. and de France, G. (1999), The influence of the non-linear stiffness behaviour of railpads on the track component of rolling noise, *Proceedings of the IMechE, Part F: Journal of Rail and Rapid Transit*, **213**, 233–41.

Thompson, D.J., Jones, C.J.C., Waters, T.P. and Farrington, D. (2007), A tuned damping device for reducing noise from railway track, *Applied Acoustics*, **68**, 43–57.

Thompson, D.J. (2008), *Railway Noise and Vibration: Mechanisms, Modelling and Means of Control*, Elsevier, Oxford, UK.

Vincent, N., Bouvet, P., Thompson, D.J. and Gautier, P.E. (1996), Theoretical optimization of track components to reduce rolling noise, *Journal of Sound and Vibration*, **193**, 161–71.

Wetta, P. and Demilly, F. (1995), Reduction of wheel squeal noise generated on curves or during braking, *Proceedings 11th International Wheelset Congress*, Paris, France, 301–6.

Wu, T.X. and Thompson, D.J. (2000), Theoretical investigation of wheel/rail non-linear interaction due to roughness excitation, *Vehicle System Dynamics*, **34**, 261–82.

Wu, T.X. and Thompson, D.J. (2002), A hybrid model for the noise generation due to railway wheel flats, *Journal of Sound and Vibration*, **251**, 115–39.

Wu, T.X. and Thompson, D.J. (2003), On the impact noise generation due to a wheel passing over rail joints, *Journal of Sound and Vibration*, **267**, 485–96.

Copyrighted Materials
Copyright © 2009 Woodhead Publishing Retrieved from www.knovel.com

17

Adhesion and friction modification

U. OLOFSSON, Royal Institute of Technology (KTH), Sweden

Abstract: Trains operate within the limitations imposed by the friction between the railway wheel and rail surfaces. Inadequate friction causes poor adhesion during braking, which is a safety issue as it leads to extended stopping distances. Inadequate friction is also a performance issue as it affects traction and thus limits the tangential force that can be developed in curving. Delays occur if a train passes over areas of poor adhesion while in service. In this chapter are models and methods to analyse and control the adhesion between railway wheel and rail discussed. Furthermore, recent research results are also presented and discussed.

Key words: adhesion, friction modification, railway, wheel, rail.

17.1 Introduction

Trains operate within the limitations imposed by the friction between the railway wheel and rail surfaces. Inadequate friction causes poor adhesion during braking, which is a safety issue as it leads to extended stopping distances. Inadequate friction is also a performance issue as it affects traction and thus limits the tangential force that can be developed in curving. Delays occur if a train passes over areas of poor adhesion while in service.

Because the tiny contact zone (roughly 1 cm^2) where steel wheel meets steel rail is an open system, it is exposed to dirt and particles as well as natural lubrication such as high humidity, rain and leaves. All of these can seriously affect the contact conditions and the friction forces transmitted through the contact.

Acceleration and braking usually require a coefficient of friction (the ratio of the tangential load to the normal load) of about 0.2. However, modern power cars and locomotives demand a higher friction coefficient. The friction between the wheels and rail also plays a major role in other wheel–rail interface processes such as rolling contact fatigue (RCF), rail corrugation and noise generation. If the coefficient of friction is too high, most types of surface damage occur more frequently. Friction coefficients above 0.4 increase the chance of surface fatigue of wheels and rail (see Chapter 7). A decreasing coefficient of friction with increasing sliding velocity can cause squeal and rail corrugation (see Chapters 16 and 11, respectively). Measurements taken on test tracks have shown that proper management of wheel–rail adhesion can result in fuel savings of around 30 % (compared to consumption in dry conditions) (Reiff and Creggor, 1999).

The Swedish National Rail Administration has estimated the annual cost of poor rail–wheel adhesion in European as approximately 1 billion Kronor (€90 million) (Berggren, 2003). In the UK and Sweden, where the problem is reported on in depth, more detailed analysis shows an annual cost of £50 million in the UK (Adhesion working group, 2001) and 100 million SEK in Sweden (Forslöv, 1996).

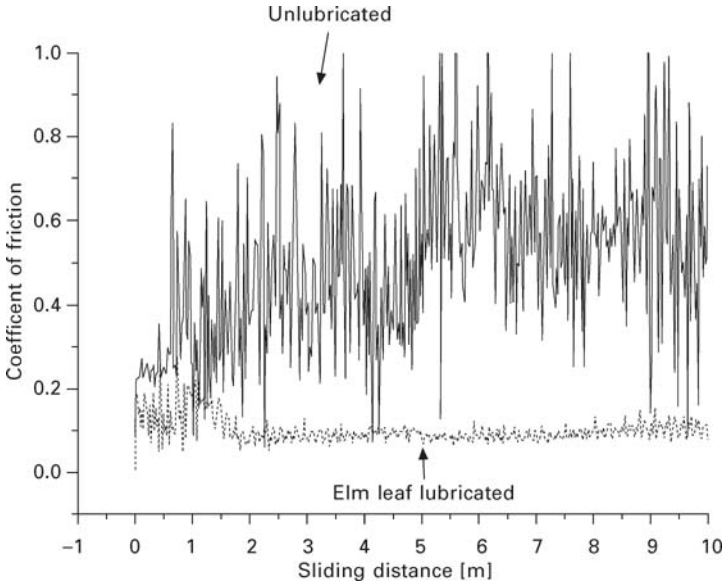
A great deal of research on adhesion loss was carried out in the UK during the 1970s using both laboratory and field tests (Collins and Pritchard, 1972; Broster *et al.*, 1974; Beagley and Pritchard, 1975; Beagley, *et al.*, 1975a, b). These studies identified factors that contribute to adhesion loss, such as slippery coatings formed by leaves, humidity and oil and wear particles. Recent adhesion problem analyses, together with metrological data from the UK, reveal that approximately one third of the overruns due to adhesion problems occurred when no rail contamination had been reported. However, rail contamination was reported in 43 % of the overruns (Veal, 2008).

Rail services all over the world are disrupted by fallen leaves. It has been estimated that a mature tree has between 10 000 and 50 000 leaves, and that thousands of tonnes of leaves fall onto railway lines every year. The leaves are usually swept onto the track by the slipstream of passing trains. The photograph in Fig. 17.1 shows the effect of a passing train on leaves. The wheels crush the leaves, which then form a hard, black,

17.1 Leaf residue and blackish layer on top of rail, from the Brommaplan test track, Stockholm, Sweden. (Photograph taken by Ulf Olofsson KTH)

slippery layer that coats the rail and is very difficult to remove (Olofsson and Sundvall, 2004). Personal experience has shown that it is impossible to keep one's balance when walking on a rail with such a coating, even when wearing shoes with rubber soles. This shoe test implies that the coefficient of friction is below 0.1. The experience is similar to walking on smooth wet ice in winter.

A recent literature review from the Rail Safety and Standards Board in the UK (Fulford, 2004) reveals that, although much field work has been done on low railway wheel–rail adhesion, there is a lack of controlled laboratory experiments. Recently, however, Olofsson and Sundvall (2004), Olofsson (2007), Gallardo-Hernandez and Lewis (2008) and Cann (2006) presented results from controlled experiments. In Olofsson and Sundvall (2004) a pin-on-disc machine with controlled temperature and humidity was used to study how the environment affects the friction between railway wheel and rail steel under leaf-lubricated conditions. The results show that the coefficient of friction decreases when the relative humidity increases, and decreases even more when the leaves act as a lubricant. When leaves act as a lubricant, the coefficient of friction reaches the same levels as when using ordinary lubricating oils (Fig. 17.2). Olofsson (2007) analysed the surface layers formed between the blackish leaf layer and the steel surface and found a chemically reacted, easily sheared surface layer containing elements



17.2 Measured coefficient of friction versus sliding distance for leaf-lubricated and unlubricated tests. (From Olofsson and Sundvall, 2004)

such as P and Ca. Gallardo-Hernandez and Lewis (2008) used a twin-disc machine to study the effect of leaves acting as a lubricant, and the use of sand to remove the blackish layer formed by the crushed leaves.

Cann (2006) studied the problem of leaf residue and loss of adhesion in the wheel–rail contact using a ball-on-disc test device. Analysis of the black film formed by the leaves at the contact showed that it was due to a chemical reaction between water-soluble components of the leaves and the steel surface. Pectin was found to be a major contributor to the formation of the thin but highly viscous layer on the rail surface. The Swedish National Rail Administration in their leaf lubrication study (Forslöv, 1996) chemically analysed the top of rail layer on sections with a known history of adhesion problems. By scraping away the content of the top of the rail and subjecting it to chemical analysis, they found that this third-body layer contained iron particles, applied lubricants in the form of oil and grease, and natural contaminants such as leaves and sand. Poole (2006) used a full-scale wheel-on-rail rig to produce the blackish leaf layer and examine its shear stress properties. The shear stress was found to be inversely proportional to the humidity level. Analysis of the chemical content of the blackish layer revealed that lignin, cellulose and pectin were the major sources of the adhesion of leaves to the rail.

17.2 The coefficient of friction, adhesion and braking distance

Friction force can be defined as the resistance encountered by one body moving over another body. This definition covers both sliding and rolling bodies, for even pure rolling nearly always involves some sliding, and thus the two classes of motion are not mutually exclusive. Friction force operates parallel to the direction of motion. If solid bodies are in contact, the static friction force is equal to the tangential force required to initiate sliding between them. The kinetic friction force is the tangential force required to maintain sliding. Kinetic friction is generally lower than static friction.

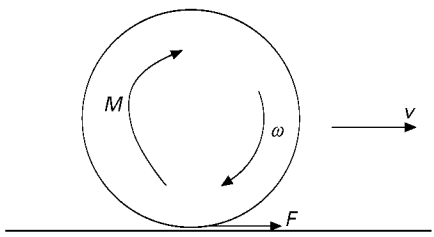
For sliding bodies the friction force, and thus the coefficient of friction (friction force divided by normal force), depends on three different mechanisms in dry and mixed lubricated conditions: deformation of asperities, adhesion of the sliding surfaces and ploughing caused by deterioration particles and hard asperities (Suh and Sin, 1981). For most metal pairs, the maximum value of the coefficient of friction ranges from 0.3 to 1.0 (Czichos, 1992). The ploughing component of the coefficient varies from 0 to 1.0 and the adhesion component from 0 to 0.4 (Suh, 1986) .

It is generally recognised that friction due to rolling of non-lubricated surfaces over each other is considerably less than dry sliding friction of the same surfaces (Harris, 1991). The rolling resistance, or the rolling friction

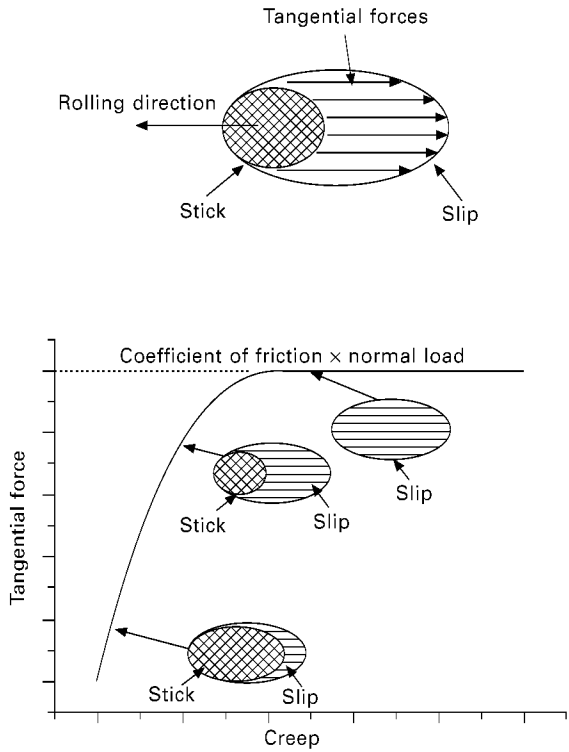
force, is mainly a cumulative result of hysteresis loss and slip. For the steel wheel–steel rail contact, the rolling coefficient can be of the order of 0.01 to 0.001. However, when plastic deformation occurs, rolling resistance may increase sharply. In rolling motion, new material is continually indented in front of the wheel and the material elastically recovers behind the wheel. A small fraction of the energy needed to indent the material will not be recovered due to material relaxation. This hysteresis loss depends on the hardness of the surface and the stress level. An asymmetrical contact pressure distribution is generated so that the resultant of the contact pressure distribution is in the direction of movement. To reduce the rolling resistance (i.e. for a low μ-value), a large wheel radius, a high modulus of elasticity and a low load are desirable. If plastic deformation occurs at the wheel–rail contact, the asymmetry of the contact pressure distribution increases, which leads to an increase in the rolling resistance. The effect can be compared to riding a bicycle in sand – all the energy expended to indent the sand is lost.

Consider the wheel in Fig. 17.3. If the wheel is rolling freely, $v = r \cdot \omega$. To achieve a positive tangential force F at the wheel, we need a larger velocity, i.e. $r \cdot \omega > v$. This implies that we have a velocity difference $\Delta v = r \cdot \omega - v$. The relation between the velocity difference and the velocity can be defined as the creep ($\Delta v/v$). In a railway context, 'adhesion' can be used to define the friction that can be made available to transfer tangential forces between a driving railway wheel and the rail. Sometimes 'traction' is used for a driving wheel and 'adhesion' is used for a braking wheel, but in this document 'adhesion' is used for both situations. The available adhesion is limited by the coefficient of friction. If a driving railway wheel applies a tangential force larger than this limit, the wheel will spin causing severe damage to the rail.

As shown in Fig. 17.4, the contact area between a wheel and rail can be divided into stick (no slip) and slip regions. Longitudinal creep and tangential forces arise due to the slip that occurs in the trailing region of the contact patch. With increasing tangential force, the slip region increases and the stick region decreases, resulting in a rolling and sliding contact. When the tangential force reaches its saturation value, the stick region disappears, and the entire



17.3 Schematic of a wheel. [r = wheel radius (m), m = torque (Nm), v = velocity (m/s), ω = rotational speed (rad/s), F = tangential force (N)]



17.4 Relationship between tangential force and creep at the wheel–rail contact.

contact area is in a state of pure sliding. The maximum level of tangential force depends on the capacity of the contact patch to absorb the adhesion. This capacity is expressed in the form of the friction coefficient, μ.

A wide range of full-scale studies of wheel–rail friction and adhesion has been conducted, using a variety of instruments ranging from a bogie-on-roller rig, to a hand-pushed rail tribometer and instrumented trains. Table 17.1 presents data measured by the author using a hand-pushed tribometer and shows the variety and scatter in the coefficient of friction. Table 17.2 shows examples of available friction in the form of the adhesion coefficient for a wider range of conditions (Moore, 1975). The Swedish National Rail Administration used an instrumented train to map the wheel–rail adhesion at six well-known adhesion problem locations (Forslöv, 1996). The coefficient of friction varied from 0.4 for dry rail down to 0.1 with a blackish leaf layer on the rail.

Olofsson and Telliskivi (2003) compared coefficients of friction measured on a track with a hand-held tribometer and those measured in a laboratory with a pin-on-disc machine. In tests of pure, non-lubricated sliding, both

Table 17.1 Friction coefficients measured with a salient system tribometer

Condition	Coefficient of friction
Sunshine dry rail, 19 °C	0.6–0.7
Recent rain, 5 °C	0.2–0.3
With a lot of grease on rail, 8 °C	0.05–0.1
Damp leaf film on rail, 8 °C	0.05–0.1

Table 17.2 Examples of rail–wheel adhesion coefficients (Moore, 1998)

Condition of rail surface	Adhesion coefficient
Dry rail (clean)	0.25–0.30
Dry rail (with sand)	0.25–0.33
Wet rail (clean)	0.18–0.20
Wet rail (with sand)	0.22–0.25
Greasy rail	0.15–0.18
Moisture on rail	0.09–0.15
Light snow on rail	0.10
Light snow on rail (sand)	0.15
Wet leaves on rail	0.07

methods yielded much the same results, varying between 0.5 and 0.6. In tests of a full-scale lubricated rail, the coefficient of friction was lower – down to 0.1 – depending on the amount of applied lubrication. The pin-on-disc tests in Olofsson and Sundvall (2004) with crushed leaves in the contact region yielded coefficients of friction down to 0.1. The coefficient of friction with the leaf layer decreased with increasing relative humidity.

By contrast, Harrison *et al.* (2000), using the same type of hand-held tribometer, reported a coefficient of friction that was typically 0.7 under dry conditions and varied between 0.25 and 0.45 under lubricated conditions. Using a twin-disc tribometer Gallardo-Hernandez and Lewis (2008) measured coefficients of friction ranging from 0.6 for dry conditions to 0.2 for wet conditions, and down to 0.05 for leaf-layer lubricated conditions. Using an instrumented bogie on a test vehicle, Naganse (1989) measured coefficients between 0.2 and 0.4 for dry conditions, 0.05 and 0.2 for wet conditions and 0.025 and 0.1 for leaf-layer lubricated conditions. The above review thus shows that full-scale instrumented tests results are lower than laboratory results under controlled conditions (e.g. for humidity and velocity). It is well known that velocity and humidity affect the coefficient of friction and thus the available adhesion. The influence of the velocity v (m/s) on the adhesion α can be seen from the well known equation by Curtius and Kniffler:

$$\alpha \approx \frac{7.5}{44 + 3.6v} + 0.161 \qquad [17.1]$$

Japanese railways use different equations for the adhesion depending on whether the rail is dry or wet Ohyama (1991):

$$\alpha \approx 0.265 \frac{1 + 1.451v}{1 + 1.879v} \qquad \text{normal rail surface} \qquad [17.2]$$

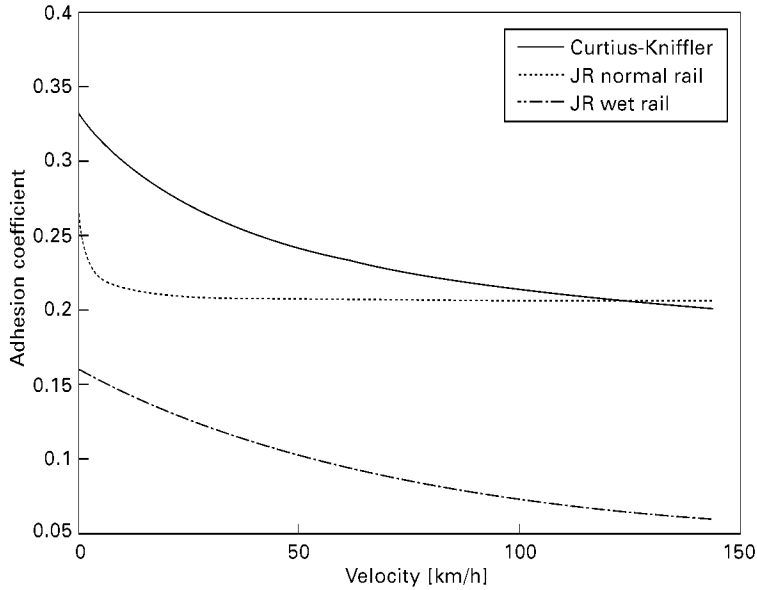
and

$$\alpha \approx \frac{3.78}{22.6 + v} \qquad \text{wet rail surface} \qquad [17.3]$$

Figure 17.5 shows the empirical equations (eqs 17.1–17.3) graphically. We clearly see the influence of increasing velocity as well as how a change from dry to wet rail can significantly reduce the adhesion.

'Braking distance' means the distance a train travels from a certain given velocity to when the train is standing still. In UIC (leaflets no. 540–549) a graphical method is used to determine the braking distance, which depends on braking weight. A train's braking weight is a measure of its braking ability (a higher braking weight gives a larger braking ability).

Here we use an analytical model to predict the braking distance when the adhesion between railway wheel and rail is varied (Andersson and Berg, 2001). The first assumption is that the retardation is constant during braking. The braking distance s can then be written as:



17.5 Adhesion coefficient as a function of velocity for Eqs 17.1–17.3.

$$s = \frac{v_0^2}{2r_0} \qquad [17.4]$$

where v_0 velocity at start of braking (m/s) and r_0 is the constant retardation (m/s^2).

As an example, let us consider a vehicle with a braking torque on all eight wheels (four wheel sets), as in Fig. 17.6, with, F_b = braking force for the vehicle (N), F_1 = braking force for wheel (1) (N), m_b = vehicle mass (kg), v = vehicle velocity (m/s) and, D_b = rolling resistance (N). The total braking force for the vehicle is:

$$F_b = \sum_{i=1}^{n} \frac{\alpha_1 \cdot m_b \cdot g}{n} \qquad n = 8 \qquad [17.5]$$

where α = adhesion coefficient wheel 1 and g = acceleration of mass (m/s^2). The equation of motion for the vehicle in the travelling direction is:

$$F_b = m_b \cdot r_x - D_b \qquad [17.6]$$

The retardation can then be written as:

$$r_x = (F_b - D_b)/m_b \qquad [17.7]$$

Figure 17.7 shows how the braking distance varies due to a change in the adhesion. For comparison, the rolling resistance was varied by changing the slope by ±20 %. The data for the vehicle analysed was as follows:

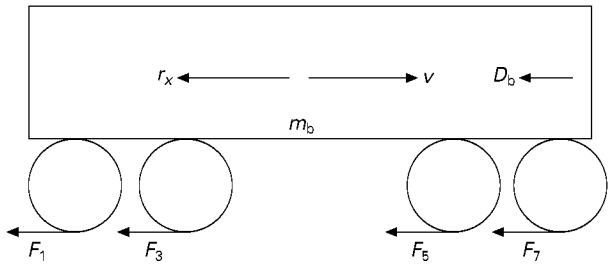
m_b = 40 000 kg

v_0 = 70 km/h

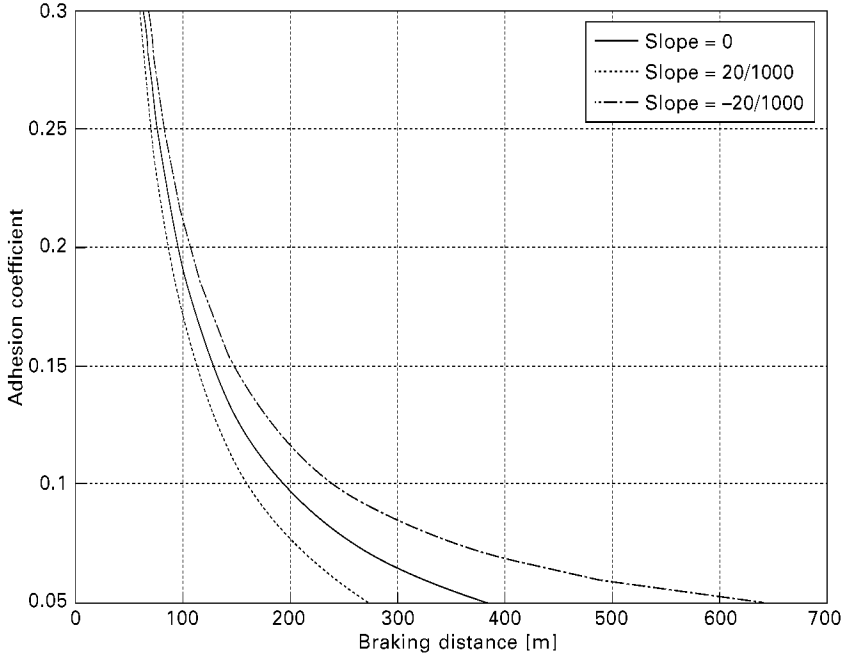
α_1 = 0.05–0.3 constant for all wheels

$D_b = \alpha_1 \cdot m_b \cdot g \cdot \beta$, where β is the slope,
$\beta \in \langle -20/1000, 0, 20/1000 \rangle$

This example shows how a reduction in the adhesion coefficient can greatly increase the braking distance.



17.6 Schematic figure with forces acting on a braking vehicle.



17.7 Example of braking distance as a function of adhesion coefficient and varying slope.

17.3 Friction modification

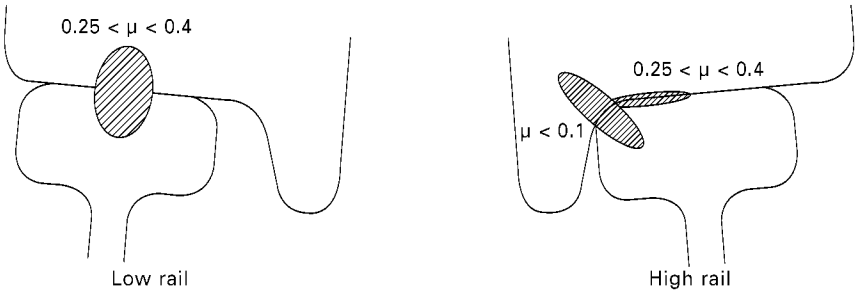
It is clearly advantageous to control friction through the application of friction modifiers to the wheel–rail contact, although the process has to be carefully managed. The aim of friction management is to maintain friction levels in the wheel–rail contact to give (IHHA, 2001):

- 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 loss of adhesion problems occur).

Ideal friction conditions in these contact regions for high and low rails are shown in Fig. 17.8 (Sinclair, 2004). These are similar to values quoted for Canadian Pacific (Roney, 2001).

Friction modifiers can be applied to the wheel–rail contact to generate the required coefficients of friction. These modifiers can be divided into three categories (Kalousek and Magel, 1999):

- low coefficient friction modifiers (lubricants), used to give coefficients down to 0.1 at the wheel flange–gauge corner interface;



17.8 Ideal friction coefficients in the wheel–rail contact.

- high coefficient friction modifiers with intermediate friction coefficients of 0.2 to 0.4, used in wheel tread–rail top applications;
- very high coefficient friction modifiers (friction enhancers), used to increase adhesion for both traction and braking.

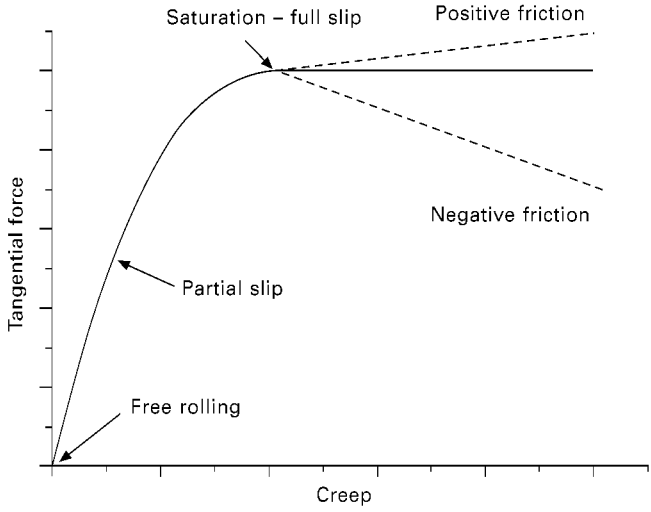
Low coefficient modifiers can be solids, liquids (oil) or greases. The primary application of these modifiers is in reducing friction and wear and preventing seizure at the wheel flange–rail gauge corner contacts, particularly in curves, where the contact conditions can be quite severe. This reduction is clearly beneficial for energy saving and equipment maintenance. The action of the lubricants is to transport chemicals (additives) to the contact. These react chemically with the surface and form an easily sheared surface layer that can also prevent direct metal-to-metal contact, and thus prevent seizure. The proper choice of lubricant is important since both the amount and type of lubrication applied affect the wear rate and the active wear mechanism, as highlighted by Sundh *et al.* (2008). Fuel savings of around 30 % (compared to dry conditions) have been reported for measurements taken on test tracks (Reiff and Creggor, 1999). Other studies carried out in the field have shown improvements of a similar order of magnitude (Allen *et al.*, 1985; Samuels *et al.* 1987). Laboratory (Sundh and Olofsson, 2008); Clayton *et al.*, 1989) and field (Waara, 2001; Olofsson and Nilsson, 2002) tests have all shown the wear-reducing benefits of lubrication at the wheel–rail contact. The Alvsjo test track (Olofsson and Nilsson, 2002) is a good example: the worn rail area for the lubricated case was approximately one sixth of the worn area for the unlubricated case.

Lubrication of the wheel–rail contact with low friction modifiers can cause problems if it is poorly managed. If over-lubrication occurs and lubricant migrates onto the railtop, there may be a loss of adhesion that can cause wheel slip and loss of braking. Interference with ultrasonic flaw detection has been reported at over-lubricated sites (Sinclair, 2004). Over-lubrication can also cause an increase in RCF crack growth on the rail gauge corner (Cannon and Pradier, 1996). This happens either because pressurisation of

the cracks leads to increased growth or because reduced wear means that the cracks are truncated less. However, full-scale test results from tight curves show that well-maintained lubrication can reduce both the wear rate and the propagation rate of surface cracks (Olofsson and Nilsson, 2002).

High friction modifiers usually consist of liquid-borne products that contain particles with different physical properties (Li *et al.*, 2008). After application either with a brush or by spraying, they dry as a thin film that is intended to alter the surface shear stress and thus keep the friction coefficient in a range between 0.2 and 0.4. Often these products also aim to increase the coefficient of friction with increasing sliding velocity, whereas this coefficient usually decreases with increasing sliding velocity (Fig. 17.9). A decreasing coefficient of friction can cause squeal and rail corrugation (see Chapters 16 and 11, respectively). Field tests with these products confirm that they can reduce the noise level (Eadie *et al.*, 2002, 2003) and the rolling resistance (Matsumoto *et al.*, 2002; Tomeka *et al.*, 2002). However, they have the drawbacks in that they can affect rail insulation causing signalling problems (Lewis *et al.*, 2003) and that they are effective for only a limited number of wheel passages.

The main very high friction modifier used on railway networks worldwide is sand. While applying sand is effective and easy, it can cause complex and costly problems for rolling stock and track infrastructure. It has been shown to substantially increase the wear rates of both wheel and rail materials (Lewis and Dwyer-Joyce, 2006). Gallardo-Hernandez and Lewis (2008) showed how sand can be used to remove the blackish layer formed by crushed leaves and thus restore the adhesion level. Another approach to recovering



17.9 Behaviour of friction modifier.

adhesion is to clean the rails. A number of methods of doing this have been developed, including using high-pressure water-jets (Fig. 17.10) and blasting with Sandite (a mixture of sand and aluminium oxide particles). All of these methods require maintenance trains. There are very few such trains and it can be difficult for them to obtain access to the track. Moreover, even though there may be no visible remains of the hard, black, slippery coating after cleaning with high-pressure water jets, the improvement in the coefficient of friction may be insignificant. (Tests performed in the Stockholm local network in 2005 used a water pressure of 800 bar and a water temperature of 30 °C.) Measurements performed with a hand-held tribometer before and after the use of high-pressure water jets show that the coefficient of friction increases by only between 0.05 and 0.1 (Nilsson, 2006).

Very high positive friction modifiers that enhance the coefficient of friction to between 0.4 and 0.6 are also available in the form of solid components that have different physical characteristics depending on their purpose. The active components can be embedded in a polymeric matrix and then formed into a rectangular stick of material that can be applied directly to the wheel tread. When the material in the stick transfers to the wheel, the resin oxidises under the high temperatures at the wheel–rail interface and leaves a thin film of the dry friction modifier on the wheel. Some of the drawbacks of the very high friction modifiers are the same as for high friction modifiers, that is, they can affect the rail insulation, causing signal problems, and are effective for only a limited number of wheel passages. However, promising

17.10 High-pressure washing of rails. (Photograph by Lazlo Scabo, Stockholm Public Traffic)

results regarding rail insulation have been presented in Lewis *et al.*, 2008). The wear rate of these products also needs further investigation.

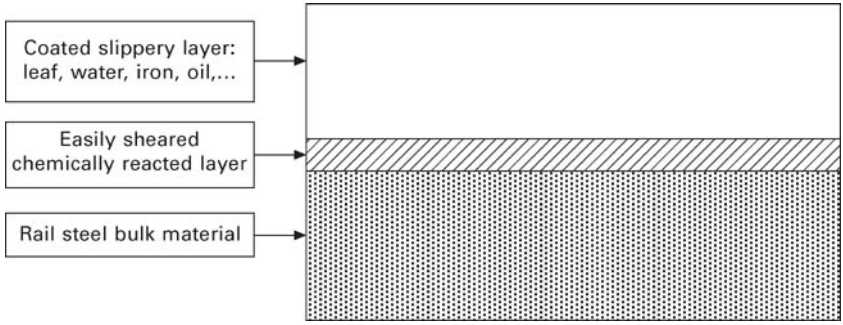
17.4 Possible models for low friction at the wheel–rail contact

Various models have been put forward to explain low coefficients of friction in situations where there is high contact pressure. If the surfaces are separated by a fluid, then the coefficient of friction is determined by viscous forces in the fluid layer. A lubricant can be used to separate the surfaces, as shown in the classical experimental works of Tower (1883) on hydrodynamic sliding bearings and in the analytical work of Reynolds (1886). Reduction to coefficient values determined by viscous forces can occur when there is plenty of lubricating oil on the railhead. It can also occur when there is phase transformation, as when metals slide at high speeds (above 100 m/s) or when cross-country skiing on snow. In both cases, frictional heat raises the local temperature in the interface to melting point and a viscous fluid is formed.

The second explanation for the low coefficients of friction is based on the classical theory of Bowden and Tabor (1950), who studied the effect of a hard bulk providing a small area of real contact and a thin, easily sheared surface layer. The surface layer can be formed by additives in lubrication oil or by an oxide layer on the contacting surfaces. During boundary lubricated conditions in roller bearings and gearboxes, the lubricating oil does not separate the surfaces; instead the lubricant acts as a transporter of chemical substances such as sulphur and phosphorus that form an easily sheared surface layer (Olofsson and Dizdar, 1997). Certain alloys, such as the cobalt-based alloys called stellites, also offer low friction properties because they continually generate an easily sheared surface layer (Persson, 2005). Stellitic materials are frequently used in valves because of their resistance to seizure.

Models have also been developed for contacts under low pressure. Plastic materials like PTFE (polytetrafluoroethylene) and HDPE (high-density polyethylene) can generate very thin films oriented in the sliding direction, decreasing the coefficient to as low as 0.05 (Hutchings, 1992). Since this is a low contact pressure phenomenon, it will not be considered here.

The preceding review of the literature suggests a model in which leaves crushed beneath rail wheels transport chemical substances to the wheel and rail surfaces in the same way as the lubricating oil used under boundary lubricated conditions transports chemical substances. In both cases, the chemical substances react with the surface and form an easily sheared surface layer. Chemical analysis of the coated slippery layer shows that it contains a mixture of iron and iron oxide, water, leaf debris and oil (Forslöv, 1996; Cann, 2006). Olofsson (2007) performed elemental depth profiling of rail



17.11 Schematic of proposed multi-layer low-adhesion model on rail steel surface. Crushed leaves form both a coated slippery layer and an easily sheared, chemically reacted surface layer. (Olofsson, 2007)

surface layers. The results showed that a chemically reacted, easily sheared surface layer containing substances such as P and Ca can be formed on a rail under the blackish slippery layer. The multi-layer model presented in Fig. 17.11 shows the crushed leaves forming both a coated slippery layer and an easily sheared, chemically reacted surface layer on the rail. Note that both layers probably form on the railway wheel as well as on the rail, but for sake of clarity only the rail is shown in Fig. 17.11.

17.5 Future trends

Further development is necessary to be able to control as well as model the adhesion between railway wheel and rail. This will involve studies of the surface layers of the wheel and rail, and perhaps also surface modification of both by coating their surfaces with a surface layer (Heinsch *et al.*, 2003). Another promising direction is the further development of high friction modifiers (Lu *et al.*, 2005). However, the greatest challenge from both a scientific and applied point of view lies in the development of reliable models that can be used to predict the adhesion between railway wheel and rail (see Chapter 4 and Butcher *et al.*, 2006).

17.6 References

Adhesion Working Group (2001), *Managing Low Adhesion*, London, UK.

Allen, R.A., Mims, W.E., Rownd, R.C., and Singh, S.P. (1985), Energy savings due to wheel rail lubrication: Seaboard system test and other investigations, *Trans ASME, Journal of Engineering for Industry*, **107**, 190–96.

Andersson, E., and Berg, M. (2001), *Railway Systems and Rail Vehicles*, Division of Rail Vehicles, KTH, Stockholm, Sweden (in Swedish).

Beagley, T.M., and Pritchard, C. (1975), Wheel/rail adhesion: the overriding influence of water, *Wear*, **35**, 299–313.

Beagley, T.M., McEwen, I.J., and Pritchard, C. (1975a), Wheel/rail adhesion: the influence of railhead debris, *Wear*, **33**, 141–52.

Beagley, T.M., McEwen, I.J., and Pritchard, C. (1975b), Wheel/rail adhesion: boundary lubrication by oily fluids, *Wear*, **33**, 77–88.

Berggren, A-C. (2003), *Final Report on Leaf Lubrication 2003*, Swedish National Rail Administration, Börlange, Sweden (in Swedish).

Bowden, F.P., and Tabor, D. (1950), *The Friction and Lubrication of Solids*, Parts I and II, Oxford, UK, Clarendon.

Broster, M., Pritchard, C., and Smith D.A. (1974), Wheel/rail adhesion: its relation to rail contamination on British railways, *Wear*, **29**, 309–21.

Bucher, F., Dmitriev, A.I., Ertz, M., Knothe, K., Popov, V.L., and Psakhie, S.G. (2006), Multiscale simulation of dry friction in wheel/rail contact, *Wear*, **261**, 874–84.

Cann, P.M. (2006), The "leaves on the line" problem: a study of leaf residue film formation and lubricity under laboratory test conditions, *Tribology Letters*, **24**(2), 151–8.

Cannon, D.F., and Pradier, H. (1996), Rail rolling contact fatigue: research by the European Rail Research Institute, *Wear*, **191**, 1–13.

Clayton, P., Danks, D., and Steele, R.K. (1989), Laboratory assessment of lubricants for wheel/rail applications, *Lubrication Engineering*, **45**(8), 501–6.

Collins, A.H., and Pritchard, C. (1972), Recent research on adhesion, *Railway Engineering Journal*, **1**, 19–29.

Czichos, H. (1992), Presentation of friction and wear data. in Blau, P.J. (ed.), *Friction, Lubrication and Wear Technology*, *ASM Handbook*, **18**, ASM International, Materials Park, OH, USA.

Eadie, D.T., Kalousek, J., and Chiddick, K.C. (2002), The role of high positive friction (HPF) modifier in the control of short pitch corrugation and related phenomena, *Wear*, **253**, 185–92.

Eadie, D.T., Santoro, M., and Kalousek. J. (2003), Railway noise and the effect of top of rail liquid friction modifiers: changes in sound and vibration spectral distributions, *Proceedings 6th International Conference on Contact Mechanics and Wear of Rail/Wheel Systems*, Gothenburg, Sweden, 10–13 June, 503–10.

Forslöv, L. (1996), *Wheel Slip due to Leaf Contamination*, Swedish National Rail Administration TM 1996 03 19, Borlänge, Sweden, (in Swedish).

Fulford, C.R. (2004), *Review of Low Adhesion Research*, Rail Safety and Standards Board, London, UK.

Gallardo-Hernandez, E., and Lewis, R. (2008), Twin disc assessment of wheel rail adhesion, *Wear*, **265**(9–10), 1309–1316.

Harris, T.A. (1991), *Rolling Bearing Analysis*, Wiley, New York, USA.

Harrison, H., McCanney, T., and Cotter, J., Recent development in COF measurements at the rail/wheel interface, *Proceedings 5th International Conference on Contact Mechanics and Wear of Rail/Wheel Systems*, Tokyo, Japan, 25–27 July, 30–34.

Heinsch, E.J.M., Franklin, F.J., Nielson, J.C.O., Ringsberg, J.W., Weeda, G.J., Kapoor, A., and Josefson, B.L. (2003), Prevention of RCF damage in curved rail through development of the infra-star two-material rail, *Fatigue and Fracture in Engineering Materials and Structures*, **26**, 1007–17.

Hutchings, I.M. (1992), *Tribology: Friction and Wear of Engineering Materials*, Edward Arnold, London, UK.

IHHA (2001) *Guidelines To Best Practices For Heavy Haul Railway Operations: Wheel and Rail Interface Issues*, International Heavy Haul Association, Virginia Beach, VA, USA.

Kalousek, J., and Magel, E. (1999), Modifying and managing friction, *Railway and Track Structures*, **5**, 5–6.

Lewis, R., and Dwyer-Joyce, R.-S. (2003), Wear at the wheel/rail interface when sanding is used to increase adhesion, *Proceedings of the IMechE, Part F: Journal Rail and Rapid Transit*, **220**(1), 29–41.

Lewis, R., Dwyer-Joyce, R.-S., and Lewis, J. (2003), Disc machine study of contact isolation during railway track sanding, *Proceedings of the IMechE, Part F: Journal of Rail and Rapid Transit*, **217**, 11–24.

Lewis, R., Gallardo, E.A., Cotter J., and Eadie, D.T. (2008), The effect of friction modifiers on wheel/rail isolation, *Proceedings 2008 IEEE/ASME Joint Rail Conference*, 22–24, April, Wilmington, DE, USA, on CD.

Li, Z., Arias-Cuevas, O., Lewis, R., and Gallardo-Hernandez, E.A. (2008), Rolling-sliding laboratory tests of friction modifiers, in *Leaf Contaminated Wheel–Rail Contacts, International Joint Tribology Conference*, Miami, FL, USA, 20–22 October, on CD.

Lu, X., Cotter, J., Eadie, D.T. (2005), Laboratory study of the tribological properties of friction modifier thin films for friction control at the wheel/rail interface, *Wear*, **259**, 1262–69.

Matsumoto, A., Sato, Y., Ono, H., Wang, Y.J., Yamamoto, M., Tanimoto, M., and Oka Y. (2000), Creep force characteristics between rail and wheel on scaled model, *Wear*, **253**, 199–203.

Moore, D.F. (1975), *Principles and Applications of Tribology*, Elsevier, Amsterdam (1998).

Nagase, K. (1989), A study of adhesion between the rails and running wheels on main lines; results of investigations by slipping adhesion test bogie, *Proceedings of the IMechE, Part F: Journal of Rail and Rapid Transit*, **203**, 33–43.

Nilsson, R. (2006), Stockholm Local Traffic, personal communication, January.

Ohyama T. (1991), Tribological studies on adhesion phenomena between wheel and rail at high speed, *Wear*, **144**, 263–75.

Olofsson, U. (2007), A multilayer model of low adhesion between railway wheel and rail, *Proceedings of the IMechE, Part F: Journal of Rail and Rapid Transit*, **221**, 385–9.

Olofsson, U., and Dizdar, S. (1997), Surface analysis of boundary lubricated spherical roller thrust bearings, *Wear*, **215**, 156–64.

Olofsson, U., and Nilsson, R. (2002), Surface cracks and wear of rail: A full scale test and laboratory study, *Proceedings of the IMechE, Part F: Journal of Rail and Rapid Transit*, **216**, 249–64.

Olofsson, U., and Sundvall, K., Influence of leaf, humidity, and applied lubrication on friction in the wheel–rail contact: pin-on-disc experiments, *Proceedings of the IMechE, Part F: Journal of Rail and Rapid Transit*, **218**, 235–42.

Olofsson, U., and Telliskivi, T. (2003), Wear, friction and plastic deformation of two rail steels: full-scale test and laboratory study, *Wear*, **254**, 80–93.

Persson, D.H.E., *On the Mechanism Behind the Tribological Performance of Stellites*, PhD Thesis, Uppsala University, Uppsala, Sweden.

Poole, W., *Characteristics of Railhead Leaf Contamination: Summary report*, October AEA Technology Rail, Derby, UK.

Reiff, R., and Creggor, D. (1999), Systems approach to best practice for wheel and rail friction control, *Proceedings International Heavy Haul Association Specialist Technical Session*, Moscow, Russia, 14–17 June.

Reynolds, O. (1886), On the theory of lubrication and its application to Mr Beuchamps Tower's experiments including an experimental determination of the viscosity of olive oil, *Philosophical Transactions of the Royal Society of Lond*, **177**, 157–234.

Roney, M.D. (2001), 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 Beach, VA, USA.

Samuels, J.M., and Tharp, D.B. (1987), Reducing train rolling resistance by on-board lubrication, *Proceedings 2nd Rail and Wheel Lubrication Symposium*, Memphis, TN, USA, 2–4 June.

Sinclair, J. (2004), Friction modifiers, *Vehicle Track Interaction: Identifying and Implementing Solutions*, IMechE, Seminar, 17 February.

Suh, N.P. (1986), *Tribophysics*, Prentice-Hall, Englewood Cliffs, NJ, USA.

Suh, N.P., and Sin, H.C. (1981), The genesis of friction, *Wear*, **69**, 91–114.

Sundh J. and Olofsson U. (2008), Seizure mechanisms of wheel/rail contacts under lubricated conditions using a transient ball-on-disc test method, *Tribology International*, **41**, 867–74.

Sundh, J., Olofsson, U., and Sundvall. K. (2008), Seizure and wear rate testing of wheel/rail contact under lubricated conditions using a transient and a standard ball-on-disc test method, *Wear*, **265**, 1425–30.

Tomeka, M., Kabe, N., Tanimoto, M., Miyauchi, E., and Nakata M. (2002), Friction control between wheel and rail by means of on-board lubrication, *Wear*, **253**, 124–9.

Tower B. (1883), First report on friction experiments, *Proc Instn Mech Engrs* 632–59.

UIC sub-section 54 – brakes, leaflets no. 540–549, International Union of Railways, Paris, France.

Veal, A. (2008), Identification of weather related cases of poor rail adhesion, *Adhesion Working Group Conference*, 29 April, UK.

Waara, P. (2001), Lubricant influence on flange wear in sharp railroad curves, *Industrial Lubrication and Tribology*, **53**(4), 161–8.

Copyrighted Materials
Copyright © 2009 Woodhead Publishing Retrieved from www.knovel.com

18
Wheel–rail isolation

R. LEWIS, University of Sheffield, UK

Abstract: This chapter outlines problems with track circuit-based train detection systems that rely on good wheel–rail conductance to operate safely. Contaminants such as leaves, ballast dust, rust and substances such as friction modifiers forming a third-body layer in the wheel–rail contact may be able to cause train wheels to be isolated from the rail, thus causing the train to be 'lost' by the signalling system. Field testing can clearly be used to assess likely problems, but this is time-consuming and it is difficult to maintain the levels of control required. Dynamic laboratory tests have been developed based on a twin-disc simulation of a wheel–rail contact, as well as static tests using actual wheel and rail specimens, that have been used to study the effect of sand, leaves and solid stick friction modifier (applied to wheel treads) on isolation. It has been found that sand and leaves could isolate, but that the friction modifier does not. Critical loads for breaking down sand and leaf layers and threshold amounts of contaminant that lead to isolation have also been identified. The tests are relatively severe, particularly for sanding, where far more of the sand is entrained to the test contact than would be into an actual wheel–rail contact, but they can give a good guide to where isolation problems may occur.

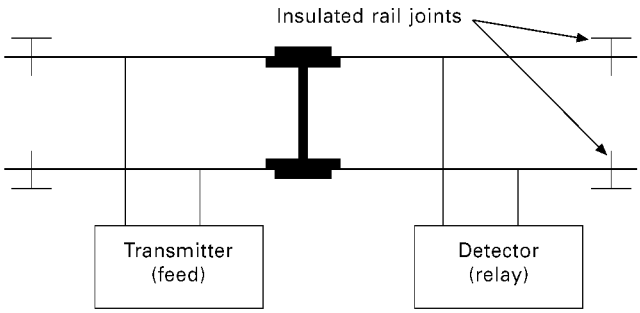
Key words: wheel–rail contact, contaminants, isolation, train detection.

18.1 Introduction

18.1.1 Track circuits

Track circuits are devices designed to continuously detect the absence of a train from a particular section of track. Their designed failure mode is to indicate the presence of a train and therefore they cannot be used to detect whether a train is present. A clear track circuit can be used to allow a train to safely progress.

A track section is electrically defined by insulated joints, as shown in Fig. 18.1. An electrical energy source (transmitter) is connected, via a series impedance, across one end of the track circuit. At the other end is a detector. If there is no train within the boundaries of a track circuit the detector picks up the electrical energy from the transmitter. It in turn energises a repeater circuit, which tells the signalling system that the section of track is clear. If a train is present on the track section, the rails will be short-circuited and the detector will no longer be able to sense the electrical energy from the transmitter. It therefore changes state and the signalling system is informed that the section of track is occupied.



18.1 Track circuit schematic.

It can be seen that any short-circuit, caused by a train or otherwise, or a break in the circuit will fail the track circuit and inform the signalling system that the track is occupied, so a good degree of *fail-safe* is incorporated. The system, however, relies on good wheel–rail electrical contact to work. If this is inhibited in any way, for example by contamination on the railhead, train identification may be prevented. An example of where this could be potentially serious is if the rear of a train crosses into another track circuit and the wheels are isolated from the rail. This will not be detected by the track circuit and may lead to a collision occurring if another train is allowed onto that section of track.

18.1.2 Wheel–rail isolation problems

Contamination present on the railhead or on the surface of the wheel tread could cause the wheel to become isolated from the rail which would inhibit train identification. In Section 18.2 more information is given about some of the contaminants, both naturally occurring and those deliberately applied to the wheel–rail contact. Recent design changes to railway vehicles have led to these becoming an increasing problem. Disc brakes are gradually replacing tread braking systems. Tread brake pads roughen the wheel tread surfaces which allows them to cut through some contaminants. The smoother surfaces that now result are unable to do this. Railway vehicles are gradually becoming lighter in an effort to reduce energy requirements and material usage, etc. which leads to lower wheel–rail contact loads and leaves trains too light to break through, for example, rust layers that may form on infrequently used track.

Third bodies present on the railhead or wheel surface, whether they be natural (rust, leaves, ballast fragments) or deliberately applied (friction modifiers, sand, lubricants, etc.) could compromise the wheel–rail electrical contact and lead to loss of train detection.

18.1.3 Types of track circuit

Basic DC track circuits are commonly battery-powered at low voltages (1.5–12 V DC) to protect against line power fail. In almost all railway electrification schemes, however, one or both of the rails is used to carry the return current. This prevents use of the basic DC track circuit because the substantial traction currents overwhelm the very small track signal currents. To overcome this problem, AC track circuits are used. These, typically, use frequencies in the range from 91 Hz up to 250 Hz. The relays are arranged to detect the selected frequency and to ignore DC and AC traction frequency signals.

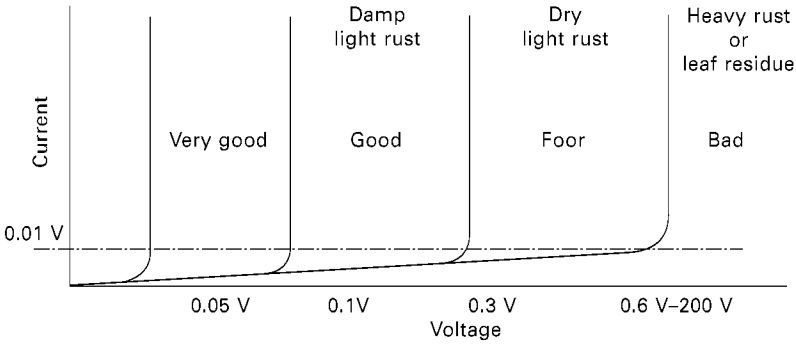
Jointless track circuits are also used in some regions. These operate at much higher frequencies (1500–3000 kHz). The advantage of jointless circuits is that they negate the need for insulated joints which are prone to failure.

In the UK Railway Standard on track circuits (Railway Safety, 2001), there is some discussion on railhead contamination and its effect on train isolation. Rust films in particular are addressed. It is stated that light rust films act as a semi-conductor, in that they exhibit high resistance until a particular threshold voltage is applied when they break down completely. The break-down voltage increases with the extent of the contamination; very heavy rust films and leaf residue are very hard to break down using increasing voltages, as shown in Fig. 18.2.

18.2 Third bodies in the wheel–rail contact

There are a number of third bodies that could affect the wheel–rail electrical contact and compromise the operation of the track circuit. They can be split into two types (Descartes *et al.*, 2005):

1. **Climatic bodies**. These include leaves and the oxide layer that builds up on the rail surface, i.e. substances occurring naturally.

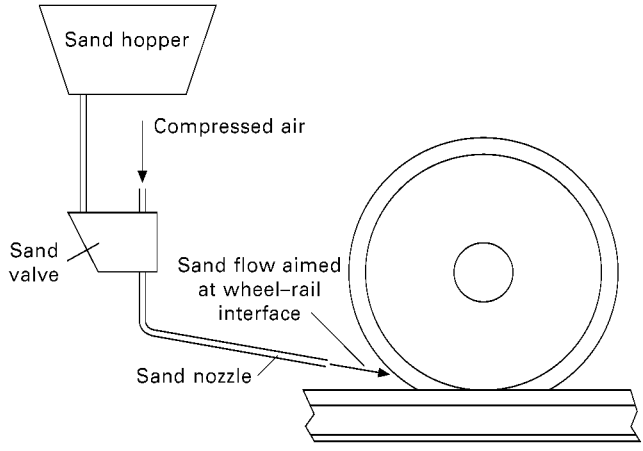


18.2 Contaminant film breakdown voltages. (Railway Safety, 2001)

2. **Operational bodies**. Such as oil (deposited by passing trains), friction modifiers, grease, sand (used as a friction enhancer), i.e. substances applied to the wheel or the rail.

Leaves are a particular problem, particularly during Autumn. Leaves that have fallen from trackside trees are whipped up by the turbulence of passing trains, carried along in their slipstream, swirling around the vehicle underside, and are drawn into the wheel–rail contact where they are crushed under the wheels (Pritchard and Tanvir, 1973). Eventually a hard, glazed, black leaf film layer forms on the railhead. The leaves on their own and this glazed layer can form an electrical insulating barrier.

The leaves also cause low adhesion between wheel and rail (Nagase, 1989; Olofsson and Sundvall, 2004; Gallardo-Hernandez and Lewis, 2008), which can cause problems for braking and traction. To improve adhesion in both situations, sand or Sandite (a mixture of sand and metal particles with water and a binder) is applied to the wheel–rail contact. The sand grains are usually 1–1.5 mm in size and supplied to the contact at rates of around 7 g/m at 10 mph, although variable rate sanders are available. 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 Fig. 18.3). 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. This can lead to excessive sand being applied. Sandite is applied to the railhead by special maintenance trains or by trackside applicators.



18.3 Sanding apparatus.

Other types of friction modifiers are also applied to the wheel–rail contact to control friction, minimise curve noise and lateral forces and reduce wear (Eadie *et al.*, 2005, 2006). Friction modifiers are divided into three categories (Kalousec and Magel, 1999):

1. **low coefficient friction modifiers** (lubricants), used to give friction coefficients less than 0.2 at the wheel flange/gauge corner interface;
2. **high friction modifiers** with intermediate friction coefficients of 0.2–0.4; used in wheel tread–railtop applications;
3. **very high friction modifiers** (friction enhancers, e.g. sand and Sandite); used to increase adhesion for both traction and braking.

The two main methods used to apply friction modifiers to the wheel–rail contact are *wayside applicators* and *onboard applicators*. Wayside applicators are mounted next to the track and apply friction modifier to the rail gauge corner–railhead. There are three types: mechanical, hydraulic and electronic. Onboard applicators apply liquid or solid friction modifiers directly or by spraying on to the wheel flange/tread, which is then transferred to the rail gauge corner or railhead. Complex control systems are used in the application process to avoid the application of friction modifier at inappropriate locations. *Mobile applicators*, which are essentially railway vehicles designed to apply friction modifier to the gauge corner of the track/railhead, are also sometimes used.

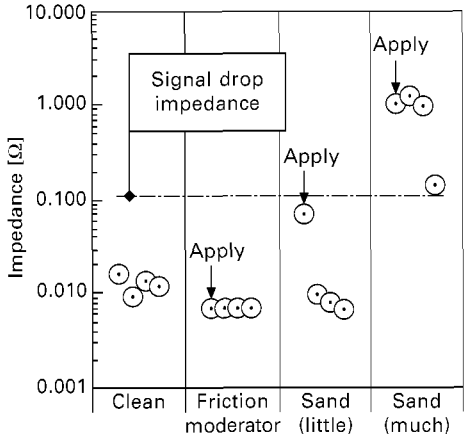
Work has been carried out that has identified the presence of a permanent third-body layer in the wheel–rail contact (Descartes *et al.*, 2005). This is thought to be made up of wear particles from the wheel and rail and contaminants such as those outlined above (sand, leaves, etc.). Clearly this layer will also have an effect on conductance. Some modelling of the layer to estimate its electrical resistance has been carried out and will be discussed further in Section 18.5.

In the following sections, test methods developed to assess isolation are described as well as tests carried out on sand, leaves and solid friction modifier applied to wheel treads.

18.3 Testing for isolation

Field tests can be used to assess isolation. Some results from these are discussed in Chapter 26 (Japanese experience) and are shown in Fig. 18.4. Sand is clearly shown to cause a problem, but friction modifier does not. Clearly this will provide the most appropriate evidence of whether a problem is likely to occur. However, it is impractical to carry out large-scale test programs in the field as these are costly and time-consuming and it is often difficult to gain access to track.

Laboratory isolation tests have been developed, based on simulations of



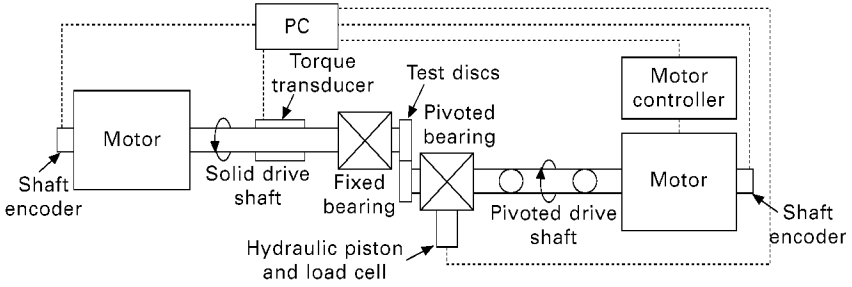
18.4 Impedance measurements.

the wheel–rail contact conditions, that allow measurement of the contact conductance (Lewis *et al.*, 2003, 2008; Lewis and Masing, 2006). Using these methods, controlled tests can be carried out with a range of contact loads/pressures, slips, etc. and different contaminants to establish isolation thresholds. This type of test is impossible in the field. Careful consideration of the results must be carried out, however, when transferring the results to the conditions that will be seen in the field, particularly for the cases where isolation is found to occur. As will be seen in later sections, the laboratory tests are thought to be relatively severe. The laboratory tests are described in the following sections.

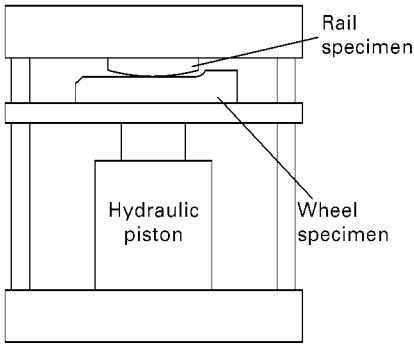
18.3.1 Dynamic test method

A test has been developed to study isolation in a dynamic twin-disc simulation of a wheel–rail contact. A schematic of the test machine used to carry out the testing is shown in Fig. 18.5. This machine was originally designed for studying rolling contact fatigue (RCF) phenomena (Garnham and Beynon, 1991; Fletcher and Beynon, 2000), but has been used extensively for studying other wheel–rail contact issues such as wear and adhesion (see for example Lewis & Dwyer-Joyce, 2004); Gallardo-Hernandez and Lewis, 2008).

The test discs (one manufactured from wheel steel which acts as the driving disc and one made from rail steel which is the brake) are hydraulically loaded together and driven at controlled rotational speeds by independent electric motors. The slip ratio required is achieved by adjustment of the rotational speeds. The top disc can be electrically isolated from the rest of the rig to enable measurement of voltage across the discs to be measured during the tests. Specific contaminant feed mechanisms were designed as appropriate



18.5 Twin disc test set-up.



18.6 Static test apparatus for assessing wheel–rail isolation.

for the tests carried out with sand and solid stick friction modifier, as will be explained in later sections.

18.3.2 Static test methods

The worst-case scenario for isolation is when a train has come to a standstill as there is no tractive force to help remove contaminant from the wheel–rail contact. A static test employing actual wheel and rail sections has been developed to investigate the situation where a train has come to a standstill while contaminant is present on the track (Lewis and Masing, 2006). This allows testing using the actual wheel–rail contact geometry and wheel and rail surface roughness, etc. The test apparatus is shown in Fig. 18.6. The wheel and rail sections are hydraulically loaded together. Contaminants can be placed in the contact area prior to loading. Tests carried out with sand and leaves will be outlined in later sections. Static tests have also been carried out using the twin-disc test set-up described in Section 18.3.1 (Lewis *et al.*, 2008). In these tests the discs were loaded together using the hydraulic jack while stationary.

18.3.3 Track circuit simulations

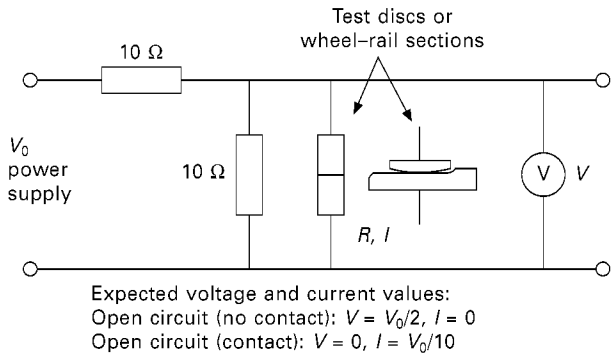
As mentioned in Section 18.1.3, a range of track circuits are used across the world's railway networks. In the bulk of the work carried out using the test methods described here (sand and leaves – see Sections 18.4.1 and 18.4.2), a simplified version of the TI21 track circuit was used. This circuit is typical of those used in the UK. The circuit consisted of a 2 kHz AC voltage source, V_0, connected in series with a 10 Ω resistor, in series with the disc contact. Another 10 Ω resistor was connected in parallel with the disc contact. The resistors were used to replicate the transmitter and receiver resistances found in the TI21 track circuit. The RMS voltage, V, was logged using data capture apparatus with samples taken at 0.1 s intervals.

In order to provide a means of assessing the likelihood of isolation occurring for all types of track circuit, it was necessary to characterise the resistance of a contact and relate it to sand flow rate. The resistance, R, across the discs can be calculated by:

$$R = \frac{10}{\left(\frac{V_0}{V}\right) - 2} \qquad [18.1]$$

As the voltage, V, approaches its open circuit value ($V_0/2$), however, the resistance across the discs becomes infinite. This makes assigning an average value for the contact resistance for a given sand flow rate impossible. In order to overcome this, the conductance, G, was considered rather than resistance (where $G = 1/R$).

For the work using solid stick friction modifier (described in Section 18.4.3), a 8 kHz track circuit was set up. The circuit shown in Fig. 18.7 was modified to achieve this. The primary change was the reduction of the 'transmitter' resistance to 1 Ω to reflect that of the audio frequency track



18.7 Simplified TI21 electrical circuit used for determining voltage across the wheel–rail sections and twin-disc contact.

circuit being replicated. A signal generator was used to provide the 8 kHz signal and a power amplifier boosted the signal to 1 A. A data acquisition system was used to log input voltage; rms voltage across the discs and rms current. Data were logged every 0.1 s. The impedance of the circuit when the discs were out of contact was approximately 9000 mΩ.

18.4 Effects of contaminants on isolation

18.4.1 Sand

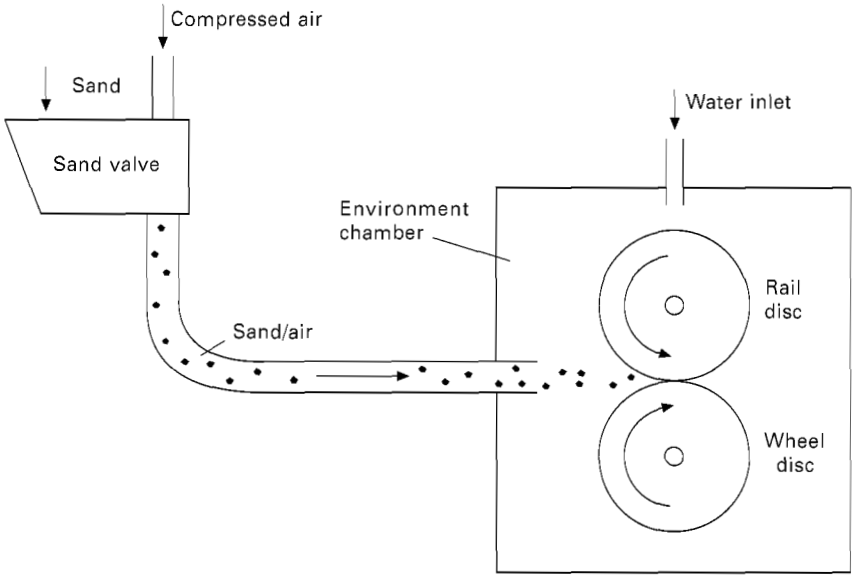
Both dynamic and static tests have been carried out to determine the effect on isolation of sand (Lewis *et al.*, 2003; Lewis and Masing, 2006). The initial driver for this work was the requirement for design changes to a particular railway vehicle. At present, sanding devices must be mounted no further forward than the third axle in order to prevent all wheelsets from being isolated by sand. There was a demand to move a sander to the front axle of a certain train which required evidence that sand would not cause isolation before it could go ahead.

Dynamic tests were carried out using the twin-disc set-up described in Section 18.3.1. Test discs were loaded and rotated to achieve surface speeds of 2 mph and 0.5 mph with a mean contact pressure of 1500 MPa and a slip of 20 % (typical for a driving wheel experiencing loss of adhesion). Tests were carried out applying sand to both a dry disc interface and a wet disc interface (water has been shown to be one of the more frequent causes of adhesion loss – Beagley and Pritchard, 1975). Sand was applied using a sand valve (as shown in Fig. 18.8) at various flow rates.

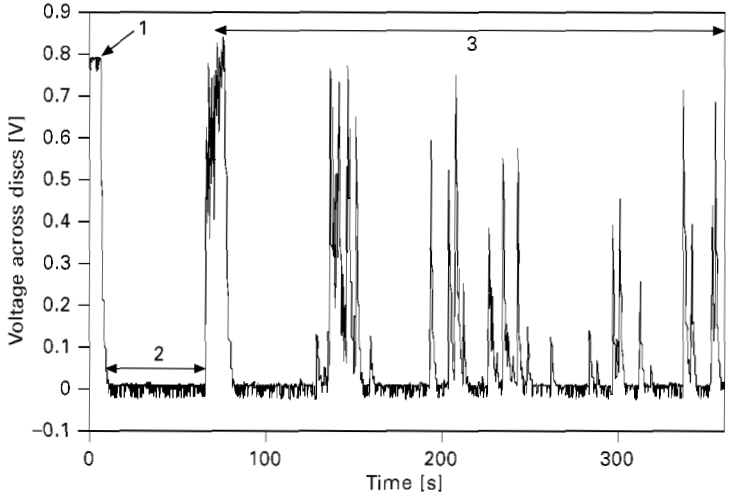
Figure 18.9 shows rms voltage plots for a contact run with dry sand. The three stages indicated correspond to: (i) discs out of contact, no sand; (ii) discs in contact and under load, no sand; (iii) discs in contact and under load with sand application at intervals. Figure 18.10 shows the same data plotted separately for each sand flow rate. Also shown is the expected 'open circuit' voltage (0.8 V, discs out of contact).

For disc operation without sand the signal is stable (with some small amounts of noise). For sand flow rates of 0.5 kg/min and above, the voltage is almost continuously above the closed circuit value, whilst for sand flow rates below 0.5 kg/min, the voltage changes intermittently, but tends towards its closed circuit value. The intermittent voltage signal is probably caused by a non-uniform flow of sand particles into the contact. Whilst the sand is fed into the contact inlet region directly at a uniform rate, it appears that sand enters the contact itself fairly unevenly.

Figure 18.11 shows rms voltage plots for a contact run with sand and water using a constant input voltage and varying the sand flow rate (2 mph, 1500 MPa, 20 % slip, 1.6 V at different sand flow rates). For sand flow rates



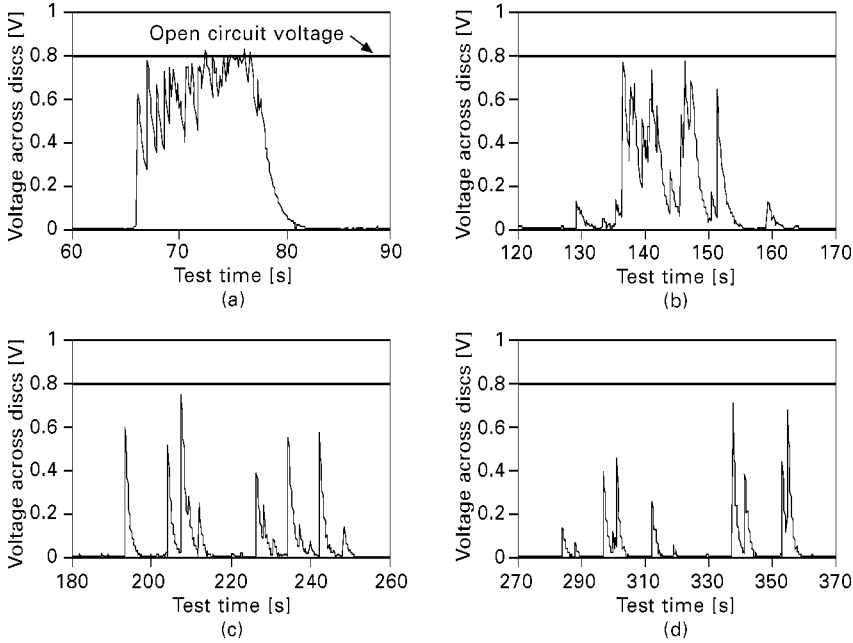
18.8 Sand application to the twin-disc contact.



18.9 RMS voltage for a test run dry with dry sand (0.5 mph, 1500 MPa, 20 % slip, input voltage of 1.6 V and sand flow rates of 0.75, 0.52, 0.25, 0.17 kg/min (left to right)).

above 0.4 kg/min almost complete open circuit voltage was recorded; and for rates below 0.4 kg/min voltage changes were intermittent.

With a wet disc contact, the voltage was above the closed circuit value for longer, and greater complete disc isolation was seen than with dry tests at the

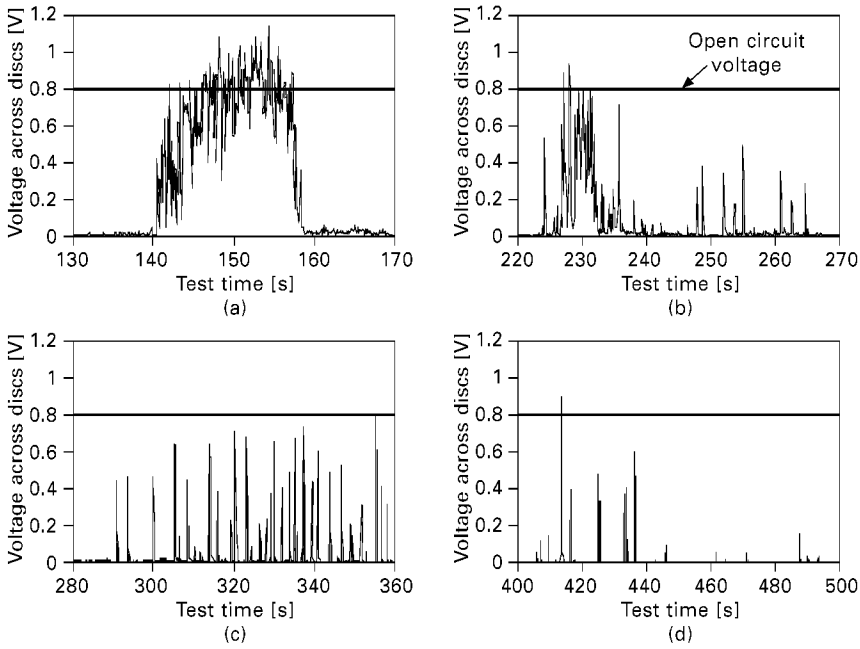


18.10 RMS voltage plots for a test run dry with dry sand (0.5 mph, 1500 MPa, 20 % slip, input voltage of 1.6 V and sand flow rates of: (a) 0.75 kg/min; (b) 0.52 kg/min; (c) 0.25 kg/min; (d) 0.17 kg/min).

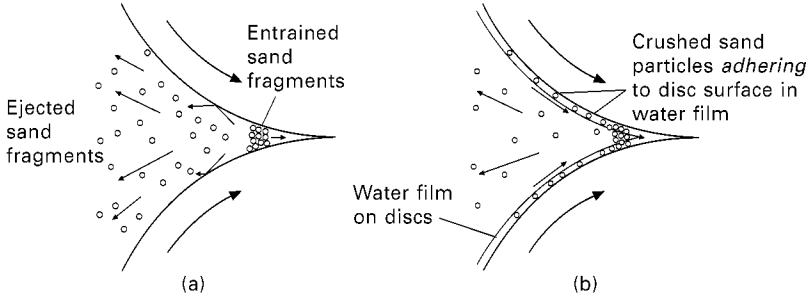
same sand flow rate. This may be because the water makes the sand clump together and also adhere to the rail surface better. Figure 18.12 illustrates how crushed sand particles may be ejected from a dry contact, but pulled into a wet contact after adhering to the water film on the disc surfaces.

Figure 18.13 shows how the average conductance (calculated from voltage reading using Eq. 18.1) varies with sand flow rate for wet and dry tests. It is clear that, as seen with the voltage plots, a transition occurs at a sand flow rate of 0.40 kg/min for tests run at 2 mph, sand flow rates below 0.40 kg/min giving much better conductance at the contact than those above. For dry tests, better conductance occurred at a surface speed of 0.5 mph than at 2 mph. This suggests that sand entrainment was greater at the higher speed. For the wet tests, surface speed had no effect on conductance. With wet discs, it is possible that the sand particles were pulled into the contact in the water film on the discs. This effect is probably overriding any speed effects.

Static testing was also carried out using the test set-up described in Section 18.3.2. The amount of sand in the contact was varied and different contact loads were used. Clear transitions were again seen, as shown in Fig. 18.14, where results for tests with a 60 kN load are illustrated. The wet sand gives



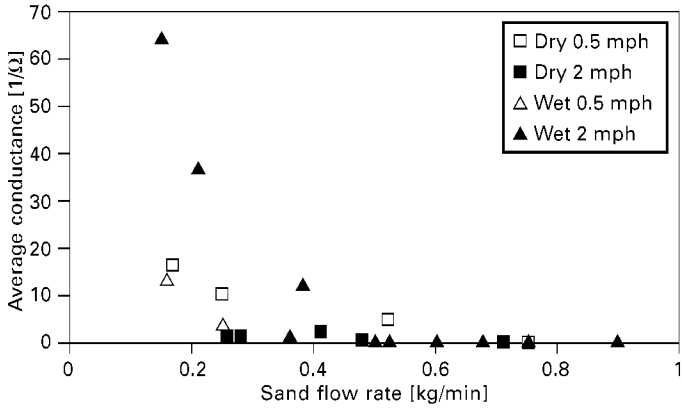
18.11 RMS voltage plots for a test run with water and sand (2 mph, 1500 MPa, 20 % slip, input voltage of 1.6 V and sand flow rates of: (a) 0.68 kg/min; (b) 0.38 kg/min; (c) 0.21 kg/min; (d) 0.15 kg/min).



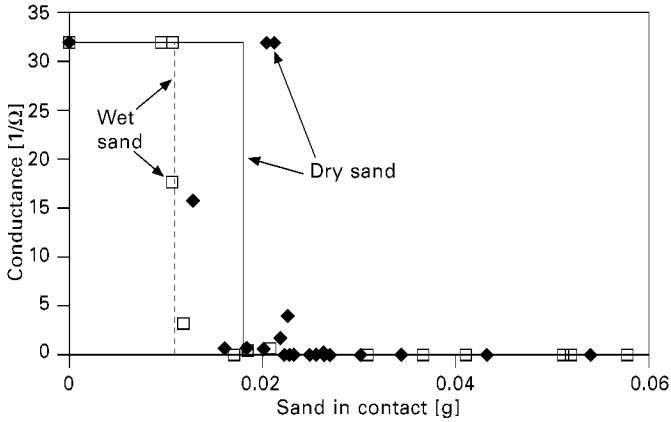
18.12 Particle entrainment in to: (a) dry disc contact; (b) wet disc contact.

conductance earlier as the water causes the sand to cling together and present a greater barrier to conductance.

Critical sand flow rate clearly depends on the disc speed as sand entrainment varies with speed. Given that sand entrainment at 2 mph is greater than that at 0.5 mph, at higher disc speeds still, it could further increase giving a much lower critical sand density.



18.13 Average conductance against sand flow rate for wet and dry tests.



18.14 Conductance against amount of sand in the model wheel–rail contact for dry and wet conditions at 60 kN.

It is the quantity of sand per unit area covering a surface which will determine the conductivity. So it is preferable to use mass/area as a means to compare the critical flow rates rather than mass/distance or the amount of sand fed per unit time. Calculating the critical mass/area of sand for the 2 mph disc test gives 0.75 kg/m^2 (for a *track* width of 10 mm). The critical mass/area of sand spread on actual rail at 10 mph (assuming a track width of 50 mm and using 7.5 g/m) is 0.15 kg/m^2 (critical sand concentrations for test and rail conditions are summarised in Table 18.1).

These calculations ignore dispersion of the sand, so actual sand rates entering the contact will be much lower. In the dynamic test, more sand enters the contact than would in an actual contact. This is mainly because the

Table 18.1 Critical sand concentrations for test and rail conditions

Test	Critical sand concentration (kg/m^2)
Static	0.3
Dynamic (2 mph)	0.75
Rail (10 mph)	0.15

disc geometry is smaller than a wheel–rail contact and both are in rotational motion and the sand nozzle was placed closer to the interface. The critical amount of sand in the static test leading to isolation equates to 0.3 kg/m^2. This figure, as well as that calculated for the twin-disc contact, are above that used in practice, which suggests that above 10 mph train identification should be unhindered.

There are a number of idealisations inherent in the test methods, particularly the dynamic technique. Results are therefore only to be taken as a guide to what happens in the full-size wheel–rail interface. However, as mentioned previously, it is suggested that the test method used here represents a severe case. Both the geometry and feed method will tend to entrain more sand particles into the contact, and the electrical circuitry with its high sampling rate and relatively low inductance will be more sensitive to transient contact resistance fluctuations.

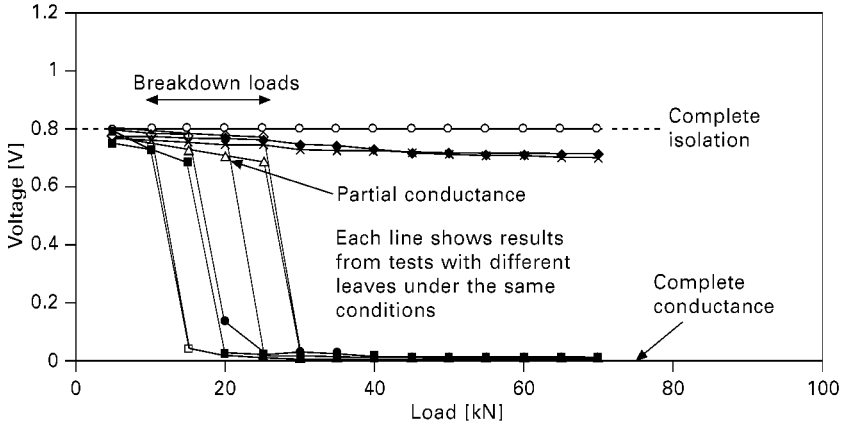
Given these limitations, it is likely that, at their present stage of development, they are best used as a means to qualitatively assess the relative effects on electrical isolation of different contaminants.

18.4.2 Leaves

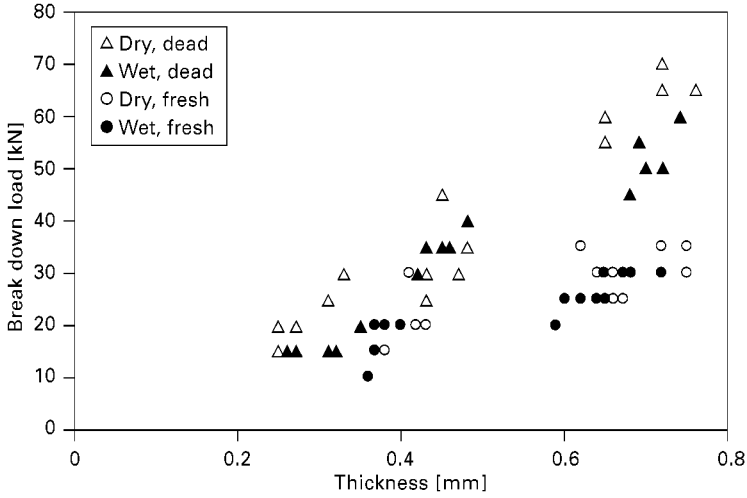
The effect of leaves in the contact was studied using the static test technique. Voltage across the wheel and rail sections was measured as load was increased (see Fig. 18.15). Breakdown loads, below which isolation occurred, for different leaf layer thicknesses indicated that lightweight rail vehicles may be isolated if leaves are present on the railhead (see Fig. 18.16). These data also show that dry leaves are far more likely to isolate than wet leaves and fresh leaves are easier to break down than dead leaves.

18.4.3 Solid friction modifier

In order to duplicate the solid stick friction modifier transfer mechanism found in the field, a bespoke applicator was designed which would apply a customised solid stick onto the rotating surface of the wheel disc (see Fig. 18.17). Similar to field applications, it was applied under a constant force

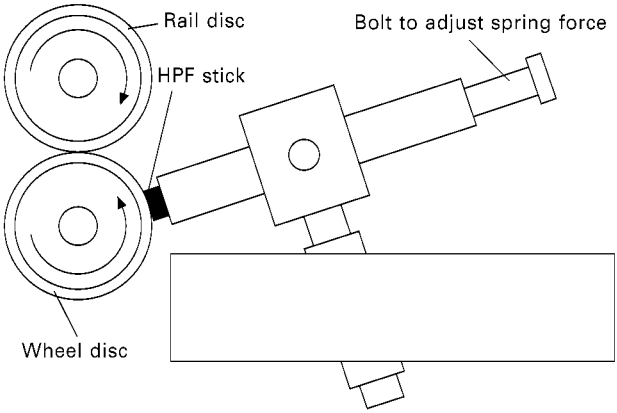


18.15 Voltage reading against load for a single dead leaf layer.

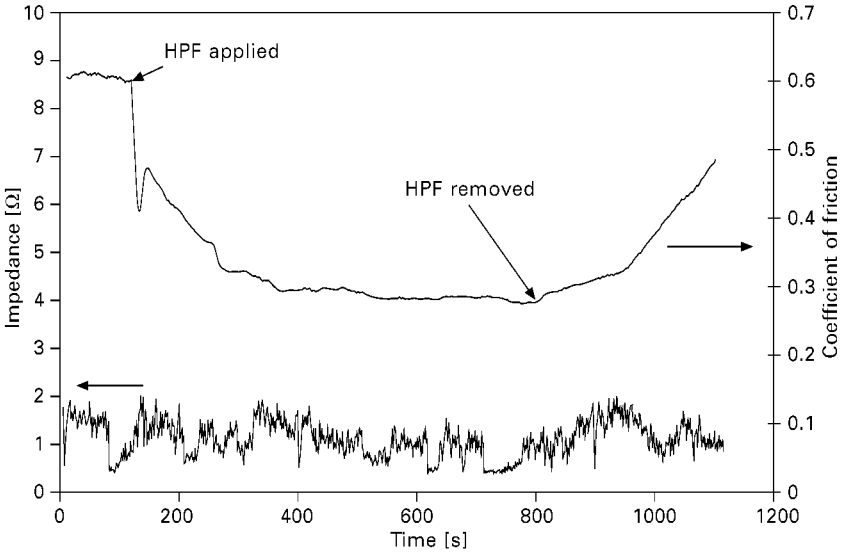


18.16 Leaf thickness against break down load.

(spring). The diameter of the solid stick was 10 mm in order to ensure that the product transferred across the entire surface of the rail disc. Tests were carried out at a contact pressure of 900 MPa and slip values of 0.1, 1 and 3 %. The friction modifier used (high positive friction – HPF) comprises the active components embedded in a thermoset polymeric matrix. No conductance improving medium is included in the matrix. Figure 18.18 shows data from the 3 % slip test, which clearly show that the friction modifier has no affect on the contact impedance. The same was seen in all other tests. For static tests, six points were marked on discs made from wheel and rail material to ensure contact was always made at the same place. During the



18.17 Disc arrangement and friction modifier applicator.



18.18 Friction and contact impedance during 3 % slip test.

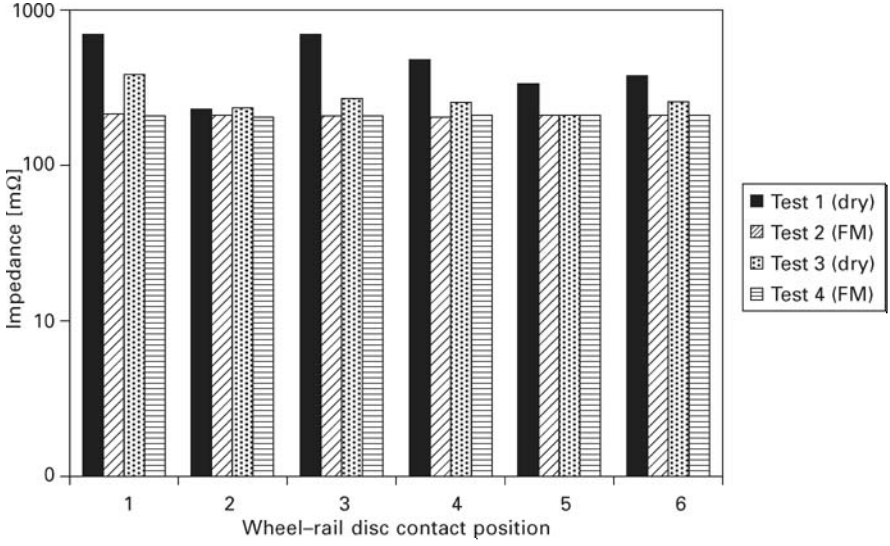
measurements the same two points on each disc were brought into contact. The discs were conditioned at the start of each test by running them at 3 % slip with a contact pressure of 900 MPa until a stable friction coefficient was reached. The dry impedance measurements were taken at this stage while loading the wheel and rail discs together statically at the six contact positions with a pressure of 900 MPa.

The discs were then reconditioned by running dry before the solid stick friction modifier was applied to generate a surface film. The conditioning was

stopped when the friction coefficient had re-stabilised. The impedance was then measured under static load at the same six points. The whole process was then repeated, so in total there were four sets of measurements taken; two dry and two with friction modifier.

Static conditions perhaps represent the worst-case scenario in terms of isolation as, with traction present under the dynamic conditions, there is a mechanism to remove material present in the contact, which means metal-to-metal contact is more likely to occur. Figure 18.19, however, shows that the static impedance measurements taken at the six disc contact positions are relatively similar for dry and friction modifier conditions. In fact the friction modifier values were lower than those for dry conditions. This may have been due to the presence of debris on the disc surfaces. The key result, however, is that introduction of friction modifier to the contact surfaces does not increase the contact resistance.

To sum up, dynamic twin-disc testing under typical wheel–rail contact pressures and slips has shown that the application of HPF does not affect contact impedance. Impedance was shown to increase with decreasing slip in the contact. This is as would be expected as higher slip will give greater metal-to-metal contact. Static tests have also shown that, even with no traction in the contact that would enhance metal-to-metal contact, the presence of HPF does not affect impedance. This means that HPF application under the conditions applied here would not affect train identification.



18.19 Measured static impedance values for dry and friction modifier (FM) conditioned wheel–rail discs.

18.5 Modelling approaches

The disc interface has been modelled using two approaches (Lewis *et al.*, 2003), which are outlined in Sections 18.5.1 and 18.5.2. The first approach assumes full disc separation by a contaminant layer. The second is based on a calculation of the amount of metal-to-metal contact likely with partial disc separation by sand particles. Here both are outlined for sand contamination. Recent work (Descartes *et al.*, 2008) has proposed the use of a new electromechanical modelling approach for determining the resistance of third-body layers in a wheel–rail contact. This is discussed in Section 18.5.3.

18.5.1 Wheel–rail surfaces separated by sand

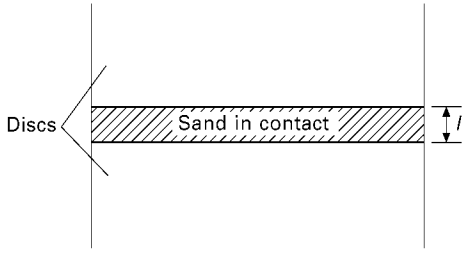
For the situation when the two discs are separated by a thin layer of sand (with a thickness, l, equal to the size of one fractured sand fragment) (as shown in Fig. 18.20), the contact resistance can be given by:

$$R = \frac{\rho l}{A} \qquad [18.2]$$

where ρ = resistivity (of the sand layer and A = area. The size of a crushed sand fragment will be dictated by the fracture toughness and the size of any flaws in the material. An estimate of the minimum fragment size after crushing in the disc contact can be obtained from the stress surrounding the particle, α, and the size of the largest flaw, a, in the material of fracture toughness K_{IC}:

$$K_{IC} = Y\sigma\sqrt{\pi a} \qquad [18.3]$$

where Y is a constant depending on the crack geometry (e.g. Y = 1.12 for an edge crack or 0.6 for a semi-circular flaw). When a sand particle is in the disc contact, it can be assumed that it will be subjected to a maximum stress equal to the hardness of the disc material (2.9 GPa for the rail). Using a fracture toughness value of 1.5 MPa$\sqrt{m}$ for the sand gives a maximum flaw



18.20 Wheel–rail discs separated by sand.

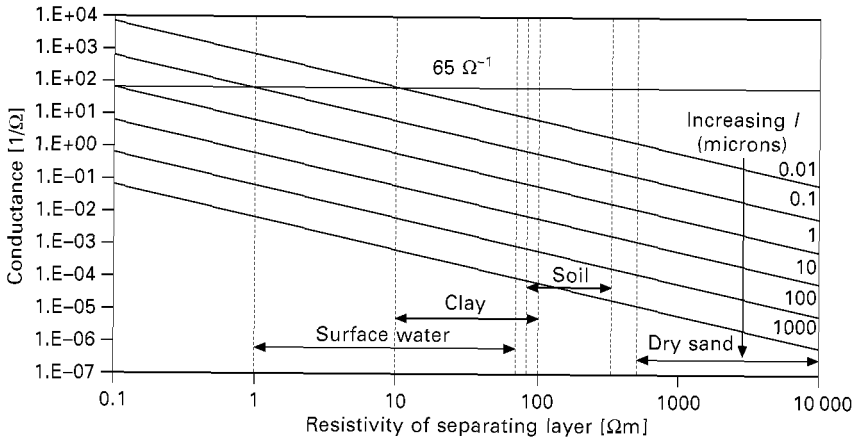
which will not propagate to fracture. Thus a crude estimate for the smallest possible surviving fragment can be obtained as 0.1–0.2 μm.

Figure 18.21 shows how conductance at the twin-disc interface changes with resistivity of the separating layer between the discs for varying layer thickness. Resistivity values for a number of materials are plotted to indicate how conductance varies with different separating layers and to determine how water alone affects conductance. The maximum conductance calculated for the tests carried out is also shown (65 Ω^{-1}).

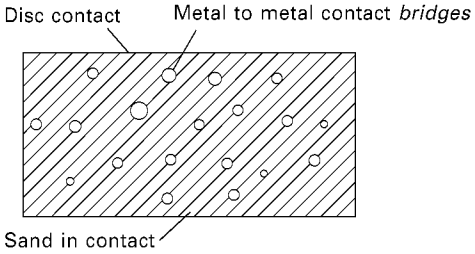
As can be seen, resistivity for particular materials can vary across orders of magnitude so only approximate conductance values can be determined. It is clear though that, if the disc surfaces are fully separated by a layer of sand of only one grain thickness (~0.1 μm), there is negligible conductance. However, when the surfaces are separated by a layer of dry sand of similar thickness mixed with sufficient water, it would seem possible that the conductance is much greater and could reach the highest levels recorded during testing (up to 65 Ω^{-1}). Indeed, during testing it was seen that, at low sand flow rates-higher conductances were seen with wet tests than with those run dry.

18.5.2 Wheel–rail surfaces in partial contact

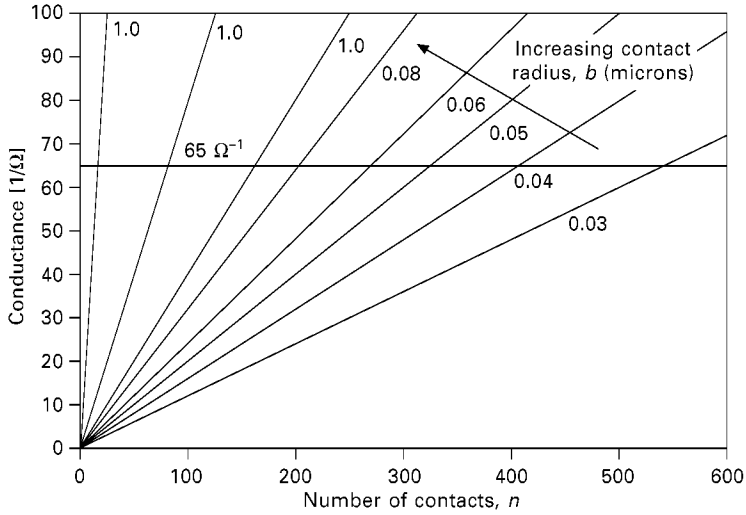
Bowden and Tabor (1950) showed that the electrical resistance of the interface between two contacting metal bodies could provide a measure of the real area of contact at the interface. The assumption was made that, if two surfaces are supported on n equal *bridges* of radius b (see Fig. 18.22), the contact resistance when the *bridges* are relatively far apart is given by:



18.21 Conductance against resistivity for varying layer thickness.



18.22 Schematic of disc contact showing metal to metal contact bridges.



18.23 Number of contacts (bridges) against conductance for varying contact radius.

$$R = \frac{\rho}{2bn} \qquad [18.4]$$

where ρ is the resistivity of the contacting metal bodies (Ωm). It is possible to calculate, for the disc contact, how many *bridges* there are when sand is present from the resistance or conductance data presented in Fig. 18.22, using Eq. 18.4.

Figure 18.23 shows the relationship between the number of *bridges* and the conductance at the disc interface (calculated using $\rho = 5 \times 10^{-7}$ Ωm – (typical value for steel). The maximum conductance calculated for the tests carried out is also shown (65 Ω^{-1}). Assuming that a contact will not be smaller than a crushed sand fragment (0.1–0.2 μm) gives the lowest *bridge* radii to be expected as 0.05–0.1 μm. This means that the the largest number of

contacts will be approximately 300 (looking at values below the 65 Ω^{-1} line). With larger contact *bridges* ($b \geq 0.1$ μm), far fewer contacts would be expected.

These numbers seem reasonable; however, they represent only a very small proportion of the total contact area. It is clear that, for conductance to occur, only a very small amount of metal to metal contact is necessary. In other words, practically, isolation will occur only when the surfaces are completely separated by sand.

18.5.3 Electrodiscrete element method

Descartes *et al.* (2008) have proposed a new method for modelling the isolating effects of a third-body layer in the wheel–rail contact. This involves creating a 'mesh' of particles representing constituents of a third-body layer, some of which are conducting and some isolating. Calculations have been carried out for two moving surfaces and have shown that the technique has some potential, but needs further development.

18.6 Conclusions

The potential for wheel–rail isolation is a safety risk that clearly needs more investigation. The use of friction modifiers and enhancers is growing, for example, and their safe use must be proven before they go out into the field. The test methods outlined here, while severe, offer the only available solution at present. Modelling of the contact resistance is rather simplistic at present, although the electrodiscrete element method outlined above shows some promise.

18.7 References

Beagley, T.M., Pritchard, C., 1975, Wheel/rail adhesion – the overriding influence of water, *Wear*, **35**, 299–313.

Bowden, F.P., Tabor, D. (1950), *The Friction and Lubrication of Solids*, Clarendon Press, Oxford, UK.

Descartes, S., Desrayaud, C., Niccolini, E., Berthier, Y., 2005, presence and role of the third body in a wheel/rail contact, *Wear*, **258**, 1081–90.

Descartes, S., Renouf, M., Fillot, N., Gautier, B., Descamps, A., Bertier, Y., Demanche, Ph., 2008, A new mechanical-electrical approach of the wheel/rail contact, *Wear*, **265**, 1408–16.

Eadie, D.T., Santoro, M., Kalousek, J., 2005, Railway noise and the effect of top of rail liquid friction modifiers: changes in changes of sound and vibration spectral distributions in curves, *Wear*, **258**, 1148–55.

Eadie, D.T., Oldknow, K., Maglalang, L., Makowsky, T., Reiff, R., Sroba, P., Powell, W., 2006, Implementation of wayside top of rail friction control on North American

heavy haul railways, *Proceedings 7th World Congress on Railway Research*, Montreal, ac, Canada, 4–8 June, on CD.

Fletcher, D.I., Beynon, J.H., 2000, Development of a Machine for Closely Controlled Rolling Contact Fatigue and Wear Testing, *Journal of Testing and Evaluation*, **28**(4), 267–75.

Gallardo-Hernandez, E.A., Lewis, R., 2008, Twin disc assessment of wheel/rail adhesion, *Wear*, **265**, 1309–16.

Garnham, J.E., Beynon, J.H., 1991, The early detection of rolling-sliding contact fatigue cracks *Wear*, **144**, 103–16.

Kalousek, J., Magel, E., 1999, Modifying and managing friction, *Railway and Track Structures*, **5**, 5–6.

Lewis, R., Dwyer-Joyce, R.S., 2004, Wear mechanisms and transitions in railway wheel steels, *Proceedings of the IMechE, Part J: Journal of Engineering Tribology*, **218**, 467–78.

Lewis, R., Masing, J., 2006, Static wheel/rail contact isolation due to track contamination, *Proceedings of the IMechE, Part F: Journal of Rail and Rapid Transit*, **220**(1), 43–53.

Lewis, R., Dwyer-Joyce, R.S., Lewis, J., 2003, Disc machine study of contact isolation during railway track sanding, *Proceedings of the IMechE, Part F: Journal of Rail and Rapid Transit*, **217**, 11–24.

Lewis, R., Gallardo, E.A., Eadie, D.T., Cotter, J., 2008, The effect of friction modifiers on wheel/rail isolation, *Proceedings 2008 ASME/IEEE Joint Rail Conference*, Wilmington, DE, USA, 22–23 April.

Nagase, K., 1989, A study of adhesion between the rails and running wheels on main lines: results of investigations by slipping adhesion test bogie, *Proceedings of the IMechE, Part F: Journal of Rail and Rapid Transit*, **203**, 33–43.

Olofsson, U., Sundvall, K., 2004, Influence of leaf, humidity and applied lubrication on friction in the wheel–rail contact: pin-on-disc experiments, *Proceedings of the IMechE, Part F: Journal of Rail and Rapid Transit*, **218**, 235–42.

Pritchard, C., Tanvir, M.A., 1973, *A Basic Study of Wheel/Rail Adhesion – 9: Further Observations on Leaves*, Report No. TM TRIB 6, British Rail Research, Derby, UK.

Railway Safety, 2001, *Railway Group Standard for the Sanding Equipment Fitted to Multiple Units and On-Track Machines*, GM/RT2461, Issue 1, London, UK.

Copyrighted Materials
Copyright © 2009 Woodhead Publishing Retrieved from www.knovel.com

19
Air-borne particles from the wheel–rail contact

M. GUSTAFSSON, VTI – Swedish National Road and Transport Research Institute, Sweden

Abstract: Inhalable particles in the ambient air are considered a major threat to public health. Although generally environmentally friendly, railway traffic is, in some environments, a significant particle-source of particles primarily emanating from the rail–wheel contact and brake systems while secondary sources are pantographs and voltage lines. This chapter deals with the emission, dispersion and effects of these wear particles in relation to their sources. Finally, possible measures to present high particle concentrations are discussed and future issues outlined.

Key words: inhalable particles; railway traffic wear; air quality; health effects; measures against particles.

19.1 Introduction

Inhalable particles, so-called PM_{10}, are today regarded to be amongst the most serious public-health environmental problems. Road traffic is one important source, but rail traffic also produces PM_{10} which, in certain environments can cause high, potentially harmful concentrations. Most of these particles emanate from the wear of wheels, rails and brakes. This chapter is concerned with current knowledge about inhalable particles from rail traffic. The available research data do not allow us to restrict the account to the contact between wheels and rails alone. Most studies have been carried out in station environments where the contribution from braking can be considerable, and several attempts at source apportionment have been made. The literature basis is international, but with a deeper insight in Swedish research, not yet published internationally.

Initially a brief background is given to the current interest in wear particles from railway traffic. In the section on concentrations we describe what concentrations of particles that have been recorded in different railway environments and say something about studies of the dispersion patterns of the particles. Then we describe how the understanding of the properties and behaviour of the particles has been linked to the different possible sources in the railway. The dispersion section gives an overview of data on how particles disperse in rail road environments, while the following section gives a survey of the, still few, investigations of the possible health effects

of railway particles. Finally we make a survey of known and possible actions to lower the concentration of particles in railway environments where they are high.

19.2 Background

Particle pollution from rail-bound traffic has attracted attention as a problem mainly in tunnel environments, but also as dust fall-out near railways above ground. Most often the problems have been observed as a sharp 'smell of brakes' at underground stations or in similar places. Dust problems around railway stations have also had effects on cars parked nearby that have had their paint-finish damaged by dust fall-out (Holm, 1999).

As the importance of particle pollution for public health has received increased attention, and as an international demand has arisen for lowered concentrations of inhalable particles, the focus of studies of particle pollution from railways has also been directed towards the inhalable fraction PM_{10} (particles smaller than 10 μm). The directives of the EU (European Council, 1996) state that the PM_{10} concentration can exceed 50 $\mu g/m^3$ 35 days per year, and that the annual average must not exceed 40 $\mu g/m^3$. This norm is for outdoor air, i.e. not in road or railway tunnels. It is planned to supplement the current norm with one for $PM_{2.5}$.

Karlsson *et al.* (2005) published cell studies showing that the iron (FE)-rich particles from the underground have a considerably greater potential to cause DNA fractures than particles from a street environment. Another study complicated the picture, however, when it was shown on the same particle material that the street particles were more highly inflammatory than the underground particles (Lindbom *et al.*, 2006; Gustafsson *et al.*, 2008). Ongoing exposure studies with healthy test subjects in the Stockholm underground do not, however, indicate that particles in the underground have negative health effects (Larsson, 2008).

19.3 Concentrations

Numerous studies have been made in which particle concentrations have been measured in underground stations, in underground trains and in railway environments above ground. It is difficult to compare the results, as the measurements often differ with regard to the particle fractions studied, instrument technology and measurement strategy. Most studies of particle concentrations from rail traffic have been done in underground environments, and measurements in environments above ground are rare. Gehrig *et al.* (2007) noted a measurable contribution from the railway traffic to the PM_{10} concentration in Zürich. The contribution was measurable only in the immediate vicinity of the track area. Iron dust was judged to be the only

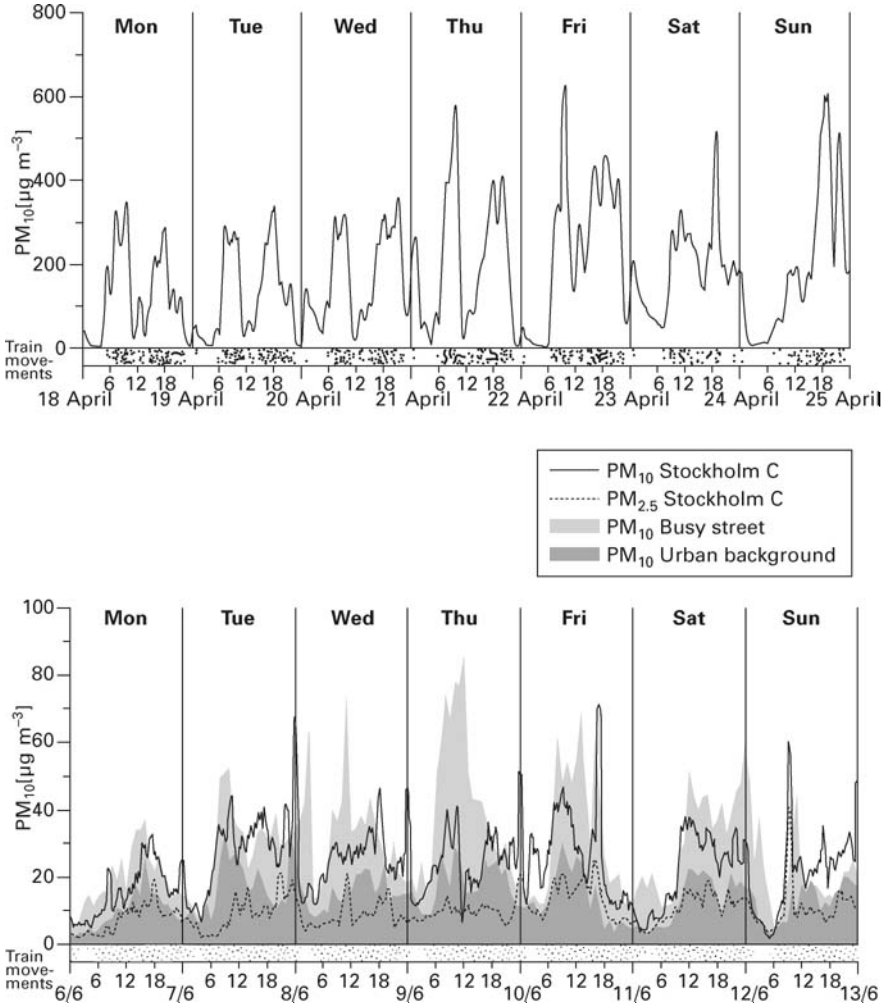
significant contribution to the PM_{10} concentrations in the environmental air, and was measured near an intensively frequented railway at about 1 $\mu g\ m^{-3}$ as a long-term average. The contribution 120 m from the track was only 25 % of the level 10 m from the track. In Gustafsson *et al.* (2006) the concentrations of PM_{10} and $PM_{2.5}$ were measured at four stations above ground and two stations below ground. The measurements were carried out during summertime above ground and during the spring in the tunnel stations. The average concentrations are shown in Table 19.1.

These concentrations can be compared to concentrations measured in the Stockholm underground ($357 \pm 185\ \mu g/m^3$) (Johansson and Johansson, 2003) and at the Liseberg station in Gothenburg ($120\ \mu g/m^3$) (Sjöstedt, 2005). In the underground systems of Helsinki (Aarnio *et al.*, 2005), London (Seaton *et al.*, 2005) and New York (Chillrud *et al.*, 2004) the $PM_{2.5}$ concentrations have been measured to be 47–60, 270–480 and $62\ \mu g/m^3$, respectively. Nieuwenhuijsen *et al.* (2007) observed in a compilation of existing measurements of particles in underground systems in 12 cities around the world that many certainly had higher particle concentrations in the tunnel environment than in the street environment, but that several cities also reported lower concentrations (for instance Hong Kong, Guangzhou and Mexico City). The causes of these differences may be found in the material in wheels and braking systems and in differences in ventilation.

The variation in concentration over time depends in general on the traffic situation; thus there is a diurnal and a weekly variation. The patterns are most noticeable where the railway source dominates strongly, such as in a tunnel environment, but can also be discerned in railway environments above ground (Fig. 19.1). In Gustafsson *et al.* (2006) it was shown that the daily variations in concentration at the tunnel station Arlanda C (Fig. 19.1) were very great, with hourly averages from a few $\mu g/m^3$ between about 2 and 5 a.m., when there is no traffic, to about 600 $\mu g/m^3$ during periods in the mornings and afternoons. The diurnal variation was very similar on both weekdays and Saturday–Sundays and clearly depended on the traffic intensity. When individual 24-hour periods were studied at Arlanda C the

Table 19.1 Mean PM_{10} concentrations at six station environments in Sweden (Gustafsson *et al.*, 2006)

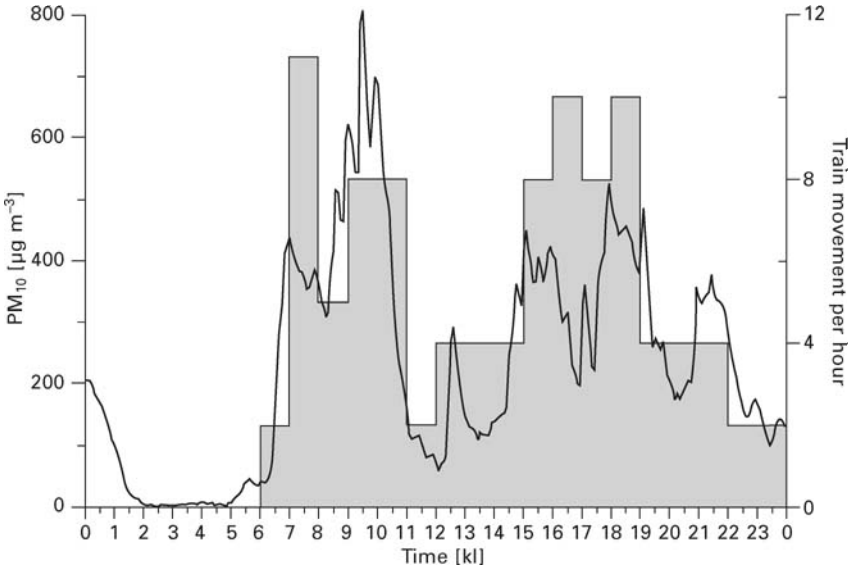
City	PM_{10}
Arlanda C (tunnel)	$237 \pm 75\ \mu g/m^3$
Arlanda S (tunnel)	$88 \pm 22\ \mu g/m^3$
Stockholm C	$25 \pm 7\ \mu g/m^3$
Linköping	$23 \pm 5\ \mu g/m^3$
Lund	$19 \pm 4\ \mu g/m^3$
Eslöv	$19 \pm 4\ \mu g/m^3$



19.1 PM_{10} concentration at Arlanda Central (a tunnel station) during one week (top) and PM_{10} and $PM_{2.5}$ concentrations at Stockholm Central Station during a summer week. PM_{10} at a busy street (Hornsgatan) and at an urban background site for comparison (bottom). At the bottom of the diagram train movements are plotted. (Gustafsson *et al.*, 2006)

manifest co-variation between particle concentration and train traffic was clearly evident (Fig. 19.2).

In measurements at Stockholm Central Station (Gustafsson *et al.*, 2006) there was also a clear diurnal variation (Fig. 19.1). It was judged to be difficult to separate the train-induced variation from the variation in particle concentrations from car traffic, which has a similar pattern. Bukowiecki *et al.*



19.2 PM_{10} concentration and trains per hour at Arlanda C. (Gustafsson *et al.*, 2006)

(2007) measured particle concentrations along a section where no braking occurred, in order to more clearly expose particle contributions from rail and wheel rather than from braking. It was not possible to calculate the contribution from rail traffic to the PM_{10} concentration in the air, but the comparison to train type and frequency data showed that 75 and 60 % of the Iron (Fe) and manganese (Mn) mass concentrations related to cargo and passenger trains, respectively, were found in the coarse fraction (2.5–10 µm). At the same measuring location, Gehrig *et al.* (2007) demonstrated that the contribution of rail traffic to the PM_{10} concentration in the air was approximately 4.3 µg m^{-3}, of which 65 % was Fe, 23 % aluminium (Al) and 10 % calcium (Ca). The two latter elements were considered to emanate from abrasion and re-suspension of particles from the macadam of the rail embankment.

Apart from PM_{10} and $PM_{2.5}$, some investigations have also studied the concentrations of ultra-fine particles in, above all, underground rail environments. In Paris, Mazoue and Parfait (2006) measured concentrations of ultra-fine particles, but concluded that they are the result of air exchange with the polluted city air, as the frequency distributions were very similar. The underground station has an evacuation fan, something that would make the exchange easier. Gustafsson *et al.* (2006) could establish that ultra-fine particles were related to the train traffic at the underground stations at Arlanda Airport. The concentrations were affected by the train traffic, but certain trains contributed marked concentrations of ultra-fine particles while others affected the concentration very little. Norman and Johansson (2005) also measured

concentrations of ultra-fine particles in the underground system in Stockholm, which were generally somewhat higher than at Arlanda. Particularly high concentrations were noted at nighttime and could be attributed to diesel-engined service trains. The origin of the daytime concentrations was unclear, however. A possible source, according to Norman and Johansson (2005), could be electrical discharges at the train's contact with the conductor rail. If electrical discharges are the source also for the ultra-fine particles observed in the Arlanda tunnels, this indicates that there must be very great differences in how many electrical discharges different train sets cause.

In contrast to the references above, Braniš (2006) considers that the results of his measurements in the underground system of Prague indicate that the outdoor air probably is an important source of the high particle concentrations in the underground. The measurements were carried out after the underground system had been cleaned up after the floods of 2002, which, according to the author indicates that re-suspension has probably given a very minor contribution. The concentrations are consistently lower than in other underground studies, an average of 102.7 $\mu g/m^3$ on the platform and somewhat higher on the trains (113.7 $\mu g/m^3$). The correlation between the PM_{10} concentrations on the platform and on the trains was very high ($r^2 = 0.96$). The correlation with the outdoor air was also relatively high, with a coefficient of determination of 0.82 for the platform and 0.77 for the trains. The author suggests that these facts, taken together, indicate a common source of particles in the tunnel system but also a strong influence from the pollution of the street traffic.

19.4 Particle properties and sources

The operation of trains on the railway is the one dominant process that emits particles to the surrounding air by wear of wheels, rails, brakes, conductor rails and voltage lines, but also by re-suspension of dust from the embankment. In a survey study of particle concentrations in the underground of Stockholm, Johansson *et al.* (2001a) found that the concentrations of PM_{10} and $PM_{2.5}$ were five and ten times higher, respectively, in the underground than in an urban street environment. In a follow-up study the concentrations were as much as ten to twenty times higher (Norman and Johansson, 2005). In Johansson *et al.* (2001a) $PM_{10-2.5}$ constituted 45 % of PM_{10}. In continued studies (Christensson *et al.*, 2002) of particle concentrations in the underground of Stockholm it was found that PM_1 constituted about 10 % of PM_{10}. Norman and Johansson (2005) measured the same fraction to be 15 % of the mass, while $PM_{2.5-1}$ constituted as much as 50–60 % and $PM_{10-2.5}$ about 25–35 %. As particles larger than 1 μm more or less disappeared when train traffic ceased and the concentrations instead were dominated by particles smaller than 1 μm, it was concluded that the train traffic contributed most all to particles larger

than 1 μm. The total concentration of PM_{10} was also considerably lower at night. A similar pattern was demonstrated by Okinaga *et al.* (2000) in the underground of Tokyo.

Tokarek *et al.* (2002), studying numbers of particles in an underground station in Paris, concluded that 99 % of the number of particles there were smaller than 2.5 μm and that 50–60 % were smaller than 0.35 μm. In contrast to the concentrations of mass, Norman and Johansson (2005) established that the concentration of numbers was about five to ten times lower in the underground than in the street environment, because the contribution from car exhausts is considerably lower. An important health-related conclusion is that in the underground one is exposed to a considerably lower number of particles smaller than 0.5 μm, but at the same time a considerably greater mass concentration of particles larger than 1 μm, compared with a street environment.

Christensson *et al.* (2002) also carried out comparative studies of the composition of chemical elements in PM_{10}, PM_1 and the trains' brake shoes, in order to find out whether the last were an important source of the particles. The contents of the various elements in PM_{10} and PM_1 were in a linear relationship to one another, indicating that the sources for the fractions are the same and/or have the same process of formation. About 62 % of PM_{10} consisted of iron oxides. Other metals identified were Ca, Al, barium (Ba) and copper (Cu). The Cu concentrations especially were distinctly enriched in comparison with geological material. Mn was also enriched in PM_{10}. Ba in the form of barium sulphate and barytes has been related to brakes in the underground of Tokyo in earlier studies (Furuya *et al.*, 2001). Copper and antimony (Sb) were especially studied to find whether braking dust was an important source. Earlier studies on road traffic have shown that these elements can be used as tracing elements for braking dust (Sternbeck *et al.*, 2001). PM_{10} in the underground contained considerably higher concentrations of Cu and Sb than PM_{10} in a road environment, however, and also considerably higher concentrations than in the brake shoes of the Stockholm underground trains. For those reasons, the brakes were not considered to be a likely source. As the ratio between Fe and Mn (Fe:Mn) in the rails was very similar to that in PM_{10} (112:1), the rails were deemed to be a likely source. The current collectors and conductor rails, on the other hand, had a different quotient (200:1). Al, Ca, Mg and titanium (Ti) presented a strong mutual correlation. Together, brakes and rails were estimated to explain 75 %, of which the brakes were estimated to account for about 15 %. Ten per cent was deemed to be geological material.

Christensson and Anker (2004) carried out a more detailed study of the composition of the inhalable fractions in the Stockholm underground. Electron microscopy showed that small and sub-micron particles were common, while particles larger than 10 μm were unusual. At daytime the

particles were considerably more numerous than at night, morphologically more heterogeneous, often agglomerated and with a high Fe content. It was considered not clear whether these were mechanically generated or the result of condensation. An analysis indicated that condensation and reaction to crystalline inorganic ferrous particles is an important process, but the thesis could not be verified. The concentration of organic particles at daytime is so high that one or several sources of these must exist in the underground. They consist largely of plastic, paint, rubber, cellulose and exhausts from outdoor air. About 15 % of the particles are of organic origin (this is true for both PM_{10} and PM_1). Still, about 15 % of the particles are chemically unidentified. In contrast to Christensson *et al.* (2002), the most prevalent elements were Fe, sulphur (S), silicon (Si), Cu and zirconium (Zr). Of the iron not quite 60 % is magnetite, a little over 35 % is hematite and 5–10 % metallic.

Ripanucci *et al.* (2006) studied PM_{10}, respirable fraction (RF), respirable combustible fraction (or respirable organic fraction) and the contents of organic, metallic, siliceous and fibrous material in the particles in the underground of Rome. The underground trains have an emergency system which sprays sand on the tracks to increase friction. The system is activated when the maximum permitted speed at a section is exceeded for a certain time. About 300 tonnes of sand is used every year for this purpose. Ripanucci *et al.* (2006) found that the particle concentration at the platforms and in the tunnels was about three times as high as outside the tunnels, the highest diurnal average being 479 $\mu g/m^3$. The RF made up between 61 and 85 % of PM_{10}. The organic RF, on its part, made up between 19 and 52 % of the total RF. Analysis in an electronic microscope of settled dust showed that the dominating elemental contents were Fe, Si, Ca, potassium (K) and Al, with lower presence of Cu and Mn. There were also traces of chlorine (Cl), lead (Pb), zinc (Zn), Mg, nickel (Ni), Sb, cobalt (Co), arsenic (As) and chromium (Cr). Compared with the outside air, the enrichment of Fe was +1346 %, Mn +1800 % and Cu +752 %. To a lower extent Cr (+128 %), Sb (+500 %), Ni (+131 %) and Zn (+77 %) were enriched.

In studies carried out in the tunnels under Arlanda Airport (Gustafsson *et al.*, 2006), it was shown that the size distributions are somewhat different between Arlanda Central (C), frequented by a variety of long-distance trains, and Arlanda South (S) which is used only by the commuter train Arlanda Express. The coarse particles had a mass maximum around 5–7 μm at Arlanda C and 2–3 μm at Arlanda S. Compared with the size distribution of the particles in the underground, in the above-mentioned study by Norman and Johansson (2005), the particles at Arlanda thus seem to be somewhat larger. Why the particles at Arlanda S are somewhat smaller than those at Arlanda C is not clear, but the authors think that it most likely depends on the respective break-up properties in the main sources.

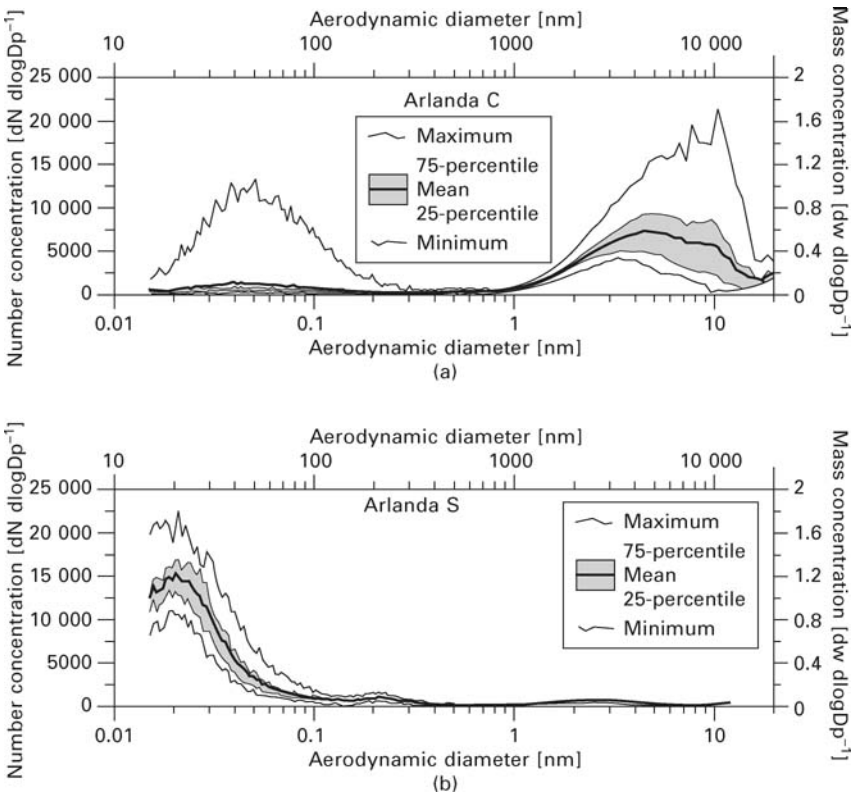
In a chemical elements analysis of the collected particles at Arlanda C and S (Gustafsson *et al.*, 2006), Fe dominated, but with evident contributions of among others Mn, Zn, Cu, Ni, Sb and also molybdenum (Mo) (the latter only at Arlanda C, however). Most of these metals have been mentioned in the studies already accounted for. Already Trattner *et al.* (1975) could show that the dust in the underground of Newark, USA, contained Mn, Fe, Ni, Cu, Zn and lead (Pb). Aarnio *et al.* (2005) showed that iron was the most highly enriched element in the underground system of Helsinki but that also Mn, Cr, Ni and Cu were enriched in $PM_{2.5}$, something that agrees well with the high concentrations of these metals at the Arlanda stations (Gustafsson *et al.*, 2006). In the tunnel at the Liseberg station in Gothenburg, Fe is also what dominates overwhelmingly, followed by Ca, Cu, Mn and Zn and, in lesser concentrations, K, Ti, Cr and Ni (Sjöstedt, 2005).

Norman and Johansson (2005) also studied the size distribution in a more precise allocation in the interval 10–120 nm. There was a frequency maximum at 60–80 nm, which is coarser than the ultra-fine particles observed by Gustafsson *et al.* (2006) in the railway tunnels under Arlanda Airport. These particles had a size distribution with maximum numbers around 20–50 nm at Arlanda C and around 10–20 nm at Arlanda S (Fig. 19.3).

In a study by Johansson (2005), the main purpose was to estimate the contributions of various sources to the particle concentrations, something that is of great interest for this literature study. In Johansson's study it was estimated that about 60 % of PM_{10} consisted of iron oxides and metallic iron. Ca, Ba, Cu and Al were also identified, but it was also judged that 10–20 % of the particles were of organic material, which was believed to emanate from human clothing, etc. Birenzvige *et al.* (2003) estimated the organic part to be < 1 % at an underground station in Washington, DC.

Johansson (2005) analysed the chemical composition of brake shoes (two different types), wheels, current collectors and conductor rail in the underground trains (Figs 19.4 and 19.5). The compositions of the various sources are apparently very different. According to calculations carried out using so-called chemical mass balance, it was concluded that only three of the analysed materials gave significant contributions to the concentrations of PM_{10}. The major contribution comes from the wheels, which were estimated to contribute as much as 70 % of the particles, while the brake shoes contributed a little less than 30 % (16 and 10 %, respectively). Relatively great unreliability is, however, attached to both the regression analysis and the chemical analyses.

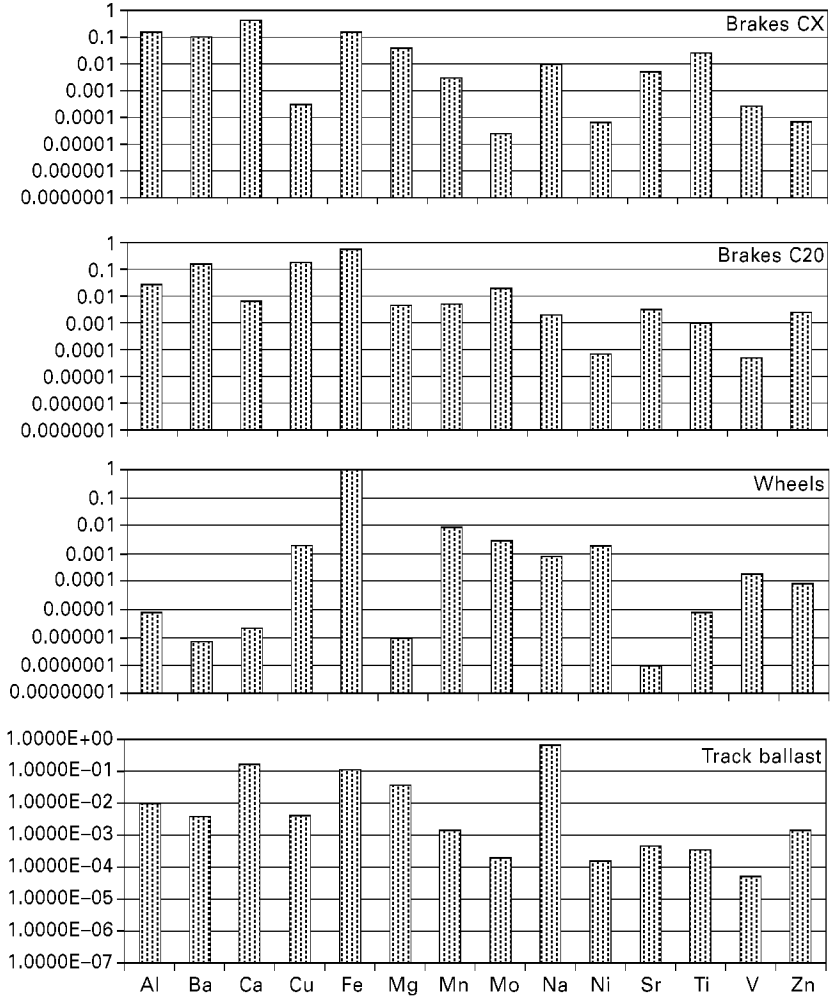
It has also been attempted to model the generation and sources of particles by measuring the particle concentrations underneath a moving underground train at the same time as speed, acceleration, moment response, curvature of the rail, braking and environment (tunnel or above ground) were registered and calculated (Furusjö and Christensson, 2005). The analysis was carried



19.3 Time excerpt (6–7 h p.m.) of number and mass distributions at Arlanda Central (a) and Arlanda South (b). (Gustafsson *et al.*, 2006)

out with the multivariate PLS (partial least square) method. The factors which were most important, according to the model, were environment followed by mechanical braking and speed. The large-scale changes could be modelled, but only 49 % of the variation, indicating that other factors play an important role. This is especially so for particularly high concentrations, which are believed to partly depend on meeting trains.

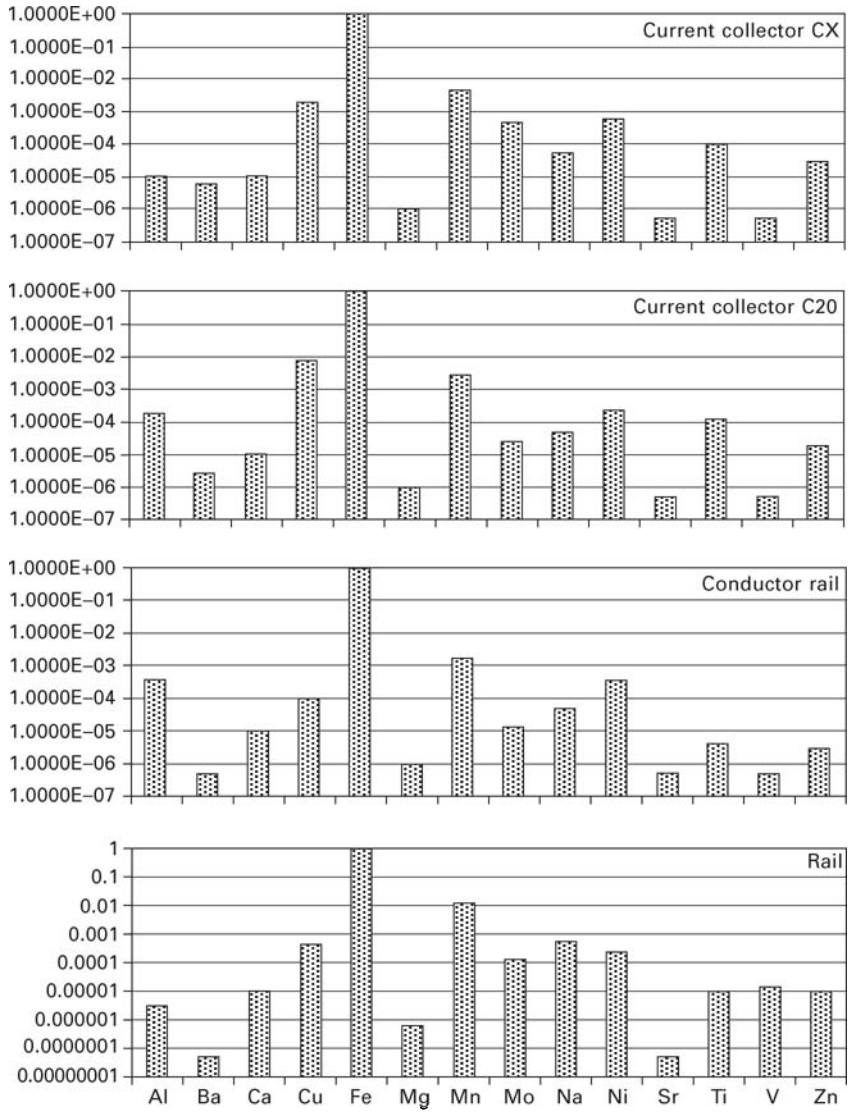
Gustafsson *et al.* (2006) suggested the hypothesis that in the railway tunnels underneath Arlanda Airport Fe, Mn and probably Ni have a common source (wheels and/or rails) and Cu and Zn another (brakes). This is based on the fact that a linear regression of the elements in the samples from the two stations at Arlanda shows that Fe and Mn have a very high regression coefficient (1.0), and the same is true for Cu and Zn (0.98). The regression coefficient, between Fe and Cu, and between Fe and Zn, were not as high (0.73 and 0.68, respectively). This can indicate, according to the authors, that Fe and Mn have one common source and Cu and Zn another. The Fe/Mn source was deemed to be of a similar composition at the two stations,



19.4 Relative composition in brakes, wheels and embankment (track ballast) in the underground in Stockholm. In certain cases a minimal detectable metal content has been presumed. CX and C20 refers to train type. (Johansson, 2005)

while the Cu/Zn source appeared to be stronger at Arlanda S than at Arlanda C. The quotients between Fe and Mn and between Fe and Ni, respectively, were the same at the two stations. Cu and Zn displayed a greater proportion at Arlanda S but were more or less the same at the other stations. Ni is comparatively more prevalent at the stations above the ground.

The results in Johansson (2005) support this hypothesis, as Fe, Mn and Ni (together with Vanadium – V) are estimated to largely emanate from the wheels, while Cu and Zn (together with Mo and Ba) almost exclusively originate from one of the two analysed brake shoes (C20). An interesting



19.5 Relative composition in current collectors, conductor rail and rails in the underground in Stockholm. In certain cases a minimal detectable metal content has been presumed. (Johansson, 2005)

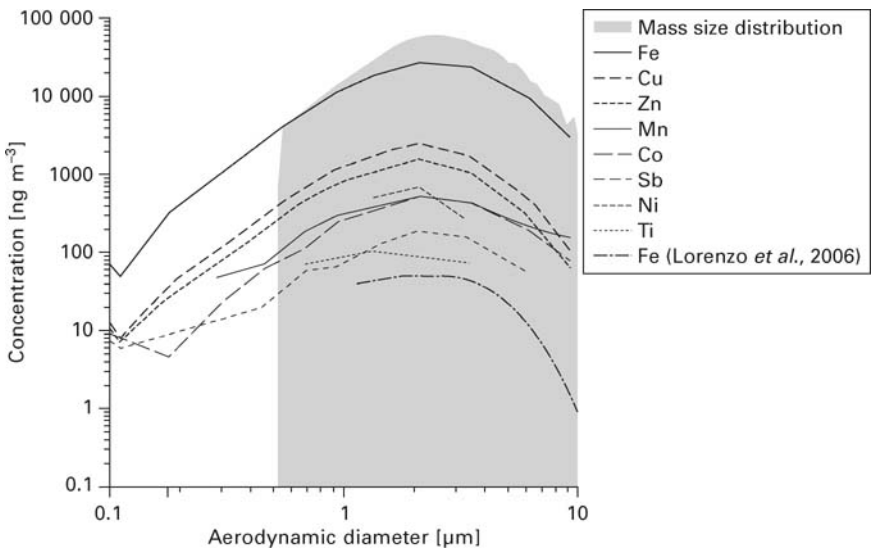
observation is that the dust on the embankment in the underground largely consists of sodium (Na), something that is not expected against the background of the geological material in the track ballast.

Gustafsson *et al.* (2006) also studied the elemental composition of different size fractions in air at Arlanda South. The results show that the elements Mn, Fe, Co, Ni, Cu and Zn are assembled in a group with very similar size

distributions (Fig. 19.6). These elements are found above all in somewhat coarse particles with a peak around 2–3 μm, which agrees well with the distribution of the mass concentration. It is plausible to assume that these six metals at Arlanda S have a common source and/or a common formation process. Sulphur has its highest relative concentration in the sub-micron fraction, while Si, Al and Cr between them have similar distributions with a peak in the fraction around 1 μm (Fig. 19.6). Si and Al are common in minerals, but their source in this tunnel environment is unknown.

Lorenzo *et al.* (2006) calculated the size distribution of PM_{10} next to a railway in Zürich with the help of an electron microscope. This is, as far as we know, the only existing study of size distributions next to a railway above ground where normally no braking occurs. The size distribution for the Fe particles is very similar to what was measured by Gustafsson *et al.* (2006) in a tunnel station environment, with a maximum between 2.5 and 4 μm (Fig 19.6). Particles dominated by Ca had a similar size distribution, while Al particles were somewhat more coarse and Si particles were mainly found in $PM_{2.5}$ (Lorenzo *et al.*, 2006).

According to Sitzman *et al.* (1999) the high content of Fe and Si, as well as Ca- and K-rich particle types, in dust from the underground in London emanates from brake shoes. These contain much Fe and glass fibre and have



19.6 Elements with a concentration distribution agreeing with the mass distribution from the tunnel station Arlanda S (partly from Gustafsson *et al.*, 2006). For comparison the size distribution of ferrous particles from a measuring spot 10 m from a railway above ground been added (Lorenzo *et al.*, 2006).

a binding agent containing Ca and K. Sitzman *et al.* (1999) estimate that not quite 70 % of the particles in PM_{10} emanate from brakes, wheels and rails. Weckwerth (2001) is of the opinion that Cu in the city environment of Cologne in the vicinity of the tramway can be traced to the tramway's voltage lines.

Glivberg (2004) presents a comprehensive analysis of the composition of different brake linings used on trains and carriages. Ten of the most common brake linings in Sweden were analysed with regard to both organic and inorganic components. Glivberg concluded that Fe and Cu are the dominating metals, but that there is much variation between the linings. Compared with brake linings for cars, the average proportion of Cu is low. High proportions of Zn (up to 8 %) were found in most linings. It is remarkable that arsenic (As) has been identified in several linings, and also that only about half of the contents of the linings have been identified. The latter is believed to be because many substances which are not perceived to be environmentally harmful have not been included in the analyses (compounds with Si, Mg, Al and Na). Referring back to the discussion of different main sources for Fe, Mn, Ni and Cu, Zn, respectively, Glivberg (2004) shows that the proportion Fe to Mn is in the same order of size for all the ten investigated brake linings for rail vehicles. The contents of Cu and Zn, however, can vary several tenth-power steps between linings.

With measurements in the underground of New York, Chillrud *et al.* (2004) could show that above all Fe, Mn and Cr were enriched in $PM_{2.5}$, but also Al, Cu, Sn, Ag and Sb. The most highly enriched metals were deemed to emanate from steel. It is surprising that Ni was not enriched, indicating that a certain type of steel low in Ni was used.

In Paris, it has been observed that the different rail systems give different particle concentrations and different content distributions in the particle contents of the air (Jouannique *et al.*, 2006). The underground (Métro with iron wheels) and RER (suburban commuting trains with iron wheels) have approximately the same concentrations of PM_{10}, while the underground with tyres (Métro roues à pneus) generally exhibits lower, but still rather high, particle concentrations (about 100 $\mu g/m^3$ as a median value). If the elemental concentrations in PM_{10} are compared, the differences are greater. The underground with tyres has the lowest concentrations of Fe, Ni, Cr, Mn, Zn and Ba but the highest concentration of Cu. The ordinary underground has the highest concentrations of Pb and the lowest concentrations of Cu, while the commuting trains had the highest concentrations of Ni, Fe, Cr and Mn. Tokarek *et al.* (2002) state that 98 % of the particle mass at an underground station in Paris consists of 'ash'. The largest fraction consists of 52 % of 'iron ash', 9 % Cu and traces of Zn, Ma, Pb, S and Si. It is not clear what is meant by ash in this reference.

An article by Weckwerth (2001) describes the emission of particles to

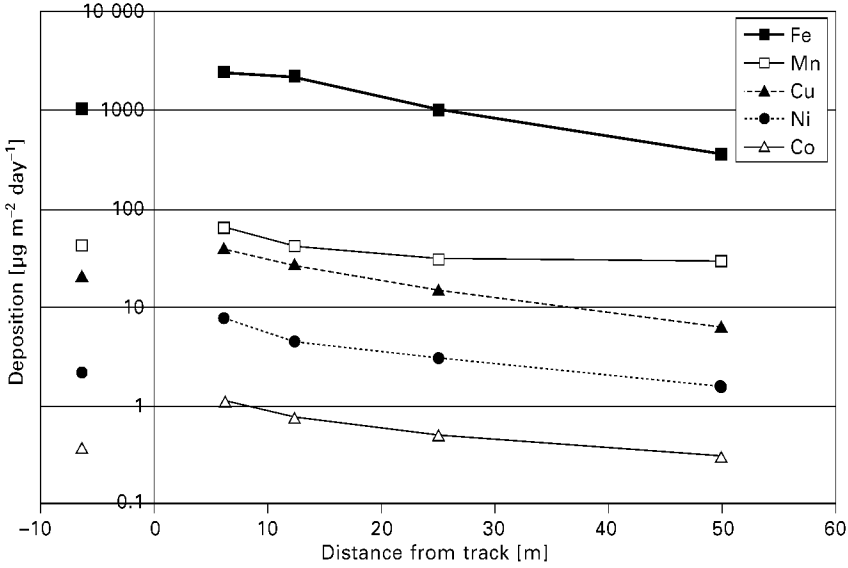
urban air. The article describes the potential to identify different sources of particles. A combination of grain size and colour could be used to determine the source of particles, for instance diesel soot consists of small black particles which can cause measurable blackening of filters. The quotients of various elements can also be used. The fact that the quotient Cu/Sb in brake linings is manifestly different from the same quotient in the Earth's crust can be used as a clear indication of brake linings and can be traced relatively far from the source of emission. Other elements or combinations of elements can only be traced near the source; Cu from tramways, Platinum (Pt) from catalytic converters and As from rusting rails. Small particles are often the result of combustion, while larger particles can be caused by friction. Of the larger particles (> 2.5 μm), the proportions in urban air are estimated to be 30 % from diesel, 3 % from tyres and 0.3 % from brake linings.

19.5 Dispersion

In studies in the underground system in Stockholm, for instance (Norman and Johansson, 2005), and at Arlanda (Gustafsson *et al.*, 2006) it has been estimated that a large part of the dust that has been measured has been directly emitted. This is based mainly on the absent effect of tunnel wash. No more detailed studies have been made, however, in which moisture in tunnel walls and embankment has been followed up. Ripanucci *et al.* (2006), however, concluded on the basis of their studies in the underground of Rome that re-suspension was a significant factor and that the generation of dust could be reduced by more efficiently removing dust from the tracks.

Gustafsson *et al.* (2003) studied the dispersion of metals from the railway by measuring deposition along profiles perpendicular to the railway at three localities in Sweden. The studies were carried out at a section with no braking in the province of Småland, where the railway was supposed to be the dominating source, at a section with some braking near the city of Linköping and at the Main Southern Line in the northern parts of the city of Lund. Considering the prevalent dispersal conditions, it was concluded that at least 15 metals that could be related to the railway were dispersed up to about 50–100 ms before the deposition had gone down to background level. The metals were Fe, As, Pb, Co, Cu, Cr, Mn, Ni, V, gallium (Ga), germanium (Ge), niobium (Nb), Mo, rhodium (Rh) and Sb. Several of these metals have also been identified in dust in tunnel environments. Others are more unexpected, for instance Ga, Ge, Nb and Rh. Examples of the deposition profiles for the most distinct metals in Lund are shown in Fig. 19.7.

The dispersion of the metals in the air was manifestly related to weather and traffic conditions. Drier and warmer weather in combination with frequent traffic (such as the measurements in Lund) gave higher deposition and more



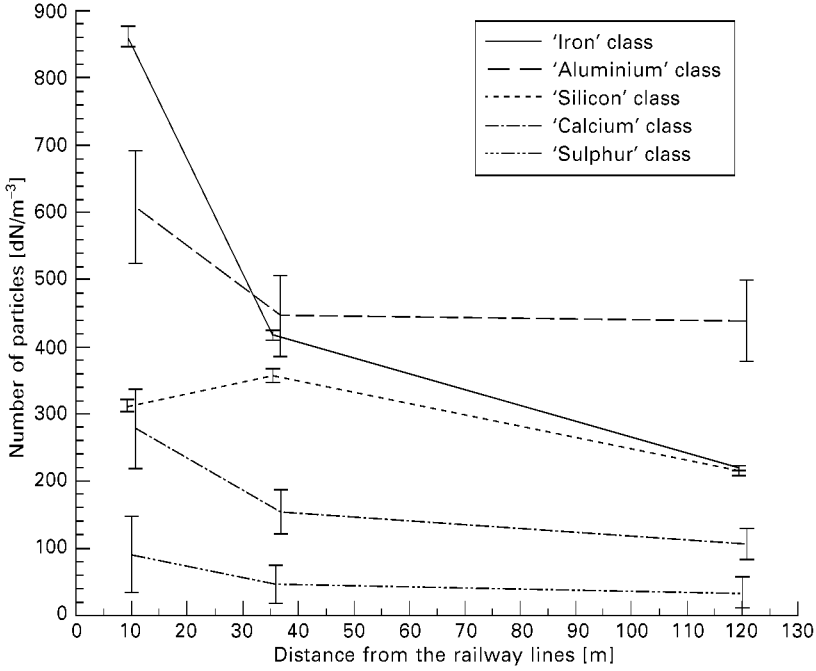
19.7 Deposition gradients for the metals iron, cobalt, copper, manganese and nickel downwind the Main Southern Line in Lund during a measurement 19 June until 5 July 2002. (Gustafsson *et al.*, 2003)

distinct relationship to the railway than moist and cold with little traffic. It was also clear that the prevailing wind direction affected the dispersion, so dispersion by wind was considered an important process.

Lorenzo *et al.* (2006) studied different classes of particles in PM_{10} at three distances from a heavily frequented railway section in Zürich, Switzerland. The most distinct relationship to the railway was evidenced by particles in the Fe class, but also particles dominated by Al and Ca evidenced higher concentrations near the railway (Fig. 19.8). The Fe particles were estimated not to have gone down to background level 120 m from the track, which was the most distant measuring site, whereas Al and Ca were estimated to have gone down to background level already at the measuring site 36 m from the track.

19.6 Health effects

Karlsson *et al.* (2005) and Lindbom *et al.* (2006) studied the cell effects of PM_{10} from a street environment, pure road-wear, underground travel environment and diesel particles. Karlsson studied DNA damage in cells, whereas Lindborn and others studied the capacity of particles to induce inflammatory effects. The studies yielded different pictures of the health effects of underground particles. While Karlsson *et al.* (2005) showed that



19.8 Distance profiles for the five particle classes. The error bars are calculated from the standard deviation of three sampling sites at 10 m distance from the railway lines. The values at 10 m are average values from the same three sampling sites. (Lorenzo *et al.*, 2006)

particles from the underground have a considerably greater capacity to cause damage to DNA, Lindbom *et al.* (2006) demonstrated that the same kind of particles was the least inflammatory of the tested particles. A possible explanation as to why the underground particles are more harmful to DNA may be that they mainly consist of magnetite (Fe_3O_2), while the particles from road traffic mostly consist of hematite (Fe_2O_3) (Karlsson *et al.*, 2005). Magnetite has been shown to emit more free radicals than hematite (Fubini and Mollo, 1995). Exposure studies in the underground of Stockholm (Larsson, 2008) have not demonstrated any apparent health effects with healthy test subjects. The exposure studies display the inflammation-inducing properties of the particles and are, to that extent, an affirmation of the cell studies in Lindbom *et al.* (2006) and also Gustafsson *et al.* (2008).

Chillrud *et al.* (2005) have summed up the health effects of Fe, Ma and Cr, which are among the most highly enriched metals in underground environments. According to the author, exposure to Fe is suspected to have negative health effects because of its capacity to form free radicals. Changes in iron metabolism have been tied to diseases such as Parkinson's, Alzheimer's and multiple sclerosis. There is no epidemiological evidence,

however, that a relatively low exposure to air-borne Fe is related to these diseases. Welders have been shown to run greater risk for pneumonia, probably because Fe may act as a growth factor for bacteria. Manganese poisoning is known to cause Parkinsonism, while other health effects of Mn have been mainly related to pesticides containing Mn or to inorganic Mn dust or steam among steelworkers. Cr is a well-known carcinogen, studies from Cr industries indicate a connection with cancer in the respiratory tracts. Based on the total Cr concentrations observed among students commuting by underground, it can be calculated that life-long underground commuters may have an enhanced cancer risk to a degree of approximately 1; 100 000. People who work in the underground are likely to run a greater risk. There is great uncertainty, however, in the estimation, of how carcinogenic Cr in underground air is. Among other things, the oxidation phase of Cr plays a major part, as it has been demonstrated that Cr(VI) is considerably more carcinogenic than Cr(III) (for instance, Gorell *et al.*, 1999; Birgersson *et al.*, 1999). When welding stainless steel, practically all the Cr in the welding smoke is Cr(VI), something that indicates a dependence on temperature. Whether Cr in railway particles is emitted at high temperatures is not known.

Wear particles are a pollution in themselves, but can also act as conveyors of other pollutants in the form of gases or particles. For instance, Tomiyasu *et al.* (2004) state that particles from diesel exhausts, which are considerably smaller than wear particles, can be adsorbed on the surface of larger Fe particles emitted from a railway and in this way be carried down into the respiratory tracts.

Another important aspect when estimating the health risks of high particle concentrations in railway tunnels is the exposure time. Most persons who are exposed to high concentrations in tunnels are so for a very limited time. The understanding of short-term effects of particles on human health is not so comprehensive as that related to the epidemiological long-term studies on which the air quality directives. There is, therefore, much uncertainty about the effects of these particles, and there is great need to have them elucidated.

19.7 Measures

To act on particle emission at the source involves affecting the wear properties of the material, techniques or the manner in which the trains are run in order to reduce or minimise the emissions. In the existing studies, only a few actions to limit the dispersion of air-borne particles have been presented. Actions to reduce the very emission from the sources in the system seem not to have been discussed. Such actions could be related to the properties of the material in brakes, wheels and rails, and also to a development of techniques and changes in how and where trains are run and braked. At present, only occasional analyses of material have been carried out, and

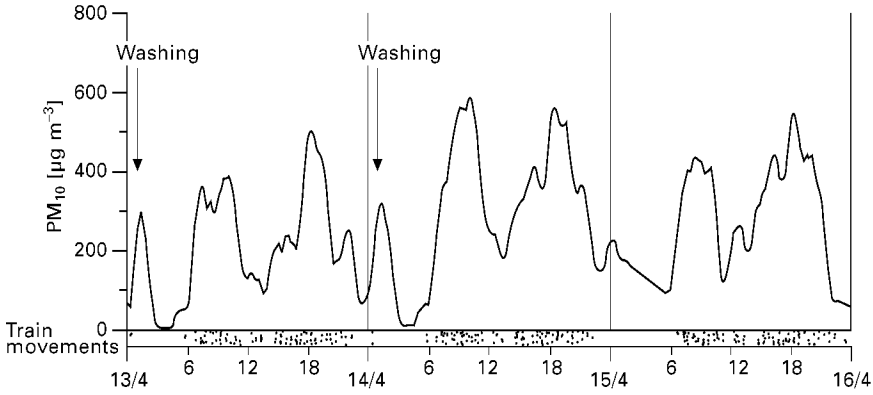
then with the main purpose of identifying substances potentially harmful to health or the environment.

In situations where sand is used to enhance the effect of acute braking Ripanucci *et al.* (2006) suggest that, in order to reduce the silicon content of the dust, some other substance be used. At the same time, a braking system is suggested which would reduce the need for acute braking to a minimum, in order to reduce exposure to the particles, and that personnel be trained to avoid acute braking as far as possible. Gustafsson *et al.* (2006) demonstrated that at Arlanda S, frequented only by the commuting trains between Arlanda and Stockholm, the particle concentrations (PM_{10}) were considerably lower than at Arlanda C which is frequented by long-distance trains. This is so despite the fact that there are many more train movements every day at Arlanda S. One reason can be differences in the braking systems, the commuting trains employing electromagnetic brakes while most of the long-distance trains have wheel brakes.

Thus, understanding of how to affect the formation of particles is limited. Therefore, studies have mostly concerned means to prevent the particles from reaching the inhaled air. Because whirled-up dust was estimated to be an important factor in tunnel environments with high concentrations, a couple of studies have been carried out to study the effects of flushing the tunnels with water. Johansson *et al.* (2001b) measured the concentrations of PM_{10} and $PM_{2.5}$ at an underground station in Stockholm. The walls and embankment were flushed and the particle concentrations were followed up. The results show that the flushing had little effect. At daytime (7 a.m. until 7 p.m.) the concentration of PM_{10} was reduced by 13 % and that of $PM_{2.5}$ by 10 %. It was not investigated how long the effect lasted.

Gustafsson *et al.* (2006) measured PM_{10} at the long-distance train station under Arlanda Airport. The two following nights, the tunnel was flushed by a cleaning vehicle. The effect on the PM_{10} concentration was nil. On the other hand, there was a strong particle maximum during the very flushing operation, probably the result of the emission of diesel particles from the diesel-driven cleaning vehicle but possibly partly also the effect of re-suspension (Fig. 19.9).

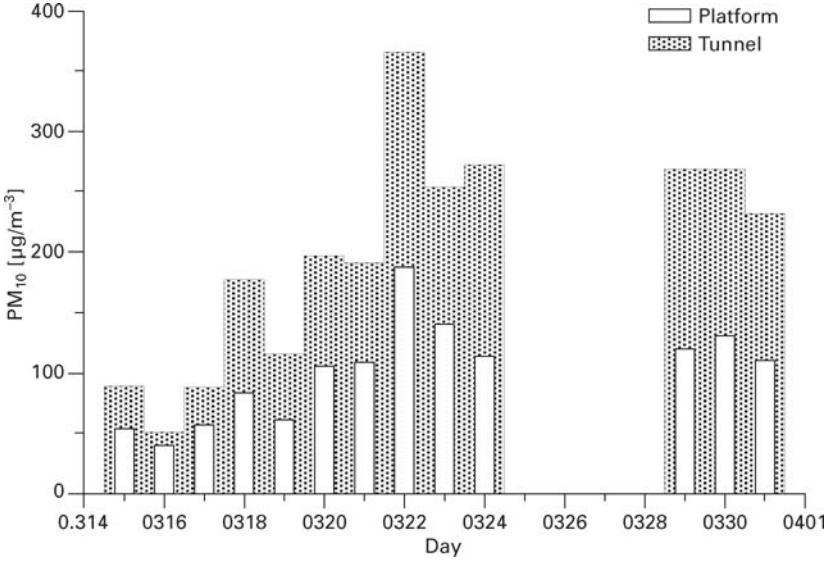
Another method to prevent the dispersion of particle emissions from one volume of air to another is to close off the train area from the platform. At the Liseberg underground station in Gothenburg, glass walls have been used between the platform and the tunnel (Fig. 19.10). In his study of the particle concentrations at this station Sjöstedt (2005) demonstrated that the concentrations on average were 120 $\mu g/m^3$ at the platform and about 250 $\mu g/m^3$ in the tunnel, during weekdays with normal traffic. This indicates that the glass wall, combined with excess pressure at the platform, reduced the average particle concentrations by a little more than half (Fig. 19.11). The correlation between the PM_{10} concentrations in the tunnel and at the



19.9 The effect on the PM_{10} concentration of tunnel washing at Arlanda Central. When the washing is carried out, the particle concentrations increase considerably because of particle emission by the diesel engine and resuspension. Note the otherwise very similar diurnal variation. (Gustafsson *et al.*, 2006)

19.10 The platform at the Liseberg underground station, Gothenburg. Note the glass wall towards the track. (Photo: Svante Sjöstedt, Environment Administration, City of Göteborg)

platform is quite high (r = 0.97), confirming that it is the train-produced particle concentrations that are reduced thanks to the glass wall. Chemical elemental analyses also indicate that the aerosols are more or less identical. For $PM_{2.5}$ the difference is not quite as evident. The $PM_{2.5}$ concentrations



19.11 The effect of the glass wall on PM_{10} concentration at the platform at the Liseberg underground station, Gothenburg. (Data from Sjöstedt, 2005)

in the tunnel vary from 30–80 µg/m^3, at the platform from 20–50 µg/m^3, indicating a reduction by the glass wall of somewhat over 30 %. The correlation between $PM_{2.5}$ at the platform and in the tunnel is not as high as for PM_{10} ($r = 0.76$), but it is still significant.

When long-distance or cargo trains pass the station without stopping, air is being pressed onto the platform, increasing particle concentrations. There is also a certain air exchange when the doors to the tunnel are opened. There are plans to install filters and fans in order to further reduce the influx of tunnel air onto the platform (Sjöstedt, 2005).

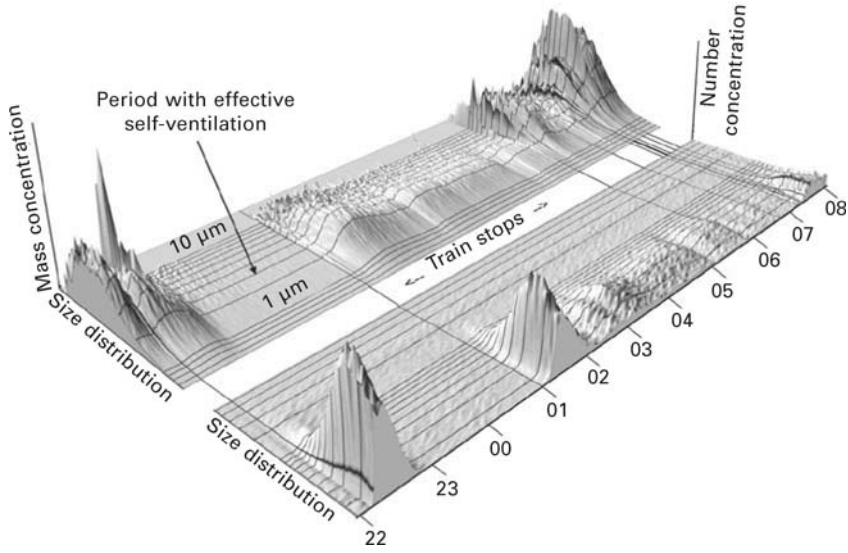
Ripanucci *et al.* (2006) suggested that to reduce particle dispersion in the underground of Rome one should install a better system to collect dust at track level. As the authors are of the opinion that re-suspension is very important in the underground, one can compare this to road traffic, where dust-binding on streets and roads has comparatively good effects in reducing PM_{10} concentrations. We have not been able to find any such studies on tunnels in the literature.

Studies on filter techniques as a measure against high particle concentrations are scarce. Tokarek and Bernis (2006) tested an electrostatic precipitator on a Paris subway station, concluding that it would take about 20 filters placed in the top of a non-mechanically ventilated station to permit a reduction to half an initial concentration in particles of 230 µg/m^3. Results also showed that, after approximately one year, the efficiency of the precipitator had

fallen by 15 % and that cleaning of the precipitation cartridges is needed to maintain their effectiveness.

One example of a very tangible effect of self-ventilation was documented by Gustafsson *et al.* (2006) at the Arlanda Central station underneath Arlanda Airport. The particle concentrations there accorded very closely with the changing traffic intensities during the day and night. During the traffic-free period of the night the concentrations fell generally, but on certain nights there were periods with exceptionally low concentrations. As it could be shown that all particle fractions fell simultaneously, one could conclude that the low concentrations were the result of ventilation (Fig. 19.12).

The tunnels underneath Arlanda are not mechanically ventilated, so probably self-ventilation occurred, generated by a gradient between the tunnel and outdoor temperatures and wind conditions outside the tunnels. A simple analysis of the impact of meteorology indicates that southerly winds were somewhat more common during the nights with particularly low particle concentrations. It is probable, however, that temperature differences play a greater role. During the night shown in Fig. 19.12 it was at most 12° warmer at the platform than in the air outside the tunnel. Unfortunately, there are no extended time series of temperature measurements in the tunnel environment, so a more careful analysis of the ventilation effect could not be made.



19.12 Mass distribution for coarse particles and numbers distribution for finer particles from late evening until morning at Arlanda Central. Note the period with very low concentrations of all particle sizes. (From Gustafsson *et al.*, 2006)

The investigations by Jouannique *et al.* (2006) of the properties of particles from different rail traffic systems in Paris are interesting, as they indicate that the system with rubber wheels, dominating steel does give somewhat lower concentrations of PM_{10} than the wheel variant, but the difference is small. On the other hand, the composition of the particles is completely different. Whether this difference is important from a health perspective is not clear.

19.8 Future trends

An important prerequisite for possible future actions against particles from railway traffic is to determine to what degree they are injurious to health. It is true that epidemiological studies show that inhalable particles, regardless of their origin and composition, have a negative impact on health. Nevertheless, it can be expected that there are differences in health effect between particles with different chemical and physical characteristics, differences that it should be possible to demonstrate in, for instance, toxicological studies. Studies so far do not provide an unambiguous answer as to whether particles specifically from railways are more or less hazardous to health than other particles. Cell studies are a good point of departure and can indicate whether particles can be expected to be dangerous to health. Studies that have been made indicate, however, that the results will vary depending on which type of cells and what effects are studied. Exposure studies of underground particles on healthy test subjects have not revealed any negative effects; studies on asthmatic persons are currently underway. Thus, much work remains to be done before the health effects of particles can be confidently estimated.

The sources of inhalable particles from electrified railway traffic are mainly rails, wheels, braking system, pantographs and voltage line. The two last mentioned are believed to give a marginal contribution to the mass concentration of particles. A small number of source apportionment studies show that the distribution varies depending on where the measurements have been carried out. In certain cases, the contact between wheels and rails is considered to be the main source, while other studies consider the brakes to be most important. In certain tunnel environments, the railway sources are believed to be of minor importance; high particle concentrations are considered to originate rather from street traffic. Local conditions, with reference to the character of the traffic, ventilation conditions and other affecting sources, have an effect on the concentrations and composition of particles. If effective action is to be taken to prevent the production of particles, it is therefore important that concentrations, composition and sources are known.

If the problem can be handled at the source rather than treating particle concentrations that already have arisen, this is to be preferred. Unfortunately, that is likely to be both more difficult and more expensive. The wear particles

formed at the wheel–rail contact can be reduced by lubricating the contact, reducing the load and the creep (relative sliding speed). Alternatives to wheel-braking with brake linings, or improved, less particle-generating materials in traditional braking systems, can reduce the particle contribution from braking. Driving manner may also reduce the formation of and exposure to particles, by performing the main braking further from the station area and possibly also more carefully.

In practice, ventilation and filtering can be used to handle particle concentrations when they already exist, but this carries with it large investments and great expenditure of energy. It should, however, be possible to design such solutions in future tunnel projects. To separate platforms from the track area and induce over-pressure at the platform to stop the air from the track area entering the platform should be one possible and relatively effective action.

19.9 References

Aarnio P, Yli-Tuomi T, Kousa A, Makela T, Hirsikko A, Hameri K, Räisänen M, Hillamo R, Koskentalo T and Jantunen M (2005), The concentrations and composition of and exposure to fine particles (PM2.5) in the Helsinki subway system, *Atmospheric Environment*, **39**(28), 5059–66.

Birenzvige A, Eversole J, Seaver M, Francesconi S, Valdes E and Kulaga H (2003), Aerosol characteristics in a subway environment, *Aerosol Science and Technology*, **37**(3), 210–20.

Birgersson B, Sterner O and Zimerson E (1999), *Kemiska hälsorisker: Toxikologi i kemiskt perspektiv*, Berlings, Arlöv, Sweden.

Braniš M (2006), The contribution of ambient sources to particulate pollution in spaces and trains of the Prague underground transport system, *Atmospheric Environment*, **40**, 348–56.

Bukowiecki N, Gehrig R, Hill M, Lienemann P, Zwicky C N, Buchmann B, Weingartner E and Baltensperger U (2007), Iron, manganese and copper emitted by cargo and passenger trains in Zürich (Switzerland): Size-segregated mass concentrations in ambient air, *Atmospheric Environment*, **41**, 878–89.

Chillrud S N, Epstein D, Ross J M, Sax S N, Pederson D, Spengler J D and Kinney P L (2004), Elevated airborne exposures of teenagers to manganese, chromium, and iron from steel dust and New York City's subway system, *Environmental Science and Technology*, **38**, 732–37.

Chillrud S N, Grass D, Ross J M, Coulibaly D, Slavkovich V, Epstein, Sonja D, Sax N, Pederson D, Johnson John D, Spengler D, Kinney P L, Simpson H J and Brandt-Rauf P (2005), Steel dust in the New York City subway system as a source of manganese, chromium, and iron exposures for transit workers, *Journal of Urban Health: Bulletin of the New York Academy of Medicine*, **82**, 33–42.

Christensson B and Ancker K (2004), *Luftburna partiklars sammansättning*, IVL Svenska Miljöinstitutet U1056, SL Infrateknik AB, Stockholm, Sweden (in Swedish).

Christensson B, Sternbeck J and Ancker K (2002), *Luftburna partiklar – partikelhalter, elementsammansättning och emissionskällor*, A22147 SL Infrateknik AB, Stockholm, Sweden (in Swedish).

European Council (1996), Council Directive 96/62/EC of 27 September 1996 on ambient air quality assessment and management, *OJ*, **L296**, 21-11-96, 55–63.

European Council (2008), Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe, *OJ*, L152, 11, 62008, p. 1–44.

Fubini B and Mollo L (1995), Role of iron in the reactivity of mineral fibers, *Toxicology Letters*, **82–83**, 951–60.

Furusjö E and Christensson B (2005), *Multivariat utvärdering av partikelmätningar under tunnelbanetåg*, U1782, IVL Svenska Miljöinstitutet, Stockholm, Sweden (in Swedish).

Furuya K, Kudo Y, Okinaga K, Yamuki M, Takahashi S, Araki Y and Hisamatsu Y (2001), Seasonal variation and their characterization of suspended particulate matter in the air of subway stations, *Journal of Trace and Microprobe Techniques*, **19**(4), 469–85.

Gehrig R, Hill M, Lienemann P, Zwicky C N, Bukowiecki N, Weingartner E, Baltensperger U and Buchmann B (2007), Contribution of railway traffic to local PM10 concentrations in Switzerland, *Atmospheric Environment*, **41**, 923–33.

Glivberg G (2004), *Analys av bromsbelägg*, Miljösektionen, Rapport 2004:1 Banverket, Borlänge, Sweden (in Swedish).

Gorell J M, Johnson C C, Rybicki B A, Peterson E L, Kortsha G X, Brown G G and Richardson R J (1999), Occupational exposure to manganese, copper, lead, iron, mercury and zinc and the risk of Parkinson's disease, *Neurotoxicology*, **20**, 239–47.

Gustafsson M, Blomqvist G, Franzén L and Rudell B (2003), *Föroreningsnedfall från järnvägstrafik*, VTI meddelande 947, VTI, Linköping, Sweden (in Swedish).

Gustafsson M, Blomqvist G, Dahl A, Gudmunsson A and Swietlicki E (2006), Inandningsbara partiklar i järnvägsmiljöer, VTI Rapport 538, VTI, Linköping, Sweden (in Swedish).

Gustafsson M, Blomqvist G, Gudmundsson A, Dahl A, Swietlicki E, Bohgard M, Lindbom J and Ljungman A (2008), Properties and toxicological effects of particles from the interaction between tyres, road pavement and winter traction material, *Science of the Total Environment*, **393**, 226–40.

Holm E (1999), *Miljökontrollrapport. Utredning avseende sotnedfall på bilar vid Västergatans anslutning till stambanan i Lund*, EMK Energi Miljö Konsult AB, Malmö, Sweden (consultancy report in Swedish).

Johansson C and Johansson P-A (2003), Particulate matter in the underground of Stockholm, *Atmospheric Environment*, **37**, 3–9.

Johansson C (2005), Källor till partiklar i Stockholms tunnelbana, SLB Rapport 6:2005. SLB-analys, Environment and Health Administration, Stockholm, Sweden (in Swedish)

Johansson C, Johansson P-Å and Sjövall B (2001a), *Partikelhalter i Stockholms tunnelbana*, SLB Rapport 2:2001 SLB-analys, Environment and Health Administration, Stockholm, Sweden (in Swedish).

Johansson C, Johansson P-Å and Sjövall B (2001b), *Partikelhalter i Stockholms tunnelbana – effekterna av spolning*, SLB Rapport 6:2001 SLB-analys, Environment and Health Administration, Stockholm, (in Swedish).

Jouannique V, Campagna D, Ihaddadene K, Hatton-Vaugier I, Trouillon B, Piro S and Mattei N (2006), Caractérisation de la qualité de l'air inhalé par les agents dans les enceintes ferroviaires de la RATP – résultats de la phase 2, *Proceedings 2nd conference on Environment and Transport – Transport and Air Pollution*, Reims, France, 12–14 June, INRETS, Bron, France.

Karlsson H L, Nilsson L and Möller L (2005), Subway particles are more genotoxic than

street particles and induce oxidative stress in cultured human lung cells, *Chemical Research in Toxicology*, **18**(1), 19–23.

Larsson B-M (2008), Inducerar exponering av friska försökspersoner i tunnelbanemiljö akuta luftvägseffekter? Lecture at *Transportforum 2008*, Linköping, Sweden, 9–10 January (in Swedish), presentation available at http://www.vti.se/templates/Page_8336.aspx, accessed March 2009.

Lindbom J, Gustafsson M, Blomqvist G, Dahl A, Gudmundsson A, Swietlicki E and Ljungman AG (2006), Exposure to wear particles generated from studded tires and pavement induces inflammatory cytokine release from human macrophages, *Chemical Research in Toxicology*, **19**, 521–30.

Lorenzo R, Kaegi R, Gehrig R and Grobéty B (2006), Particle emissions of a railway line determined by detailed single particle analysis, *Atmospheric Environment*, **40**, 7831–841.

Mazoue S and Parfait C (2006), Mass and submicron size distribution of the aerosol in a Parisian metro station, *Proceedings 2nd Conference on Environment and Transport – Transport and Air Pollution*, Reims, France, 12–14 June, INRETS, Bron, France.

Nieuwenhuijsen M J, Gómez-Perales J E and Colvile R N (2007), Levels of particulate air pollution, its elemental composition, determinants and health effects in metro systems, *Atmospheric Environment*, **41**, 7995–8006.

Norman M and Johansson C (2005), Karaktärisering av partikelförekomsten vid Mariatorgets tunnelbanestation, ISLB Rapport 1:2005, SLB-analys, Environment and Health Administration, Stockholm, Sweden (in Swedish).

Okinaga K, Takahashi S, Tsugoshi T, Kudo Y, Furuya K and Araki Y (2000), Characterization of suspended particulate matter in the air in subways and corresponding above-ground areas, *Journal of Japan Society for Atmospheric Environment*, **35**(1), 12–20.

Ripanucci G, Grana M, Vicentini L, Magrini A and Bergamaschi A (2006), Dust in the underground railway tunnels of an Italian town, *Journal of Occupational and Environmental Hygiene*, **3**(1), 16–25.

Seaton A, Cherrie J, Dennekamp M, Donaldson K, Hurley J F and Tran C L (2005), The London Underground: dust and hazards to health, *Occupational and Environmental Medicine*, **62**, 355–62.

Sitzmann B, Kendall M, Watt J and Williams I (1999), Characterisation of airborne particles in London by computer-controlled scanning electron microscopy, *Science of the Total Environment*, **241**, 1–3.

Sjöstedt S (2005), *Partikelmätningar på Lisebergsstationen. Miljövetenskap*, MSc Thesis, Göteborgs universitet, Göteborgs, Sweden (in Swedish).

Sternbeck J, Sjödin Å and Andreasson K (2001), *Spridning av metaller från vägtrafik*, B1431 IVL, Svenska Miljoinstitet. Stockholm, Sweden (in Swedish).

Tokarek S and Bernis A (2006), An example of particle concentration reduction in Parisian subway stations by electrostatic precipitation, *Environmental Technology*, **27**, 1279–87.

Tokarek S, Gabay D and Bernis A (2002), A method to reduce the particle pollution in the air of parisian subway stations, *Advances in Air Pollution*, **11**, 561–70.

Tomiyasu B, Suzuki K, Gotoh T, Owari M and Nihei Y (2004), TOF-SIMS measurement for the complex particulate matter in urban air environment, *Applied Surface Science*, **231–232**, 515–19.

Trattner R B, Perna A J, Kimmel H S and Birch R (1975), Respirable dust content of subway air, *Environmental Letters*, **10**(3), 247–52.

Weckwerth G (2001), Verification of traffic emitted aerosol components in the ambient air of Cologne (Germany), *Atmospheric Environment*, **35**(32), 5525–36.

Copyrighted Materials
Copyright © 2009 Woodhead Publishing Retrieved from www.knovel.com

20

Maintenance of the wheel–rail interface

S. L. GRASSIE, Stuart Grassie Engineering Ltd, Germany

Abstract: The two fundamental components in maintenance of the wheel–rail interface are control of friction and control of the wheel and rail profiles. Appropriate design of the wheel–rail interface, with attention to these factors, can greatly reduce maintenance requirements. Some basic guidance to prevention of damage is contained here, but the majority of this chapter comprises information regarding equipment, techniques and expectations for maintenance of wheels and rails. 'Best practice' is to undertake routine, preventative maintenance that restores transverse profile, removes longitudinal or circumferential irregularities and either removes surface cracking or otherwise forestalls its development, while removing minimal material from the wheel or rail surface. The ability to undertake preventative maintenance can be regarded as a sign of a well-designed wheel–rail interface. If damage is sufficiently severe, 'corrective' maintenance is required. Rail grinders and wheel lathes are appropriate for both preventative and corrective maintenance. Both milling and planing are appropriate for several types of corrective re-profiling of rails but, if such work is required routinely, it is symptomatic of a deeper underlying problem.

Key words: railways, vehicle–track interaction, wheel–rail interface, friction, rail maintenance, rail re-profiling, wheel maintenance, wheel re-profiling, rail grinding.

20.1 Introduction

The two fundamental components of maintenance of the wheel–rail interface are control of friction and control of the wheel and rail profiles. Both longitudinal (circumferential in the case of wheels) and transverse profiles are significant, albeit for different reasons.

It is an obvious but often neglected point that appropriate design of the wheel–rail interface can significantly reduce the requirement for subsequent maintenance. It is very much easier and, ultimately, also less expensive to maintain a friction level and both wheel and rail profiles that are appropriate for a particular railway system than to maintain those that are inappropriate. An obvious example of this is that, on a heavily curved railway, wheel and rail profiles are desirable that promote curving, in particular by giving a high 'rolling radius difference' for allowable lateral displacements of a wheelset. If bogies on such a railway also help wheelsets to align themselves radial to a curve, i.e. if the bogies have some 'self-steering' capability, contact between wheel flanges and the rail may be avoided in many curves, and lubrication

will be unnecessary. On the other hand, if interaction of vehicle and track is not considered systematically, a railway could acquire relatively expensive, self-steering bogies, but achieve no significant benefit from them because wheel and rail profiles are either inappropriate or not adequately maintained to enable the bogies to steer. Conversely, on railways that are predominantly straight, contact often deviates little from a narrow band on the wheel tread, thereby concentrating wear in this narrow area. In many such cases, wheels are re-profiled surprisingly frequently in order to eliminate the 'hunting' instability and bad ride that result from the development of excessively conformal profiles. Wheel and rail profiles that give lower conicity spread wear over a wider contact area and, by doing so can greatly extend wheel life (although in the short term such low-conicity profiles may give poorer ride in the vehicle). Similar benefits are achieved by adopting a maintenance regime that deliberately varies the rail profiles to spread wear over the wheel treads.

The key motivation for maintenance on freight and heavy-haul railways is economics: a well-maintained railway is also one that operates more economically. On a passenger railway system this is also true, but maintenance in this case is critical to reduce risk of derailment and thereby improve passenger safety. The consequence of one example of inadequate rail maintenance on a mixed passenger railway is shown in Fig. 20.1. While this example (the derailment of a passenger train at Hatfield in England in October 2000) is one

20.1 Hatfield derailment.

of the best-known examples of catastrophic failure as a result of inadequate maintenance on such a railway, it is far from unique.

This chapter considers maintenance that should be undertaken to improve passenger safety and also that which is undertaken primarily to improve the economic well-being of the railway or to reduce the railway's impact on its passengers and neighbours. Maintenance that improves safety is related primarily to avoidance, detection and/or treatment of cracks (initiated either at the rail surface or internally) that could lead to a rail break. Maintenance that reduces costs and the railway's impact is related primarily to improvement of wheel and rail profiles, both longitudinally and transversely.

20.2 The importance of friction

An appropriate friction level is important to maintenance because the coefficient of friction limits the tangential force that can be sustained by a wheel–rail contact. If the coefficient of friction is too low, this limits the traction and (more importantly) braking rates of a vehicle, and similarly limits the tangential force that can be developed in curving. The importance of a sufficiently high coefficient of friction is exemplified by the chaos that occurs in many temperate countries during Autumn, when leaves reduce the coefficient of friction to levels of $<<0.1$, which are lower than can be achieved with the very best of lubricants. However, if the coefficient of friction is high, most types of surface damage of rails and wheels occur more readily since almost all depend on either the tangential force or the traction ratio (the ratio of tangential force to normal load) or both. It should be apparent from the discussion of rail wear and fatigue that traction ratios of greater than about 0.4 promote surface damage of wheels and rails, and the coefficient of friction should accordingly be less than this level. If one were to consider maintenance alone, it is very much more important to avoid coefficients of friction that are too high than too low, although if coefficients of friction were low it would be difficult to run any railway.

There are many circumstances in which movement of a vehicle either requires or gives rise to high tangential forces. For example, acceleration and braking require a modest level of friction (usually a coefficient of friction of no more than 0.2 although higher levels are increasingly demanded by modern 'high-traction' power cars and locomotives). It is accordingly desirable (and, for purposes of braking, critical) to ensure that these levels of friction exist. A modest level of friction is also required to sustain the tangential forces that are required for vehicle stability. On the other hand, when a vehicle traverses a curve, constraint of the wheelsets in a bogie prevents these rolling freely around the curve, thereby giving rise to tangential forces that increase with severity of curvature and increasing constraint of the wheelsets by the primary suspension. Reducing wheel–rail friction in these circumstances

can lead to a bogie sliding around the curve. Allowing high friction to exist inevitably gives rise to high tangential forces, which can in turn exacerbate some types of surface damage and lead to increased maintenance.

If curves are sufficiently severe, there is almost inevitably some sliding between wheel and rail. In order to reduce wear and, accordingly, maintenance, lubrication should be applied to one or other of the sliding surfaces. Whether lubrication is applied to the wheel or to the rail is largely a matter of choice, maintainability of equipment and, ultimately, economics. 'Best practice' in this case is that lubrication is required in sufficient volume to reduce wear to acceptable levels. Although wear can also be reduced using hard materials, an order of magnitude greater reduction is possible with lubrication.

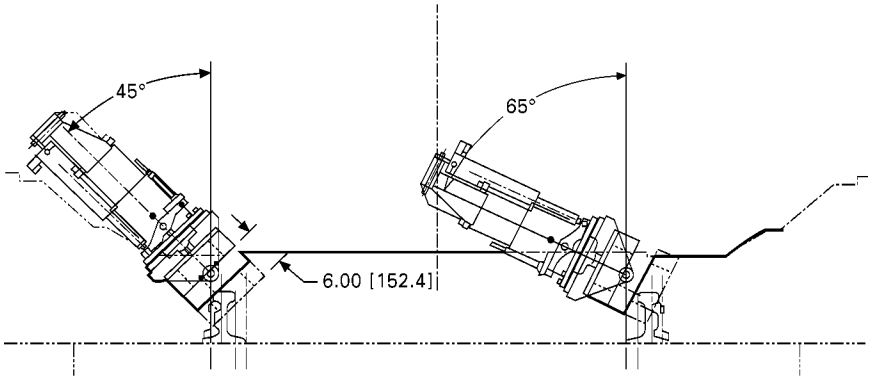
20.3 Routine metal removal and maintenance of transverse profile

20.3.1 Rails

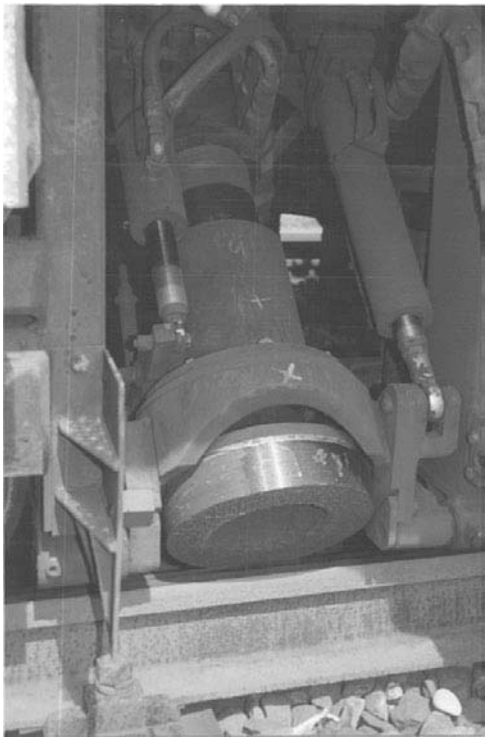
Maintenance of the rail's transverse profile and routine metal removal is the most effective and most widely used treatment of surface-initiated rolling contact fatigue (RCF). Since rail breaks can and sometimes do develop from RCF, e.g. Fig. 20.1, this is accordingly safety-critical maintenance. The most common treatment is to 'unload' the area that is cracked or in which cracks are likely to appear, usually without removing the cracks entirely. Indeed, while it is possible to demonstrate that cracks have been removed using manual techniques (magnetic particle inspection or dye penetrant) or specialist equipment,[1] this is time-consuming and expensive. Provided that cracks have not yet 'branched' into the rail (Section 1.10) they can be unloaded by grinding a suitable profile and their growth arrested.

The rail's transverse profile should be maintained for other reasons. Wear and plastic flow cause broadening of the running band between wheel and rail, poorer steering of the vehicle and increased wear of wheel and rail as a result essentially of 'scrubbing' of surfaces in contact over a broad area. In extreme cases, the greater conformity that results from wear and plastic flow leads to localised 'hunting' of vehicles. Hunting gives rise to bad ride and can increase lateral wheel–rail forces sufficiently to shift the track. If localised hunting is not controlled, the rails wear periodically at a wavelength that is sufficiently long to be difficult to correct other than by re-railing or quite sophisticated re-profiling.

A controllable transverse profile is ground using stones that rotate about an axis normal to the rail, each of which produces a facet at a specified angle (Figs 20.2 and 20.3). Grinders used on mixed-traffic railway systems typically have at least 32 grinding modules, e.g. Fig. 20.4. Grinders with as few as four modules can be used for 'special purpose' work, such as grinding of



20.2 End view of grinding train showing how modules are tilted relative to rail to give facets at different angles.



20.3 Grinding module, showing typical annular grinding stone. Usually contact is made with the rail on the leading or trailing end of the stone, so that grinding scratches are oriented almost transverse to the rail. In the orientation shown, the stone would cut a facet on the gauge corner of the rail with almost longitudinal grinding scratches.

20.4 Rail grinding train of a typical size used for mainline railway work in Europe.

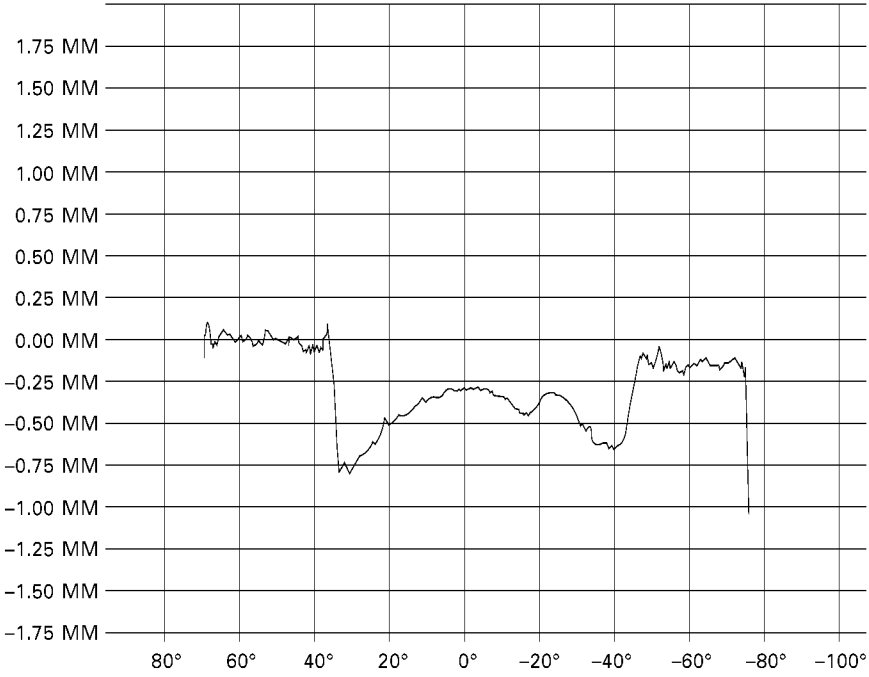
20.5 Ground rail after carrying a day of traffic, showing the individual facets that give an approximation to a curved profile.

switch and crossing work and removal of defects at specific sites that may be difficult to access with a larger machine. The desired rail profile and depth of metal removal are achieved by grinding sufficient overlapping facets of 10 mm width or less on the rail crown and a few millimetres on the gauge

corner (Fig. 20.5). Larger facets may be ground on the field side of the rail with no loss of function because the principal reason for removing material is to ensure that there is no contact between the outside of the wheel and the rail. Several passes of a grinding train are commonly required (known as 'multi-pass grinding'). However, if the profile has changed relatively little and only superficial cracks are to be removed, it is sometimes possible with larger machines (32 or more modules) to finish the rail in a single pass with the modules positioned in a way that is known, from previous monitoring, to remove the desired amount of metal and leave rails with the desired profile.

Pre- and post-grind profile measurements in Fig. 20.6 (made with a Miniprof instrument, Section 20.3.3) show metal removal around the crown of the high rail in a curve that has been ground (four passes with a 32-stone machine) at an interval of about 25MGT to treat RCF. There is a minimum of about 0.2 mm of metal removal at the rail centre, with slightly more required to restore the profile towards the field and gauge sides of the rail. The small change that has occurred in the transverse profile as a result of traffic demonstrates that this profile is appropriate for the location. Conversely, if there were a large change in profile after 25 MGT of traffic, this would indicate either that the profile was inappropriate or that there are other problems such as inadequate lubrication. Usually the edges between grinding facets are deformed sufficiently after the passage of only a few wheels to give almost a continuously curved profile. If these edges are still pronounced after a few days of traffic, the facets and edges between them are probably too severe for the axle loads and rail hardness.

Usually a client specifies a desired or reference transverse profile and a tolerance to which this should be achieved. There are several methods of measuring the profile to ensure that the specification has been met (Section 20.3.3). A method of quantifying objectively the quality of transverse profile is provided by the European Standard for re-profiling of rails,[2] to which the reader is referred for details of the required analysis. This European Standard explicitly accepts that that in practice it is difficult, and often impossible, to grind rail within a tight tolerance of a specified transverse profile throughout a typical site of several hundred metres length. This is difficult because the initial transverse rail profile can vary significantly as a result of factors such as welds and dynamic behaviour of traffic that have caused rails to wear and flow to different extents. Grinding produces a more consistent profile, both transversely and longitudinally and, in general, the more metal that is removed, the more consistent the final profile. In recognition of this slight inconsistency, the European Standard suggests tolerance ranges and the fraction of measurements that can be expected to lie within those ranges (Fig. 20.7). Class R grinding as defined by the European Standard[2] is a reasonable standard both to require and to achieve for a typical mixed



20.6 MiniProf measurements superimposed to show metal removal.

The percentage of measurements in which fhe range of deviation is less than 0.6 mm, 1.0 mm and 1.7 mm shall be calculated and shall not be less than the values given in Table 6(a) for the Class specified.

Table 6(a)—Minimum proportion of measurements within the specified range

Range of deviation (mm)	0.6	1.0	1.7
Class Q	90 %	95 %	98 %
Class R	75 %	85 %	98 %
Class S	No requirement	No requirement	75 %

The maximum positive and negative deviations shall not exceed those in one of groups A, B or C, as specified in the contract

Table 6(b)—Minimum proportion permitted deviations

Range of deviation (mm)	Permissible maximum deviations (mm)		
	Group A	Group B	Group C
≤ 0.6	+ 0.6/ – 0	+ 0.3/ – 0.3	+ 0/ – 0.6
> 0.6: ≤ 1.0	+ 1.0/ – 0	+ 0.5/ – 0.5	+ 0/ – 1.0
> 1.0: ≤ 1.7	+ 1.7/ – 0	+ 0.8/ – 0.9	+ 0/ – 1.7

20.7 Requirements for acceptable transverse profile following grinding, from reference[2].

traffic line that is ground routinely. The most stringent class Q is difficult to achieve unless there is very little variation in the initial, pre-grind rail shape. It is accordingly more appropriate to dedicated high-speed lines or to track which is ground routinely, e.g. in which curves and straights are ground at intervals of no more than about 30 MGT and 60 MGT, respectively. On a line that is ground infrequently it may be difficult to achieve even class R grinding.

Correlation of a measured profile with the profile to which it should have been ground, i.e. the 'reference profile', is conveniently assessed by superposing the measurement at two points on the reference profile, as shown in Fig. 20.8 (from reference 2). These reference points are essentially the rail centre and gauge point (A and B_1) or rail centre and gauge corner (A and B_2). It is strongly recommended that reference point B_1, on the gauge face, be used for the low rail in curves and for straight track. Although this requires more metal to be removed in order to obtain the desired profile, and also requires a grinding train in which the modules can be moved to angles lower than 45° (Fig. 20.2), use of reference point B_2 on the gauge corner leads essentially to narrowing of the gauge. This is illustrated in (Fig. 20.9)

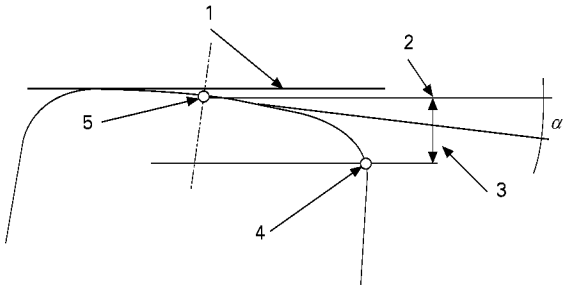
Reference point B_2, on the gauge corner, should be used to grind and assess the quality of grinding of the high rail in curves in which there is flange contact. Use of point B_1 is impractical in such cases because, for some heavily worn rails, this leads to removal of enormous amounts of metal for no practical benefit.

20.3.2 Wheels

The most important reason for maintaining the transverse wheel profile is safety. If the wheel flange becomes too thin, it is possible to 'split' points, i.e. for the wheel flange to run between the stock rail and the wing rail, thereby derailing the wheelset and also the train. 'Thin flange' is accordingly a safety-critical criterion. Examples of the transverse profile of wheels that have been removed from service for thin flange are shown in Fig. 20.10. The flange thickness is found from the transverse profile as illustrated in Fig. 20.11.[3] For profiles commonly used in the UK and continental Europe, 4–6 mm of wear is allowable on each flange at the position designated in Fig. 20.11 depending on the initial profile (see e.g. Table A.1 of reference 3).

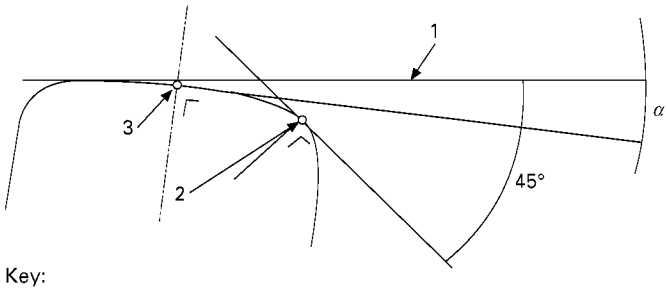
A second safety-critical criterion that dictates the need for wheel maintenance is flange height (Fig. 20.11). If the tread wears sufficiently, thereby essentially increasing the flange height, the flange can eventually interfere with trackwork such as joint bars. This condition must clearly be avoided, which leads to a limit on flange height of about 36.5 mm for common wheel profiles.

There is increasingly a limit on 'hollow wear' or 'false flange' on wheel



Key:
1 reference line
2 line through A parallel to reference line
3 14 mm
4 reference point B_1
5 reference point A

(a)



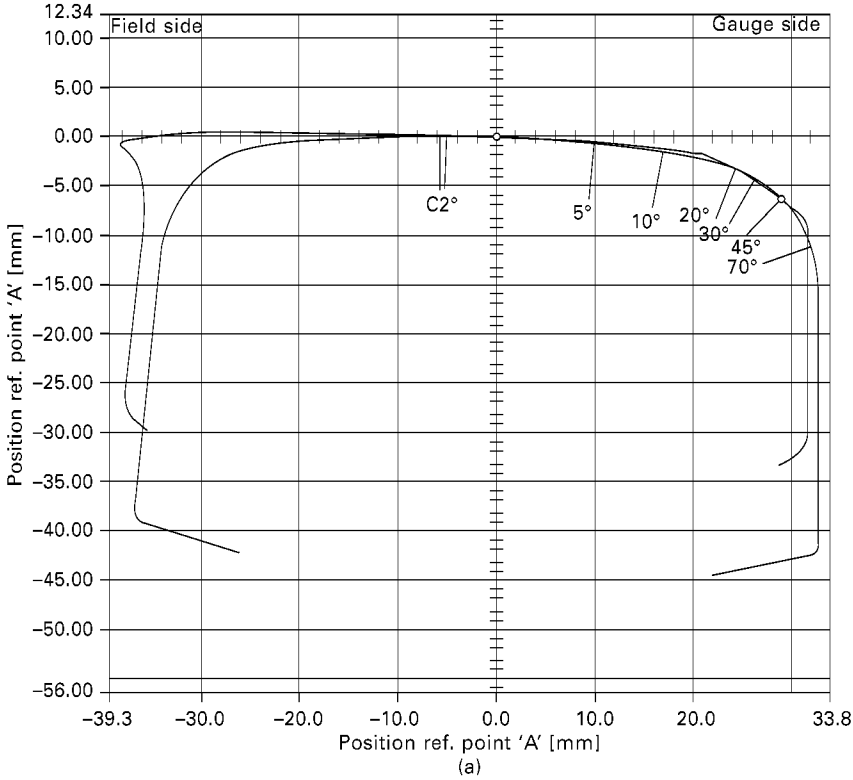
Key:
1 reference line
2 reference point B_2
3 reference point A

(b)

20.8 Definiition of reference points A, B_1 and B_2 at which a measured profile should be aligned with a reference profile, from reference[2].

treads (Figs 20.12 and 20.13): a maximum of 2 mm hollow wear is typical, e.g. references 3 and 4. Where such a limit exists, it is a clear recognition that a railway system has recognised the importance of treating the wheel and rail together as a system, since 'false flange' causes relatively little damage to wheels but is often the main cause of severe plastic flow on low rails in curves. If hollow wear is allowed to develop, it can give rise to premature hunting of vehicles and poor ride.

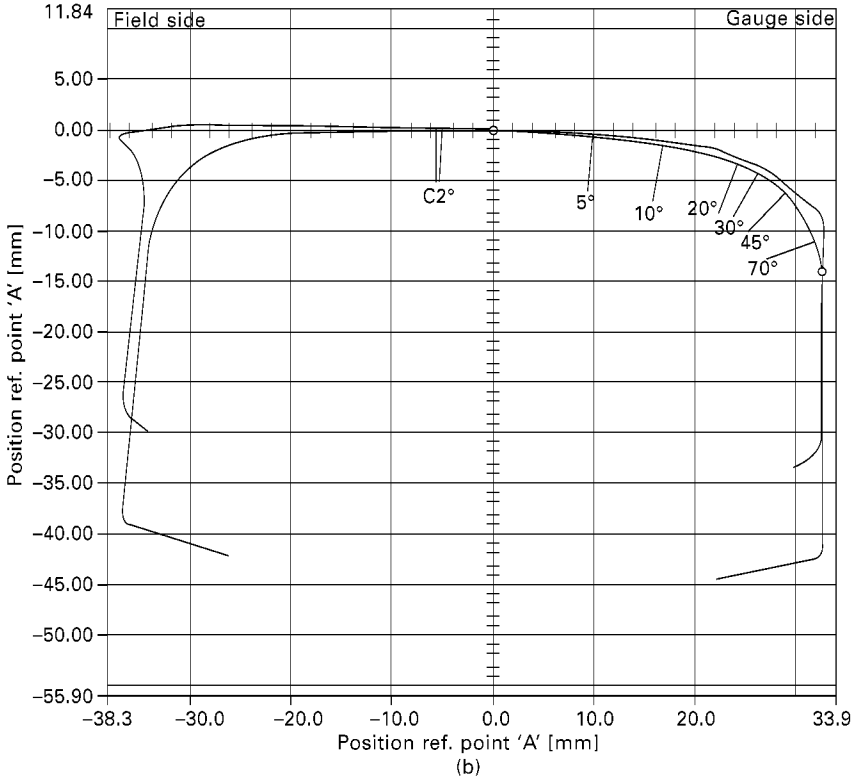
Wheelsets are re-profiled in purpose-built wheel lathes. In order for this to become a routine operation that can be separated from overhaul of the bogie itself, so-called 'underfloor wheel lathes' are increasingly common. These enable the wheelset to be re-profiled without being removed from the train. Typical output of an underfloor wheel lathe is that a full train of eight carriages (32 wheelsets) can be re-profiled in a working day. The tolerance that is demanded on transverse profile is similar to that required on the rail



20.9 Transverse profile of a partially ground rail. In Fig. 20.9a, the measured profile is aligned with the reference at the gauge corner, and grinding appears almost complete in this area. However, when the measured profile is aligned at the gauge point (Fig. 20.9(b)), there is clearly significantly more material to remove. If the profile was left as shown in Fig. 20.9(a) this would effectively narrow the gauge.

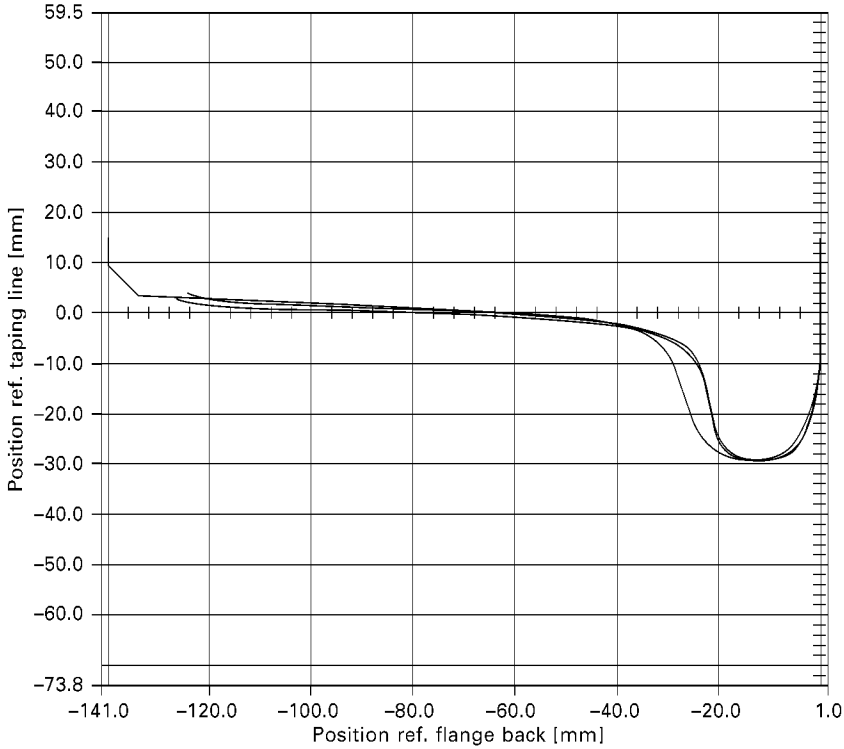
(Section 20.3.1): a deviation of +/–0.25 mm from the specified profile.[3] Even if wheels have worn asymmetrically, as occurs on some railway systems with curvature in one sense, the asymmetry should be removed when they are re-profiled.

The back-to-back distance on wheelsets is as critical as the transverse profile itself to the performance of both the wheelset and the vehicle of which it is part. For standard gauge track (1435 mm) the typical back-to-back distance on wheelsets is 1360 mm. The wheel diameter is measured at a specified distance from the flange back, known traditionally as the 'taping line' (shown as the 'tread datum position' in Fig. 20.11). If the taping line is at 70 mm from flange back, the distance between taping lines on opposite



20.9 Cont'd

wheels of the same wheelset is 1500 mm. For a wheelset running centrally in track, each taping line is at 32.5 mm from the gauge face, which is slightly towards the gauge side of typical rails with a head of 70 mm width. If the gauge is tight or when wheels wear to have thick flanges, the contact points on wheels move towards the root of the wheel flange. These conditions increase conicity of wheel–rail contact and are conducive to hunting of vehicles. The back-to-back distance on wheelsets and flange thickness clearly also limit the allowable lateral movement of a wheelset, or the 'flangeway clearance'. If there is inadequate flangeway clearance, this limits the ability of a wheelset to roll freely around a curve by moving outwards and creating a rolling radius difference between the wheels. In this case, a greater tangential force is required for curving, and wear of the wheel flange and gauge face of the rails is increased when these come into contact. The need for maintenance of wheels and rails is accordingly reduced by ensuring that these basic dimensions of track and wheelsets are compatible.

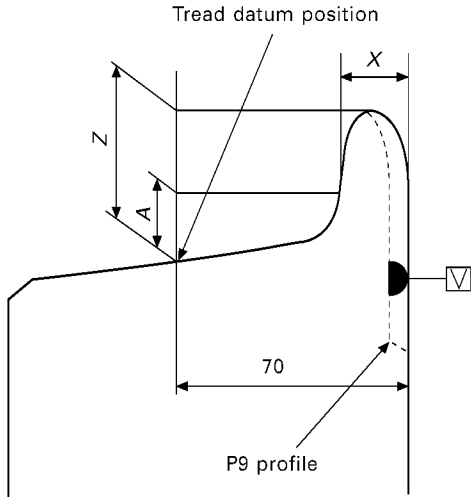


20.10 Profile of wheels removed from service for 'thin flange'. Note pronounced wheel flange wear and relatively little tread wear.

20.3.3 Transverse profile measurements

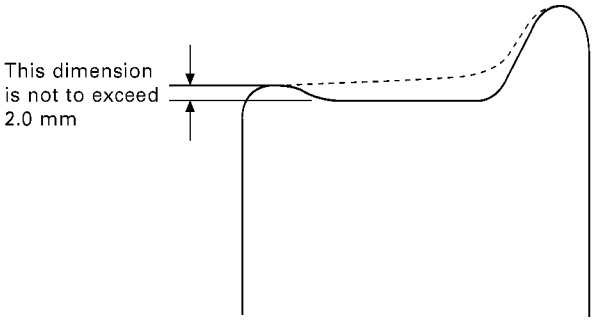
When rails are re-profiled, a simple instrument such as a 'bar gauge' (Fig. 20.14) enables an experienced user to assess *subjectively* how well a rail has been ground or where grinding is required. Measurements are required to provide *objective* quantification of the depth of metal removal or quality of transverse profile. An excellent hand-held instrument that is suitable for small surveys, for scientific purposes, as a reference and for quality control purposes is the Miniprof (Fig. 20.15), but this cannot sensibly be used routinely for grinding. Modern grinding trains usually have laser-based profile measuring equipment whose accuracy is poorer than that of a Miniprof but satisfactory for everyday grinding.

The accuracy of even the best equipment that is currently available for measuring wheel and rail profiles is no better than 0.05–0.1 mm, which is similar to the elastic deflection of wheel and rail at the point of contact. The



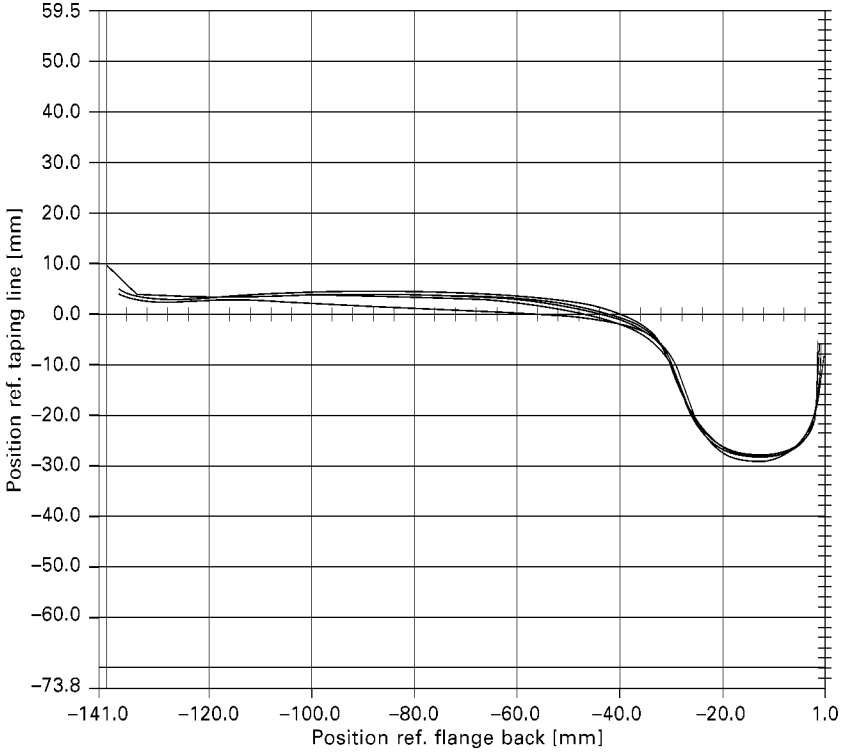
Key: X = Flange thickness (worn)
Z = Flange helght (worn)
A = Flange thickness datum – 13 mm for UK profiles
– 10 mm for EN profiles

20.11 Measurement of flange height and thickness. (Fig. A.2 of reference 3 Flange height and thickness – worn)



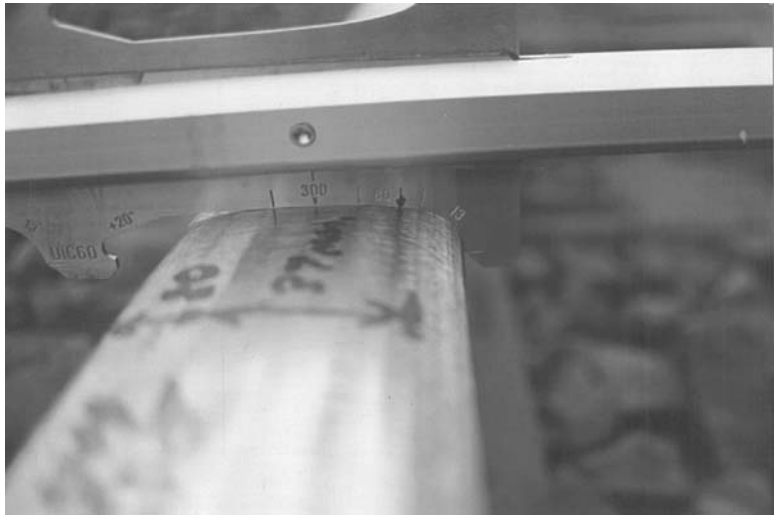
20.12 Measurement of 'false flange' or hollow wear. (Fig. 7 of reference 3)

accuracy of the equipment that is used routinely for such measurements, which is based almost universally on the 'laser triangulation' technique, is less good than this. Such variations, which are well within the tightest tolerance proposed in the European Standard for re-profiling of rails (a range of 0.6 mm) and typical tolerances on re-profiling of wheels (a range of 0.5 mm),



20.13 Measurements illustrating 'hollow' wear of wheel treads, i.e. wear that has been concentrated in a narrow area. Note that there is little flange wear, which is consistent with the fact that wear of this type is associated with a railway system that is predominantly straight.

can either put a point into contact or take it out of contact. Consequently, despite the availability of such equipment, there remains no substitute for an experienced eye assessing, after the passage of traffic, whether re-profiling of a rail has achieved its desired objective and modifying the way the rail is ground if it is not. This limitation is illustrated in Fig. 20.16, which shows measurements and photographs taken on three occasions of the same point on a rail in a mainline curve of modest radius (1000–1500 m). The measurements, in the bottom of the figure, show the difference between the profile (measured with a Miniprof, Fig. 20.15) and the desired profile. This varies little over the approximately seven month interval between the first and last measurements. However, the appearance of the rail does vary significantly, with relatively little contact on the gauge shoulder of the rail

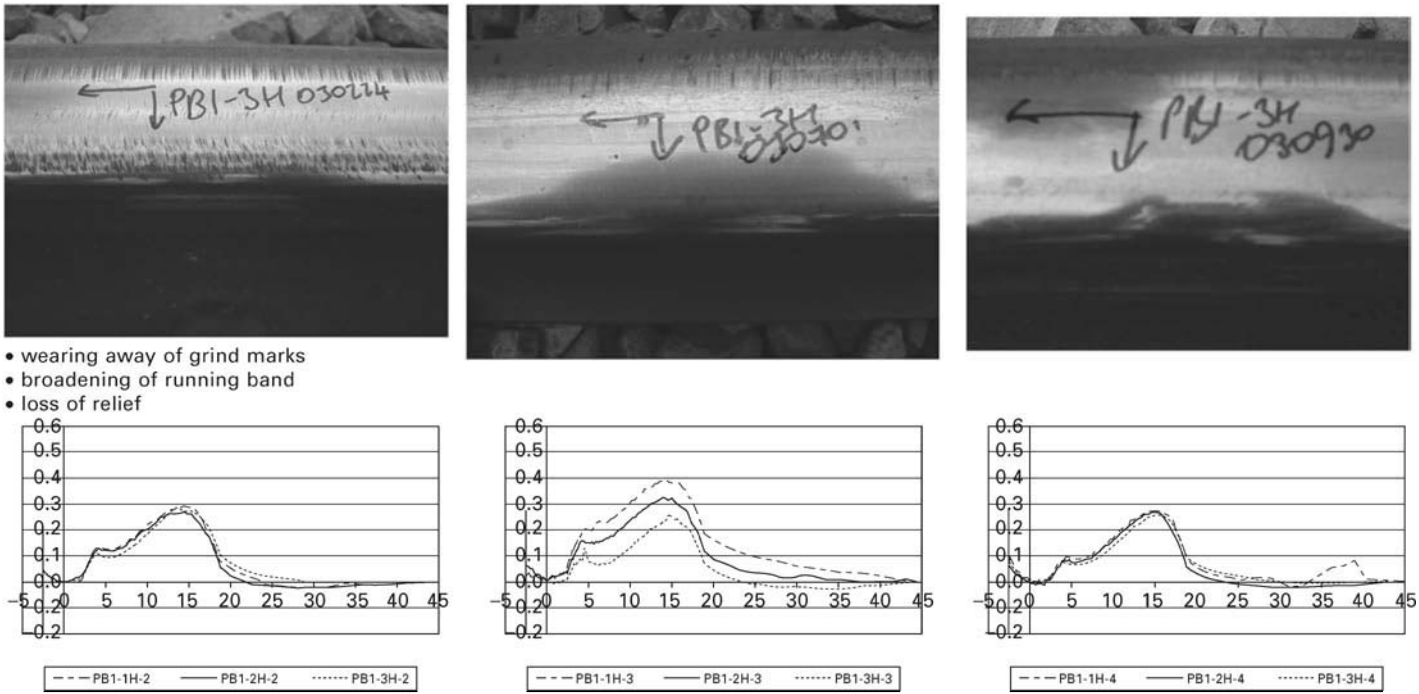


20.14 Illustration of measurement of transverse rail profile using a 'bar gauge' holding a template of the desired profile.

20.15 MiniProf instrument for measuring transverse profile.

in the first photograph, but a broad, shiny running band covering the full rail crown seven months later.

For wheels, measurements of flange thickness and height are increasingly taken from measurements of the full transverse profile, in which software locates the reference points shown in Fig. 20.11. Monitoring of this type is



20.16 Measurements of transverse profile and corresponding photographs for the same position on a rail over a 7-month period. There is little difference in the profile measurements, but a significant difference is visible in the contact band on the rail. There is no substitute for looking at the effects of different profiles on the contact position between wheel and rail.

undertaken using either hand-held instruments (e.g. a version of the 'Miniprof' instrument shown in Fig. 20.15) or equipment that routinely measures wheels as a train passes by, in a maintenance workshop or elsewhere.[5] Such a semi-automatic system is sensibly used to monitor wheelset wear rates and predict maintenance requirements. Measurements of wheelset condition were traditionally made using a 'finger gauge' that located the points in Fig. 20.11 and provided a direct manual measurement of thin flange and flange height. Such gauges are still an excellent and inexpensive method of finding 'bad wheels', but they are necessarily limited.

20.3.4 Frequency of maintenance

Grinding, milling and planing trains exist to re-profile rails. Grinding is the most appropriate of these three technologies for routine maintenance since it is possible to remove shallow depths of metal whereas milling and planing require a greater depth of cut than is desirable for routine maintenance thereby limiting the rail life.

Routine maintenance is often undertaken to a fixed cycle whose period can be defined in terms of the passage of traffic: 20MGT and 50MGT are reasonable intervals to use for maintenance of curves and straight track, respectively, until site-specific information is obtained. Monitoring should be undertaken to ensure that the maintenance interval is satisfactory, i.e. it is sufficiently short to keep problems under control, but not so short that the costs of maintenance are excessive and the rail is ground more than is required: 'over-grinding' should be avoided. Grinding such as this is usually regarded as 'preventative' because it is undertaken regularly before significant damage appears. Indeed, it is undertaken essentially in order to prevent the appearance of significant damage.

So-called corrective maintenance is undertaken according to some measured or observed condition of the rail, e.g. excessive corrugation or noise, or excessively tight gauge. Routine maintenance can also be tied to routine monitoring of the rail condition and the establishment of limits on corrugation, RCF or changes in transverse profile above which the rail should be ground. There are obvious attractions of condition-based maintenance provided that there are adequate measurements of the conditions on which a decision as to whether or not maintenance should be undertaken is based, and that such measurements are undertaken sufficiently frequently. Since this is seldom the case, it is strongly recommended that cyclical, preventative maintenance be undertaken. The period for such cyclical maintenance should be determined from limited, but nevertheless comprehensive, monitoring on the railway system.

On passenger railways, rail grinding is usually limited to a speed of a few kilometres per hour in order to obtain a satisfactorily low level of longitudinal

irregularities (Section 20.4.1). Nevertheless, even if multiple pass grinding is undertaken, the productivity is usually greater than 1 km of finished rail per hour. Productivity of several km per hour is possible when rails are maintained 'preventatively', so that relatively light metal removal is required with little change in the transverse profile. Particularly high productivity is possible where the orientation of grinding modules, i.e. the 'grinding pattern', can be specified in advance and a train with sufficient modules used to ensure that single-pass grinding gives a profile that is sufficiently close to that desired and the facets are sufficiently small. In such conditions, a train with more than 48 modules can complete more than 10 km of rail per grinding hour. Other re-profiling operations are usually at lower speed, e.g. milling is typically undertaken at a speed of 500 m/h, although in this case also only a single pass is necessary.

Wheels are usually re-profiled when bogies are overhauled, and more frequently if their condition, e.g. thin flange, demands. It is reasonable to expect a wheel re-profiling interval of no less than 200 000 km. If wheels require re-profiling more frequently, this is a symptom of underlying problems in the wheel–rail interface such as inappropriate profiles or poor lubrication.

20.4 Routine maintenance of longitudinal/ circumferential profile

20.4.1 Rails (corrugation removal and noise control)

Rails on mixed passenger railways should be maintained so that the depth of short wavelength corrugation (wavelength of less than about 100 mm) rarely exceeds 0.1 mm. Corrugation of this severity is already sufficient to cause complete loss of wheel–rail contact, high impact loads and extreme noise. Since the depth of corrugation of all wavelengths develops approximately exponentially, it is strongly recommended that maintenance be undertaken sufficiently frequently to limit severity of corrugation to well below that at which loss of contact occurs. The maintenance interval or limits on severity depend on costs and operational conditions for each railway.

If routine grinding is undertaken on a mixed passenger railway primarily to control corrugation, it is essential that the transverse profile be controlled to tolerances commensurate with conditions for the line. The recommendations given in Section 20.3 regarding maintenance of transverse profile and metal removal accordingly apply also to sites that are maintained to control corrugation.

Short and medium wavelength corrugation ($\lambda < 300$ mm) is removed essentially because the diameter of the grinding stone is larger than that of the irregularity and its inertia is sufficient to resist it being pushed out of the way of the irregularity. Longer wavelength corrugation ($\lambda \approx 300$–1000 mm)

is removed by holding the grinding modules to a chord, such as that formed by a carriage holding the modules or by the vehicle itself between bogie centres if there is no separate grinding carriage. The shorter the wheelbase of the datum chord, the shorter is the wavelength of corrugation that can be removed. How quickly a corrugation is removed depends essentially on how well a module can be made to cut into the crest of a wave rather than following it. Since the datum chord follows the rail to some extent, moving up and down with the waves, it is difficult to remove longer wavelength corrugation (300–1000 mm), and ever more difficult the longer the corrugation.

Although it is difficult to remove these irregularities, it may nevertheless be economically justified in order to reduce dynamic loading, associated damage to vehicles and track and longer-term maintenance costs. The expense of removing these long waves is particularly well justified if the root cause of the problem can be addressed, so that once the corrugation is removed it recurs slowly if at all.

The European Standard for re-profiling[2] suggests an objective method of quantifying the standard of longitudinal profile of a ground rail (Fig. 20.17). (The reader should refer to reference 2 for detail of the analysis and assessment procedure as well as a definition of terms such as 'moving average.) This requires calculation of the so-called 'moving average' of corrugation amplitude within specified wavelength ranges. Peak-to-peak and RMS averages are both permissible but have different acceptance limits (Fig. 20.17). For grinding to be acceptable, the moving average should be less than a specified limit over a specified fraction of the grinding site. For example, if Class 1 grinding were specified, the moving average of RMS amplitudes of medium wavelength corrugation (100–300 mm) should exceed 0.012 mm over no more than 5 % of a grinding site (Tables 20.2 and 20.4, Fig. 20.17). Pre- and post-grind corrugation measurements (from an RLA, Section 20.4.3) are shown in Fig. 20.18 that demonstrate this has been achieved over one 200 m grinding site.

Although track can be ground to give the low levels of residual irregularity suggested by the European Standard,[2] a more discriminating approach is to require a low level of irregularities in that wavelength range in which corrugation typically occurs. For example, if short wavelength corrugation were the principal problem, it is strongly recommended that the requirements of Class 1 grinding be satisfied for this wavelength range, if not an even more stringent requirement that ensures minimal residual corrugation or grinding 'signature' at the wavelength characteristic of a site. To demand tight tolerances in all wavelength ranges is of questionable value, but is unquestionably expensive.

While the requirements of Class 1 grinding for long wavelengths (300–1000 mm) are achievable where long wavelength corrugation is not a problem, they are extremely difficult to achieve where this is the principal problem.

Table 1 — Window lengths

Wavelength range [mm]	10–30	30–100	100–300	300–1000
Window length [m]	0.15	0.5	1.5	5

Table 2 — Moving average of RMS amplitude limits

Wavelength range [mm]	10–30	30–100	100–300	300–1000
Limit of moving average of RMS amplitude [mm]	0.004	0.004	0.012	0.040

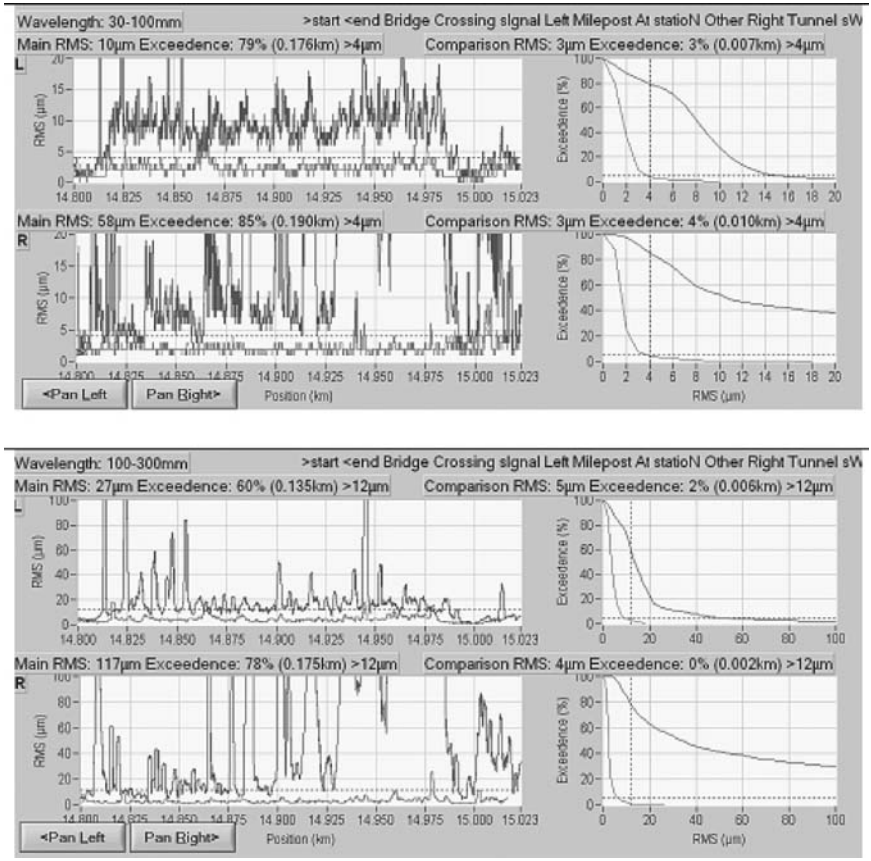
Table 3 — Moving average of peak-to-peak amplitude limits

Wavelength range [mm]	10–30	30–100	100–300	300–1000
Limit of moving average of peak-to-peak amplitude [mm]	0.010	0.010	0.030	0.100

Table 4 — Acceptance criteria for longitudinal profile expressed in terms of allowable percentages of track exceeding moving average RMS or peak-to-peak amplitude limits

Wavelength range [mm]	10–30	30–100	100–300	300–1000
Class 1	5 %	5 %	5 %	10 %
Class 2	No requirement	10 %	10 %	No requirement

20.17 Requirements for acceptable longitudinal profile following grinding (from reference 2)



20.18 Pre- and post-grind corrugation measurements over 200 m grinding site, showing that corrugation has been brought within requirements of reference 2 in 30–100 mm and 100–300 mm wavelength ranges.

Experience suggests that a more appropriate requirement is for the RMS average amplitude of residual irregularity to exceed 0.065 mm (rather than 0.040 mm, Fig. 20.17) over no more than 10 % of the length of track that has been ground. A corresponding limit for the peak-to-peak amplitude is 0.160 mm.

If grinding is undertaken primarily to control noise, particularly so-called 'acoustic grinding', limits on residual irregularities are more stringent than those proposed in reference 2. The requirement in this case is essentially that the one-third-octave spectrum of residual irregularities lies below the level for a 'smooth rail' specified in ISO 3095.[6] Such low levels of residual irregularity are extremely difficult to achieve by conventional methods.

Supplementary treatment, usually with a 'shuffle block' grinder, is normally required.

20.4.2 Wheels: removal of circumferential irregularities

Irregular wheels cause high dynamic loads that are certainly damaging for the track and can also exacerbate damage to components of the vehicle. Examples of damage are fracture of concrete sleepers and pulverisation of ballast, both of which clearly increase the cost of maintenance. Whereas quasi-periodic corrugation is fairly common on rails, on wheels it is more common to find discrete irregularities such as wheelflats and localised spalling, which give rise to impact loads. Since it is the impact loads that cause the damage, the best way of controlling wheel irregularities is to have a means of monitoring dynamic loads routinely, and removing 'bad wheels' from service when a threshold of load is exceeded. This approach is adopted in reference 3, for example, where it is stated that if a total vertical load of more than 350 kN is detected from a wheel, it should be inspected. A decision on the 'minimum action' to take thereafter depends then on the length of the irregularity and the speed of the train in which it runs; e.g. for a passenger vehicle operating at almost any speed, a vehicle should be removed from service immediately if a wheelflat of longer than 60 mm is detected and within 24 hours if the flat length is in the range 40–60 mm, with further recommendations for the maximum speed at which a vehicle can be moved to a site where remedial work can be undertaken.[3]

Wheelflats occur because brakes have locked the wheel and prevented it from rolling. There is accordingly not only mechanical damage that causes a measurable geometrical irregularity but also thermal damage. Thermal damage is usually associated with cracks, which lie beneath the surface of the wheel tread. When wheels are re-profiled to remove flats, it is essential that the depth of cut is sufficient to remove the material that is irreparably thermally damaged. It is common when such a wheel is re-profiled to find that the subsurface damage spreads more widely than is apparent on the surface.

20.4.3 Measurement of rail corrugation and circumferential wheel irregularities

Corrugation can be measured most simply using a straight edge and feeler gauges. This is often sufficient to demonstrate that track should be ground. However, to demonstrate that residual irregularities exist on the rail to the levels proposed in the European Standard[2] is a much more demanding requirement, even although these levels are routinely achieved by grinding trains. When the European re-profiling standard[2] was drafted, it was thought

that the accuracy of equipment on re-profiling trains would be inadequate to demonstrate that the low levels of irregularity demanded by the standard had in fact been achieved. It was accordingly proposed that two types of equipment would be required: so-called 'reference' equipment, whose accuracy would be sufficient to measure the low levels of irregularity, and 'approved' equipment, which would encompass the types of equipment typically used on re-profiling trains. Procedures were proposed to validate both types of equipment, and, in particular, to ensure that the error of 'approved' equipment was always conservative, i.e. the level of irregularity would always be less than that shown by typical train-based equipment.

The Corrugation Analysis Trolley (CAT, Fig. 20.19) is a 'reference instrument' within the meaning of the European Standard,[2] and indeed the calibration procedure that was developed for the CAT is the basis for the validation procedure for such instruments.[7] Instruments that have been developed to measure acoustic rail roughness should also be satisfactory as 'reference' instruments for measurement of residual corrugation. There is very little published information on train-based 'approved' equipment to measure residual longitudinal irregularities. The 'Rail Corrugation Analyser' (RCA, Fig. 20.20), is one instrument for which a comparison with a reference instrument has been published[8] and whose accuracy is only slightly poorer than the reference instrument (less than 1 micron RMS) when used to measure at a speed of 10 km/h. The measurements in Fig. 20.18 were obtained using an RCA on a grinding train.

Circumferential wheel irregularities are rarely measured after re-profiling: it is assumed that the machining procedure itself contains irregularities within acceptable limits. The length of circumferential irregularities, such as wheelflats (Section 20.4.2), is measured simply using a ruler. If measurements of the circumferential irregularities are required for acoustic purposes, these are usually taken using purpose-built equipment or by rotating a wheelset in a lathe and measuring the profile with a displacement transducer pressed against the wheel tread.

20.5 'Corrective' maintenance

When a rail's condition has deteriorated sufficiently that it can no longer be restored to its desired condition by routine maintenance, corrective maintenance is required. The principal differences between 'corrective' and 'routine' maintenance are:

- more material must be removed from the rail;
- usually each site must be treated on a 'case-by-case' basis, e.g. grinding modules are positioned on the basis of pre-grind measurements rather than to a routine pattern, as should be possible with routine maintenance.

20.19 CAT (corrugation analysis trolley): a reference instrument for measuring rail corrugation.

The quality of finished rail, e.g. fit of the transverse profile and amplitude of residual corrugation, should in principle be no worse than that of rails that are routinely maintained. It is, of course, more expensive to obtain this quality because the re-profiling train requires significantly more time and cannot be as easily scheduled. If a railway system commits itself to moving towards routine, essentially planned maintenance from the position of corrective maintenance, an interim stage is possible for which the term 'preventive gradual' maintenance was coined.[9] This strategy offers less potential benefit on a mixed-traffic railway than on the heavy-haul railways where it was developed because there is a finer line between 'corrective' and 'routine' maintenance when rails are less damaged.

Specific problems that require corrective maintenance are the following.

- **Wheelburns** – Light wheelburns are corrected essentially by localised

20.20 RCA (rail corrugation analyser): a train-based instrument for measuring rail corrugation (shown here attached to a Unimog vehicle and measuring embedded rail).

re-profiling that is tapered out sufficiently gradually to give insignificant irregularity. An undamaged section of rail should be inserted if the wheelburn is sufficiently severe. This should be worn similarly to the rail into which it is inserted, with localised grinding undertaken where necessary, in order to ensure satisfactory longitudinal and transverse profile.

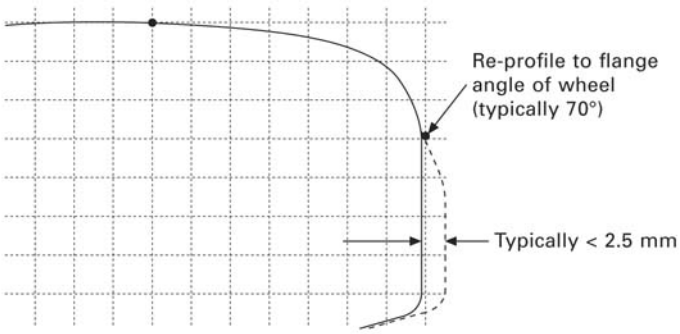
- **Irregular welds or joints** – Welds that are dipped or humped are rectified using maintenance equipment that is essentially a rail-mounted three-point bending jig. A reasonable goal to achieve is a dip or hump of less than 1.5 milliradians (0.4 mm under a 1 m straight edge).
  Irregularities in the lateral plane can also give rise to problems including poor ride resulting from lateral 'jerks'. To avoid bad ride and other problems, welds should be finished to give a lateral irregularity of no worse than 1.5 milliradians.
- **Change of transverse profile or initial rail inclination** – There are several circumstances in which a railway may wish to make a 'one-off'

change in the transverse rail profile. These circumstances include re-profiling of vertically inclined rail in switches and contacts, re-profiling curves to improve curving behaviour or reduce RCF (These goals are usually contradictory for reasons that are not explained here), and re-profiling sections of straight track to alleviate localised 'hunting'. Any such change in profile should be maintained.

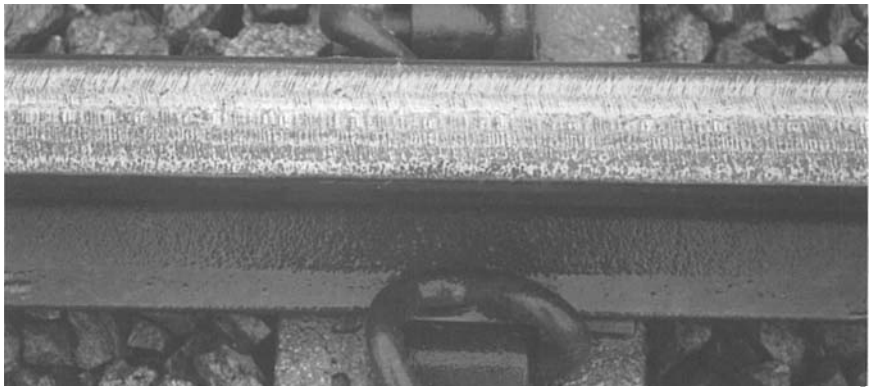
- **Correction of tight gauge** – Tight gauge is undesirable but usually not critical. However, on some railway systems it has been found that tight gauge has given rise in places to vehicle 'hunting' when speeds have been increased, which is experienced by passengers as bad ride. Re-profiling, as illustrated in Fig. 20.21, is an excellent, site-specific way of increasing gauge (up to about 2.5 mm per rail or 5 mm overall) and fixing this problem.

Rail grinders are the most widely used equipment for corrective maintenance of rails. Indeed, a grinder is essentially the only general-purpose equipment for both corrective and preventive rail maintenance. However, since corrective maintenance usually requires the removal of a substantial amount of metal from the rail surface (0.3 mm or more), milling and planing trains are more appropriate for corrective work than for routine preventive maintenance.

An unusual form of rail maintenance on mixed passenger lines is the use of grinding to improve adhesion during periods of leaf fall. There has been only modest quantification of the effects of leaf-fall grinding and the reasons for its undoubted success, e.g.[10] Leaf-fall grinding is undertaken at a relatively high speed with modules distributed to cover the full rail crown, remove minimal metal and leave quite a rough surface (Fig. 20.22). The high roughness is typically retained for several weeks.



20.21 Illustration of rail profile that should be ground in order to widen gauge.



20.22 Surface of a rail that has been ground to improve adhesion in the 'leaf-fall' season.

20.6 Future trends in maintenance of the wheel–rail interface

When railway systems realise the importance of the wheel–rail interface and devote attention to ensuring that this interface is adequately 'designed', this will give less requirement for maintenance. Therefore, one of the most significant practical developments in maintenance of the wheel–rail interface is that there should be less of it but, at the same time, the life of the wheel and rail should increase. Factors such as those discussed in Section 20.1 should be satisfactorily addressed to ensure that there is less maintenance: in particular, ensuring that levels of friction are not routinely higher than necessary to sustain traction and braking requirements, and that lubrication is adequate where there is inevitable sliding of the wheel against the rail, in particular between the wheel flange and gauge face, and between wheels and check rails in curves. Wheel and rail profiles should also be used that are more thoughtfully designed for individual railway systems, with rail profiles varied within a system in order to distribute wear more evenly on wheel treads, encourage vehicle curving behaviour in order to minimise gauge face and wheel flange wear, etc. These changes in design of the wheel–rail interface must then be maintained routinely in order to reap the benefits. The evidence from those railway systems where this has been done is that the payback is substantial. The evidence to date has come mainly from heavy-haul freight railways where conditions of operation would otherwise be impossible, but there are notable examples of mixed–traffic railways and metro systems where the lessons have already been learned.

Re-design and optimisation of the wheel–rail interface is a bold step for a railway system to take, but this will become more common as the benefits

become more widely appreciated. Two enormous steps can then be taken to reduce overall maintenance: the use of harder, more wear-resistant materials and adoption of a preventive maintenance policy. The most critical impediment at present to use of harder materials is fatigue, which is exacerbated by the use of inappropriate wheel and rail profiles. Although it is a common and traditional belief amongst railway engineers that hard wheels damage their rails and vice versa, damage merely reflects the inappropriateness of the profiles. Some railway systems are happy to suffer rapid wear and plastic flow rather than take the bold step of realising that this reflects a problem with the vehicle–track system. There will be incremental steps in development of wheel and rail materials, in particular to increase hardness and wear resistance without sacrificing fatigue-resistance.

There is already vast knowledge of how to make vehicles more 'track-friendly', i.e. to reduce the impact that these have on the track, which coincidentally reduces the requirement for both track (and usually also vehicle) maintenance. At present much of the knowledge that would help optimise vehicle–track interaction and reduce long-term maintenance is not used because the most that railway systems demand is a snapshot of how a train behaves on their track. It is typical, for example, to demand calculations simply of how a new train operates on new track with perfect wheel and rail profiles, without applied traction and braking, and without consideration of how conditions change as a result of the forces generated by the system. Where this knowledge can be brought into the railway and applied in a more routine manner than is currently possible, it will be possible to benefit from it.

Research that has been undertaken in the UK following the Hatfield derailment shown in Fig. 20.1 has highlighted the fact that imperfections in track geometry (particularly lateral geometry) excite vehicle dynamic behaviour, thereby giving rise to high tangential forces between wheel and rail, e.g.[11] These forces are associated with localised damage, and increased maintenance, of the track. In some cases, it has been possible to show that wheel maintenance has also increased as a result of this behaviour. These findings have provided an excellent theoretical justification for what many permanent way engineers had realised and practised for a considerable time: the effort of removing 'triggers' of dynamic behaviour will repay itself in reducing long-term maintenance. This lesson was again learnt at an early stage on heavy-haul railways, whose track geometry would often be the envy of passenger railways running at very much higher speeds. As this lesson permeates through the railway community, the hidden costs of poor track geometry on railways of all types will be better understood, the triggers for dynamic behaviour eliminated and maintenance correspondingly reduced.

There is a constant improvement in wheel slip and slide control on vehicles, with the goal (for the vehicle engineer) of being able to operate, if necessary, close to the point of slip but without actually slipping. If coefficients of friction are high, this inevitably introduces the possibility of allowing high tangential forces to exist without the wheel actually slipping. Since high tangential forces exacerbate most forms of surface damage, which may in the past have been removed by high wear that resulted from slip, progress in this area may well give rise to new forms or new levels of surface damage. This may require new forms of maintenance, or a new compromise to be reached between the desirability of high traction and the offsetting cost of increased wheel and rail maintenance. No such compromise can be reached without sustaining a sensible dialogue across the wheel–rail interface.

When damage is contained by a better design of the wheel–rail interface, maintenance can be better planned and can become preventive, thereby extending the life of the components of both vehicle and track. To assist planning of the preventive maintenance, measurements of the rail and wheel condition that are now undertaken as quasi-scientific exercises (if at all) will become more commonplace. Track recording cars will increasingly measure transverse rail profile and surface condition, while wheels will be measured routinely in maintenance yards. Software will increasingly be used to translate these rivers of data into useful information for those who maintain the track and wheels on the vehicles.

20.7 Sources of further information and advice

Inadequacies in the wheel–rail interface are apparent where stresses are high and become apparent most quickly where there are high levels of traffic. It is accordingly understandable that heavy-haul railways have led research on wheel–rail interface problems. An excellent 'best practice' manual in this area is published by the Heavy Haul Association (IHHA).[12] Network Rail have also sponsored a general 'best practice' manual on the wheel–rail interface with a specific concern for the mixed-traffic railway systems that are typical of Europe.[13]

There are several series of conferences that impinge to a greater or lesser extent on the wheel–rail interface. The series on contact mechanics and wear of rail–wheel systems, which are held at an interval of three years, contain particularly pertinent material, much of which is specifically on the subject of maintenance. Proceedings of the CM conferences are published in the journal *Wear*, e.g.[14,15] The IHHA holds conferences at a four-yearly interval.[16,17] Because loads and traffic density are very much higher on heavy-haul railways than elsewhere, subjects that are discussed at IHHA conferences today often become the topics of interest elsewhere tomorrow.

The European Community in particular has been extremely active in producing standards, of which reference 2 is but one example. While these provide material that is useful in showing what the railway industry expects and what contractors can achieve, their purpose is more to provide a 'level playing field' than to encourage 'best practice'. Standards are usually also the result of many years of discussion amongst and compromise between different interests, with both the benefits and the flaws that are a consequence of this approach.

20.8 References

1. Schoech W, Heyder R and Grohmann H-D (2006), Contact geometry and surface fatigue – guidelines for appropriate rail maintenance, *Proceedings 7th International Conference on Contact Mechanics and Wear of Rail/Wheel Systems*, Brisbane, Qld, Australia, 24–26 September, 23–9.
2. CEN (2006), *Railway applications – track – Acceptance of work – Part 3: Acceptance of rail grinding, milling and planing work in track*, EN 13231-3:2006, European Committee for Standardization, Brussels, Belgium.
3. Rail Safety and Standards Board (2008), Railway Group Standard GM/RT 2466, Issue 2, *Railway Wheelsets*, (see www.rgsonline.co.uk).
4. IHHA (2001), *Guidelines to Best Practices for Heavy Haul Railway Operations: Wheel and Rail Interface Issues*, International Heavy Haul Association, Virginia Beach, VA USA, Section 2.5.3.4.
5. See http://www.kldlabs.com/wheel.html
6. CEN (2005), *Railway applications – Acoustics – Measurements of noise emitted by railbound vehicles*, EN ISO 3095:2005, European Committee for Standardization, Brussels, Belgium.
7. Grassie SL, Saxon MJ and Smith JD (1999), Measurement of longitudinal rail irregularities and criteria for acceptable grinding, *Journal of Sound and Vibration*, **227**, 949–64.
8. Grassie SL (2005), Rail corrugation: advances in measurement, understanding and treatment, *Wear*, **258**(7–8), 1224–34.
9. Stanford J, Magel E and Sroba P (2001), Transitioning from corrective to preventive grinding on the BNSF railroad, *Proceedings 7th International Heavy Haul Conference*, Brisbane, Qld Australia 10–14 June, 493–501.
10. Grassie SL and Baker PR (2000), Grinding to alleviate leaf fall problems, *Proceedings Seminar, Adhesion – Money on the Line*, Institution of Mechanical Engineers, London, UK 1st June.
11. Evans JR and Dembosky MA (2004), Investigation of vehicle dynamic influence on rolling contact fatigue on UK railways, *Vehicle System Dynamics*, **41**(Supplement), 527–36.
12. IHHA (2001), *Guidelines to Best Practices for Heavy Haul Railway Operations: Wheel and Rail Interface Issues*, International Heavy Haul Association, Virginia Beach, VA, USA.
13. *Best Practices Guide for the Wheel-Rail Interface on Mixed Service Railways*, (2009), to appear.
14. Ekberg E, Ringsberg JW, Lunden R (guest eds) (2005), Contact mechanics and wear of rail/wheel systems, CM2003, June 2003, *Wear*, special edition, **258**(7–8).

15. Grassie SL (guest ed.) (2008), Contact mechanics and wear of rail/wheel systems, CM2006, September 2006, *Wear*, special edition, **265**(9–10).
16. IHHA (2001), *7th IHHA Conference, Brisbane, Australia*, IHHA Inc, Virginia Beach, VA, USA.
17. IHHA (2005), *8th IHHA Conference, Rio deJaneiro, Brazil*, IHHA Inc, Virginia Beach, VA, USA.

Copyrighted Materials
Copyright © 2009 Woodhead Publishing Retrieved from www.knovel.com

21

Models for infrastructure costs related to the wheel–rail interface

E. ANDERSSON, Royal Institute of Technology (KTH), Sweden, J. ÖBERG, Banverket, Sweden

Abstract: A model for determination of costs for track deterioration is presented. In particular, the model is able to discriminate between vehicles with different characteristics operating on the track, as well as incorporating operating data (speed, cant deficiency, etc.) and track geometry. The model is implemented in an Excel™ environment. Its use is exemplified by a Swedish case of mainline passenger and freight traffic. Some results are presented on the influence of vehicle characteristics, track geometry, track lubrication and speed. The model predicts that axle load, radial steering ability, unsprung mass and track lubrication are decisive for track deterioration and its associated costs.

Key words: cost, model, railway, infrastructure, track, deterioration, maintenance, renewal.

21.1 Introduction

The wheel–rail interface is subject to tangential and normal forces, as well as impact and vibration, when rail vehicles run on a track. This results in deterioration, i.e. in several types of wear and damage, influencing system performance and limiting life of components on both vehicle and track. Part of this deterioration is also relevant for safety. In order to maintain system performance and safety, the different components must be subject to maintenance or renewal. Further, the conditions of the system must be inspected and evaluated on a regular basis. All these actions result in costs.

This chapter presents a method for estimating costs related to maintenance and renewal of the *track* resulting from the above-mentioned actions necessary to maintain system performance. The cost is estimated as a result of different types of rail vehicles running over the track, with vehicles having different design characteristics, track loading and speed. In this context only *variable costs* are considered, i.e. *traffic-dependent* costs directly related to train operations and tonnage, not the fixed initial costs to lay and maintain the track in working order. The variable cost as defined above is sometimes also called *marginal cost*.

The method was developed by Banverket (Swedish railway infrastructure manager) in co-operation with the Royal Institute of Technology (KTH) in

Stockholm. The presented cost model can be used for different purposes such as:

1. planning of infrastructure maintenance and renewal;
2. evaluate effects on infrastructure of changes in train operations;
3. evaluate costs for different track geometry and speed;
4. evaluate vehicle type-specific track deterioration cost per gross tonne-km;
5. basis for optimization of vehicle design with regard to 'track friendliness';
6. basis for track access charging for different types of rail vehicles under different operating conditions.

Regarding points (5) and (6), in order to evaluate different vehicles and their 'track friendliness', in particular if differentiated track access charges are introduced, a quantitative description of the process of track deterioration is called for.

Since the majority of all tracks are ballasted, focus is on this type of track design. The model presented does not explicitly include track construction parameters, such as type of rail or sleepers. The main reason for this is that the model is intended to focus on the overall cost related to infrastructure maintenance and the distribution of the cost on different vehicle types over a specific network. However, it would be possible to further develop the model in order to discriminate between different types of track constructions (see Section 21.4 for method outline). Costs are related to current average conditions of the track and the traffic it is subjected to. The costs are distributed over different curve radii and vehicle types, depending on their characteristics and operational parameters.

Eventually track renewal will be required for which the cost can be distributed over the years that the track is in use. The model also accounts for historic traffic by considering an average state of track. It should be noted that it is, usually but not always, the deterioration of the track that triggers a renewal. All this is accounted for by determining marginal traffic-dependent costs by performing econometric studies on real railway maintenance and operations.

21.2 Track deterioration model

As a basis for the cost estimations a track deterioration model is set up. The deterioration model considers the four mechanisms:

1. track settlement, i.e. non-uniform movement in the ballast, in particular vertically;
2. component fatigue (internal fatigue of rails, switches and crossings, sleepers, railpads, ballast, etc.);

3. abrasive wear of rails;
4. rolling contact fatigue (RCF) of rails.

These mechanisms are considered to be decisive for the traffic-dependent marginal costs for track maintenance and renewal.

21.2.1 Basic models used

Generally, a lot of research has been carried out on track deterioration and different models have been developed. Most research has been concentrated on only one or a few mechanisms of track deterioration. This could be taken as a sign of the complexity when modelling track deterioration, and no models include 'everything'. More detailed descriptions and discussions of deterioration models are given in Öberg.[1]

The most important bases of the model proposed are the UIC/ORE model[2,3] and the Rail Surface Damage (RSD) model.[4,5] They are both generally well recognized and accepted, although more detailed and specialized models exist. For the authors purpose, it is important to have simple models with few parameters, still expressing the most essential dependencies close to 'state-of the art'.

The *ORE model* considers track settlement and component fatigue. This model is the result of full-scale testing and evaluation by European railway expertise during several years in the 1980s and is frequently cited in this field. The model states that damages (1) and (2) have a dependence on vertical load by a power of 3 as an average. This is also supported by other independent models usually having a dependence of vertical load by powers between 2 and 4.

The ORE model was developed in order to evaluate the influence of increased axle load. Therefore axle load is the principal parameter in this model. The present authors have extended the ORE model to account for wheel loads instead of axle loads. This makes it possible to account for the quasi-static force contributions due to curving, i.e. the vertical load shift from one side to another. Also considered is the influence of speed and unsprung mass on the high-frequency part of the dynamic vertical wheel load, in this context up to about 90–100 Hz, or slightly above the highest possible sleeper-passing frequency for the speeds in question. This is justified by the fact that the dynamic forces between some 20 and 100 Hz are responsible for a lot of the track deterioration and are linked to the unsprung mass of the vehicle; see for example references, 6 and 7.

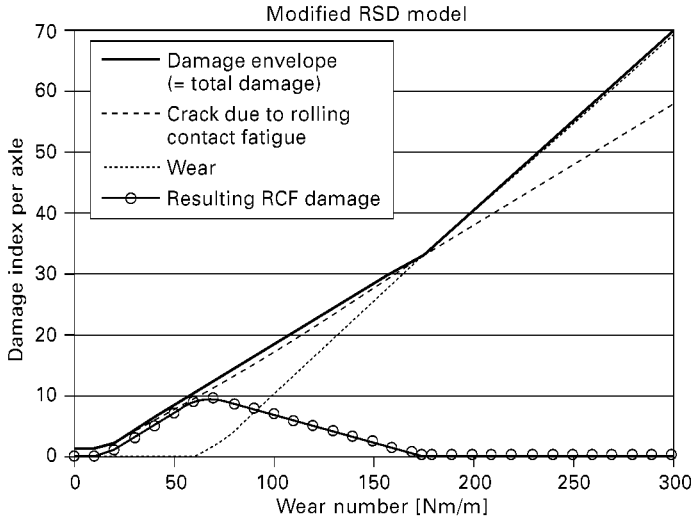
When it comes to track settlement the vertical forces are believed to be the most important, at least in normal cases where the lateral wheel–rail forces are below allowed limits. Within the ORE studies, the lateral wheel loads are left out due to limited knowledge of their influence on track settlement.

The same approach is applied in the present author's model. However, the general model is principally able to include lateral forces in the settlement mechanism in the future, by adding a term to the equations.

The *RSD model* is used for deterioration mechanism (3) and (4), i.e. abrasive wear and RCF. It was developed for the Railway Safety and Standards Board (RSSB) in the UK. The RSD model considers that higher wear numbers contribute to removal of initial surface cracks due to RCF; see Fig. 21.1. The basic parameter in the RSD model is the so-called *wear number*, sometimes called *'T-gamma'*, which is the vector product of wheel–rail creep and the resulting friction creep force. This, in turn, is approximately equivalent to the dissipated friction energy in the wheel–rail contact.

By conducting computer simulations and tests in curves covering radii from 400 to 3000 m it was originally shown that the RSD model was quite accurate and consistent with field observations in the UK. Subsequently, the model was also tested and verified by the authors according to Swedish data and conditions. Further, regarding abrasive wear the RSD model shows similarities with Pearce and Sherrat's model[8] which, in turn, has been validated against advanced simulations and measured wheel wear by Enblom[9] with good results. It should be noted that Enblom considered wear on wheels, but there is generally a strong correlation between wheel and rail wear.

The authors propose to use the RSD model to account for the *total damage* from abrasive wear and RCF. Figure 21.1 shows the modified RSD model and the total damage function.



21.1 The modified RSD model.

21.2.2 General considerations and assumptions

The main focus is on the *rail subjected to the highest forces* as a *representative for the deterioration* of the whole track and its subgrade. In curves, this is usually the outer rail. This is the one that most frequently has to be replaced due to wear and fatigue damage in curves. Further, this rail is usually also subject to the highest vertical and lateral forces causing settlement in the track ballast or in the subgrade.

Also, the inner rail – as well as the ballast and subgrade layer below this rail – is subjected to damage and deterioration. The model takes into account that, in some cases, the inner rail is more vertically loaded than the outer rail. Further, there is usually a strong correlation between the settlement and fatigue at the outer and inner rails due to high vertical forces caused by high speed, axle load or unsprung mass. There is usually also a strong correlation between risk of wear and surface fatigue on the outer and inner rails in curves. In addition, it is assumed that there is a strong correlation between the forces produced in curves and the damage to switches and crossings.

A decisive parameter for track deterioration is curve radius, influencing wear and RCF as well as vertical and lateral forces which, in turn, cause settlement and component fatigue. Therefore, track deterioration has to be evaluated for different *curve zones*. (A curve zone is a radius interval. A typical radius, or mean radius, and cant are considered in each zone. Also, typical vehicle speeds for different vehicles will depend on the actual curve zone. Since there are several curves of different radii belonging to a certain curve zone, the deterioration for a given track section will depend on each curve zone's share of the total track length.)

Different types of vehicles run at different average speeds in the different curve radii. The total deterioration depends not only on vehicle characteristics, but also on the curve radii distribution and the tonnage on the examined rail network. By calibrating the model to the annual marginal deterioration cost, each mechanism (track settlement, component fatigue, wear, etc.) is expressed as a cost component of the total deterioration for each vehicle type on the network. Part of the deterioration occurs on straight track as well. Straight track is therefore one of the generalized curve zones.

As pointed out earlier, the estimated costs are related to current average conditions of the track and the traffic it is subjected to. Usually the current rate of track degradation and deterioration is dependent on the accumulated historic traffic; more accumulated tonnage usually causes a higher rate of degradation. Traffic history, i.e. accumulated tonnage since the track was renewed or maintained, is considered implicitly by assuming a mean value for passed tonnage on the track section or network. This assumption can be justified for most of the purposes put forward in Section 21.1.

21.2.3 The principal mathematical model

The *accumulated marginal track deterioration cost per km of track* (since the track was built or maintained) within a certain curve zone R_j for rail vehicle type Z is expressed in Eq. (21.1). It consists of three terms, each beginning with a marginal cost coefficient k_1, k_2 and k_{34}, respectively. The coefficients are calibrated to reflect the share of annual average marginal cost for deterioration on each of the deterioration mechanisms; k_1 for settlement, k_2 for component fatigue and k_{34} for wear and RCF of rails. It can be seen that the additional deterioration from rail vehicle type Z is dependent on the annual tonnage from that vehicle type, T_Z. It can also be seen that the deterioration is dependent on the accumulated tonnage T_a since the track was renewed or maintained (as in ORE studies);

$$E_{a,z}(R_j, T_a) = \begin{cases} k_1 \cdot T_z \cdot \dfrac{1}{n_z} \cdot \sum_{i=1}^{n_z} Q_{tot_i}^3 \\ + k_2 \cdot (T_a + T_z) \cdot \dfrac{1}{n_z} \cdot \sum_{i=1}^{n_z} [\sqrt{Q_{tot_i}^3 + Y_{qst_i}^2}]^3 \\ + k_{34} \cdot T_z \cdot \dfrac{\sum_{i=1}^{n_z} [f(\overline{F}_v \overline{v})_i]}{m_z} \end{cases} \quad [21.1]$$

where the total dynamic vertical load

$$Q_{tot_i} = Q_{0_i} + |Q_{c_i}| + Q_{d20\,Hz_i} + Q_{d\ hf_i} = Q_{qst_i} + Q_{d_i}$$

while added subscript i relates to axle number i, Q_c relates to quasi-static and Q_d to dynamic contributions (for details, see Section 21.2.4 as well as Appendices 1 and 2): Y_{qst} = quasi-static lateral wheel guiding force on highest loaded wheel; $f(\overline{F}_v \overline{v})$ = the function shown in Fig. 21.1, i.e. relating total damage of wear and RCF to the wear number (in this context for the outer wheels); m_Z = mass of vehicle Z and n_Z is number of axles in vehicle type Z.

In Eq. (21.1) the influence of the vertical forces is given by an exponent of 3, as suggested in the ORE model for track settlement and component fatigue. In the second term the authors consider the resultant force, instead of the pure vertical, to be a measure of the combined deteriorating force (vertically and laterally) affecting the components of the track in terms of internal fatigue damage of rails, switches and crossings, railpads and fasteners, sleepers, etc. (With the use of total vertical wheel load and quasi-static lateral wheel load it is considered that dynamic forces in both directions need not coincide in time, and are most likely not to do so at higher frequencies.)

Regarding wear and RCF, the formulation is principally as in the original

RSD model. By dividing the wear and RCF-deterioration by the mass of the vehicle, the deterioration is evaluated per total number of vehicles of type Z when multiplied by the tonnage T_Z. The wear number is evaluated on the outer rail while the vertical and lateral wheel loads in some cases could be higher on the inner rail (vehicles running at cant excess).

Taking the derivative of Eq. (21.1) with respect to the annual tonnage T_Z from vehicle type Z assuming that $T_Z << T_a$ (true on average) the deterioration per gross tonne-km is deduced:

$$e_{\text{ton-km},Z}(R_j) = \begin{cases} k_1 \cdot \frac{1}{n_z} \sum_{i=1}^{n_z} Q_{\text{tot_}i}^3 + k_2 \cdot \frac{1}{n_z} \cdot \sum_{i=1}^{n_z} [\sqrt{Q_{\text{tot_}i}^2 + Y_{\text{qst_}i}^2}]^3 \\ + k_{34} \cdot \frac{\sum_{i=1}^{n_z} [f(\bar{F}_v \bar{v})_i]}{m_z} \end{cases} \quad [21.2]$$

where $e_{\text{ton-km},Z}(R_j)$ is deterioration per gross ton-km for vehicle type Z in curve zone R_j.

Model simplifications and additions

Any effects of load shift between axles on a vehicle due to acceleration and braking, or non-symmetric loading are not considered, nor are local rail surface damage due to excessive wheel slip as potentially caused by high tractive forces. For Q-forces, the average over the different axles is considered. Regarding the lateral quasi-static wheel loads Y, the leading axles in each bogie, or the leading axle in a single-axle running gear (SARG), are considered. A quasi-static average wear number is evaluated for the outer wheels on leading and trailing axles, respectively. For a bogie vehicle, the average wear numbers are valid for each of the bogies, and the number of bogies and axles consequently determine the total deterioration due to wear and RCF.

With the simplifications above, Eq. (21.3) expresses the deterioration per gross tonne-km and vehicle type (within the curve zone R_j):

$$e_{\text{ton-km},z}(R_j) \approx \begin{cases} k_1 \cdot Q_{\text{tot}}^3 + k_2 \cdot [\sqrt{Q_{\text{tot}}^2 + Y_{\text{qst}}^2}]^3 \\ + K_{34} \cdot \frac{\sum_{i=1}^{n_z} [f(\bar{F}_v \bar{v})_i]}{m_z} \end{cases} \quad [21.3]$$

For the whole railway network, the total annual accumulated deterioration cost for all vehicles Z and all curve zones j, becomes:

$$\Delta E_a \approx \sum_z \left[\begin{array}{l} k_1 \cdot \sum_j T_z \cdot \frac{L_{R_j}}{L_t} \cdot Q_{\text{tot}\,j}^3 \\ k_2 \cdot \sum_j T_z \cdot \frac{L_{R_j}}{L_t} \cdot [\sqrt{Q_{\text{tot}}^2 + Y_{\text{qst}}^2}]^3 \\ + k_{34} \cdot \sum_j T_z \cdot \frac{L_{R_j}}{L_t} \cdot \frac{\sum_{i=1}^{n_z} [f(\bar{F}_V \bar{V})_i]}{m_z} \end{array} \right] \quad [21.4a]$$

where L_{R_j}/L_t is the share of each curve zone in relation to the total track length examined. Eq. (21.4a) can also be expressed in condensed form as in Eq. (21.4b) by incorporating summations over vehicle types and curve zones:

$$\Delta E_a = k_1 \cdot \bar{E}_1 + k_2 \cdot \bar{E}_2 + k_{34} \cdot \bar{E}_{34} \quad [21.4b]$$

It is also possible to use the traffic volume, i.e. gross tonne-km, per vehicle type and year instead of tonnage in the equations presented in this chapter. The cost coefficients would then be evaluated for the whole track section or network, and hence the deterioration cost would reflect the cost for the whole section. Since it is often useful to know the cost for the whole section or network (and year), this approach is followed when implementing the equations into practical use.

22.2.4 Model of vertical dynamic wheel load

In many cases, the total dynamic vertical wheel loads have not been evaluated from vehicle tests. This is particularly the case for high-frequency contributions above a frequency of 20 Hz. Therefore an appropriate model for these loads is useful. Such a model should be based on and be consistent with field test data for a number of representative vehicles.

The axle load is a most decisive parameter for track deterioration. Dynamic wheel–rail forces are often processed by a 20 Hz low-pass filter, for example as prescribed in UIC code 518[10] and the corresponding EN 14 363. However, these forces may have large dynamic contributions with frequency content above 20 Hz. The magnitudes of the dynamic contributions are approximately proportional to speed and are dependent also on the unsprung mass of the vehicle. It is proposed to filter low-pass at 90 Hz or just above the sleeper passing frequency (if higher). A model for the so-called P2-force is adopted, using a speed and unsprung mass dependence which is approximately relevant also for continuously welded rails. The origin is given by Jenkins *et al.*[6] See further Appendix 2.

The model for vertical dynamic loads is based on track force measurements on four different Swedish rail vehicle types. Part of the

model (in the frequency range up to 20 Hz) is further verified with test data for another four vehicles, only available for the 20 Hz dynamics. The model accounts for geometrical track irregularities (track roughness) by a speed-dependent coefficient. Table 21.1 shows forces from measurements and calculations.

It is seen that the model in most cases predicts total dynamic vertical forces with good accuracy, i.e. within 1–3 %, as compared with measurements. However, note that the 'track friendly' freight wagon B has some 5 % lower total forces than predicted by the model; according to the supplier this is due to a more advanced vertical primary suspension. This case demonstrates the limitations of the model; for special vehicles real test data should be used, or advanced computer simulations. However, the model could be used for an approximate estimates applied to the majority of conventional vehicles.

21.3 Computational tools and input data

The practical use of the model requires a computational tool as well as input data on vehicles, track and operational conditions. Input data are briefly presented in this section, with special reference to a case on the national Swedish rail network where focus is laid on deterioration costs resulting from different types of vehicles, in particular as a basis for part of future track access charging. Practical use of the model also requires a cost calibration based on estimated marginal costs of maintenance and renewal. The calibration procedure is described in Section 21.4.

Table 21.1 Comparison of vertical wheel loads – test data vs model calculations

Vehicle	Q_0	Q_c	$Q_{d20\,Hz}$	$Q_{d\,hf}$	Q_{tot}
Rc-locomotive	96	20	? ***(22)***	? ***(27)***	165 ***(165)***
X2 light loco	90	33	? ***(10)***	? ***(26)***	160 ***(160)***
OTU-motor coach	76	19	10 ***(8)***	20 ***(22)***	126 ***(125)***
Freight wagons	110	17	? ***(24)***	? ***(13)***	160–170 ***(164)***
Passenger vehicle A	68	20	7 ***(7)***	–	96 ***(95)***
Passenger vehicle B	87	24	9 ***(10)***	–	117 ***(121)***
Freight wagon A	123	15	29 ***(27)***	–	165 ***(167)***
Freight wagon B (track friendly)	125	25	18 ***(27)***	–	168 ***(176)***

Notes:

1. Quasi-static force contribution due to curving Q_c is calculated with knowledge of cant deficiency and vehicle parameters (see also Appendix 2).
2. Bold italic values within brackets are calculated with the general model and, as can be seen, are usually consistent to measured test data when available.

Sources: Wrang,[7] Andersson and Stichel[11] and Öberg *et al.*[12]

21.3.1 The computational tool DeCAyS

The track deterioration model derived has been built into a software package: DeCAyS (Deterioration Cost Associated with the Railway Superstructure). DeCAyS has been specially developed by the authors to assist with the calculations of track deterioration and cost. It is completely built up in MS Excel™ which allows easy use and analysis of results. Vehicle, traffic and track data are entered into the system together with parameters for cost calibration. A large part of DeCAyS also holds the simulation results on wear numbers and lateral guiding forces for different vehicles and operating conditions; see Section 21.3.2.

21.3.2 Simulation of lateral wheel–rail forces and wear numbers

As mentioned in Section 21.2.2 the RSD model for wear and RCF of rails is based on the wear number, being an approximate measure of friction energy dissipation in the wheel–rail interface. However, it is almost impossible to measure energy dissipation at a single wheel. Therefore computer simulations are needed.

Simulations of approximate friction energy dissipation (the 'wear number') and lateral guiding forces Y from a vast number of cases are performed in the multi-body simulation system GENSYS.[13] The purpose of the simulations is to achieve a database for examination of relations between vehicle characteristics on one hand, and the wear number and lateral guiding force on the other. The database is built up by multi-dimensional matrixes.

Simulations are time-consuming and the database generated is large. It is therefore desirable to generalize the description of vehicles in order to limit the number of parameter combinations as far as possible. The principle is that parameters essentially influencing the resulting wear numbers are properly taken into consideration; see below. Parameters with a very limited influence, or with limited variations in operational vehicles, are set to fixed values. The important parameters are represented by a limited number of discrete values. Intermediate values are dealt with by interpolation in the generated multi-dimensional matrices.

In the general model for four-axle bogie vehicles different *axle load* and *radial steering capability* are independent parameters. The steering capability for these vehicles is, in turn, defined by the *longitudinal stiffness* for the wheelset guidance in conjunction with the *axle distance* in the bogie. Also, different *curve radii* and values of *wheel–rail friction* and four *wheel–rail profile combinations* (including one with two-point contact) are entered as independent parameters. Results from variation of the two latter parameters are combined into a representative average and is later verified by authentic

measured average wear data as function of curve radius. The cant deficiency (lateral track-plane acceleration) is of modest importance for wear and RCF, so this parameter is set constant for each curve zone, representing an average for most passenger vehicles and locomotives, although with some exceptions. However in principle, the model and the methodology allows these parameters to be taken fully into consideration if desired.

Vehicles with special features regarding steering capability and suspension are dealt with separately. These classes of vehicles include standardized two-axle, bogie freight wagons as well as iron ore wagons and locomotives with three-axle bogies. Another special class of vehicle is high-speed tilting trains, usually running at very high cant deficiency. A typical cant deficiency is set for each class of vehicle. In some cases simulation results for these special vehicles are checked against test data on lateral guiding forces, when available.

21.3.3 Track data

Some data in this section are entered into DeCAyS; other data are given as reference information only.

Data entered into DeCAyS

In the example Swedish case mentioned at the beginning of Section 21.3, a set of data is entered regarding track. However, data entered are more detailed than shown here. Relevant track geometry, in particular occurrence of different curve zones, is summarized in Table 21.2.

Table 21.2 Track geometry data on Swedish national railway network

Curve zone [m]	250–400	401–550	551–900	901–1500	1501–10 000	> 10 000
'Mean' radius [m]	325	475	725	1200	2000	1 000 000
Typical cant [mm]						
• passenger traffic	150	150	135	120	40	0
• freight traffic	150	145	120	90	20	0
• iron ore line	125	80	65	40	15	0
Curve distribution						
• passenger traffic	0.8 %	1.6 %	7.6 %	9.6 %	19.3 %	61.2 %
• freight traffic (incl. iron ore line)	1.2 %	2.9 %	13.8 %	8.6 %	11.5 %	62.0 %

Notes:
1. Curve distributions are weighted according to actual traffic volume in gross tonne-km.
2. Cant figures are approximate averages for the network.

Curves with radii less than 550 m are usually equipped with track-borne rail lubrication on the gauging corner of the outer rail. However, all curves are not lubricated, and curves with larger radii are sometimes also lubricated. On lubricated outer rails, the average friction coefficient μ_0 (at large creepage) is anticipated to be 0.25. On inner rail, a friction coefficient of 0.35 is assumed. For wider curves without artificial lubrication, a mix of two friction conditions is considered: dry ($\mu_0 = 0.45$) and humid ($\mu_0 = 0.25$).

Reference information

There is a mix of 50 kg and 60 kg rails, with a predominance of 50 kg. Most tracks, if weighted by actual traffic volume, have continuous welded rails (CWR) resting on concrete sleepers. Permissible axle load is mostly 22.5 tonne, but some lines are upgraded to 25 tonne. The iron ore line allows for 30 tonnes axle load. The total route length is 10 000 km, of which about 1900 km is double or quadruple track. Sidings on stations and yards are not included in these figures. The routes include some 11 400 switches.

21.3.4 Vehicle and operational data

All data in this section are entered into DeCAyS, although in much more detail and for each vehicle type. On the track presented in Section 21.3.3 the total annual traffic volume is 65 000 Mega gross tonne-km (Mgtkm). This volume is divided by different vehicles as follows:

Passenger	19 200 Mgtkm
of which	
• locomotives	1400
• cars	5400
• multiple units	7200
• high-speed tilting trains	5100
• other	100
Freight	45 500 Mgtkm
of which	
• locomotives	4500
• standardized wagons	33 600
• iron ore wagons	7200
• other	200

Axle load on most locomotives is 19.5 tonne, with some exceptions. Multiple units usually have axle loads between 10 and 16 tonnes, the lighter being trailer cars. About 25 % of the freight tonnage is performed at more than

21 tonne of axle load. Iron ore transport is performed at 25 or 30 tonnes axle load.

Roughly 20 % of passenger cars, 70 % of multiple units and 100 % of high-speed trains are equipped with bogies having quite soft wheelset guidance, allowing approximate radial steering in normal curves ($R > 600$ m). All standardized freight wagons and some iron ore wagons have the same features.

21.4 Calibration of model and cost

21.4.1 Principles

In order to relate the model output to reality in terms of traffic volumes, track deterioration and hence cost, the model must be calibrated. There are principally two different levels of application of the calibration:

1. Only interest in relative measures of track deterioration, for example predicting changes in deterioration due to changes in traffic or track construction.
2. The need for setting cost levels reflecting traffic-dependent maintenance cost – for example in terms of track access charging with respect to different vehicle concepts.

In (1) the calibration aims at setting the correct proportions of the different deterioration mechanisms by relating the cost coefficients k_1, k_2 and k_{34} to each other (see Sections 21.2.3 and 21.4.2). Moreover, in (2) the cost for the different deterioration mechanisms must be known or estimated, and in this context the *traffic-dependent costs* (or *marginal costs*) are being considered.

With known data on vehicle types, operational conditions and track geometry, the traffic volume in gross tonne km of each vehicle type will determine the accumulated deterioration from all vehicles on the network. The calibration then results in a top-down distribution of cost on the different vehicle types. These costs could in turn be summarized to give the cost for a certain track section (depending on track geometry and traffic), curve zone, type of traffic (passenger, freight), etc.

First, the total cost to be distributed is divided into three categories, each representative for the maintenance (and principally also renewal) cost of track due to the different deterioration mechanisms. By iterative hand calculation, guessing the initial values or use of algebraic relations (see next section), the coefficients are adjusted so that the total cost given by the model as well as the costs on the different deterioration mechanisms concur with reality. This exercise can be performed over a specific line or a whole network. As costs and traffic may change over time, the calibration is made for a reference year with known cost and traffic data. Cost calibration can

be updated if conditions change and new data are available, preferably on a regular basis.

Cost levels

The track deterioration model distributes *traffic-dependent* (marginal) cost which is not easily determined. Increased track maintenance due to an additional vehicle-km differs depending on vehicle type, track type and age, traffic, volume and historic traffic etc.

Studies on average marginal cost of track deterioration have been performed by Andersson[14] at the Swedish National Road and Transport Research Institute (VTI), among others. In Andersson's work econometrical models are applied to traffic data, track data and cost data supplied by Banverket. The regression models are complex and not dealt with here, but the output consists of change in cost due to change in traffic at an average level, i.e. the average marginal cost for all type of vehicles and track expressed per gross tonne-km.

21.4.2 Calibration procedures

The calibration is currently performed over all Banverket's railway network, hence generalizing track data to the average. The reason for this is mainly the lack of detailed cost data for different track sections in combination with aggregated data on traffic volumes. In principle, however, it is possible to determine the deterioration at a more detailed level with more detailed knowledge of cost and traffic data.

As an average, marginal cost (excluding renewal) is 0.0029 SEK/gross tonne-km (0.0003 Euro/gross tonne-km) in the example case presented here. This figure is a cautious estimate since only current maintenance is included. Recent studies by Andersson[14] point at a total marginal cost in the range of 2.5–3 times the value used here, if distributed costs for periodic track renewal are also included.

In the case used for calibration, the total annual traffic volume is about 65 000 Mega gross tonne-km, distributed over the whole network. According to an experience-based cost analysis at Banverket, it can generally be said that about 25 % of marginal cost is due to track settlement and 35 % to component fatigue with the remaining 40 % allocated to wear and RCF of rails.

Curve distribution is taken from the majority of the Swedish rail network (where about 90 % of the total annual traffic volumes in gross tonne-km are performed) and weighted against the annual traffic volume in gross tonne-km at each line or track section; cf Section 21.3.3 This method implies that the geometry of tracks with heavy traffic is weighted higher than sections with less traffic.

Mathematical description

The calibration procedure can mathematically be described by expressing the annual accumulated deterioration ΔE_a in terms of settlement (Eq. 21.1), component fatigue (Eq. 21.2), RCF and abrasive wear of rails (Eqs 21.3 and 21.4) as in Equation (21.5):

$$\Delta E_a = \Delta E_{a1} + \Delta E_{a2} = \Delta E_{a34} \quad [21.5]$$

By introducing coefficients a, b and c for the percentage of cost on the different deterioration mechanisms, as well as $\overline{C}$ for average marginal cost per gross tonne-km, Eqs (21.6a–c) are given:

$$\Delta E_{a1} = a \cdot \Delta E_a = \overline{C} \cdot \Sigma\, T_z \cdot a = k_1 \cdot \overline{E}_1 \quad [21.6a]$$

$$\Delta E_{a2} = b \cdot \Delta E_a = \overline{C} \cdot \Sigma\, T_z \cdot b = k_2 \cdot \overline{E}_2 \quad [21.6b]$$

$$\Delta E_{a34} = c \cdot \Delta E_a = \overline{C} \cdot \Sigma\, T_z \cdot c = k_{34} \cdot \overline{E}_{34} \quad [21.6c]$$

where $\overline{E}_1$, $\overline{E}_2$ and $\overline{E}_{34}$ are the different terms in Eqs (21.4a and b) with the cost coefficients set to 1 for each mechanism.

Three equations and three unknowns (cost coefficients) give a unique solution which can be solved according to Eq. (21.7):

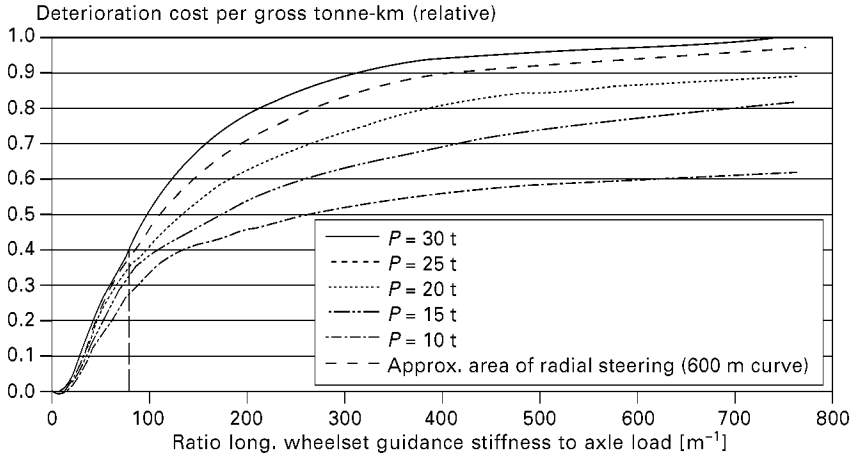
$$\begin{bmatrix} k_1 \\ k_2 \\ k_{34} \end{bmatrix} = \begin{bmatrix} \overline{E}_1 & 0 & 0 \\ 0 & \overline{E}_2 & 0 \\ 0 & 0 & \overline{E}_{34} \end{bmatrix}^{-1} \cdot \begin{bmatrix} a \cdot \Delta E_a \\ b \cdot \Delta E_a \\ c \cdot \Delta E_a \end{bmatrix} = \begin{bmatrix} \overline{E}_1 & 0 & 0 \\ 0 & \overline{E}_2 & 0 \\ 0 & 0 & \overline{E}_{34} \end{bmatrix}^{-1} \cdot \overline{C} \cdot \Sigma\, T_z \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} \quad [21.7]$$

In the case presented here $\overline{C}$ is defined over all the network. If the cost coefficients are to be determined for a specific railway line or track section, Eq. (21.7) may be indexed so that every variable except average marginal cost $\overline{C}$ depends on traffic and track section. The geometry of the different track sections must of course be defined in Eq. (21.4) and matched against the respective cost coefficient.

21.5 Examples of results

There are different focuses on results depending on the purpose with investigations. Primarily, the model for deterioration is developed in order to differentiate among vehicle types according to their tendencies to cause deterioration in the tracks. By so, different vehicle concept may be compared with regard to evaluated costs of track deterioration.

As an example, the cost for wear and RCF deterioration of the rails increases with stiffer wheelset guidance. This is usually the case if forced wheelset steering is not used. Figure 21.2 plots the deterioration cost for four different axle loads, with the ratio of longitudinal stiffness in wheelset



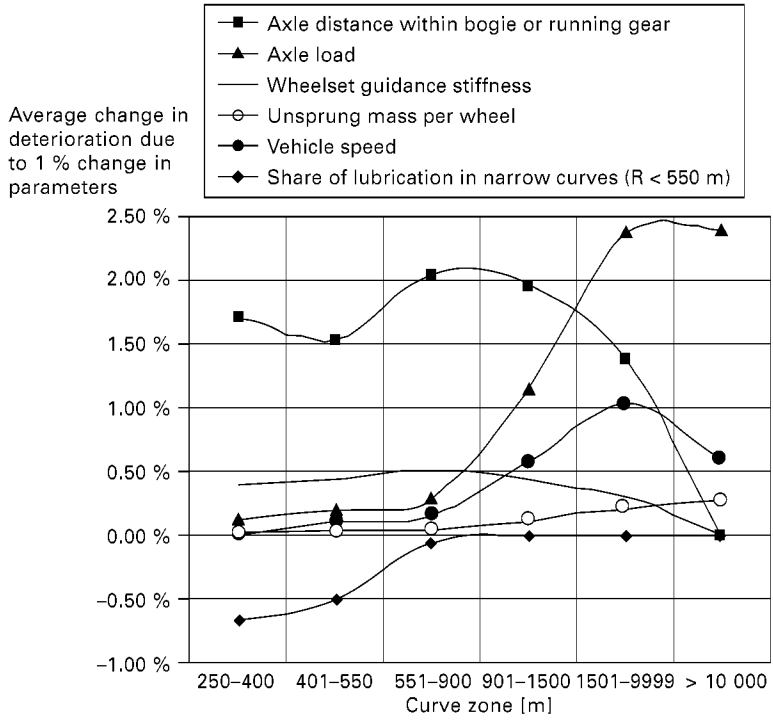
21.2 Deterioration per gross tonne-km for wear and RCF increases with stiffer wheelset guidance. The example vehicle is a 4-axle locomotive with 2.7 m axle distance in bogies. Within the rectangular region in the bottom left corner, radial steering is approximately achieved in curves with radii ≥ 600 m.

guidance to axle load on the horizontal axis. A larger value indicates less ability for the wheelsets to yaw and steer themselves radially in curves. Consequently they cause higher wear and RCF deterioration. Within the marked rectangle, radial steering is approximately achieved for curve radii down to about 600 m on dry or medium dry wheel–rail interface.

Another way of presenting model output is based on the importance of the different vehicle parameters and their effect on cost. Figure 21.3 shows the elasticity of cost defined as percentage change in deterioration cost per gross ton-km when a single parameter is changed by 1 %. The elasticities are given for different curve radii and clearly show – in this context applied to Swedish mainline traffic – a large sensitivity on change in parameters concentrated to curves haning a modest curvature (600–2000 m radius). The reason is a large proportion of such curves combined with a considerable deterioration in them.

From Figure 21.3 it can be concluded that vehicle features and parameters influencing wear and RCF are most importance in curves, while parameters influencing track settlement and component fatigue are most important in wider curves or on straight track. Also, track lubrication is essential for reducing wear of rails. All this is well known, but the model makes it possible to quantify the resulting effect on costs.

The share of costs for wear and RCF depends on the curve distribution, but also on track lubrication and vehicle characteristics. Generally, there are large differences among vehicle types with respect to their tendency to



21.3 Deterioration per gross ton-km expressed as cost elasticity. Example of passenger traffic on Swedish national rail network.

cause track deterioration. In particular, axle load, wheelset steering ability and unsprung mass are important features and parameters.

21.6 Conclusions and future trends

By predicting the progress of track deterioration many benefits for the infrastructure manager can be achieved; for instance, better planning of investments and maintenance, and giving incentives to vehicle operators by reflecting marginal deterioration cost in differentiated track access charges.

Based on earlier research by UIC/ORE and more recent findings from the UK on wear and RCF damage on rails, a 'state-of-the-art' model for track deterioration has been formulated and implemented in a software package called DeCAyS. When using DeCAyS it can be seen that there are large differences among vehicle types regarding their tendency to deteriorate the tracks:

- Settlement and component fatigue are proportional to the third power of the actual wheel load and the deterioration, therefore, increases

progressively with increased axle load. Also unsprung mass is also important.

- Wheelset steering capability is crucial for wear and RCF damage. If a 'stiff' bogie is replaced by a flexible bogie, the saving in cost could be in the order of 2 to 10 times per tonne-km. Also, most standardized freight wagons in Europe have a low or moderate amount of wear thanks to their flexible wheelset guidance.
- Heavy freight wagons with flexible wheelset guidance cause mainly settlement and component fatigue.

The results presented should be viewed as an example only since, to a large extent, they reflect Swedish conditions to which the model has been applied. The model and the methodology are, however, valid in general. It is possible to quantify the resulting effect on costs.

Much of the model input demands historic data on costs and traffic volumes at a detailed level. Despite the comprehensive 'state-of-the-art' model, there is still potential for further improvement and validation. This requires detailed and reliable statistical data on both cost and traffic. It could then be possible to adjust the traffic dependent deterioration cost more specifically to different types of track.

21.7 References

1 Öberg J. (2006), *Track Deterioration of Ballasted Tracks – Marginal Cost Models for Different Railway Vehicles*, TRITA-AVE Report 2006:88, Division of Rail Vehicles, Royal Institute of Technology (KTH), Stockholm, Sweden.

2 ORE (1988), *Dynamic Vehicle/Track Interaction Phenomena, From the Point of View of Track Maintenance*, Report no. 3, final report, ORE Question D 161, Utrecht, the Netherlands, 41–7.

3 ORE (1987), *Dynamic Effects of 22.5 t Axle Loads on the Track*, Report no. 4, ORE Question D 161.1, Utrecht, the Netherlands, 35–42.

4 Booz Allen & Hamilton (2005), *Review of Variable Usage and Electrification Asset Usage Charges: Final Report. 2005*, Report to the Office of Rail Regulator, McLean, VA, USA, 54–68.

5 Burstow M.C. (2004), *Whole Life Rail Model Application and Development for RSSB – Continued Development of an RCF Damage Parameter*, Report AEATR-ES-2004-880, Issue 2, AEA Technology, Derby, UK.

6 Jenkins H. H., Stevenson J. E., Clayton G. A., Morland G. E., Lyon D. (1974), The effect of track and vehicle parameters on wheel/rail vertical dynamic forces, *The Railway Engineering Journal*, **3**(1), 2–16.

7 Wrang M. (2006), *OTU Analysis of Data From Vehicle Dynamics Tests*, MiW Consult, Stockholm, Sweden.

8 Pearce T. G. and Sherratt, N. D. (1991), Prediction of wheel profile wear. Wear, **144**, 343–51.

9 Enblom R. (2006), *On Simulation of Uniform Wear and Profile Evolution in the Wheel – Rail Contact*, TRITA-AVE, Report 2006:83, Division of Rail Vehicles, Royal institute of Technology (KTH), Stockholm, Sweden papers C and D.

10 UIC (2005), *Testing and approval of railway vehicles from the point of view of their dynamic behaviour – Safety – Track fatigue – Ride quality*, Code 518, International Union of Railways, Paris, France.
11 Andersson E. and Stichel S. (1999), Model for Track Deterioration; Vehicle Parameters and Reference Values. Division of Rail Vehicles, Royal Institute of Technology (KTH), Stockholm Sweden (in Swedish).
12 Öberg J., Andersson E. and Gunnarsson J. (2007), *Track access charging with respect to vehicle characteristics*, Report F07-4947/TR99 LA-BAN 2007/31 (2nd edn), Banverket, Borlänge, Sweden.
13 Persson, I. (2006), *Using GENSYS.0603*, DEsolver, Östersund, Sweden.
14 Andersson M. (2007), *Empirical Essays on Railway Infrastructure Costs in Sweden*, Doctoral Thesis in Infrastructure, TRITA-TEC-PHD 07-002, Royal Institute of Technology (KTH), Stockholm, Sweden.

21.8 Appendix 1: Notation

b_0	Half lateral distance between left and right wheel contact points (m)
$\overline{C}$	Average marginal cost (SEK) per gross tonne km
ΔE_a	Total annual accumulated deterioration of track given as a marginal cost for all deterioration mechanisms (SEK)
$E_{a,Z}(R_j,T_a)$	Total accumulated deterioration of track within a certain curve zone for vehicle type Z, given as a marginal cost for all deterioration mechanisms due to accumulated tonnage T_a (SEK)
$e_{ton_km,Z}(R_j)$	Deterioration per gross tonne-km for vehicle type Z within curve zone R_j (SEK/gross tonne-km)
$\overline{F}_v$	Total creep (friction) force (N)
$\overline{F}_v \cdot \overline{v}$	Wear number (approximate friction energy dissipation) (Nm/m)
$f(\overline{F}_v\overline{v})$	The function shown in Fig. 21.1, relating wear and rolling contact fatigue (RCF) to the friction energy dissipation (in this context for the outer wheels)
g	Gravitational acceleration (9.81 m/s^2)
h	Total height to centre of gravity (m)
I	Cant defiency (mm)
i	Index for axles
j	Index for curve zone (1,2,3,...)
K_s	Dynamic coefficient in vertical wheel load model
K_Z	Vehicle coefficient in vertical wheel load model
k_1, k_2, k_{34}	Cost coefficients (calibrated against average annual marginal cost for deterioration). Subscript 1 for track settlement, 2 for component fatigue and 34 (3 and 4 combined) for abrasive wear and rolling contact fatigue of rails

L_{R_j}	Length of curve zone R_j with mean radius R (m)
L_t	Total length of track section or network (m)
m_Z	Mass of vehicle type Z (kg). This refers to the mass in working order with an addition of 50 % passengers for each passenger vehicle, and the total weight of each freight wagon respectively
$m_{u,w}$	Unsprung mass per wheel including half of axle, brake discs, bearings, gear boxes and (if applicable) part of unsprung traction motor on the axle etc.[3] (kg) (by simplification the unsprung mass per wheel could be set to half of the unsprung mass per axle)
n_Z	Number of axles on vehicle type Z
P	Static axle load (tonne)
Q_c	Quasistatic wheel load contribution due to curving (i.e. load shift) (N)
Q_d	Vertical dynamic wheel load contribution (up to 90 Hz usually) i.e. $Q_{d20\,Hz} + Q_{d\,hf}$ (N) (see below)
$Q_{d20\,Hz}$	Vertical dynamic wheel load contribution processed by a 20 Hz low-pass filter (N)
$Q_{d\,hf}$	Vertical high-frequency dynamic wheel load contribution with frequency content above 20 Hz and up to about the highest expected sleeper passing frequency (usually 90 Hz) (N)
Q_{qst}	Vertical quasi-static wheel load i.e. $Q_r \pm Q_c$ (N)
Q_{tot}	Total vertical wheel load (usually low pass filtered by 90 Hz) i.e. $Q_{qst} + Q_{d20\,Hz} + Q_{d\,hf}$ (N)
Q_0	Static wheel load (N)
Q_r	Representative vertical wheel load including passengers (Appendix 2) (N)
R_j	Curve zone no. j with mean radius R (m)
T_a	Tonnage or traffic volume carried by the track since built or maintained (gross tonne-km)
T_Z	Annual tonnage or traffic volume for vehicle type Z (gross tonne-km)
V	Representative vehicle speed in service (km/h)
$\overline{\nu}$	Total creepage (-), i.e. relative sliding velocity divided by vehicle speed
Y_{qst}	Quasi-static lateral wheel load or 'guiding force' (N)
y	Lateral displacement of vehicle centre of gravity at representative cant deficiency (m)
Z	Subscript for 'vehicle type'
μ_0	Wheel–rail coefficient of friction (at large creepage)

21.9 Appendix 2: Vertical dynamic wheel load model

Data from track force testing on four rail vehicles have initially been used to calibrate the model. The model has later been refined using more recent unofficial proprietary test data on additional vehicle types. The initial vehicles are an Rc-locomotive, X2 (X2000) light-weight locomotive, Oeresund train four axle unit (OTU) and a loaded freight wagon with standard running gear. The OTU vehicle in particular has undergone extensive testing. Known vehicle data are summarized in Table 21.3.

There are usually uncertainties regarding the distribution between the two dynamic contribution forces $Q_{\mathrm{d20\,Hz}}$ and $Q_{\mathrm{d\,hf}}$ since these measurements have been conducted according to the older method, that is, the vertical track forces were low-pass filtered at 90 Hz. However, the sum of the two dynamic contribution forces is known since it can be determined by calculating the quantities $Q_{\mathrm{tot}} - Q_0 - |Q_{\mathrm{c}}|$.

Dynamic contribution forces are evaluated at the 99.85 percentile which means that 99.85 % of measured data are lower than the stated forces, as set out in UIC code 518. The values given are taken as mean values of the 99.85 percentile.

The quasi-static wheel load contribution is determined according to Eq. (21.8).

$$|Q_{\mathrm{c}}| = \frac{Q_{\mathrm{r}}}{b_0} \cdot \left(\frac{|I|}{2000 b_0} \cdot h + y \right) \qquad [21.8]$$

Table 21.3 Data on vehicle types for calibration

Vehicle type	Rc	X2	OTU	Freight wagons
Static axle load P [t]	19.5	18.3	15.5	22.5
Vehicle speed V [km/h]	140	180	162	90
Cant deficiency I [mm]	100	245	135	70
Unsprung mass per wheel $m_{\mathrm{u,w}}$ [kg]	1450	950	800	750
Height to total centre of gravity h [m]	1.6	1.2	1.35	1.6
Lateral displacement of total centre of gravity y [m]	0.05	0.08	0.07	0.04
Quasi-static wheel load contribution due to curving Q_{c}[kN]	20	33	19	17
Dynamic wheel load contribution up to 20 Hz $Q_{\mathrm{d20\,Hz}}$ [kN]	?	?	10	?
High-frequency dynamic wheel load contribution 20–90 Hz $Q_{\mathrm{d\,hf}}$ (kN)	?	?20	?	
Total measured wheel load Q_{tot} (kN)	165	160	126	160–170

where b_0 depends on track gauge. For a standard track gauge of 1435 mm, b_0 equals 0.75 m. The representative wheel load is given by:

$$Q_r = m_z/(2n_z) \cdot g \quad [21.9]$$

A model approximately satisfying data in Table 21.3 becomes:

$$Q_{d20\ Hz} = 0.80 \cdot K_z \cdot K_s \cdot P \cdot (V + 760) \quad [21.10]$$

$$Q_{d\ hf} = 1.32 \cdot K_s \cdot V \cdot \sqrt{m_{u,w}} \quad [21.11]$$

where:

K_Z = 0.20 for passenger vehicles;
0.40 for locomotives and freight wagons (which regularly have a stiffer suspension than the previous vehicle types).
Values could differ from these if the vehicle is non-typical, having 'track friendly' bogies, etc. In these cases, the actual measured or simulated forces may be used.

K_s = 0.0036 on tracks for speeds > 120 km/h;
0.0042 on tracks for speeds ≤ 120 km/h (for freight trains a mean value of K_s = 0.0039 has been used).

These values are determined for Swedish conditions.

Copyrighted Materials
Copyright © 2009 Woodhead Publishing Retrieved from www.knovel.com

Part II

Industrial context – managing the wheel–rail interface

22

Managing the wheel–rail interface: Railway infrastructure maintenance in a severe environment: The Swedish experience

P.-O. LARSSON-KRÅIK, Banverket, Sweden

Abstract: This chapter comprises a short introduction to and overview of the challenges that face engineers and researchers working in the field of heavy haul railway operations in a cold climate. If these problems are to be overcome, experts from different engineering disciplines need to work together in order to find solutions for the next generation of heavy haul operation, which will see heavier, longer and faster trains with new train control systems and operating philosophies. Malmbanan runs from the port in Narvik, across the Swedish/Norwegian border via the mines in Kiruna and Malmberget, to the port in Luleå on the coast of the Gulf of Bothnia. The 100 year old ore line runs through rough terrain including high mountains, peat, terraces situated on the fjords in Norway and numerous short bridges and culverts, and part runs above the Arctic Circle. Over the course of its life, the line and its ore traffic have been continuously upgraded. Today 750 m long trains with 30 tonne axle loads are in operation. If cost-efficient operation is to be maintained, engineers must continue to seek new and effective maintenance and operating systems for both infrastructure and rolling stock.

Key words: railway, heavy-haul, maintenance, cold climate, Malmbanan.

22.1 Introduction

Most of the world's iron ore is mined in enormous open pit mines, while LKAB (Luossavaara-Kiirunavaara Aktiebolag – an international high-tech minerals group of producing upgraded high-tech iron ore products for the steel industry and supplying industrial minerals products to other sectors) operates its mines deep under ground, which is both more expensive and more complicated than open-pit mining. In addition, the LKAB railroad traverses the Arctic Circle with its long winters characterised by extreme cold and frequent heavy snowfall, again, a far different and, arguably, more challenging operating environment than its global counterparts in Brazil, South Africa and Australia.

The ore line (Malmbanan) runs from the port in Narvik (on the coast of the Norwegian Sea), across the Swedish/Norwegian border via the mines in Kiruna and Malmberget, to the port in Luleå on the coast of the Gulf of Bothnia. The railway between the two ports is 473 km long, and apart from

39 km in Norway, the line in Sweden for which Banverket is responsible is known as the Malmbanan. It runs through rough terrain including high mountains, peat, terraces situated on the fjords in Norway and numerous short bridges and culverts. A considerable part of Malmbanan is situated north of the Arctic Circle, and for approximately 70 km it runs beside the second deepest (168 m), lake in Sweden, Lake Torneträsk. This area has a varied topography, geomorphology, geology and climate, as well as flora and fauna. The area is subarctic and mountainous with a sharp gradient between maritime and continental climate. The Torneträsk valley was the area for the main flow of ice during glaciation and deglaciation fed by valley glaciers, and the landscape has been heavily reshaped by the processes. At higher altitudes in the western parts of the area there are still some glaciers. The average annual temperature in this area is approximately –1.0 °C. July is the warmest month (mean about +11 °C) and January the coldest (mean about –12 °C). Annual precipitation at this part of Malmbanan varies between about 1000 mm in the west close to Norwegian border and 400 mm in the east.

The climate conditions in the area create an increased risk of slab and snow avalanches. To reduce the risk of derailment due to avalanches, bridges and artificial tunnels have been built in high-risk track sections. The bridges are designed and situated to permit avalanches to pass under the railway and the artificial tunnels are designed to allow them to pass over the railway. These 14 avalanche protection galleries are situated between the stations of Björkliden and Riksgränsen. In low-risk avalanche sections there are also instrumented arrays of sticks placed along the track to indicate if an avalanche has begun. These stick systems are a part of the signalling system and designed to stop the train before it enters the hazard area. The system also informs the train controllers (dispatchers) that this track section is currently at high risk of avalanches.

22.2 General description of Malmbanan (Swedish ore line)

Construction of the standard gauge rail line of Ofotenban (Björnefjäll–Narvik in Norway) and Malmbanan (Riksgränsen–Luleå), the Swedish ore line, began in 1898 and took into consideration the extreme northern climate and remote location. This involved the building of 25 tunnels and 125 bridges through rugged terrain characterized by numerous sharp curves and steep elevation changes. The line, completed and operational by 1902, was later electrified with 15 kV 16 2/3 Hz, completed on 19 January 1915, and remains so to this day. The single track line was originally constructed for only 14 tonnes axle loads. Work on upgrading the ore line for a 30 tonne axle load was completed

in 2000. Overall, this means that each car will be able to carry 100 tonnes instead of 80, and that each train set will consist of 68 cars instead of today's 52. The total train weight has increased from 5200 to 8160 tonnes. Unlike its global counterparts, the train operator MTAB neither owns nor manages the track infrastructure; rather it is under the control of two national railway authorities, Banverket in Sweden, and Jernbaneverket in Norway.

22.2.1 Operations on the line

The main part of rail traffic consists of iron ore transports. There is some passenger traffic as well as goods traffic. The finished products from the mine are transported from the ore processing plants to customers by rail and by ship via the shipping ports at Narvik and Luleå. Rail traffic on Malmbanan between Luleå and Narvik is managed by LKAB's Swedish subsidiary Malmtrafik i Kiruna AB (MTAB) and its Norwegian subsidiary Malmtrafikk AS (MTAS). LKAB also operates the ore harbours in Narvik and Luleå. The Narvik ore harbour can accommodate vessels of up to 350 000 dwt. From the harbour in Luleå, products are delivered mainly to customers in the Baltic Sea region.

After the pioneering separation in 1988 of Swedish State Railway's rail infrastructure from train operations, LKAB wanted direct control over the transport by rail of its principal product. This was achieved by setting up MTAB as a subsidiary train operating company in Sweden. In 1996 Norway followed Sweden's example and established Jernbaneverket to own and manage the national rail infrastructure, including the isolated line to Narvik. Iron ore trains to Narvik are now managed within Norway by Malmtrafikk AS, a subsidiary of MTAB. LKAB had two objectives: one was to control the cost of the operation, and the other was to develop the Malmbanan into an efficient heavy-haul railway, raising the axle load from 25 to 30 tonnes, procuring new wagons and locomotives and operating longer and heavier trains in order to improve productivity. Rail operations in Norway are managed by MTAB's subsidiary, MTAS, which has its own organisation.

22.2.2 Harbours

LKAB has two shipping harbours: Narvik and Luleå. Work is being carried out systematically to improve the logistics from mine to customer. Among other things, this means that customers are offered systematised transport and that stocks in the harbours are kept to a minimum. Continuous, rolling discharge has been implemented in the harbours. Trains are unloaded while in motion, and a train set of 52 cars can be unloaded in 30–40 mins.

22.2.3 Terminals

MTAB has terminals for loading and discharge of iron ore products and additives in Kiruna, Malmberget, Svappavaara, Narvik and Luleå. The Southern Circuit (Malmberget–Luleå) has been upgraded for a 30 tonne axle load since 2000, and The upgrade also includes the terminals in Malmberget and Luleå.

22.3 Locomotives and cars

LKAB uses its own locomotives and cars. Rolling stock includes regular-traffic and terminal locomotives and various types of ore cars. The cars that are currently in operation on the ore railway carry a payload of between 80 and 100 tonnes. With its current fleet of locomotives and cars, MTAB has a mine-to-harbour freight capacity of more than 23 million tonnes per year. This corresponds to about 7000 fully loaded ore trains per year.

Each day, five or six trains make the 220 km, five to six hour journey from Malmberget to Luleå. Most products from Kiruna and Svappavaara are transported to Narvik for further delivery to customers in continental Europe and the rest of the world. Daily, 11–13 trains make the 170 km run between Kiruna and Narvik.

Loading and unloading arrangements had to be replaced by new facilities to respond to the increased transport volume on the track. Banverket's task was to meet LKAB's future requirements. Increasing capacity for higher traffic and heavier axle loads takes time, so planning and construction started well in advance of introduction. This work started in 1998, and the project will be completed in 2010. For LKAB, the benefit from the 30 tonne axle loads that became available to Luleå on October 2000 is that the payload per car is increased by 25 % from 80 to 100 tonnes. The extension of sidings on the southern section of Malmbanan to accept 750 m long trains means that the standard train formation to Luleå is now 68 cars. Agreements were signed with Bombardier Transportation (formerly Adtranz) for delivery of nine ore train locomotives, and with South African Transwerk for delivery of a set of 100 tonne ore cars. The first regular-traffic locomotive, of the IORE type, was delivered in 2000. The following eight engines arrived between 2002 and 2004. LKAB's new regular-traffic locomotives have replaced the Dm3 locomotives as the standard locomotive. The nine IORE engines have been operating on Malmbanan since 2004. The Malmberget–Luleå section of the line was first to be operated by the upgraded 30 tonne IORE and train set of 100 tonne cars, which were delivered from South Africa.

In March 2004, LKAB decided not to exercise its option to purchase more ore cars from Transwerk. Instead, LKAB, together with the Swedish company K-Industrier, has developed a new, larger ore car adapted for a

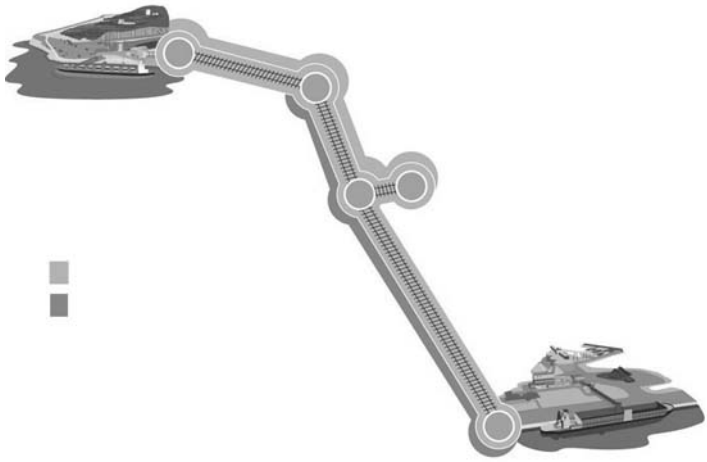
maximum 32.5 tonne axle load and with a maximum 108 tonne payload. Today, this car is operated with a 30 tonne axle load and 100 tonne payload The first car was delivered in Autumn 2005, and about five cars per week roll out from the Kiruna wagons assembly shop at LKAB in Kiruna. Once the Kiruna–Narvik section of the line has been fully upgraded for a 30 tonne axle load and all terminals, including the new Narvik ore harbour, are ready to receive the new cars, the capacity increase will have been completed.

22.3.1 Location of mines

Luossavaara-Kiirunavaara AB (LKAB) was founded in 1890, but the history of the ore fields dates back to the 1660s. It was then that the first known samples of the Gällivare ore were taken. The ore mountains of Kiirunavaara and Luossavaara were mentioned in writing for the first time in 1696. Many were aware that the mountains in the north bore great wealth, and many business ventures were started but were doomed to failure. It was not until the 1870s, with the advent of the Thomas process, a new method of producing steel from phosphorus-rich ore, that the ore deposits became commercially viable. When the railway to the ore fields was built and subsequently reached Narvik, large-scale extraction of the valuable natural resource could begin. Iron ore mining has driven the development of northern Sweden from a sparsely populated mountain and woodland region to a modern industrialized region. It is not only Kiruna and Gällivare/Malmberget that have thrived on mining. For the harbour towns of Narvik and Luleå, construction of Malmbanan brought a surge in economic development. In the 1800s, Luleå was still little more than a small town with a couple of thousand inhabitants. Only when the ore trains begin to roll did the town begin to grow. The railway brought electrification and the establishment of steelworks and other industries. To this day, the iron ore from the mines in Kiruna and Malmberget is the lifeblood of an entire region, even on the Finnish side of the Bay of Bothnia, where iron ore products from LKAB feed the steel industry in Finland.

Iron pellets are transported from the mines of Kiruna and Malmberget (Gällivare) to the ports of Narvik and Luleå as well as to the steel manufacturing company SSAB steel plant in Luleå. Some products necessary for the production of pellets are transported in the return direction from the ports in Luleå and Narvik. Otherwise, return trains are empty. The railway system that the mining company LKAB operates on is briefly described in Fig. 22.1. LKAB uses its own locomotives and cars. Rolling stock includes regular-traffic and terminal locomotives and various types of ore cars. In Fig. 22.2 the new IORE locomotive is shown.

There are workshops and service sheds for rolling stock maintenance in Kiruna, Malmberget and Narvik. Operations are controlled from an ore transport centre in Kiruna. Advances in rolling-stock technology mean

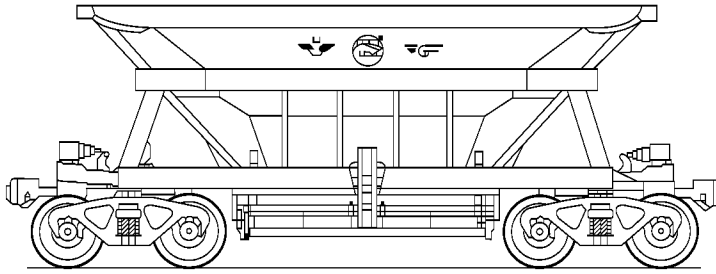


22.1 The LKAB iron ore operation, split between the North Circulation and the South circulation. [Courtesy – Thomas Nordmark, LKAB]

22.2 The new locomotive IORE. [Courtesy – Tomas Nordmark, LKAB]

that the chain of ore logistics can be utilised with greater efficiency, with locomotives that can haul longer, heavier trains. The fleet of rolling stock includes regular-traffic and terminal locomotives as well as various types of ore cars.

The ore car is a four-axle heavy-haul freight wagon specially designed for the transportation of iron ore. This so-called BoBo vehicle contains a car body sitting on two two-axle bogies. A side view of an ore wagon is shown in Fig. 22.3. The bogies are so-called three-piece bogies. This means that



22.3 Four-axle ore wagon with three-piece bogies. [Courtesy – Ansel Berghuvud, Luleå university of technology]

the bogie frame consists of two side frames and a bolster coupled together into a bogie frame.

22.3.2 Car properties

Inertia properties

The car body mass including the payload (80 000 kg) is 90 000 kg. The adopted inertia properties of the car body are:

- car body mass 90 000 kg
- mass moment of inertia in roll 130 000 kgm^2
- mass moment of inertia in pitch 400 000 kgm^2
- mass moment of inertia in yaw 400 000 kgm^2
- vertical location of centre of gravity above railhead 2.49 m
- longitudinal distance between bogie centres 5.2 m

Typical characteristics in simulation of the ore car for inertia properties of one bogie bolster are:

- bolster mass 800 kg
- mass moment of inertia in roll 330 kgm^2
- mass moment of inertia in pitch 27 kgm^2
- mass moment of inertia in yaw 330 kgm^2
- vertical location of centre of gravity above railhead 0.56 m

Typical characteristics in simulation of the ore car for inertia properties of one side frame are:

- side frame mass 311 kg
- mass moment of inertia in roll 29 kgm^2
- mass moment of inertia in pitch 121 kgm^2
- mass moment of inertia in yaw 117 kgm^2
- vertical location of centre of gravity above railhead 0.60 m
- longitudinal distance between wheelsets 1.70 m

Typical characteristics in simulation of the ore car for inertia properties of one wheel set are:

- wheelset mass 1500 kg
- mass moment of inertia in roll 940 kgm^2
- mass moment of inertia in pitch 100 kgm^2
- mass moment of inertia in yaw 940 kgm^2
- wheel radius 0.485 m

Centre bearing

The bolster has a centreplate bearing and side bearings that couple the bogie to the carbody. A spherical centre plate bearing that allows relative rotational motion in roll, pitch and yaw between the car body and the bolster is used in the investigated bogie. The side bearings restrict the relative roll motion between the car body and the bolster. The adopted characteristics of the centre bearing are:

- vertical location of centre bearing above railhead 0.81 m
- friction moment that leads to sliding motion 9.4 kNm
- viscous damping 100 kNs/m
- mechanical play, roll 0.004 rad

Secondary suspension

The secondary suspension couples the ends of the bolster to the centre of the side frame using coil springs and friction wedges. The friction wedges have a constant preload for the studied vehicles, which means that the pre-load is independent of the vertical and lateral motion in the suspension. The sliding on the wedge surfaces provides damping in the suspension. This suspension gives the bogie non-linear characteristics since it comprises both friction damping and it allows relative motion between the frame parts within mechanical displacement, limiting play in all directions of motion. This will, for instance, allow the bogie frame to shear into lozenge (parallelogram) geometry on both straight and curved track.

Typical characteristics in simulation of the ore car of the secondary suspension are:

- suspension spring stiffness, lateral 2.3 MN/m
- suspension spring stiffness, vertical 5.0 MN/m
- suspension spring stiffness, roll 106 kNm/rad
- suspension spring stiffness, pitch 106 kNm/rad
- suspension spring stiffness, yaw 54 kNm/rad
- lateral offset from bogie centre 1.0 m

- friction force that leads to sliding motion, lateral 5.5 kN
- friction force that leads to sliding motion, vertical 5.5 kN
- friction moment that leads to sliding motion, roll 2.2 kNm
- friction moment that leads to sliding motion, pitch 2.2 kNm
- friction moment that leads to sliding motion, yaw 2.2 kNm

- mechanical play, lateral 0.006 m
- mechanical play, roll 0.1 rad
- mechanical play, pitch 0.1 rad
- mechanical play, yaw 0.1 rad

Primary suspension

The primary suspension of the standard bogie in use on Malmbanan consists of a plain metal-to-metal contact between the top of the bearing box and the side frame; hence the vertical suspension is considered negligible. The horizontal primary suspension motion is restricted by friction between the side frame and the top of the bearing box. The vehicle primary suspension also has displacement limiting play in the horizontal plane.

Typical characteristics in simulation of the ore car for the primary suspension are:

- friction force that leads to sliding motion 30 kN
- mechanical play, longitudinal 0.005 m
- mechanical play, lateral 0.005 m
- mechanical play, yaw 0.2 rad

22.3.3 Vehicle maintenance practices

Operations are controlled from an ore transport centre in Kiruna. MTAB has developed its maintenance from the earlier time-based intervals to distance-based intervals with fixed maintenance procedures, based on a transponder system called ATI (Automatic Truck Identification), an identification system with a radio link that is used to count the distance of each individual ore car. MTAB is continually working on the development of their maintenance strategy to meet new and higher demands of the operation.

However, with higher demands for availability and reliability from technical systems, the next step in the evolution of maintenance is the preventive approach of condition-based maintenance. The ability to follow the actual condition of the system makes it possible to utilize it more effectivly since maintenance activities can be scheduled in advance in a more precise manner and the system operations can be planned accordingly.

22.4 Train control system

There are two traffic control centres (TCC) on the ore line, one in Sweden and one in Norway. For the Swedish part of the ore line the control centre is situated in Boden, close to Luleå, and for the Norwegian side it is situated in Narvik. Both control centres have access to information about the current status on the other side of the national border. There is some help available from stand-alone onboard driving assistance systems which keep drivers informed of the positions of other trains on the line, allowing them to adjust their speed in light of any departures from the established timetable in order to maintain optimal train running. However, at least with regard to Swedish conditions, the most interesting of these is considered to be a system that communicates with the TCCs. This makes it possible to take the daily traffic situation into consideration. This fact, together with a mixed-traffic situation on the single line, makes it even more important to find driving assistance systems that can optimise use of the route.

New technologies and standards have been launched and are ready to be implemented on the Swedish network. A new European standard for dispatch systems, the European Traffic Management System (ERTMS/ETCS) is already in existence. This system is being implemented on the Swedish network, starting with the Botniabanan line. It is hoped that the system will then also be implemented on the ore line. The ore line is also being upgraded with digital radio communication (GSM-R, the Global System for Mobile Communications – Railways) which enables communication with other systems in the TCC and the train.

22.5 Infrastructure configuration

The track gauge is 1435 mm throughout the whole of Sweden. The track configuration is an electrified single track using the block system of automatic train control (ATC). There are 116 point mechanisms along main lines, 84 point mechanisms along side lines and 37 point mechanisms along other lines. There are 46 level crossings and 21 structures with whole barriers and 11 structures with semi-barriers. Of the railway stations 13 can accommodate 750 m long trains Maximum per meter weight for track and the bridges is 12 tonnes (STVM 12). The infrastructure configuration is presented in Table 22.1.

In Table 22.2 the detailed configuration of the track is presented.

On Malmbanan, 203 km is of the rail type UIC60 (60 kg/m) and 270 km of the older rail type BV50 (50 kg/m) (see Table 22.3). In 2007 a reinvestment project was begun on the oldest section of the track. The 50 kg/m rail profile will be replaced with the standard UIC60 rail profile.

Rail defects on Malmbanan occur due to a number of causes, usually the

Table 22.1 Configuration of the ore line

	Malmbanan Sweden (MTAB)	Ofotban Norway (MTAS)
Route length	497 km	39 km
Amount of double track	0 km	0 km
Track gauge	Standard gauge 1435 mm	Standard gauge 1435 mm
Rail	50 & 60 kg/m	54 kg/m
Maximum grades	1.7 % empty direction 1.4 % loaded firection	1.8 % empty direction 0 % loaded direction
Locomotive fleet	12 DM 3 9 655 hp 9 IORE 14 483 hp 6 diesel/electrics	AS MTAB AS MTAB AS MTAB
Ore wagon fleet 25 tonne	1100	AS MTAB
Ore wagon fleet 30 tonne	280	AS MTAB
No. of trains per day	22 (including MTAS)	12
Train consist of	1 IORE locomotive & 68 cars head end power	AS MTAB
Axle load	Maximum 30 tonnes	Maximum 30 tonnes
Size, weight of trains	750 ms, 8520 tonnes	600 ms, 5473 tonnes
Maximum speed	60 km/hr empty (to 70) 50 km/hr loaded (to 60)	60 km/h empty (to 70) 50 km/h loaded (to 60)
Cycle time	Schedule calls for 12–14 hours Average is 18–22 hours	AS MTAB AS MTAB
Total iron ore railed in 2006 by MTAV and MTAS: 29 Million tonnes		

Table 22.2 Configuration of the track

	1	2	3	4	5
Length [km]	39	128	46	106	217
Sidings	4	12	1	9	22
Tunnels	22	3	0	0	0
Bridges	7	43	2	21	52
Straight, %	21	43	52	47	65
Curves, %	79	57	48	53	35
Sharpest curve [m]	260	476	559	561	335
Max grade, %	0,0 d** 1,8 u**	1,2 d 1,4 u	1,1 d 1,2 u	1,2 d 1,4 u	1,7 d 1,4 u

(1) Riksgränsen–Narvik; (2) Kiruna–Riksgränsen; (3) Svappavaara–Kiruna; (4) Kiruna–Malmberget; (5) Malmberget–Luleå (see Fig. 22.1). d = down, u = up.

Table 22.3 Rail life for the different rail profiles

Rail life in MGT	50 kg/m – Banverket	60 kg/m – Banverket	54 kg/m – Jernbaneverket
Tangent track	1049	1335	285
1 degree curves	480–610	611–776	228
2–3+ degree curves	280	355	133

same as for other heavy-haul lines around the world. Surface-initiated defects are formed due to the increase in traffic density and axle load. Subsurface defects are caused by metallurgical defects (for example: shelling, tache ovale and longitudinal vertical crack). On Malmbanan there are defects originating from rail manufacturing (rail manufactured before 1991), installation and use (wheelflat defects) and defects caused by the exhaustion of the rail steel's inherent resistance to fatigue damage. There are also defects related to the joining of the rails and defects related to rail quality (for example: horizontal head cracks, tache ovale).

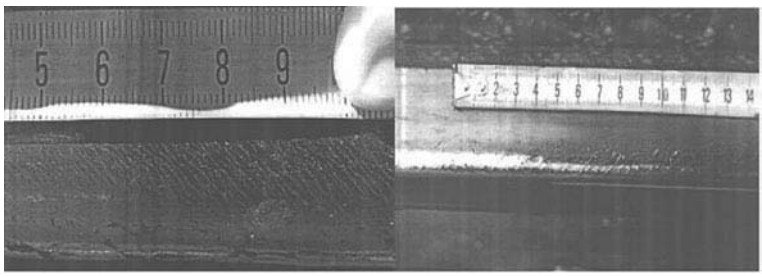
22.5.1 Head checks

Head checks are groups of fine surface cracks at the gauge corner at a distance of 0.5–7 mm from each other. Generally, contact stresses are low in the crown area as this has larger profile radius in comparison to the gauge side of rail. However, high contact stresses are generated on the gauge corner of the high rail, which generally has curve radius from 600–800 m. Head checks may also occur in tighter curves near the gauge corner of the high rail.

Head checks are surface-initiated contact fatigue defects (see Fig. 22.4), Which are developed with an angle of 30–60° to the longitudinal axis of the rail. This angle is related to the rail curve radii. Head checks are controlled using preventive rail grinding programs. Severe head checks sections are controlled using rail replacement.

22.5.2 Shelling

Banverket define shelling as a defect caused by loss of material initiated by subsurface-initiated fatigue. This type of defect has almost disappeared on Malmbanan after the introduction of the preventive rail grinding maintenance program in 1998. However, when it does appear on Malmbanan, shelling normally takes place at the gauge corner of high rails in curves. When these cracks emerge on the surface, they cause the metal to come out from the

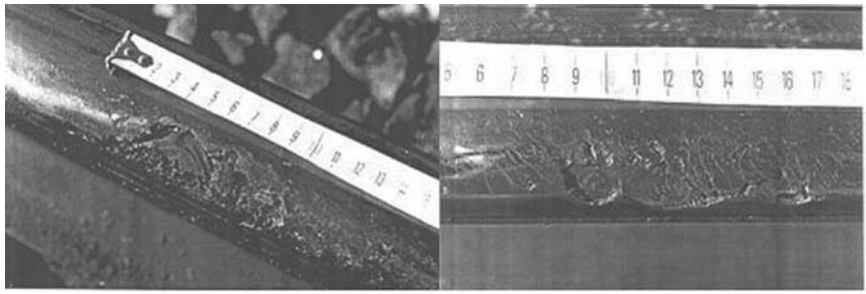


22.4 Head checks in rails.

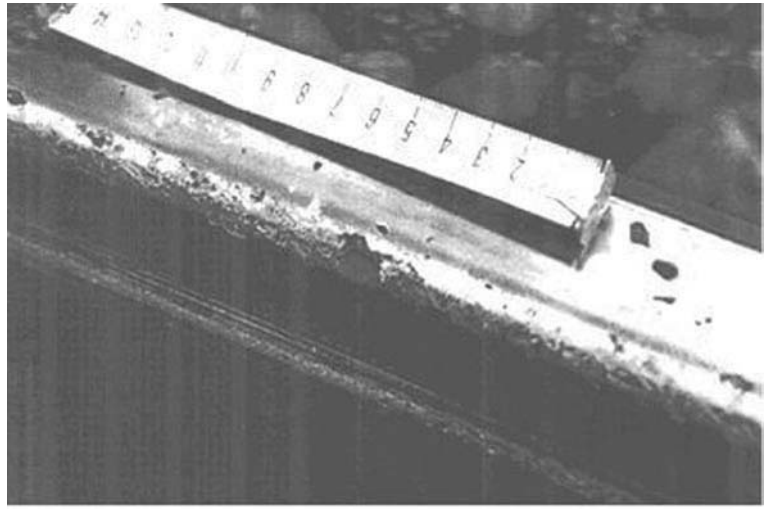
crack area. Figure 22.5 shows gauge corner shelling. It is eliminated by grinding.

22.5.3 Spalling

When a surface-initiated crack development path is intersected by other similar shallow cracks on the railhead area, a shallow chip of rail material falls out. This is known as spalling (see Fig. 22.6). This type of defect has almost disappeared on Malmbanan since the introduction of the preventive rail grinding maintenance program in 1998. However, spalling occurs at a much later stage in the crack propagation phase if it is left untreated in the grinding maintenance program. Spalling is usually more frequent in cold climates as rail material stiffness increases.



22.5 Gauge corner shelling in rails. [Courtesy – Mats Rhen and Dan Larsson, Luleå University of technology]



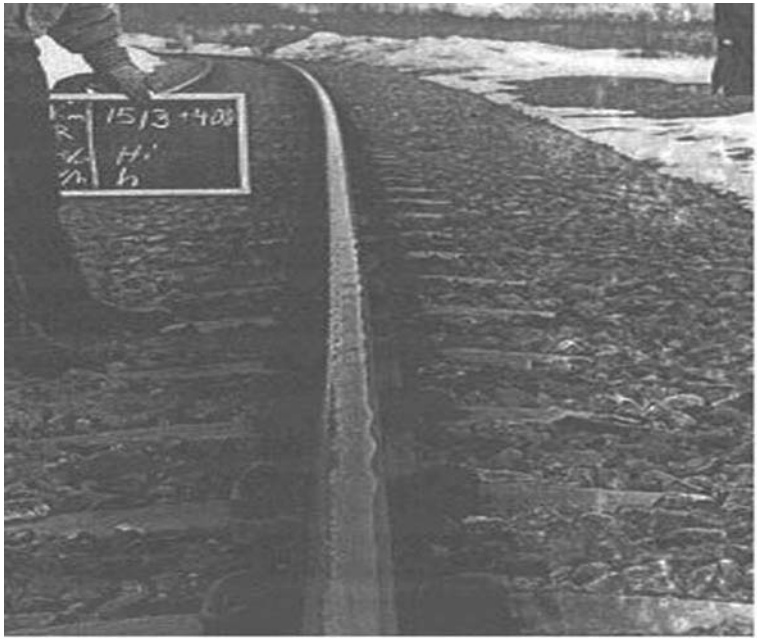
22.6 Spalling in rails. [Courtesy – Mats Rhen and Dan Larsson, Luleå University of technology]

22.5.4 Corrugation

Corrugation on Malmbanan is defined as a rail flaw consisting of the wave-like wearing of the rail tread visualised as peaks and valleys, in other words, it is a periodic irregularity of the rail surface (see Fig. 22.7). Rail corrugations are the result of a damage mechanism, such as wear, fatigue or plastic flow-operating at some characteristic frequency. Rail corrugations increase the dynamic load forces and may be responsible for loosening of rail fastenings, ballast deterioration, increase in noise and vibration levels. Corrugation can be due to several reasons and it is difficult to correlate the causes of corrugation. It is eliminated by rail grinding.

22.5.5 Plastic flow and tongue lipping

Plastic flow occurs in the railhead area, the depth of which may be up to 15 mm. It occurs on the field side of the low rail due to overloading and high friction, and it may also occur in the low rail on the curves due to overloading. Tongue lipping is also a form of plastic deformation, but it is initiated by surface cracks. These cracks partially separate a layer of material from the bulk of rail. Under high axial loads, these separated protrusions



22.7 Corrugation in rails. [Courtesy – Mats Rhen and Dan Larsson, Luleå University of technology]

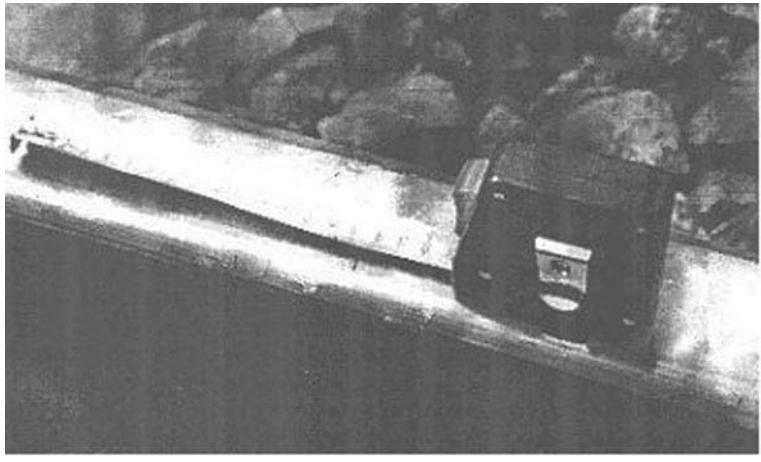
deform plastically as shown in Fig. 22.8. Tongue lipping gives an indication of the presence of cracks. This defect is eliminated by grinding which would also bring back the original rail profile.

Rail and weld degradation depends on many factors, climate, operational environment and track construction being some. Small imperfections in welds can cause cracks to initiate. A defect-free weld requires a skilled workforce, better weld material and welding techniques along with improved welding equipment. On Malmbanan, the inspection, welding and rectification process becomes a costly affair due to the presence of snow and ice during winters. Most of the defects which do not pose immediate risk of damage to rail assets or derailment risk are deferred until the end of winter.

22.5.6 Substructure

The track structure is divided into superstructure and substructure. The substructure is defined by Banverket so that it includes ballast, sub-ballast, frost insulation (if needed) embankments, embankments fill and subsoil. For old tracks there is often in sufficient thickness of each layer, hence, problems will occur when it comes to track stability and drainage.

No major reinforcement of the underground has been made since Malmbanan was built in the early 1900s. Since this railroad track still goes along the same line, there are many reasons to monitor the embankments to ensure that they are under control. In cold regions, the seasonal freezing and thawing of the ground will affect the substructure and the track quality. It all depends on how the frost penetration developeds and how problems related



22.8 Tongue lipping. [Courtesy – Mats Rhen and Dan Larsson, Luleå University of technology]

to frost are avoided. Frost related problems areas on Malmbanan are treated with insulation in order to reduce the frost penetration and, hence, reduce the seasonal changes in track quality. On Malmbanan 45 km of track has this kind of frost insulation. To ensure good dynamic condition of the track, all the different substructure layers need to be designed so that they can interact in order to reduce the static and dynamic forces from the traffic.

The minimum ballast thickness on Malmbanan is 50 cm, measured from the rail foot. Ballast used on Malmbanan is defined according to classification grade I, including 31.5/63 mm which implies that there is a rather great portion of 22.4 and 31.5 mm. The classification system is defined in the European standard EN 13450 'Aggregates for railway ballast'. A typical track structure on Malmbanan is presented in Fig. 22.9.

22.5.7 Bridges

Along the ore line there are 114 rail bridges, 12 steel bridges, seven mountain culverts and 12 road bridges. Along the line there are 722 ducts (culverts, of which 658 are made of stone).

22.9 Track structure on Malmbanan. [Courtesy – Jonas Sjöberg, Jörgen Noppa Banverket]

22.6 Electrical power system

Electrical supply is provided via the AT (automatic transformers) system. The AT system is a strong system yielding a smaller energy loss and lower resistance. There are auxiliary power systems between Boden and Björkliden consisting of two-phase or three-phase 10 kV systems. Auxiliary power is an auxiliary system which supplies the Swedish Rail Authority's own electrical facilities. The line between Björkliden and Riksgränsen is covered by regional subscription service.

22.7 Track maintenance practices

Banverket has adopted maintenance terminology according to European Standard EN 13306:2001 and EN 50126. This is to (i) ensure the availability of the item for the required function, often at optimum cost; (ii) consider the safety requirements associated with the items for both maintenance and user personnel and, when necessary, any impact on the environment; and (iii) uphold the durability of the item and/or the quality of the product or service provided when considering, if necessary, the cost.

Banverket's administrative units plan and procure operation and maintenance as well as the conversion and extension of state railway installations. The orders are placed internally from the production units, or externally from contractors and consultants. All maintenance on the Malmbanan ore line is done using the in-house contractor (Banverket Production).

The annual maintenance activities are divided between corrective maintenance, preventive maintenance and reinvestments. Corrective maintenance includes procedures such as emergency fault rectification, post-inspection measures and damage repair. In terms of rail maintenance, emergency fault rectification will be required on detection of severe defects. The concept of post-inspection measures was introduced in 1999. This concept was introduced in order to keep an account of the measures implemented immediately after inspection. Before 1999, some of these measures were included in preventive maintenance. Preventive maintenance is maintenance carried out at predetermined intervals or according to prescribed criteria and intended to reduce the probability of failure or the degradation of the functioning of an item. Preventive maintenance consists of procedures like safety and maintenance inspection, overhauling, replacement and track straightening. Reinvestments have, in principle, the same objective as preventive maintenance, i.e. restoring the condition of the track. Reinvestments are carried out when the assets become worn out or uneconomical to maintain.

Banverket uses different track monitoring guidelines for monitoring track and track components. It specifies minimum requirements for the infrastructure maintainer. The monitoring and maintenance includes the

functions of inspection and testing, assessment of inspection and test results and execution of corrective or preventive actions. The objectives are to inspect the critical elements of the track to determine its condition; record defects which might affect, or have the potential to affect, the capability of the track to safely perform its required function; carry out assessments to determine the capacity of the track; and, finally, take actions where the track is unable to carry out the required function safely (for example, where conditions are outside acceptable limits). The track maintenance plan includes activities such as non-destructive testing (NDT) of the rails, using different inspection cars every 12–15 MGT that measure rail defects, track geometry, rail profile and electric power system geometry.

Banverket uses NDT cars and visual inspection to inspect rails for identification of possible defects (see Figs 22.4, 22.5, 22.6 and 22.8). To confirm, validate and estimate potential risk, each rail failure detected by a NDT car is verified by hand-held ultrasonic equipment. The defects detected by a NDT car and verified by hand-held equipment are recorded on the spot by an inspector in the form of a report. Severe defects with high priority are immediately recommended for unplanned maintenance.

Immediate maintenance is a procedure under which emergency measures are either carried out immediately or traffic restrictions are imposed. Minimal repair under unplanned maintenance is a temporary repair. Defects having minimal repair are then scheduled for full repair at a later time.

Visual inspection is carried out separately by rail inspectors according to an inspection plan (known as planned visual inspection) and recorded in a report, stored in a database. The low-priority defects are recommended for planned maintenance which may be grinding, minimal repair, rail welding or rail section rectification/replacement. The kind of planned maintenance adopted for a particular kind of defect depends on its degree of urgency and severity.

Safety and integrity of rail sections are maintained by rail rectification, replacement and re-railing. Rail rectification is done where minimal repair is required; this may be in the form of small rail section replacement, rail welding, tamping, adjustment of rail sections, mostly on curves, and fastening of fish plates where required. Replacing rail sections greater than those covered in rail rectification is defined as rail replacement. Major overhauling of rails is defined as re-railing.

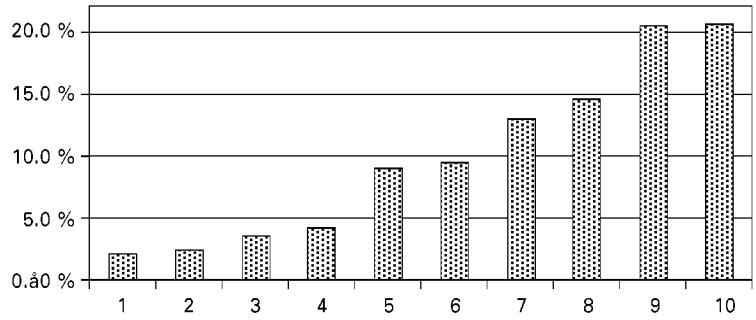
Rail replacement is based on a number of factors. Rails are often replaced based on their wear limit and fatigue status. Finding out the optimum rail replacement interval is a constant maintenance issue on Malmbanan. Weather conditions are also an important factor; it is more economical to carry out rail replacement and re-railing in Summer than in Winter.

22.8 Maintenance, wheel–rail interaction

Railway flange lubrication is used to reduce wear of the rail and wheel at curves. On Malmbanan, curves less than 600 m in radii are routinely lubricated by stationary wayside equipment. The lubricators require proper maintenance procedures, and they generally fail due to nozzle clogging and problems caused by the low outdoor temperature. The static pre-pressure of the grease container causes another problem, the oil is separated from the grease structure and the nozzle clogs, resulting in the need for thorough cleaning. Trackside lubricators are removed during the Winter and re-installed in Spring. The reason for this is climate related – the track snow remover will also remove the lubricators; hence, they are removed before the snow season starts. Maintenance supportability for these activities is achieved by giving responsibility for the lubrication program to the maintenance contractor. The extensive rail lubrication in curves less than 600 m in radii and other preventive maintenance activities are planned and executed by the maintenance contractor.

Grinding activity and railhead re-profiling with its given instance of duration and time interval is achieved using a subcontractor from outside Sweden. The required function is to reduce the influence of surface-initiated rolling contact fatigue (RCF) and reduce railhead wear. Evaluation of the yearly grinding campaign indicates that almost all problems with rail defects are under control and that RCF-initiated problems are managed with grinding and improved rail steel material combined with well-executed rail lubrication.

The switches are ground from the same standpoint as the open track. The presence of corrugations, wear, RCF defects and out-of-geometry of the railhead are parameters which dictate when and how to grind a specific switch. In 2006, 80 switches were ground on Malmbanan. The switches are ground either in solely the main track or in both main track and diverting track. Currently, on Malmbanan almost every switch is ground every year, at least the switches in the main track and most often those in the diverting track as well. At the beginning, the target profile was 'standard' profile BV50 with an inclination of 1:30 instead of the vertical profile which the switches are installed with. However, with this standard profile the result was the same as in open track, i.e. unacceptable wear and the presence of RCF defects. Therefore, an updated target profile was tested. This profile has a gauge corner relief, but not to the same extent as the previous profiles. The reason for not grinding directly to the first version of rail profile was mainly due to reduced production capacity of the switch grinder. To grind the first version of rail profile from a deformed profile would have been very time-consuming. Grinding in an annual stepwise manner was used for this reason. However, in 2007 most of the switches were ground to the new target profile.



22.10 Total maintenance cost in percent for 132 km of Malmbanan.

In some cases, the upper part of the switch blade point is exposed to cracking. The reason is believed to be a local overload when the wheels are moving from the stock rail to the switch blade. Due to wear, both natural and 'artificial' by repeated grinding, the stock rail is lowered over time with respect to the tip of the switch blade. To avoid and to minimize the risk for cracking of the upper part of the switch blade point, the switch blade is lifted a few millimetres. This results in grinding also of the top part of the blade and, after grinding with lifted blade, the transition point is moved towards the inner part of the blade.

The grinding program is based on a predictive maintenance decision. Railhead lubrication is based on a scheduled maintenance program. In Fig. 22.10, the total cost of 46 400 000 SEK (5 150 000 Euro) per annum for Malmbanan section Kiruna–Riksgränsen is divided into ten different activities and presented in percent. Among the five highest maintenance items on the Kiruna-Riksgränsen track section, items related to the wheel–rail interaction costs are switches and crossings (S&C) and rails. Hence, wheel–rail interaction represents 42.5 % of the total infrastructure maintenance cost.

The launching of grinding in 1998 was an important step in increasing the axle loads to 30 tonnes and was also one of the main strategies to prolong the renewal of existing 50 kg rails. After only a short time, measurable cost savings for both infrastructure and traffic were seen. BV are convinced that a strategic rail maintenance program has delayed necessary maintenance costs. Optimising the grinding process and continuing to develop the rail profile are possible steps to increase rail life even further. Neither grinding campaigns nor objective measurements to increase wheel life using new rail profiles seem to negatively affect the total wheel–rail system. Although it may not be possible to reproduce such savings everywhere that RCF damage occurs, some lessons from the test are of general relevance.

Copyrighted Materials
Copyright © 2009 Woodhead Publishing Retrieved from www.knovel.com

23

Managing the wheel–rail interface: Europe Metro experience on the London Underground Victoria Line

D. SCOTT, London Underground, UK

Abstract: This chapter provides an overview of the methods of managing the wheel–rail interface in use on the London Underground Victoria Line. These methods include comprehensive monitoring of the track for signs of break-downs in the wheel–rail interface and proactive management of line lubrication systems. Also within this chapter is an outline of future planned improvements and a discussion of the benefits of active management of the interface on an operational railway.

Key words: wheel–rail interface management, lubrication management, Victoria Line, London Underground, data management.

23.1 Introduction to the Victoria Line and historic wheel–rail interface issues

The London Underground Victoria Line opened in sections between 1968 and 1971. The line is 21 km end to end and is still running with a fleet of Metro-Cammell '67 Stock' trains which was the first London Underground fleet to be fitted with Automatic Train Operation (ATO). The passenger running lines are completely underground in deep tube tunnels which means that the wheel–rail interface does not benefit from environmental lubrication (water, leaf mulch, etc.).

The line has previously suffered numerous problems at the wheel–rail interface. One problem was that during the construction of the line, it had been impossible to source timber sleepers which had been properly seasoned. Under the hot dry conditions of the tunnels, many of these sleepers shrank leading to voiding and tight gauge which led in turn to unsatisfactory vehicle dynamics and high levels of wheel and rail wear. The line used trackside lubricators from its opening (as is standard for the industry); however, a significant breakdown in lubricator maintenance in the early 1990s led to exceptionally high wear rates with rolling stock wheel flanges being worn to the limits in the Standard in two to three days and sections of track becoming worn out within three months. After great effort and expense, the wear rates were brought to more acceptable levels and the fleet was fitted with solid stick wheel flange lubricators to help solve the flange wear problem. The track lubricators were abandoned in favour of the on-vehicle lubrication;

however, high levels of rail crown damage and wheel tread wear highlighted the fact that the track lubricators were also providing a level of protection to the railhead through small amounts of grease migration. The track lubricators were subsequently re-commissioned to provide this service.

The wheel–rail interface (WRI) on the Victoria Line has always been very sensitive to change. Any failure to adequately manage lubrication has resulted in extremely costly corrective works and a great deal of time before wheel–rail equilibrium has been restored. This has previously led to a fairly cyclic nature to the quality of the WRI primarily driven by shifts in management focus.

23.2 The Victoria Line upgrade

Victoria Line is currently undergoing a major upgrade as part of the Public Private Partnership (PPP). The upgrade encompasses the replacement of the existing fleet of trains with a new faster fleet of trains, the replacement of the existing fixed block signalling system with a distance to go radio (DTGR) signalling system and the supply of a new service control centre. The delivery of these assets and a faster service necessitates further works such as track upgrades and the installation of low-loss conductor rails.

Given the scale of the planned upgrade, which essentially delivers a 'new' railway to replace the old (with the exception of the track), the fact that the line has previously suffered from significant breakdowns in the WRI as outlined in Section 23.1 and that previous experiences of introducing new trains without properly managing the wheel–rail interface (specifically on the Central Line) has led to significant WRI problems, the senior management of the upgrade identified the need to set up a project that was suitably scoped to take on and manage the interface such that major problems are avoided. This project was set up to undertake active management of the interface and to trial and develop technologies to help protect against the predicted problems.

23.3 Wheel–rail interface monitoring

One of the first requirements identified by the WRI Project was to develop a comprehensive understanding of the existing WRI performance on the line and to develop a baseline of existing wear rates and general attributes of the line which could be used during future measurements of the line during the upgrade and hence make it possible to detect when aspects of the WRI had moved outside the normal steady state.

London Underground and Metronet Rail had, over the years, developed quite a substantial range of different track measurements. These measurements included those taken by the asset condition monitoring (ACM) train (a normal

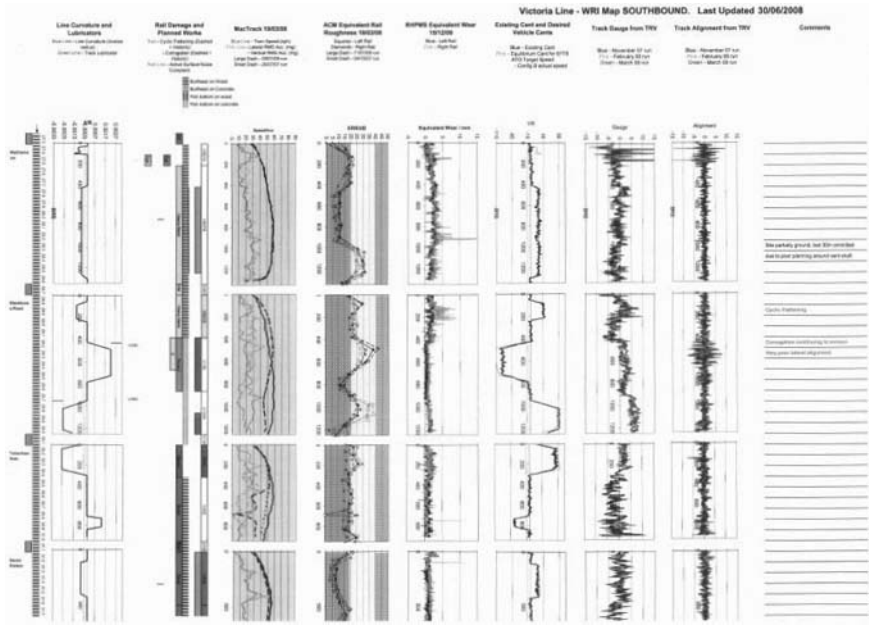
service train taken out of passenger use every three months and fitted with numerous measurement devices) and the track recording vehicle (TRV) (a bespoke vehicle designed to take detailed track geometry measurements for fault indentification purposes). The full set of regular measurements is detailed in Table 23.1.

While these measurements were extremely useful, it was found that in many cases the results of the measurements were considered in isolation and that corrective works tended to be based around addressing the 'top ten' or those which fell outside the Safety Standard limits as defined in London Underground Ltd (LUL) Standards rather than attempting to manage to a higher 'best practice' level. What was required was to bring these measurements together so that the performance of all assets which affect the WRI could be assessed as a whole and so that the inter-relationship between the measurements could be seen.

An overlaid line diagram was produced to fulfil this purpose (an extract of which is shown in Fig. 23.1) and allowed for the assessment of cause and effect. The diagram is updated regularly (every time a new data set becomes available) and includes all pertinent data sets, plus plots of the data sets from the previous two runs of each measurement set. As such, as well as providing a snapshot of the wheel–rail interface at a single moment in time, it also gives information on the rate of change of the interface and

Table 23.1 Regular track measurements undertaken on the line

Data source	Information collected	Frequency
Track recording vehicle	Cant Gauge Vertical rail alignment Lateral rail alignment 2 m twist 10 m twist Railhead and sidewear	Every 8 weeks
Asset condition monitoring train	Vehicle lateral and vertical accelerations Equivalent rail roughness In-car and under-car noise	Every 3 months
	Video inspection – wheel rail interface	Every 6 months
Resident noise complaints	Noise levels, location on track attributed to noise	*Ad hoc*
Track survey	Location and radius of all curves	One-off
Lubricator asset list	Location of all track lubricators	Revised with changes
Scoping surveys	Current track form	Revised with changes
Rail grinding scope and reports	Location of rail grinding works	Revised with changes



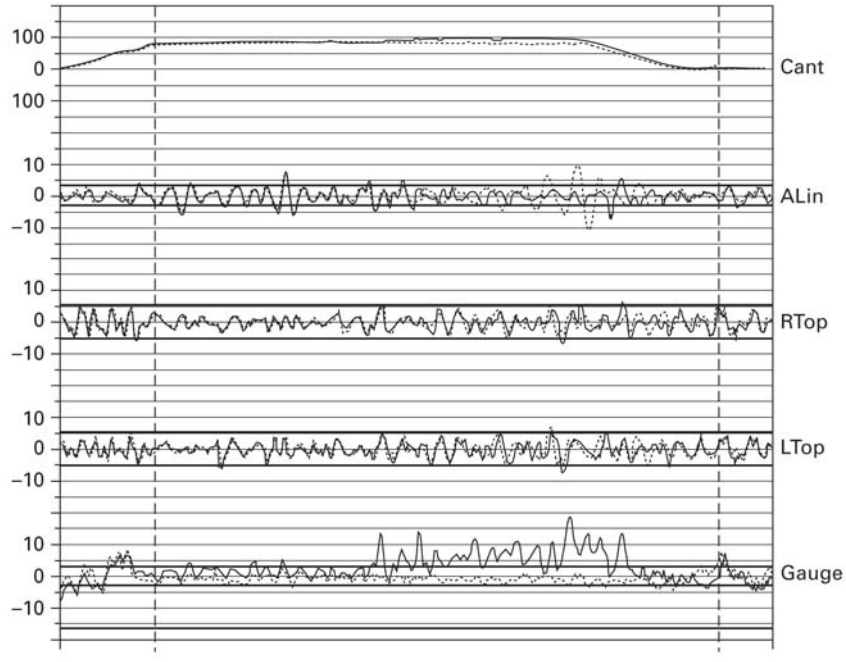
23.1 Victoria Line WRI monitoring diagram.

highlights any areas which are deviating from the steady-state performance for corrective management. The system in its current form is fairly simplistic, but it is highly effective in enabling root-cause analysis. Each data set is aligned geographically along a linear representation of the line so that the user can see the measurements from each data set that pertain to a single location on the track. In addition, the user can see information such as the location and severity of curves, the location of track-mounted lubricators, planned or completed track works, planned or completed rail grinding works and areas which are sensitive due to noise complaints from local residents. As such, the user has at a glance most of the information he/she needs to identify and diagnose a problem in the WRI.

The development and use of this method of 'holistically' viewing track and WRI data has highlighted numerous potential benefits both in terms of maintenance planning and the scoping of upgrade works. Unfortunately, the development of this method of wheel-rail interface monitoring occurred after the scopes of work for the track upgrade had been completed and passed to the contractor, thus slightly limiting the potential benefit that might have been realised had this method been developed earlier. However, in terms of maintenance planning, this system is now the primary driver behind the development of scopes of work for rail grinding, the identification of negative trends in track performance and the development of suitable scopes

for corrective works (primarily to ensure that the corrective work removes the root cause of the problem wherever possible).

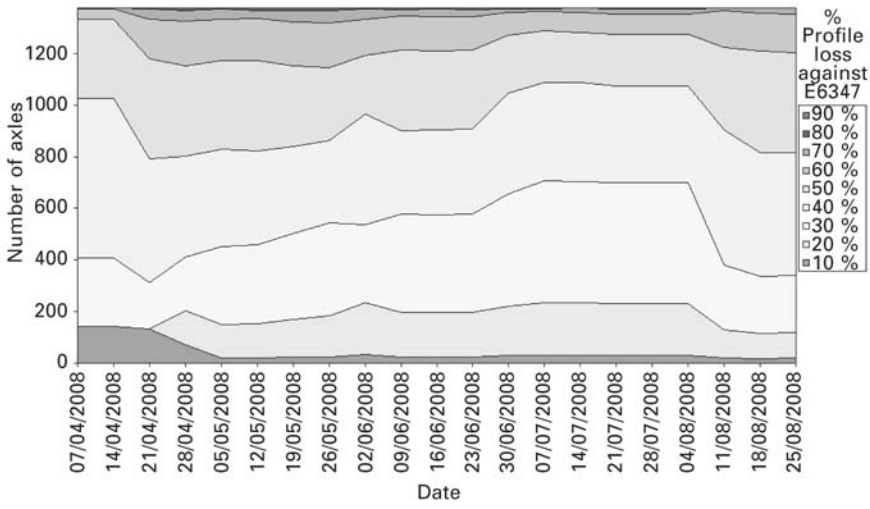
The development of this work also led to an opportunity to improve the process for reviewing completed track works against the original scope of works and assessing the benefit delivered to the business. As such, a system was developed which overlays all of the measurables from the TRV and the ACM from before the work was undertaken and immediately after work completion. Also overlaid are the appropriate quality limits required the scope of work or within the relevant Standards. An example of the overlaid TRV output is shown in Fig. 23.2. Through this analysis, it is possible to identify any non-conformances with either the Standards or the requirements of the scope of works and ensure that track of appropriate quality to ensure good WRI performance is delivered. It also provides a comprehensive baseline of the exact quality of the track at the point of work completion on which to base future assessments of track performance. By including assessments of measurables such as longitudinal roughness, vehicle lateral and vertical accelerations and in-car noise, it is possible to also assess the effects of the track works on the passenger perception of track quality. Thus it is possible to achieve a high level of confidence that the works undertaken have been delivered to within required quality tolerances and that a suitable improvement in passenger ambience and WRI has been achieved.



23.2 TRV trace from scope and benefit review report.

The fleet depot is equipped with an under-floor wheel lathe and a wheel profile monitoring system which scans each wheel on the train as it goes into the train wash. As such, the profiles of every wheel on the fleet can be measured two to three times a week. A substantial amount of work was undertaken by the Victoria Line fleet to tie the results from this measuring system into the wheel re-profiling regime to ensure that wheels are turned at the optimum point to ensure compliance with the wheel profile Standard. An offshoot of this work is that it is possible to gauge the range of wheel profiles on the fleet by deriving a value for percentage wear against limits in the Standard which is an amalgamation of the flange width and flange height measurement taken from the measurement system (derived value is a percentage where 100 % signifies that one or more of the limits in the wheelset Standard have been reached). By plotting the population of wheels in each 10 % wear band, the Victoria Line fleet can produce a curve signifying the distribution of profiles across the fleet (and hence by ensuring that none of the wheels are flagged > 100 % assure wheel profile compliance). In terms of WRI management the absolute distribution of wheel profiles is often not important but the rate of change of distribution is. Therefore, the plot shown in Fig. 23.3 was developed to show changes in the distribution curve over time. This plot is used to identify if any negative changes are occurring (i.e. rapid wheel wear) but also to identify whether any other changes which might affect the interface are occurring (i.e. the introduction of significant numbers of new wheel profiles).

By tying this plot into the overall WRI monitoring process and into the lubrication management process, it is possible to monitor both sides of



23.3 Variance in wheel profile distribution over time.

the interface and ensure that any changes that occur are suitably managed through adjustment of the lubrication system or through other changes to the system. The bringing together of these strands of information has led to huge improvements in the ability to identify potential problems early and make informed decisions as to the best course of remedial action.

23.4 Lubrication management

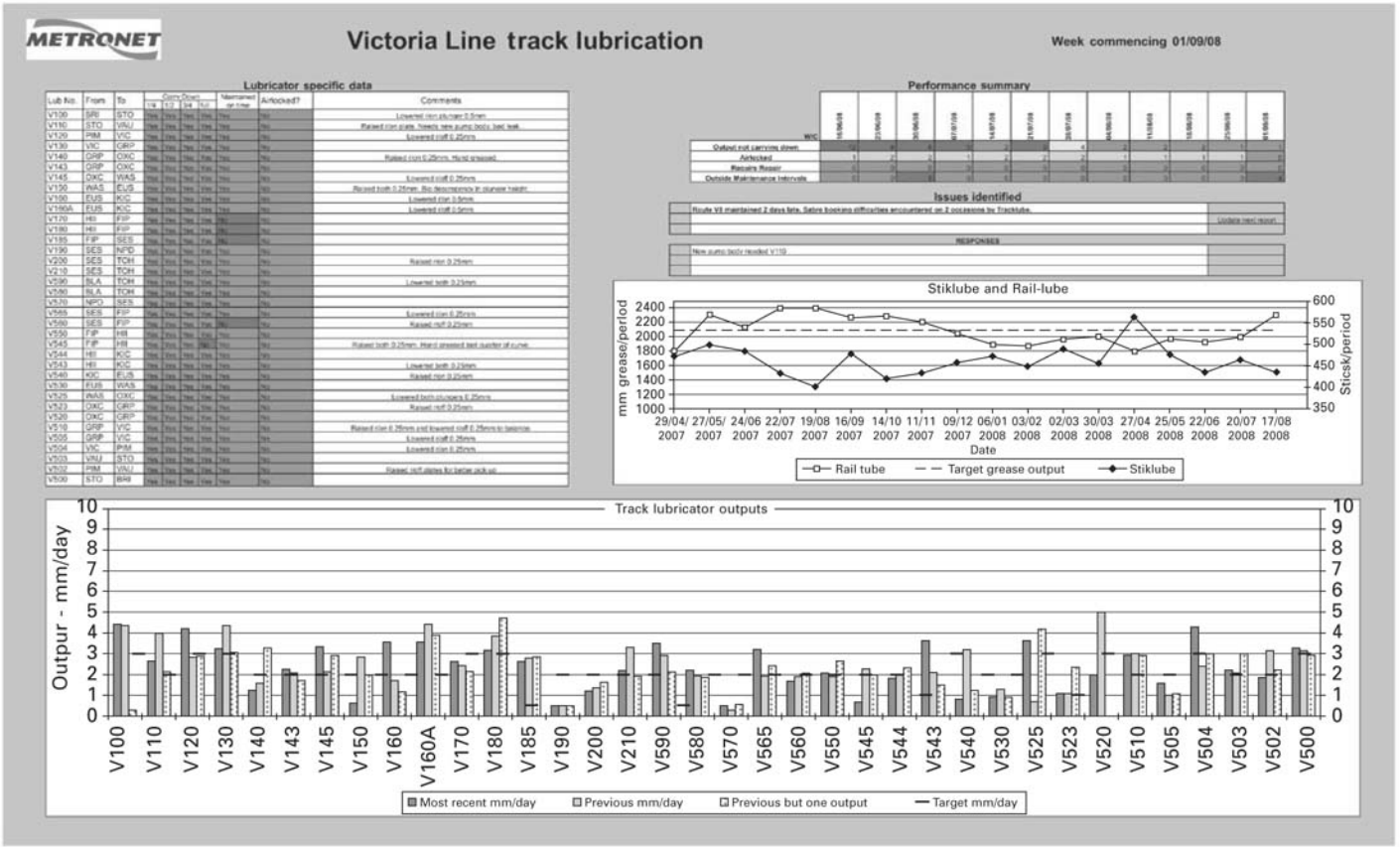
The issue of friction management was identified early on as being key to the management of the wheel–rail interface. The conclusion reached very rapidly was that while optimised friction levels do not necessarily mean that the WRI is as good as it can be, poor friction levels will mean that any breakdown in the WRI will lead to very rapid and expensive damage to assets. As such, a decision was made that the WRI project would take over responsibility for managing track lubricator upgrade and maintenance and would develop suitable management techniques for the long-term maintenance of lubrication systems. The project also introduced a regime for monitoring the consumption of the vehicle-mounted lubricators.

The first step in this was to introduce a dedicated maintenance team, because one of the main issues identified with the previous management resource was that it was shared across all track maintenance with the effect that lubrication maintenance was seen as a 'low priority' when compared with incidents such as rail breaks, etc. As such, an external contractor was brought in to provide dedicated and highly trained maintenance staff to undertake the maintenance of all track lubrication assets on the Victoria Line. The maintenance teams collected data during each 14 day maintenance cycle on grease consumption, plunger heights, grease distribution unit (GDU) heights, work carried out and presence of grease on the gauge corner throughout the applied curve. These data were converted to a measure of average daily output from each lubricator and quality of grease carry-down and were used by the maintenance team to assess what level of output adjustment was required to fix the output at an average of approximately 123 cc per day of traffic. This work was very successful at stabilising the output of all of the lubricators on the line, but it soon led to the identification of some sites where lubrication was insufficient and rapid rail damage was occurring and other areas where the lubricator position in the curve was not compatible with the fleet curving characteristics, hence leading to poor quality of grease pickup at the GDU plates. This latter problem led to the buildup of grease on the GDU plates which would occasionally be picked up by a wheel and thrown onto the bogie underframe. As a result, the positions of 11 of the 34 track lubricators were adjusted to improve the efficiency of grease dispersion. A further two lubricators were added to improve friction levels in areas identified as being deficient.

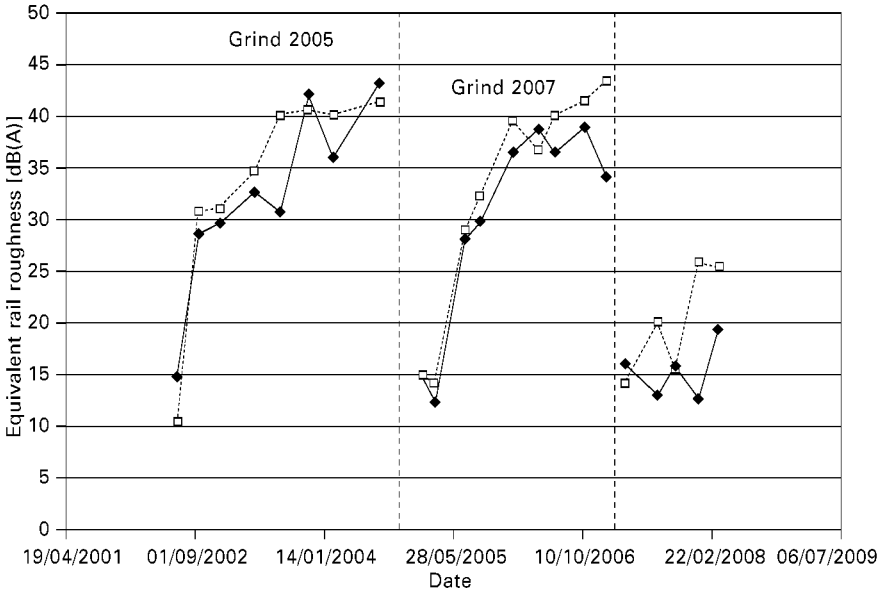
Following from this, the monitoring and management regime was developed further to prevent problems from emerging in the future. Site-specific targets were developed for each lubricator to cope with known friction problems and the length of track that each lubricator is expected to protect. The volume of metal collected with magnetic wands in each inter-station section to prevent block-joint failures is assessed by weighing the material collected and averaging over the length of track to get a rough assessment of wear rates. Direct friction measurements are taken using a tribometer trolley. The consumption of the vehicle-mounted solid lubricant is measured and plotted against the overall consumption of track grease to monitor the interactions between the two systems. A snapshot of the front page of the lubrication monitoring system is shown in Fig. 23.4. The monitoring of both the track and the wheel profiles outlined in Section 23.3 is also used to inform the decisions that are made with regard to lubrication management. For instance, if a rapid change in wheel profiles is detected as a result of the introduction of new wheels, the position and heights of GDUs throughout the line would be reviewed to ensure optimum grease pickup and distribution is maintained by a suitable proportion of the fleet without damage to the GDUs occurring.

The result of developing the management of lubrication systems to this extent is that ambient noise has reduced. Corrugation growth rates have also been found to have been reduced, most markedly so on the section of track between Victoria and Pimlico Stations on the Southbound road which has historically been one of the worst sections for corrugation on the line. Figure 23.5 shows the peak equivalent rail roughness measurements in the inter-station section. Each time the section has been ground to remove corrugation within the history of the measurement, the corrugation has been seen to have developed rapidly. The improvements in lubrication described above took place shortly after the 2007 grind for this section and the change in rate of increase in corrugation levels can be seen quite clearly. No other track works have taken place in this area; therefore, improved friction levels are the only explanation for this change in behaviour.

One lesson that has been learned though this work is that track lubrication, where an attempt is made to deliver grease to the rail gauge corner as close as possible to the contact patch without putting the lubricator in danger of being damaged, is exceptionally sensitive to changes in vehicle track dynamics. Since these changes can arise from very small changes in wheel or rail profile, vehicle suspension set-up, friction levels or track geometry, a very high level of vigilance is required to keep the lubrication systems within nominal bounds. As such, it would be almost impossible to adequately protect a new fleet of vehicles, inter-running with the existing fleet for a period of time, using a lubrication system which is predominantly track lubricator based while controlling friction levels to within a good range (i.e. not too high or low on the rail crown and not too high on the gauge corner).



23.4 Lubrication monitoring system front page.



23.5 Peak equivalent rail roughness over time.

The decision has therefore been made to develop a solid stick application of lubricant for the wheel flange and friction modifier for the railhead for the new fleet of vehicles. In consequence, it is hoped that the new trains can be set up to be largely self-protecting, thus removing a proportion of the predicted lubrication system sensitivity.

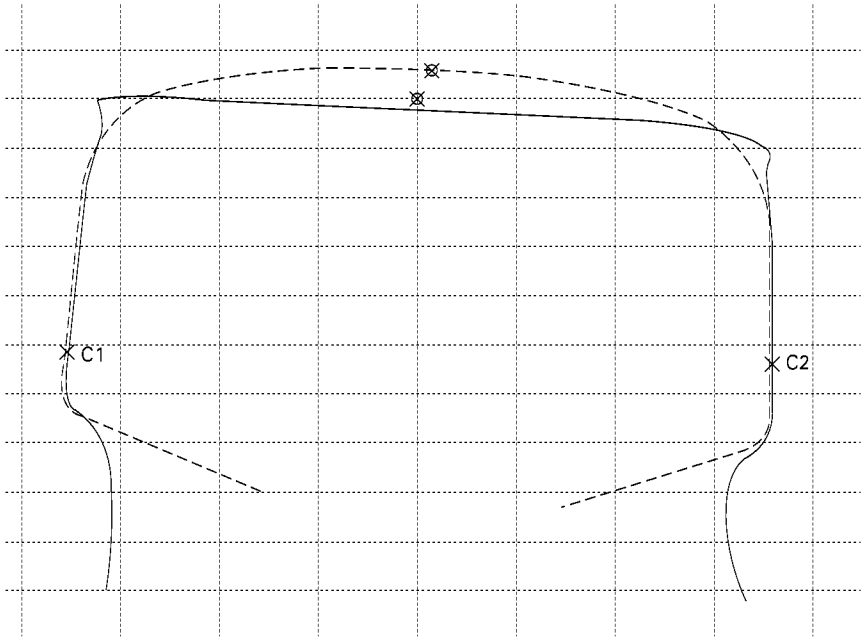
This work has shown quite clearly the importance of effective lubrication in keeping friction levels within acceptable bounds and thereby keeping wear and other damage occurring at the wheel–rail interface under control. A well-lubricated railway can often tolerate other aspects of the wheel–rail interface being sub-optimum without experiencing substantial asset damage; however, any failure in the lubrication system will result in very rapid asset damage and business cost. A phrase that has been coined for railway systems operating in this way is the 'unstable equilibrium', while everything works the system operates acceptably, but any upset to the equilibrium generates high levels of damage and substantial time and cost is required to get back to the equilibrium level.

23.5 Identified wheel–rail interface problems

As a result of this work, a number of specific WRI-related problems have been detected and the root causes identified. One of these problems is that of 'cyclic flattening' which is a cyclical loss of the crown profile of the low rail approximately every 13 m; the low rail damage often coincides with

heavy gauge corner wear of the high rail. This problem has been present in many locations on the line throughout most of its life and is one of the major drivers behind re-railing for profile loss. The loss of low rail profile occurs rapidly and to such an extent that correction, using the eight stone rail grinding machine that is the biggest currently available unit that will fit within the tube environment, is impracticable.

Figure 23.6 shows an example of the low rail profile generated by this problem. If a worn wheel profile is overlaid on this transverse profile it is seen to be almost perfectly conformant. The shape generated and the relatively low loss of rail volume leads to the conclusion that the damage is occurring primarily due to high levels of plastic flow in the contact patch. The locations of profile loss and the overlaying of wheel profiles shows that the flattening of the low rail occurs when the wheelset is in heavy gauge corner/flange contact with the high rail. The damage, profile conformance and video footage of wheel–rail interaction all show that the damage is caused by wheelset hunting behaviour through the curves which explains the 13 m cycle of damage which derives largely through the klingel wavelength[1] in addition to vehicle speed and suspension parameters. The output from the rail head profile measuring system on the TRV shows cyclic flattening quite clearly as a large-amplitude variation in railhead wear, and when there data are viewed using the WRI monitoring system in conjunction with cant



23.6 Transverse profile of 'cyclically flattened' low rail.

and vehicle speed data, it is possible to see that the phenomenon occurs predominantly in locations where the track is over-canted for the normal vehicle running speeds.

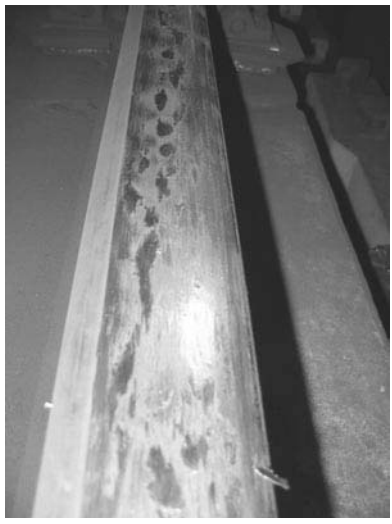
As such, the mechanism for damage occurs thus: the curve is installed with too much cant (either due to a mis-calculation of vehicle speeds through the site or because the curve starts in or adjacent to a station platform leading to the rolling stock being at low speed for a proportion of the curve). The excess cant leads to a resultant force which is trying the pull the bogie down towards the low rail, preventing the vehicle from achieving a satisfactory curving regime and exciting cyclic 'hunting' behaviour. The wheelsets oscillate laterally between the high and low rail with a wavelength of approximately 13 m for reasons discussed above. A combination of the cyclic axle wind-up generated and the wheelset yaw generated at the oscillation peaks leads to very high lateral and longitudinal forces being generated in the contact patch between the wheel and the low rail when the wheelset is at its peak (i.e. in hard flange contact with the high rail) and high contact pressures due to the sway of the secondary sprung mass. The high forces are sufficient to generate repeated plastic flow which eventually leads to the profiles seen. Once formed, the rail profiles generated start to drive the path that the wheelset takes through the curve such that the oscillation is reinforced, even if the cant excess is removed (by increasing vehicle speed for instance).

This finding has highlighted the need to ensure that track designs and vehicle speed designs are fully compatible. Future consideration ought to be given to changing the cant design methodology for ATO railways with enforced station stops such that the cant is 'tuned' to the ATO speed profile to develop a reasonably constant cant deficiency through all curves. This would have the advantage of making it easier for the vehicle to move into a steady-state curving regime (i.e. all transient dynamics are damped out) and hence this will be achieved earlier in the curve. A second benefit would be in terms of passenger comfort. In terms of the upgrade of the Victoria Line, such designs were impossible to deliver as most of the line was not being re-designed. This problem is expected to be largely removed by the higher speeds of the new fleet of rolling stock and the improved vehicle dynamics. Remaining areas where over-canting will persist will be managed by improving the friction levels.

Another issue which has been prevalent on the line is that of rail contamination. The fact that the entire operational line is underground in deep tube tunnels, and the fact that the drainage on the line is generally very good and groundwater seepage is very low, means that the line is very hot and very dry. As mentioned in Section 23.1, this has previously generated the problem of the shrinkage of poorly seasoned wood sleepers leading to tight gauge. A secondary consequence of such warm dry conditions is that there is less water present in the third-body layer of the wheel–rail contact

patch compared to most other railways. This, coupled with the lack of organic contaminants (leaf mulch, etc.) that occur on surface rails, means that an unlubricated wheel–rail contact will experience extremely high friction levels. Additionally, since the third-body layer may be quite thin, the interface is exceptionally sensitive to contamination from abrasive particulate matter. This issue was demonstrated quite clearly following the contamination of rail in a platform with concrete dust (generated while mixing concrete in a bucket nearby using a mixer attached to an electric drill). The dust was left on the rail following the works and was ground into the rail by the passage of trains following the re-opening of the line. Within the same day, significant longitudinal rail roughness had developed to the point where drivers and station staff started to complain about the levels of noise. Figure 23.7 shows the rail shortly after the contamination incident. The most noticeable features of this damage were the presence of particulate concrete matter embedded into the railhead and the 'slithers' of metallic material left present along the railhead. The rail was ground using an eight stone grinding machine after this incident to remove the contamination and the roughness. Roughness measurements taken after grinding show that good levels of roughness were achieved; however, within two months roughness and noise were noted to be returning over a much shorter length of the rail. Additionally, the 'slithers' of metallic material were noted to have returned but not the presence of concrete matter.

At the time of writing, the exact mechanism of damage at work is still under investigation, but the current theory can be described as follows. The initial contamination with concrete matter allowed a very aggressive contact



23.7 Rail damage caused by concrete contamination.

regime to be set up. Within this regime, some of the concrete acted as a stress raiser on the railhead and allowed small cracks to be formed. Under the high braking and traction loads in the platform area, these cracks propagated until they turned upwards breaking out a small piece of the railhead. These broken out pieces of railhead were subsequently caught up in the wheel–rail contact and, under the extremely high pressures generated, extruded out along the rail head to form the 'slithers' seen. The grinding carried out removed these slithers and the embedded concrete, but did not go deep enough to remove all of the cracking throughout the site. As such, in the area where the most braking and tractive load is seen, the cracks continued to propagate until material started to be removed from the railhead and extruded once more. Further work, including testing of the 'slithers' of material, will confirm whether this hypothesis is correct.

Railhead shelling associated with contamination of the railhead with concrete has also been noted elsewhere on the line, although none of the other incidents caused such widespread damage (most likely due to the lack of substantial braking and tractive loading). As a result, the working methods for contractors working adjacent to rails with materials such as concrete have been improved to ensure that either the rail is suitably protected while the work is being carried out or suitable cleaning is carried out before the start of passenger traffic. These improvements have thus far been successful in preventing any further major damage arising from rail contamination. Given that within tube tunnels concrete slab track is present throughout, vigilance must always be maintained to ensure that concrete is not allowed to contaminate the rail.

23.6 Ongoing work and future plans

Despite best efforts, a risk (however small) will always persist that the monitoring and management outlined in the previous sections fails to prevent a rapid degradation in the wheel–rail interface for whatever reason. Previous experience has shown that, under such circumstances, substantial energy is put into maintaining the profiles of both wheels and rails whilst maintaining a passenger service. Under such a regime, it would be impossible to carry out the required testing and commissioning of new rolling stock and so the upgrade of the line would be delayed until equilibrium was re-established. Obviously, this scenario is one which must be avoided. Therefore, it is planned to trial and develop the option for using harder rail grades on the underground. Currently, most re-railing on the line is carried out using 56E1 section rail with a hardness of 260 BHN (also known as grade A flat bottom rail). It is currently possible to source micro-head hardened (MHH) rail in 54E1 section (which only differs from the 56E1 section a small amount in the rail web) at a hardness of 400 BHN. This, once developed, has numerous

potential benefits in the London Underground environment. The increased wear resistance offered by the harder steel offers improvements in both corrugation and head and sidewear resistance. The higher shakedown limit[2] will reduce occurrences of heavy plastic flow. As such, under most circumstances, the 400 BHN rail would be expected to maintain its transverse and longitudinal profile for significantly longer than the 260 BHN rail (if not remove the damage mechanism completely).

Therefore, it is planned to install a trial 90 m site of 400 BHN rail in an effort to understand its performance under London Underground conditions and assess its likely eventual failure mode. If successful, the use of 400 BHN rail will be available as an off-the-shelf alternative to the regular rail steels in use. As such, should a break-down in the interface occur, it will be possible to replace damaged rail with 400 BHN rail which will hold its shape for longer under degraded conditions and hence enable wheel–rail equilibrium to be achieved far sooner than would be possible otherwise. The performance of this rail may also make it the rail steel of choice for future re-railing of curves throughout London Underground.

Vehicle design has improved greatly over the years to the point that vehicles designed now are generally less damaging than older vehicles. One important aspect of preparing for future changes in the interface is a good understanding of the future dynamic performance of the new rolling stock. To this end, substantial dynamic modelling has been carried out using computer models developed by the rolling stock manufacturer. Some of this modelling has included a like-for-like assessment of the new rolling stock against the existing over track geometry which has been derived from actual TRV measurements and using actual wheel and rail profile measurements. This modelling has shown that, under the majority of circumstances, the new rolling stock is less damaging than the existing. Further work is planned to use the computer models to test the dynamic stability of the new rolling stock under a range of different wheel and rail profiles. It is hoped that through this work it will be possible to map out the stability envelope of the vehicle due to equivalent conicity generated by the wheel and rail profiles. Once mapped, it will be possible to use this information to inform wheel and rail profile maintenance such that, as far as possible, the steering force generated at the contact patch is not allowed to generate undesired vehicle dynamics and hence asset damage.

Finally, it is hoped to be able to develop the wheel profile monitoring system in use further such that is not just assessing the wheel profile at the discrete positions identified by the Standard. By developing the output from the system further, it is hoped to be able to develop a set of metrics that will allow future WRI managers to understand exactly what each profile in the fleet distribution means in terms of wheel–rail damage and lubricator compatibility.

23.7 Conclusions

To date, substantial improvements have been made to the Victoria Line through improving the management of the wheel–rail interface. The biggest challenges to maintaining good wheel–rail conditions on the line still lie ahead; these will be generated as significant numbers of the new rolling stock start to run on the line and when the new fleet eventually moves to running at its planned higher speed. It is hoped that the level of management that has been put in place to manage the interface during this time will be sufficient to have pre-empted and solved most of the likely problems before they are allowed to manifest and to be able to detect any unexpected problems at the interface and resolve them long before they become an expensive (and potentially service affecting) problem. The approach that has been adopted to manage this problem has been to use the raft of existing measurements, bringing them together and actively using them to drive the work that is carried out on the line. In almost all cases, no new measurements were required over and above existing. The developments in wheel–rail interface management achieved have been delivered firstly by taking ownership of the issue and then through vigorous application of engineering science and good process. As new technologies and measurements become available, these can be incorporated into the process of wheel-rail interface management, but these technologies should only make the existing management of the wheel–rail interface easier and less resource-intensive. Experience shows that there is no single solution that will solve all wheel–rail interface issues and allow the railway to walk away and not need to manage the interface; invariably, the direct solution to one problem simply allows other hidden problems to surface or introduces new problems. The asset cost of a degradation in the wheel–rail interface is generally so high that failure to actively own and manage the interface leads to substantial business cost for any railway. Experience on the Victoria Line shows that management of the interface need not be expensive or be reliant on state-of-the-art technology; it merely requires the impetus to think of the railway as a complete system and to manage all aspects of the wheel–rail contact accordingly.

23.8 References

1. Wickens A H (2003), *Fundamentals of Rail Vehicle Dynamics*, Swets & Zeitlinger, Lisse, the Netherlands.
2. Bower A F and Johnson K L (1991), Plastic flow and shakedown of the rail surface in repeated wheel–rail contact *Wear*, **144**, 1–18.

Copyrighted Materials
Copyright © 2009 Woodhead Publishing Retrieved from www.knovel.com

24

Managing the wheel–rail interface: the Canadian experience

E. MAGEL and P. SROBA, National Research Council, Canada

Abstract: Canada is home to a large number of railroads, including the heaviest freight systems and the lightest transit applications and, as such, has encountered a large number and variety of wheel–rail interface problems. Canadian experience with, and solutions for, wheel shelling, rail fatigue, wheel–rail wear, bogie hunting, friction management, rail grinding and wheel–rail profiles are discussed through several case studies.

Key words: Wheel–rail interface, wheel shelling, rail grinding, friction management, rail corrugation.

24.1 Introduction

Railroading in Canada has a long and rich history that extends back further than the nation itself. Four Eastern provinces in Canada formed the Confederation of Canada in 1867, with New Brunswick and Nova Scotia being promised rail links to Ontario and Quebec. Manitoba joined in 1870 and British Columbia, on the West Coast, was persuaded to join in 1871 under the condition that within 10 years they would be linked to the rest of the country by a railway. In 1885 Canada had a railroad that crossed from the Atlantic to the Pacific, a distance of nearly 5500 km.

Canada is the world's second largest country but, with only 33 million inhabitants, it ranks as one of the most sparsely inhabited nations on the planet. However, with its rich supply of lumber, iron ore, coal and grain resources, Canada has always been a relatively prosperous nation, thanks largely to the railroads. A country of great geographical breadth and topographical variety, Canada has low mountains in the east, a vast central prairie region and the towering Rocky Mountains in the west that have been conquered, coast to coast, by Canada's two national railroads.

The Canadian Pacific Railway (CPR) was the first to cross the nation, having been formed specifically with the goal of uniting the nation. The Canadian National Railroad (CNR) was formed in 1918 by the federal government out of the remnants of several bankrupt railways, including the Canadian Northern Railway and the Grand Trunk Pacific Railway. CNR and CPR now each operate trains over 24 000 km (15 000 miles) of track from coast to coast and into the USA.

The British Columbia Railway, initially owned by the provincial government but sold in 2004 to CNR, has perhaps the most tortuous routing of all Canada's railroads. In the east, the Cartier Railway Company and the Quebec North Shore and Labrador run from the ports on the St Lawrence Seaway to the iron ore mines in the north. Additionally, Canada has more than 40 short line railways operating on about 16 000 km (10 000 miles) of track.

VIA Rail, Canada's national intercity passenger railway, possesses no mainline rail and operates about 460 trains weekly on 14 000 km (8700 miles) of primarily CN-owned track between Halifax and Vancouver. Because of Canada's sparse population, regular intercity service exists only in the Ottawa–Montreal–Toronto corridor. Other heavy commuter rail systems exist in Vancouver (West Coast Express) and Toronto (GO Train). There are no high-speed rail systems in Canada. The maximum speeds are run by VIA – about 160 kph (100 mph). Steel wheel on steel rail transit systems exist in Vancouver, Calgary, Edmonton, Ottawa and Toronto. We will encounter several of these railroads in our subsequent discussion about managing the wheel–rail interface in Canada.

24.2 Canadian Pacific Railway – wheel–rail management experience

Strategies to control wheel and rail wear on the Canadian Pacific Railway have primarily been concentrated on the curvy and heavier tonnage western coal route. These strategies have evolved through practices that balanced wear, risk management (e.g. fatigue), wheel and rail profile matching and the changing demands of the operating environment.

Some of the toughest railroading in the world is experienced on the Canadian Pacific Railway's western coal route between the mines in southern British Columbia and the port of Vancouver, where unit trains with payloads of 13 150 metric tonnes (14 500 tons), powered by three 3280 kW (4400 Horsepower) AC traction locomotives, negotiate the steep grades and sharp curves over a 1200 km (750 mile) route. Near Golden, British Columbia, the coal traffic joins up with the primary east–west mainline, which carries approximately 91 million gross tonnes (mgt) (equivalent to 100 million gross tons (MGT)) per year.

The route is predominantly single track with 46 % of the route traversing curves less than 3500 m radius (sharper than ½°) and 129 km (80 miles) of curves less than 300 m radius (sharper than 6°). Minimum curve radius is 170 m (11°). The controlling grade westbound is 1.1 % and the route passes through several tunnels, including the Mount MacDonald Tunnel, at 14.6 km (9.1 miles), the longest in the Western hemisphere. Temperature extremes in the Thompson River valley range from +43 °C (110 °F) to –34 °C (–30 °F). The rail in curves of radius 220 m and less (8° and sharper)

is predominantly 68 kg/m (136 lb/yard) head-hardened, high-strength rail (see Table 24.1) with a Brinell hardness number (BHN) of 350–390. Rail in curves is supported by 2.7 m (9 ft) long hardwood ties on 41 cm (16 in) rolled eccentric plates.

As tonnages increased on the line throughout the 1970's, rail and sleeper damage accelerated rapidly. Rail in curves with a radius less than 436 m (sharper than 4°, see conversion Table 24.2) was surviving an average of only 192 mgt (212 MGT). Severe head-checking, corrugations and rapid high rail wear forced the replacement of rail in three years or less. Gauge corner shells initiated transverse defects, with rail failures occurring under trains. Timber sleepers suffered plate cutting and wide gauge conditions were present in sharp curves.

At that time CPR used two-wear wheels of Class C metallurgy and hardness (see Table 24.3). (Having a rim with thickness of about 62 mm, sufficient for re-truing one or more times. This compares with a single-wear wheel, where the wheel is allowed to wear out and is then scrapped.)

Table 24.1 North American rail classes

Rail	Hardness specification	Hardness: typical
Standard	300 HB minimum	320 HB
Intermediate		320–340 HB
High strength (premium)	341–388 HB	360+HB
HE premium		390–405 HB

Table 24.2 AAR common wheel grades

Curvature (α)	Radius (m)	Radius (m)	Curvature α
8°	218	200	8° 44′
7°	249	300	5° 49′
5°	349	400	4° 22′
4°	436	500	3° 29′
3°	582	600	2° 54′
2°	873	800	2° 11′

Table 24.3 Comparison between degree curvature and radius in metres

	Class B	Class C
Carbon	0.57–0.67 wt%	0.67–0.77 wt%
Manganese	0.60–0.85 wt%	0.60–0.85 wt%
Phosphorus	< 0.05 wt%	< 0.05 wt%
Sulfur	< 0.05 wt%	< 0.05 wt%
Silicon	> 0.15 wt%	> 0.15 wt%
Hardness	277–341 HB	321–363 HB

These wheels were found to last 430 000 km (270 000 miles), with the prime cause (63%) of changeouts being shelled treads. Crack initiation was caused by a combination of lateral tangential forces due to frequent curve negotiation in conjunction with heat input from braking. Once the surface cracks had initiated, the moisture present from blowing snow would provide the hydraulic action that caused such cracks to grow.[1]

24.2.1 Management methods to control rail and wheel wear

CPR has been continuously improving their rail and wheel management strategies since the 1980s to improve rail and wheel life and lower their operating costs. Strong improvements have resulted from the introduction of harder and cleaner rail steels, preventive rail grinding, friction management, the introduction of improved materials to strengthen the track, improved wheel profiles and the introduction of frame-braced bogies on their coal fleet.

Harder and cleaner rail steels

Since the 1980s CPR have been upgrading their rail in sharper curves with harder and cleaner steels. This was tied in with the introduction of a preventive grinding program that allowed them to increase their railhead wear limits in curves.

In 1984, CPR introduced chromium-alloyed head-hardened rail with a section changed from 66 kg/m (132 lb/yard) to 68 kg/m (136 lb/yard). In 1987, they upgraded to deep head-hardened heat-treated alloy rail with improved yield strength and greater depth of hardening in the railhead. Also at this time, a cleaner intermediate grade rail with 325–340 BHN hardness was introduced to mild curves and tangent track. A further refinement was to change the rail profile from the standard American Railway Engineering Association (AREA) 355 mm (14 in) crown radius to a 203 mm (8 in) crown radius. In 2001, CPR introduced low-alloy hyper-eutectoid rail steels to sharp curves. These steels have demonstrated reduced wear rates and improved fatigue resistance. Average rail life on the highly curved coal route has improved from 295 mgt (325 MGT) to 617 mgt (680 MGT).

Preventive rail grinding

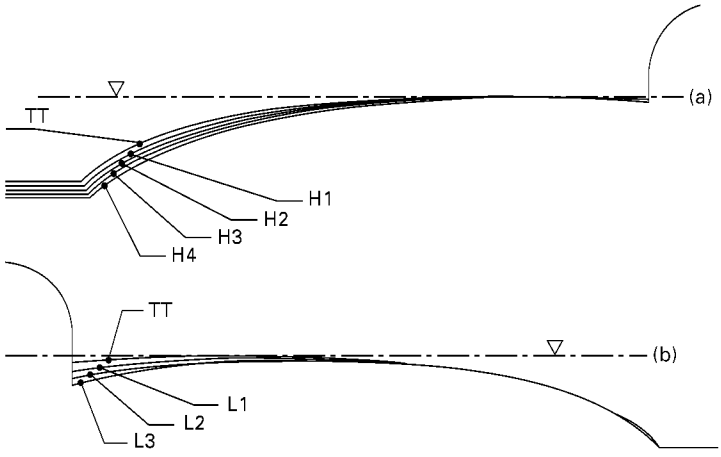
CPR, in the early 1980s, trialed a rail grinding practice developed in the Pilbara region of Western Australia, termed asymmetric grinding. This practice produced rail shapes that relieved the high rail gauge corner and the low rail field side to reduce the incidence of high rail shelling and false flange damage on the low rail. In the late 1980s CPR tested a new grinding strategy called

preventive grinding which introduced several improved rail profiles (Fig. 24.1) to curves and tangent track. The rail profile was ground to a 200 mm (8 inch) radius and grinding cycles for curves were tightened to 16 mgt (18 MGT). Along with the introduction of harder steels this strategy eliminated rail corrugations as a reason for rail replacement. With frequent management of the rail profile through preventive grinding, harder and cleaner rail steels, frequent rail inspection for ultrasonic defects and rail wear, CPR was able to increase their railhead wear limits from 25% to 35–40% head loss.

CPR introduced a next generation of optimised rail profiles for curves and tangent track in 2001 that were more appropriate for the improved rail steels being installed by CPR.[2] In comparison with earlier steels, modern rails employed harder and more fatigue-resistant metallurgies. The amount of plastic flow between grinding cycles was small so the new profiles being ground to the rail called for a decrease in the amount of metal removed from the high rail gauge corner and the field side of all the profiles. The new high rail shapes were designed to be more conformal with the distribution of passing wheels. The improved shapes contributed to an increase in the tonnage interval between preventive grinding cycles on sharp curves from 16 to 23 mgt (18 to 25 MGT).

Friction management

In 1999, CPR surveyed its existing friction conditions using a hi-rail mounted tribometer[3] and discovered that there were many opportunities to reduce high rail wear. A comprehensive field test program conducted in



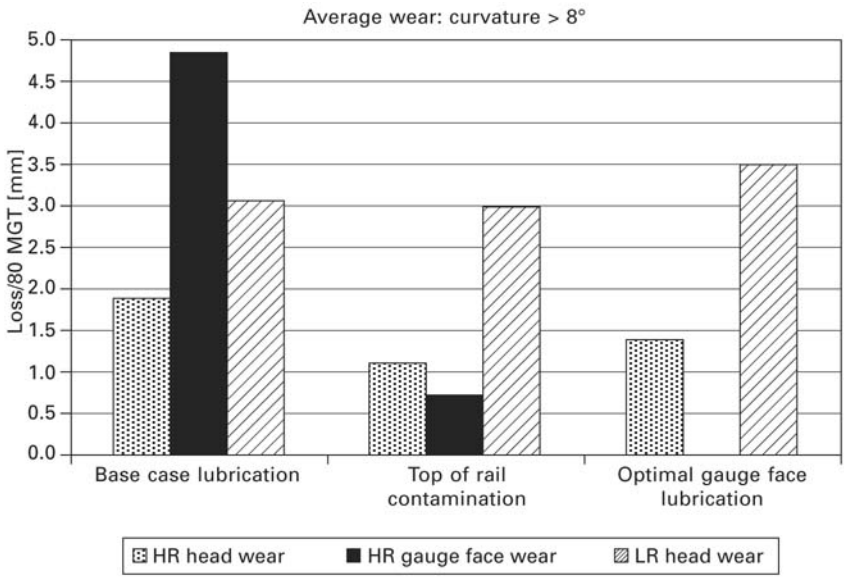
24.1 In 1991 CPR introduced a family of eight rail profiles to improve the wheel/rail interaction: (a) one tangent (TT) and four high rail profiles; (b) tangent (TT) and three low rail profiles.

their toughest western Canada subdivision defined the installation practices for new electronic lubricators and the use of a high-performance grease.[3] Rail wear on the gauge face of the high rail was virtually eliminated, but the more precise gauge face grease application resulted in a drier rail crown. Top-of-rail wear increased on both the high and low rails (Fig. 24.2).

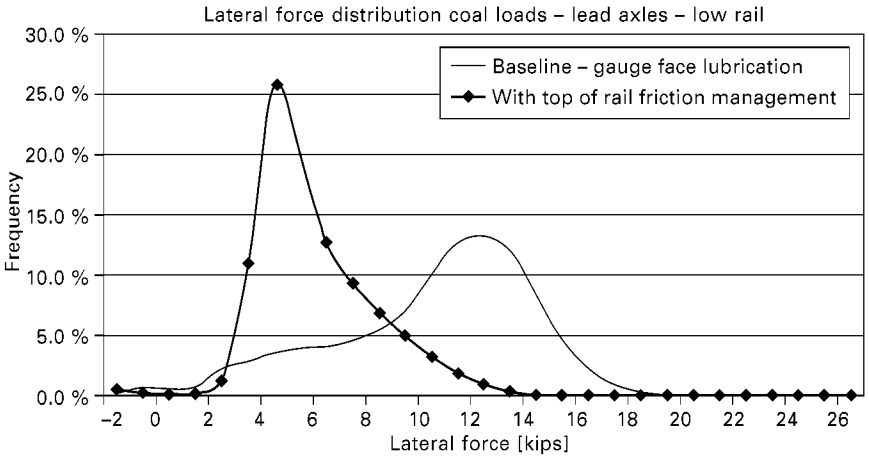
In 2004, CPR tested and perfected the implementation of a top-of-rail friction management program in the same subdivision. This defined the best practices for the future installation of top-of-rail electronic equipment and a liquid friction modifier [4]. Lateral forces in curves were reduced by between 30 and 50 % (Fig. 24.3) and top-of-rail wear was reduced by 50 %.

Track materials

It is CPR's field experience that rail deterioration increases rapidly when track gauge is permitted to progress to greater than 1448 mm (i.e. 13 mm or ½ in wide gauge). This is likely due to the greater tendency of hollow wheels to contact the field side of the low rail. The wide gauge problem was initially addressed by upgrading to 274 cm (9 foot) long hardwood sleepers



24.2 Rail wear (in mm/80 MGT of traffic) for different high rail lubrication strategies. The base case reflects the average wear rates under CP's pre-2000 lubrication practices for curves sharper than 8° (250 m). 'Top of rail contamination' refers to a practice of increasing the volume of lubricant dispensed from gauge-face units to force top-of-rail contamination by grease. The 'optimal gauge-face lubrication' scenario shows that the gauge-face wear is eliminated but the top of rail wear on both the high and low rails increases.



24.3 Lateral forces on the low rail reduced considerably when top-of-rail friction management was applied in addition to gauge-face lubrication.

in curves. The 36 cm (14 in) tie-plate was changed out for a 41 cm (16 in) plate eccentric to the field to resist overturning.

CPR is predominantly a timber-sleepered railway and became concerned with the continued use of cut spikes in the sharper curves of the western corridor under AC traction and heavier axle loads. A program was undertaken to replace cut spikes with elastic e-clips mounted on a rolled plate that is held down by five screw spikes with spring washers. This fastening system was designed to give extra safety against rail rollover and to reduce timber sleeper deterioration by greatly controlling tie-plate movement and the subsequent abrasion of the wood fibres under the rail seat. Rail life with the improved fastening system has doubled in curves previously subject to wide gauge because the rail does not rotate as freely and maintains the as-designed contact conditions with the wheel.

Wheel profiles

CPR began testing different wheel profiles in the late 1970s, with the theory that a 'worn' wheel profile, when machined onto a new wheel, would not exhibit the wearing-in that foreshortened the life of the Association of American Railroads (AAR) 1:20 coned profile used at the time (the tread is sloped at 1:20 (i.e. 2.84°), providing a 'coned' wheel tread). CPR adopted the CN wide-flange Heumann A profile when tests found that wear life (primarily limited by flange wear) was double that of the thinner flanged AAR 1:20.

AAR followed the progress of CN's wheel profile closely and, based on a different sample of wheels and rails developed and presented the AAR1B

'worn wheel profile shape' as a new interchange standard, which CPR adopted in 1986. The AAR1B profile provided a better balance between what is needed for good curving and for reduced bogie hunting. CPR wheel life increased from 273 000 km (170 000 miles) in the 1970s to 500 000 km (314 000 miles). Even so, CN found that the two-point contact conditions between their sharp curve high rails and the AAR1B wheel still resulted in high wear rates until after about 60 000 km (37 000 miles) of run-in.

In 1996 the National Research Council (NRC) Canada designed a 'worn' wheel profile to minimise creepage and contact stresses that contribute to rolling contact fatigue shelling of steel wheel treads. The NRC anti-shelling wheel (NRC-ASW) wheel profile is discussed in detail in Section 24.6. CPR continues to field trial wheels of different chemistry and steel cleanliness.

Frame-braced bogies

CPR became interested in steerable radial bogies after field tests with the Barber-Sheffel and DR-2 bogies in the early 1980s showed wheel flange wear reduced to ¼ that of control cars. Further extensive field tests with 240 cars equipped with flexible bogies confirmed the wear reduction, with the added observation that wheel tread shelling was reduced by $^2/_3$ simply because of the reduction in lateral slippage through curving. Important from an economics point of view was the fact that retrofit versions for standard bogies showed a performance that was comparable to the new flexible bogies. Frame-braced bogies that employ a rubber shear pad between the bearing adaptor and pedestal roof became the standard for CPR's coal fleet in 1989.

Recent analysis indicates that frame bracing has improved wheel performance on the coal fleet by approximately 40 %, increasing average wheel life from 314 000 km (195 000 miles) to 437 000 km (272 000 miles), despite the increase in gross vehicle weight from 119 tonnes (263 000 lbs) to 130 tonnes (286 000 lbs).

24.3 Cartier Railway Company – optimising the captive railroad

The benefits of a well-managed wheel–rail system have been amply demonstrated on Cartier Railway Company (CRC),[5,6] a railroad that connects the Mount-Wright iron ore mines with the harbour of Port Cartier on the St Lawrence Seaway. This 418 km (260 mile) railroad operates 239 tonne (263 ton) cars with axle loads of 27 tonnes (30 tons), with a total traffic of approximately 22 mgt (24 MGT) annually. The track runs through rugged mountain territory and consists of 50 % curves as sharp as 250 m radius (7°) and a maximum grade of 1.35 %.

As with many other railroads, CRC initiated its grinding program in the

1970s to address rail corrugation, but it took the addition of head-hardened rail and the adoption of a preventive grinding strategy in the 1980s to fully control the corrugation problem. Under its preventive grinding program using a 16 stone rail grinder, the CRC started with 4.5 mgt (5 MGT) intervals but, within seven years, had moved to 6.3 mgt (7 MGT), and then to 9 mgt (10 MGT) cycles. In the same time their grinding speeds increased from 10 kph (6 mph) to 13 kph (8 mph). However, CRC still suffered from contact fatigue on the rail, plastic flow on both wheel and rail and very early failure (so-called "infant mortality") of its wheels due to shelling of the tread.

In 1994, CRC tested a new wheel profile based on measured worn wheel shapes from its own cars. This profile is called the QCM-Heumann (the CRC serves the Quebec Cartier Mine) and provides additional metal in the flange root to better match the rolling radius difference required to navigate the CRC's many sharp curves. Test results in 1996 showed that the wheel profile alone, in head-to-head testing against the standard profile, provided a 60 % decrease in wheel shelling. Combined with a program of preventively re-truing wheels when they have reached approximately 2 mm (0.08 in) of hollowing, CRC has more than doubled its average wheel life since 1996 (to about 800 000 km or 500 000 miles). With better control of its wheel shapes, CRC has been able to adopt a more rigorous rail profile grinding program. This combination of improved wheel and rail profiles as well as using harder and cleaner steels reduced rail replacement from 4500 tonnes per year (4960 tons/yr) in 1986 to 500 tonnes/year (551 tons/yr) in 2001. Broken rails and detected ultrasonic defects declined to 10 % of the previous levels.

In 2004, CRC implemented revised rail profiles for their preventive grinding program, as well as a new wheel profile with a little more metal in the flange root. These profiles were designed to improve bogie steering and to spread wear over the wheel tread by providing two distinct contact bands in tangent track.

Through a rigorous program of inspection, detection and maintenance, CRC had been relatively free of derailments since 1996, but in 2003 it suffered a derailment that was attributed to a stiff bogie and high-friction conditions following a rail grinding program. The suspect vehicle had been detected the previous day using a wheelset angle-of-attack measuring system but had not yet been removed from service. This motivated them to experiment seriously with top-of-rail friction modification. Because of the difficulty in accessing and repairing track-mounted equipment in CRC's rugged territory, a vehicle-mounted system was desired. Rather than use a conventional hi-rail system, CRC asked if units could be mounted on a revenue train. The result is a friction management system mounted in the first ore car behind the locomotive of a 160 car train carrying 17 330 tonnes (19 100 tons) of ore when loaded from the mine to the port.[7] The equipment applies a thin

film of friction modifier to the rail using a spray nozzle technology. Product is applied preferentially in curves, with detection provided through a GPS-based system. Measurements at a load station on a 4° curve saw the lateral forces decline by approximately 38–45 %. CRC has had only the one (2003) derailment in the past ten years that could be attributed to problems with the wheel–rail interaction.

24.4 Several wheel–rail problems on Vancouver's Skytrain

The Vancouver Skytrain system opened in 1986 to much fanfare. An elevated guideway system was constructed with 'perfect' alignment, including track gauge with a tolerance of ± 1 mm (± 0.040). Computer-controlled, driverless cars were propelled along with linear induction motors, on body-steered bogies with solid axles terminated with small 450 mm (18″) diameter wheels. It was believed that this system would eliminate all the known wheel problems to that date, including wear and corrugation on a system with curve radii as low as 109 m (16°). Very soon after commencing revenue service, it became clear that there were problems – corrugations appeared within eight weeks of running, flange wear was noticeable, ride quality was poor and wheel–rail noise exceeded 100 dB on a system that passed roughly 15 feet from some residential buildings. Within six months of opening, over 85 % of the track and wheels developed corrugations within 10 000 km (6000 miles) of running. An initial review of the system quickly identified problems with the body-steered bogies. Adjustments to the appropriate linkages dramatically improved wheel alignment to reduce flange wear.

A problem of bogie hunting arose because the wheels running on perfectly aligned track would quickly hollow and take the shape of the rail, resulting in effective conicities exceeding 0.6. One solution proposed was to re-align the track so that some of the tangent rail was at standard gauge while other sections were set 10 mm (0.4 in) broader. Fortunately, this was not required. It proved possible, through rail grinding, to shift the running band on the rail to spread wear on the wheel tread and reduce the conformality. This solution of spreading wear through grinding has since been applied on numerous rail systems (including Edmonton, Detroit, Amtrak and Canadian Pacific) as a technique for maintaining good wheel–rail performance.

Rail corrugation on the Skytrain system was a more difficult problem to solve. It was occurring in tangent track on vehicles that had no powered wheels, but either tractive effort or lateral slip was required in all corrugation theories to that point. Acrylic replicas of the corrugated rail surface were collected and examined under high magnification using visible light microscopes. Analysis of the slip vectors suggested the spin creepage dominated the wheel–rail contact conditions (Fig. 24.4).[8] Although the problem has still

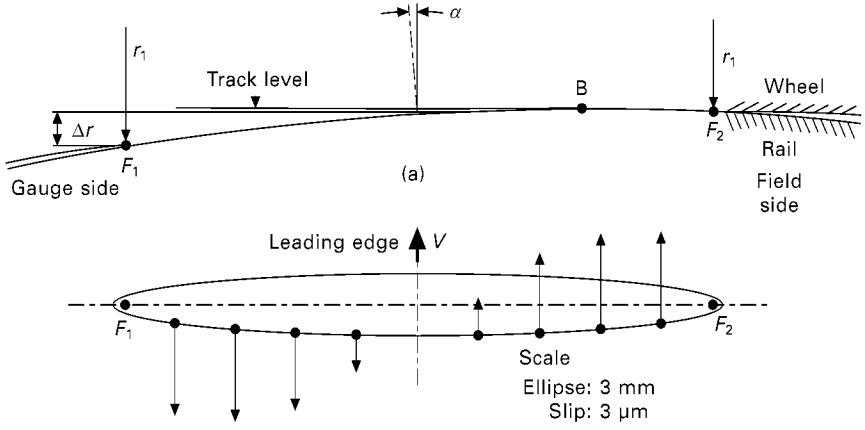
defied analytical modelling (see for example reference 9), the pioneering solution of using rail grinding to reduce spin creepage on tangent rail has since become a standard approach to dealing with short pitch corrugation on this and many other commuter systems. By dealing with the corrugation, the problem of 'roaring rail' was simultaneously eliminated.

The corrugation investigation brought to light the issue of stick–slip at the wheel–rail contact patch. Recognizing that friction conditions were controlled by the characteristics of the interfacial layer, the lead investigator searched for and found a local firm willing to partner to co-develop a new 'friction modifier' later termed high positive friction (HPF). This product would prove very effective in minimizing both corrugation and wheel squeal. The same collaboration resulted in the low coefficient of friction (LCF) sticks for controlling wheel flange and gauge face wear.

24.5 Wheel life and ride quality on Edmonton Transit[10]

North America's first light rail system opened in 1978 in Edmonton, Canada. This system includes 12 km (7.5 miles) of double track mainline with 11 stations. The system had a variety of wheel–rail challenges but most noteworthy was the problem of ride quality.

Around 1984, an instability problem developed with the light rail vehicle's articulated bogie that could be both felt and heard by the passengers, the noise sounding somewhat like a machine gun being fired. A number of treatments were devised and attempted over the next four years, including wheel re-truing every 50 000 km (30 000 miles) but, in 1988, the installation of a shock absorber within the articulated bogie proved successful in halting



24.4 Calculated slip distance due to spin in a tightly conformal contact under steady-state conditions (a) side view; (b) plan view[8].

the vibration. However, other ride quality issues then came to the forefront, including hunting at various speeds on tangent track. Field investigation and analysis of the wheel and rail contact conditions revealed that a key problem was high effective conicities for the hollowing wheels running on rail with a 300–350 mm (12–14 in) head radius. Therefore, the suggestion was to control effective conicity by grinding the rail.

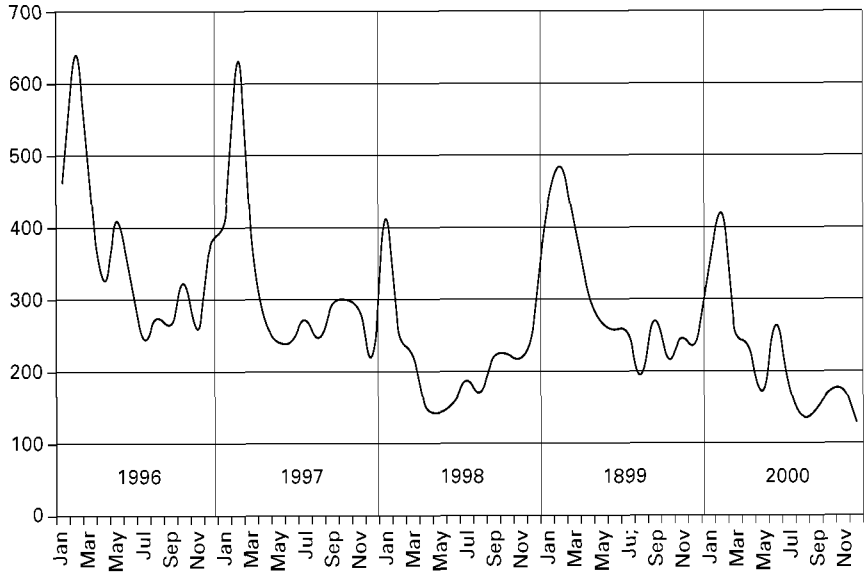
Edmonton did not initially embrace this recommendation, since a previous grinding effort had been disappointing. However, over a trial section at the end of their line, rail grinding was applied to establish a well-defined contact band. The northbound track was ground with a contact band offset to the field (outside), while the contact band on the southbound track was biased to the gauge side of the rail. MacMeter ride quality measurements and interviews with the engineering staff and vehicle operators confirmed the effectiveness of the approach. Over the next couple years, the running band shift was applied on all tangent track.

24.6 Wheel shelling on Canada's freight railroads

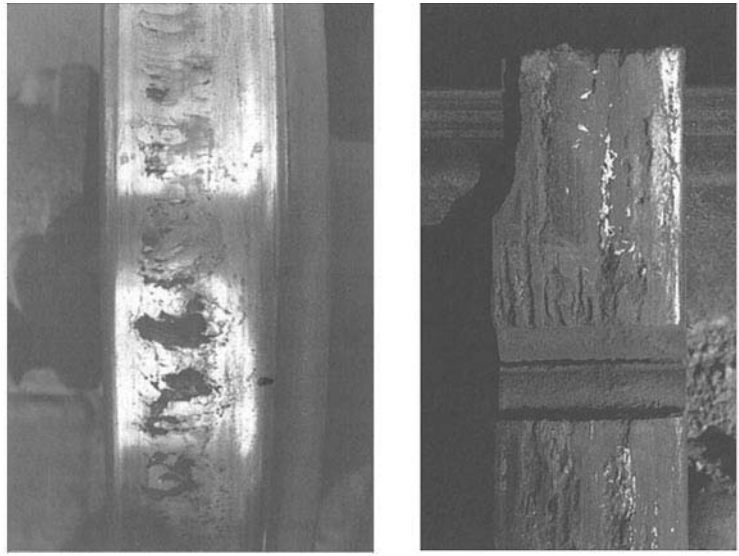
The issue of winter wheel shelling has been a topic of active research by the Canadian freight railway industry since the mid-1990s. The rate at which wheels have to be removed (either for re-truing or scrapping) because of large defects on the wheel tread is roughly five times larger in winter than in summer (Fig. 24.5). This obviously affects the availability of cars and strains the capacity of the mechanical shops to deal with the problem. In fact Canada's two class one railroads were caught short one year when they suffered an epidemic of wheel shelling but had insufficient replacement wheels and were ultimately required to make an emergency purchase from an overseas supplier.

The problem of wheel shelling has often been attributed to wheel slide. However, in Canada, the shells were often found on one wheel and not the other and, even when both wheels were shelled, the damage was not coincident (i.e. at the same 'o'clock' position). These suggested that a different mechanism was at play. One theory was that metal particles collected in the brakeshoe (called 'brakeshoe metal pickup', Fig. 24.6) were rubbing against the wheel and causing surface microscopic martensite that would crack and initiate a surface defect. Another was that extensive crack networks were developing beneath the wheel surface and allowing large chunks of metal to spall from the wheel surface. The treatments for either would be completely different.

An investigation of the brakeshoe metal pickup played through many twists and turns. Initially it was postulated that the brakeshoe metal pickup was the re-combination of oxidative wear debris and carbonaceous constituents in the micro 'tribo-furnace' of the brakeshoe.[11] Although not conclusively proven, it is now believed that when a particle of metal embedded in the



24.5 Seasonality of shelled wheel removals on the Canadian Pacific Railway.[1]



(a) (b)

24.6 Wheel shelling (a) was believed somehow to be related to brakeshoe metal pickup (b).

brakeshoe surface rubs continuously on the wheel surface it will turn red hot, burning its way into the composite brakeshoe material and also contributing to microscopic 'melt wear' of the wheel tread, the ember growing larger as it picks up more metal. Nevertheless, this metal pickup was considered unlikely to be the cause of wheel shelling.

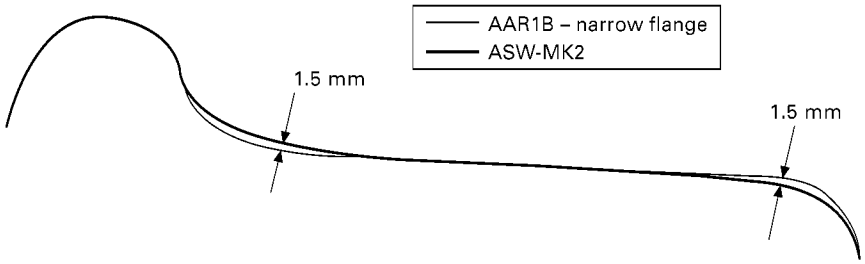
The conditions that exacerbate rolling contact fatigue are many but include metal softening at high rolling temperatures, high friction conditions, high stresses and the presence of water in surface cracks. It now appears that a complicated interaction of high stress, high friction and thermal softening under heavy downgrade braking initiates an extensive network of surface cracks. These cracks propagate under subsequent loading cycles and are accelerated by the presence of moisture at the surface. A noticeable spike in shelling coincident with a particular snow storm strongly supported the theory. Accordingly, improved wheel and rail profiles, friction management and improved braking practices are tools that the railroads use to limit their vulnerability to wheel shelling.

An anti-shelling wheel profile (called the NRC-ASW) designed for the Canadian freight railroads followed from successful field testing of a custom wheel profile designed for the Cartier Railway Company (see Section 24.3). That wheel reduced wheel shelling by 60 % when compared directly against AAR 1B wheels placed into service on trains at the same time. Encouraged by this result, the Railway Association of Canada sponsored the development of a custom wheel profile for testing on Canadian Pacific and Canadian National Railways captive grain and coal fleets.

The NRC-ASW is a 'worn' wheel profile designed to minimize creepage in curves and control contact stresses that lead to rolling contact fatigue shelling. It provides an additional 1.5 mm (0.06 in) of metal in the flange root to significantly improve steering performance and thereby reduce creepage and wear. The 1:20 cone angle in the tread contact region leaves unchanged the wheel's resistance to hunting in standard gauge, tangent track when compared with the standard AAR freight wheel shape. The field side of the wheel has been rolled-off with a 500 mm (20 in) inverse radius to further improve the wheelset steering moment, and significantly increases the time to development of a false flange. Testing on the Canadian Pacific Railway showed an increase in wheel life of 18% based on a population of 980 wheels. The second generation anti-shelling wheel (called the ASW-MK2) is a thin flange model of the NRC-ASW wheel and is more suitable for turning onto service-worn wheels (Fig. 24.7).

24.7 Canadian National Railway – rail grinding needs are site-specific

Improved rail steels, in combination with the adoption of preventive rail grinding in the 1980s, were credited with a tripling of rail life since the



24.7 The NRC anti-shelling wheel (NRC-ASW) profile compared with the industry standard AAR1B wheel shape.

1950s. However, in the mid-1990s, rail grinding in North America faced a challenge. 'No-grind' trials being conducted on premium steels in the USA found that after even 363 mgt (400 MGT) of traffic, the rail suffered only minor surface damage. Initial results gave a projected life of rail under the no-grind approach that was over 5 billion gross tons (BGT) on a sharp (220 m radius, or almost 8°) curve, while one-pass grinding would limit life to about 1.5 BGT.[12] These results led finance and purchasing personnel at railways to question the merits of rail grinding. The rail grinding budgets were under pressure.

As field tests at six sites on the Canadian National (near Kamloops, British Columbia) and 12 sites at the Norfolk Southern (NS, near Roanoke, Virginia) railroads progressed, a different picture emerged. Three grinding regimes were tested on both railroads: no-grind, aggressive or frequent grinding, and light or less frequent grinding. On CN, two of the sites were ground at CN's regular 11 mgt (12 MGT) interval and two others at an average of 19 mgt (21 MGT). On NS, four sites were aggressively ground to a non-conformal two point contact at 41 mgt (45 MGT) and four others less aggressively ground at the same interval with less gauge corner relief.

The CN sites were situated in curves with radii ranging from 545–246 m (3.2–7.1°) and contained premium Japanese rail in the 136-4 CN rail section. The low rails on the NS sites exhibited shallow spalling after 330 mgt (364 MGT) of traffic, but the high rails, whether ground or not, appeared to suffer little damage with only fine cracks appearing. The wear rates increased with the amount of grinding. On the CN sites, the frequent grinding resulted in less surface cracking, while the unground sites showed substantial gauge corner cracking and shelling. This contrasted not only with the NS results but also with results from test sites on the Union Pacific Railroad which, after supporting roughly the same tonnage, showed much less fatigue and less wear.[13]

Why the premium rail at CN so quickly deteriorated without grinding while the test rails on the other railroads survived much longer is clearly attributed to the many differences in the operating conditions. Contact mechanics shows

that the depth of the contact stress, and of crack propagation, is governed by friction at the rail surface. Under very dry conditions, the shear stress is greatest at the rail surface and it dissipates rapidly as you progress into the rail. Under lower friction conditions, the shear stress retains its larger value at a greater depth and so the crack can drive deeper into the rail. However, even more significant is the presence of moisture in a surface-breaking crack. In a process hypothesized by Way[14] and then analyzed by Bower,[15] water is drawn hydroscopically into a crack and, when pressurized by a passing wheel, can drive the cracks much deeper. Areas of intermittent rain are particularly problematic; where contaminants are washed from the rail surface, the surface friction is high, the rate of surface crack initiation is large, and then the propagation is rapid under the next incursion of water. As noted in Section 24.6, snow can serve as the source of moisture also. A further example of the hydraulic effect is outlined in reference 16 where surface-cracked rail in the immediate vicinity of a wayside applicator of a water-based friction modifier product rapidly developed spalling.

24.8 Summary and conclusions

The examples provided in this chapter have covered some of the high-profile activities, challenges and solutions that have developed in Canada since the late 1980s. There have been significant advances in our understanding of the interaction between wheel and rail profiles and their affect on fatigue, wear and ride quality. Practical implementation of friction management and the benefits that it offers with respect to forces, wear and track degradation has progressed dramatically in the last five years alone.

However, the story is far from over. As axle loads, train length and speed increase to improve throughput, the demands on the track structure, rails, wheels and car bodies, and the search for new solutions will carry on unabated. Current issues include understanding the contribution of longitudinal train forces to wheel climb derailments, rail stress and track deterioration. Quantifying the benefits of top-of-rail friction management is an ongoing effort. Tools for measuring crack growth and software to predict crack growth are needed. Understanding the role of track curvature and multiple track geometry defects on vehicle–track safety is a subject of current study. Diagnosing the required treatments for bogies that have poor performance as measured by truck performance detectors (TPDs) is yet another effort. The issue of seasonal wheel shelling remains a problem for the Canadian railroads. The relentless challenge to move freight and passengers more safely, efficiently and economically ensures that the complex interaction between metallurgy, tribology (friction, lubrication and wear), contact mechanics and vehicle–track dynamics will continue to be an important field of study for the next several decades at least. The need for new and novel solutions is unlikely to end any time soon.

24.9 References

1. D. Meyler, P. Sroba and E. Magel (2001), Reducing operating costs through improved wheel performance, *Proceedings 13th International Wheelset Congress*, Rome, Italy, 17–21 September, on CD.
2. R. De Vries, P. Sroba and E. Magel (2001), Preventive grinding moves into the 21st century on Canadian Pacific Railroad, *Proceedings AREMA 2003 Annual Conference*, Chicago, IL, USA, 9–12 September on CD.
3. P. Sroba, M. Roney, R. Dashko and E. Magel (2001), Canadian Pacific Railway's 100% effective lubrication initiative, *Proceedings AREMA 2003 Annual Conference*, Chicago, IL, USA 9–12 September, on CD.
4. P. Sroba, M. Roney, K. Adknow and R. Dashko (2005), Canadian Pacific Railway 100% effective friction management strategy, *Proceedings 8th International Heavy Haul Conference*, Rio De Janeiro, Brazil, 14–16 June, 93–102.
5. J. Kalousek and E. Magel (1998), Learning from the CRC, *Railway Track and Structures*, **4**, 27–30.
6. G. Sirois (1998), Benefits of preventive wheel maintenance, *ARM Crosstalk Wheel/ Rail Interface Seminar*, Chicago, IL, USA May.
7. J. Cotter, D. Eadie, D. Elridge, J. Bilodeau, S. Michaud, Y. Rioux, G. Sirois and R. Rieff (2005), Freight car based top of rail friction modifier application system, *Proceedings 8th International Heavy Haul Association Conference*, Brazil, 14–16 June, 69–76.
8. J. Kalousek and K. Johnson (1992), An investigation of short pitch wheel and rail corrugations on the Vancouver mass transit system, *Proceedings of the Institution of Mechanical Engineers*, **206**, 127–35.
9. A. Bhaskar, K.L. Johnson and J. Woodhouse (1997), Wheel–rail dynamics with closely conformal contact Part 2: forced response, results and conclusions, *Proceedings of the IMechE Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, **211**, (1), 11–26.
10. L. McLachlan (1992), Rail profiles can make a difference, presentation, *APTA Rapid Transit Conference*, Los Angeles, CA, USA, 13–17 June.
11. J. Kalousek, E. Magel, J. Strasser, W.N. Caldwell, G. Kanevsky and B Blevins (1996), Tribological interrelationship of seasonal fluctuations of freight car wheel wear, contact fatigue shelling and composition brakeshoe consumption, *Wear*, **191**, 210–18.
12. J. Hannafious and S. Mace (1995), Interim results of Norfolk Southern rail grinding tests, *TTCI Technology Digest*, TD95–O27, December.
13. K. Sawley and J. LoPresti (1998), Revenue-service rail grinding tests, *Railway Track and Structures*, **4**, 16–18, 33.
14. S. Way (1935), Pitting due to rolling contact fatigue, *Journal of Applied Mechanics, Transactions of ASME*, **2**, 49–58.
15. A.F. Bower (1988), The influence of crack face friction and trapped fluid on surface initiated rolling contact fatigue cracks, *Journal of Tribology*, **110**, 704–11.
16. D. Eadie, L. Maglalang, B. Vidler, D. Lilley and R. Reiff (2005), Trackside top of rail friction control at CN, *Proceedings 8th International Heavy Haul Association Conference*, Rio de Janeiro, Brazil, 14–16 June, 85–92.

Copyrighted Materials
Copyright © 2009 Woodhead Publishing Retrieved from www.knovel.com

25

Managing the wheel–rail interface: the US experience

J. TUNNA, Transportation Technology Center Inc, USA

Abstract: This chapter describes the extensive US experience in development of the wheel–rail interface. It gives details of improvements in wheel and rail materials and profiles. The management of wear and rolling contact fatigue is described. Lubrication practices and tests using friction modifiers are covered. Information is given on wayside systems used to monitor vehicle–track interaction.

Key words: wheel, rail, vehicle, track, wear, fatigue, grinding, lubrication.

25.1 Introduction

This chapter begins with an introduction to the US railroad industry. The key organisations involved in the wheel–rail interface are described. The following sections describe the evolution of wheel and rail materials and profiles, experience of wheel and rail surface damage, lubrication and friction control, and vehicle condition monitoring. The conclusion gives recommendations for further reading. In this chapter US terminology is used for railroad (railway company), car (vehicle), truck (bogie), tangent track (straight track) and tape line (tread centreline).

The railroad network in the USA is extensive. Table 25.1 summarises the size and revenue of the 560 companies that made up the industry in 2006 (AAR, 2007a). Class 1 railroads are defined as having revenue of at least $346.8M/year (2006 figure). Regional railroads operate at least 350 route miles or they have revenue between $40M/year and the Class 1 revenue threshold. Local railroads have less route miles and lower revenue than regional railroads. There are seven Class 1 freight railroads – CSX

Table 25.1 US railroad size and revenue in 2006 (AAR, 2007)

Type	Number	Route miles	Employees	Revenue [$M/year]
Class 1	7	94 801	167 581	50 300
Regional	33	16 713	7 742	1 700
Local	519	28 415	11 634	2 000
Total	559	139 929	186 957	54 000

Transportation, Norfolk Southern, BNSF Railway Co., Kansas City Southern Railway Co., Union Pacific Railroad Co., Canadian National Railway and Canadian Pacific Railway. Only the parts of the two Canadian railroads in the USA are considered Class 1 railroads.

Track safety standards in the USA are specified by the US Department of Transportation, Federal Railroad Administration–Office of Safety (FRA, 2000a). This is a government department, and the track safety standards are law. The track safety standards define nine classes of track according to the maximum operating speed. They specify track gauge, geometric quality and speed in curves, but do not have much to say concerning the wheel–rail interface. For example, there is no specification for rail profile or lubrication in curves. Remedial actions are specified for a range of different rail defects, but rolling contact fatigue is not specifically included.

The track safety standards do not include limits on rail wear. For track class 8 and 9 (maximum operating speed between 125 and 200 miles/h (200 and 320 km/h)) vehicle–track interaction limits are specified. These require running over the track periodically with instrumented wheelsets and demonstrating that limits on vertical and lateral wheel rail forces are not exceeded. Limits on carbody and truck accelerations are also specified.

The FRA also specifies freight car safety standards (FRA, 2000c). These standards give limits for wheel wear, but they do not specify any particular wheel profile. They give limits on the size of slid-flats or shelled spots on the wheel's tread, but they do not give limits for the length of rolling contact fatigue cracks.

The American Railway Engineering and Maintenance-of-way Association (formerly the American Railway Engineering Association) is an important source of advice and information for US railroads. It publishes the Manual for Railway Engineering (AREMA, 2007). The manual contains recommended practices that are used by individual railroads to develop their own particular policies and standards. Rail wear limits are not recommended by AREMA but are determined by individual railroads.

An interesting aspect of US railroads is the interchange of cars between railroads. Interchange happens when a car owned by one railroad (or private car owner) operates over track owned by another railroad. The Association of American Railroads (AAR), among other things, manages the interchange process. The Field Manual of the AAR Interchange Rules (AAR, 2007b) specifies limits for wear and damage to car components. If an inspector working for railroad A finds a limit has been exceeded on a car belonging to railroad B, then railroad A can repair the car in its workshops and send the bill to railroad B.

The Field Manual is a condensed version of AAR Mechanical Division's multi-section Manual of Standards and Recommended Practices (AAR, 2007c). The full manual contains specifications, standards and recommended

practices for rolling stock used in interchange service. It recommends wheel profiles and wear limits.

Transportation Technology Center, Inc. (TTCI) is a wholly owned subsidiary of the AAR. It conducts research and testing for the AAR and FRA, and provides consultancy to railroads and suppliers in the USA and overseas. TTCI is an important knowledge base in the USA for wheel–rail interaction. It has published several hundred technical reports on this subject. TTCI has developed the NUCARS® computer program for modelling cars. NUCARS® is used extensively to investigate wheel–rail interaction.

TTCI operates the US Department of Transportation's Transportation Technology Center (TTC) located near Pueblo, Colorado. The TTC has special test tracks that are used in the acceptance process for new cars. These test tracks aim to excite different modes of vibration in cars. Instrumented wheelsets are used to check that wheel–rail forces stay within acceptable limits. This testing has been a key factor in ensuring the safe introduction of cars with increased axleloads. In the 1970s the maximum freight car weight was 263 000 lbs (120 000 kg). The current limit for Class 1 railroads is 286 000 lbs (130 000 kg). Test are being conducted on cars weighing 315 000 lbs (143 000 kg).

TTCI also operates the Facility for Accelerated Service Testing (FAST). This is a kidney-shaped loop of track over which a loaded train operates continuously during testing periods. Over the years the axleload of the wagons has increased and is currently 78 750 lbf (350 kN). In a typical year of operation FAST will experience 150 million gross tons (mgt) (136 million gross tonnes (MGT)) of traffic. FAST is used to test rolling stock and track components (LoPresti and Kalay, 2000).

The Volpe Center performs research and development into railway, and other, transportation systems for the US government. Railway research is also performed at several universities including the University of Illinois, Texas A&M University and Virginia Technical University.

25.2 Wheel and rail materials

Both cast and wrought wheels are used by US railroads. Wheels with straight and curved plates were once in common usage, but now only curved plate wheels are recommended for freight cars. The conventional material used for wheels is carbon steel with a pearlitic microstructure. All wheels in interchange service must be heat-treated by rim-quenching and tempering. The different grades of wheel are designated Class L, A, B and C. Table 25.2 shows the carbon content and Brinell hardness for each class (AAR, 2007c). Until 1989, wheels that had not been heat-treated (Class U) were allowed. These had carbon content between 0.65 and 0.80 % and typical hardness of 243 HB.

Table 25.2 Properties of wheels in interchange service

Wheel class	L	A	B	C
Carbon content by weight (%)	0.47 max.	0.47–0.57	0.57–0.67	0.67–0.77
Brinell hardness (HB)	197–277	255–321	277–341	321–363

Rail is conventionally made from carbon steel with a pearlitic microstructure. Early rail was not heat-treated and had hardness around 230 HB. Fully heat-treated rail became available in the mid-1960s. This has a higher hardness, but it was expensive to produce. There was a period in the late 1960s and early 1970s when the wear resistance of rail was increased using various alloys. In 1983, a serious derailment was caused near Marshall, Texas due to a broken weld between an alloyed and a standard rail. This halted the use of alloyed rails and led to the improvements in heat-treatment. Head-hardened rail is now available in a range of hardness up to 380 HB.

Steel rail with a bainitic microstructure has been developed by the AAR and tested at FAST. Head-hardened rails with hardness ranging from 341–378 HB were also tested. The test results showed the high-hardness bainitic rail had a higher wear rate than the head-hardened rails that were tested (Sawley and Jimenez, 2000). Service trials produced a different result (Kirstan, 2002). The wear rates of the bainitic and head-hardened (350 HB) rail installed on the high rail in a curve were found to be similar, while the bainitic rail had a significantly lower wear rate on the low rail. In addition, the surface condition of the bainitic rail was better than the head-hardened rail after 110 mgt (100 MGT) of heavy axleload traffic.

A further series of tests began at FAST in October 2001 (Kirstan, 2004). The latest head-hardened rails from several manufacturers were tested. Average hardness had increased from 365 HB in the earlier tests to 395 HB. The increase in hardness was found to give a 12.5 % reduction in wear rate. In July 2005, the latest rail performance test began at FAST. The most modern premium rails from several manufacturers are being tested. The average hardness of these rails is 410 HB. Interim results, after 252 million gross tons (229 million gross tonnes) of traffic, show a further reduction in wear rate (Robles and LoPresti, 2007).

In 2004, two test sites (in Nebraska and West Virginia) were set up to evaluate premium rail performance, among other things, in service. The sites are 10–30 miles (16–48 km) long, include curves with several different radii and carry predominantly heavy axle load coal traffic. Seven types of premium rail from six suppliers are being tested at the site in Nebraska. The range of hardness of these rails was 360–385 HB. After 140 mgt (118 MGT) the surface hardness of all rails had increased above 400 HB. After 275 mgt (250 MGT) all the rails were performing well, with little wear and minor rail surface damage in the form of fatigue cracks and corrugations

(Li and Clark, 2007). Eight types of premium rail from four manufacturers are being tested at the site in West Virginia. After 74 mgt (67 MGT) all the rails are performing well and show little wear (Li and McDaniel, 2007).

The rail manufacturing process has improved in parallel with increases in head hardness. Today's rail from US suppliers is continuously-cast, vacuum de-gassed and ultrasonically tested before it leaves the steel mill. As a result, rail failures from internal defects have reduced and rail life has increased. Typical rail lives for Class 1 railroads are 1480 mgt (1340 MGT) in tangent track, reducing to 330 mgt (300 MGT) for small-radius curves (Stone *et al.*, 1999). The principal reason for rail renewal is no longer defects but wear or some form of surface damage.

25.3 Wheel and rail profiles

The recommended wheel profile for freight cars is the AAR1B profile. This profile was introduced in 1990 after extensive modelling and testing (Leary, 1988). The previous profile was the AAR 1:20, which had a 1:20 taper from the flange root to the rim. The AAR1B profile has a 1.5 in (37.5 mm) radius between the tape line and the flange that matched the typical worn wheel and rail profiles of the time.

The AAR 1:20 profile produced two points of contact on the high rail in curves when it was new. This generated high contact stresses and creepages, which caused the profile to quickly wear to a stable shape conformal to the rail profile. The intention of the AAR1B profile was to avoid this initial wearing in stage and reduce the wear on the rails. The AAR 1:20 profile did not have a specified flange angle. The flange angle when new could vary between 65°–70°. The AAR1B profile has a conical section on the flange with an angle of 75°. Wide and narrow flange versions of both the AAR 1:20 and AAR1B profiles are specified. The shape of the flange root of the AAR 1:20 profile was different in the two versions. The AAR1B profile has the same shape flange root in the wide and narrow flange width versions.

The AAR1B profile has a higher conicity than a new AAR 1:20 profile. Test showed that, as expected, this gave a lower hunting speed (Leary, 1988). However, it was argued that the AAR 1:20 only had a high hunting speed when new, and it soon wore to a shape with a similar hunting speed to the AAR1B.

The new AAR1B profile was tested in revenue service on a train that also had a car with AAR 1:20 profiles (Leary and Gudiness, 1992). The difference in wear rate between the two profiles was found to vary with the type of car. In general, the AAR1B profile was calculated to give a 30 % increase in wheel wear life. Benefits for rail life, fuel consumption and flange-climb derailments were also estimated.

Since it has been 20 years since the AAR1B profile was developed it is currently being reviewed (Wu *et al.*, 2006). Over that period of time there have been changes in rolling stock, rail types and maintenance and operating conditions. The review found that a new AAR1B profile would often make two-point contact on the high rail in curves. A new wheel profile has been developed that generally produces conformal contact. Tests have shown that it reduces lateral wheel–rail forces and wheel wear rates. Further revenue service tests are planned.

In addition to the tapered wheel profiles, a cylindrical profile is also accepted and is sometimes used on locomotives. The cylindrical tread has a 1:20 taper at the rim to delay its tendency to become hollow.

Research has been performed to determine a suitable limit for hollowness on the tread of a wheel (Tournay *et al.*, 2004). This resulted in a new interchange rule requiring a 4 mm limit on the height of the false flange on the rim above the deepest part of the hollow. Over time, the new rule should reduce the number of very hollow wheels in the system. The maximum allowable hollowness may then be reduced to 3 mm.

Passenger cars used the same wheel profiles as freight cars until the 1970s. Now Amtrak has its own standards for wheel profiles and the American Public Transportation Association (APTA, 2004) is recommending profiles for use in other passenger systems.

As freight axleloads have increased, wheel diameters have also increased to avoid excessive contact stress between the wheel and the rail. The recommended wheel diameter increased in 1961 from 33 to 36 ins (838 to 914 mm) when loaded car weights increased from 220 000 to 263 000 lbs (99 800 to 119 300 kg) (Hay, 1982). The latest cars, with loaded weight of 316 000 lbs (143 300 kg), are required to have 38 in (965 mm) diameter wheels (AAR, 2007c).

Rail is available in a range of sections. Traditionally, each steel mill produced different rail sections. AREMA provides drawings of the common rail sections (AREMA, 2007). Table 25.3 shows the breakdown of rail installed in the USA by weight in 2005 (AAR, 2007a). The crown radius of new rail is typically 8, 10, 12 or 14 ins (203, 254, 305 or 356 mm). The shoulders of new rail have radii of 1¼ ins (32 mm). The radius of the gauge corner varies between 3/8 and 9/16 ins (9.5 and 14.25 mm).

There is no industry standard for the acceptable limits of rail wear. Instead, each railroad sets its own limits. Most railroads use a combination of vertical and lateral wear to determine if a rail needs to be replaced. Although the limits may vary, the maximum allowable bending stresses at the top and bottom of the rail are similar (Sawley and Pasta, 2002).

The choice of rail profile after grinding is also left to individual railroads. Post-grind profiles are usually determined by modelling wheel–rail interaction to optimise contact stress, wear, curving and stability. A single railroad may

Table 25.3 Weight of rail in place (AAR, 2007)

Weight of rail (lb/yard)	Miles	Percent of total
150 and over	273	0.2
140–149	7 664	6.9
130–139	67 906	60.9
120–129	3 882	3.5
110–119	24 727	22.2
100–109	3 069	2.8
90–99	2 639	2.4
0–89	1 212	1.1
Total	111 372	100.0

use several post-grind profiles, each designed to suit a different location on the network.

It is common practice for US freight railroads to grind the field side of the low rail in curves to avoid contact with the false flange on hollow wheels (Stone *et al.*, 1999). A crown radius on the low rail between 7 and 10 ins (180 and 250 mm) is sought. Less grinding on the field side of the low rail may be necessary in future when the new rule for hollow wheels takes effect.

There are different opinions on the best post-grind profile for the high rail in curves. One theory says that the gauge corner of the high rail should be removed to avoid contact at the point where rolling contact fatigue cracks are produced. This produces two points of contact: one on the gauge face and the other on the shoulder of the high rail. Two-point contact can increase curving forces and cause severe gauge face wear. The other theory says that conformal contact should be produced since it reduces curving forces and contact stresses. The emerging consensus is that weak two-point contact, in which the contact points are close together, is the best approach since it has only a small effect on curving forces and soon wears to a conformal shape.

New rail with an 8 in (200 mm) crown radius is produced to meet the demands of railroads that have a similar radius in their post-grind rail profile. Starting with new rail that has this shape reduces the amount of material removed the first time the rail is ground.

25.4 Wheel and rail surface damage

Wheel surface damage is a major cause of wheel maintenance and replacement in the USA. A cross-industry consortium has been formed to address this problem (Cummings, 2007). Almost half the freight wheels changed out in 2006 were due to tread damage related causes. Tread damage is separated into slid-flats leading to spalling and fatigue defects leading to shelling.

Steps have been taken to prevent the formation of slid-flats. Braking tests have been reviewed and improved, and training has been given to reduce the incidence of trains being moved in yards with the brakes applied.

Considerable research and testing has been performed to investigate the causes of fatigue defects in wheels (Rajkumar and Stone, 1987a). Subsurface impurities are recognised as an important cause. US wheel manufacturers are improving the micro-cleanliness of their products. Comparative tests have been performed under controlled conditions on wheels from US manufacturers and international sources (Stone *et al.*, 2007). These tests are now continuing with a train operating in revenue service.

The requirement introduced in 1989 for all wheels to be rim quenched had a marked effect on the incidence of freight car wheel failures. The number of wheel failures each year dropped by an order of magnitude (Stone *et al.*, 1999). Rim quenching introduces residual compressive stresses on the tread of the wheel, increasing the resistance to fatigue crack initiation and growth. Research continues into ways of getting even more benefit from the process (Lonsdale *et al.*, 2007).

Before 1970, when relatively soft rail was used, corrugations on the top of the rail were a problem. Long-pitch corrugations usually occurred on the low rail in curves and had a wavelength between 200 and 600 mm. They could also be found on the high rail of curves and on tangent track (Hay, 1982).

Grinding machines to remove the corrugations began to be developed in the mid-1930s (Zaremski, 2005). These early machines focused on the top of the rail. Although their grinding stones could be turned about a longitudinal axis, this had to be done manually. A significant advance was made in 1982 when the first grinding machine with remotely controlled rotation of the grinding stones was produced. Today's production grinding machines have computer control systems and between eight and 120 grinding stones, each powered by a motor rated up to 30 hp (22.5 kW). Grinding angles up to 70° to the longitudinal axis can be achieved. After a series of mergers and acquisitions, there are three major grinding equipment manufactures and contractors supplying the US railroads today: Harsco Track Technologies, Loram, Inc. and Speno International.

Rail hardness increased in parallel with the development of grinding machines. As rail hardness increased-corrugations became less of a problem and the removal of rail surface fatigue defects became the main reason for grinding. Several passes of the grinding machine would be required to remove surface defects and restore the rail profile. This became known as corrective grinding.

As rail surface condition improved through corrective grinding, railroads began to adopt what is referred to as preventive grinding. Preventive grinding intervals are shorter than those for corrective grinding. As a result, less

material has to be removed to clean the rail, and grinding can be completed in a single pass with a large machine.

The choice between corrective and preventive grinding depends on many factors, including availability of track time, grinding train logistics and the shape of the fatigue crack growth curve. The crack depth with traffic function in turn depends on many factors including rail type, curve radius, lubrication policy, operational speed and type of traffic. If the rate of crack growth (measured as the depth from the surface) increases with traffic, then preventive grinding should increase rail life and require less track possession time. If the depth of the crack increases linearly with traffic then the corrective grinding may be a good option. Good results have been achieved with preventive grinding (Stone *et al.*, 1999) and with corrective grinding (Sawley, 1998).

Grinding is seen by US Class 1 railroads as an essential tool for rail maintenance (Sawley 1999). Most of these railroads use, or wish to use, preventive grinding. Grinding intervals as low as 10 mgt (9.1 MGT) are used. The ratio between grinding tangent and curved track varies from 1:1 to 1:4. The low rail is the main focus of grinding. Heavy gauge corner grinding to give strong two-point contact on the high rail is not common.

Most modern grinding machines are equipped with rail profile measuring systems. These are used to measure the site before grinding work begins. A pattern of grinding stone angles is selected based on the observed difference between the measured and the desired rail profiles. The profiles may be re-measured by the machine after grinding to check the desired profile has been achieved. This process can be made more efficient if the rail profiles are measured, perhaps by a system mounted on a road-rail vehicle, during a pre-grinding inspection that may take place several weeks before the grinding operation. Rail surface condition can also be recorded during the pre-grinding inspection. Software can then be used to determine the amount of material to be removed around the head of the rail (Zarembski, 2005). Careful monitoring of sites after rail grinding has been found to be very useful. The choice between corrective and preventive grinding and the choice of best post-grind profile often results from trial and error at particular sites.

Software is being developed to help the railroads maintain the wheel–rail interface (Sawley, 2001). The software calculates contact stress, contact position, contact angle, rolling radius difference, conicity and conformality between the wheel and rail profiles. One version of the software matches a library of wheel profile pairs with a pair of rail profiles. This version is used when designing a rail profile to suit the population of wheel profiles operating over a particular route. The other version matches a library of rail profiles with a pair of wheel profiles. This version is useful when evaluating potential new wheel profiles.

A version of the software has been developed for use with onboard track

geometry cars (Hou *et al.*, 2006). The track geometry car measures rail profiles and track gauge. Calculation of exceptions in maximum wheel–rail contact angle, contact position on the low rail, wheel–rail conformity on the high rail, rolling radius difference in curves, effective conicity in tangent track and contact stresses from a library of wheel profiles is made in real time.

25.5 Lubrication and friction modification

Recommended practices for wheel–rail friction control are provided by AREMA (AREMA, 2007). A distinction is drawn between gauge face lubrication and top of rail friction control. The overall recommended coefficients of friction are:

- gauge face on curves: $\mu < 0.2$;
- top of rail (curves and tangent): $0.3 \leq \mu \leq 0.4$;
- maximum difference between top of left and right rail: $\mu = 0.1$.

Lubrication of the gauge face of the high rail in curves has been common on US railroads since the 1930s. A lot of research and testing has been performed to evaluate lubricants (Steele, 1987). The performance characteristics that are important are summarised as:

- **Retentivity** – the number of wheel passes before the effect of the lubricant is significantly diminished
- **Spreadability** – how well the lubricant spreads along the rail from its point of application
- **Flowability** – how much of the lubricant migrates to the top of the rail near the point of application

Wayside lubricators are commonly used in curves for the primary purpose of reducing gauge face and wheel flange wear. On-board applicators became more popular when US railroads began to appreciate the fuel savings that also come from correct lubrication. Originally, the applicator was mounted on the locomotive, but unless it was correctly adjusted lubricant would migrate to the wheel tread and cause slip under traction. Onboard applicators were developed that were located behind the locomotives. The intention was to apply sufficient lubricant to affect the last car in the train, but not so much to affect the locomotives of the following train (Reiff and Gage, 1999).

Onboard applicators offer the possibility of controlling the coefficient of friction on the top of the rail as well as on the flange of the wheel. Since on the top of the rail the aim is to achieve a coefficient of friction in a defined band, the term 'friction modifier' is used instead of lubricant. Significant fuel savings can be achieved by reducing the friction levels in tangent track. Modelling results show that rolling resistance of a loaded freight car can be reduced by 50 % on both tangent and curved track (Reiff *et al.*, 1999).

Reducing the coefficient of friction in tangent track has also been show to raise the speed at which hunting begins (Toma *et al.*, 2002).

Experience has shown that large differences between the coefficients of friction on opposite rails are undesirable. Dry conditions on one rail and lubricated conditions on the other cause an imbalance in longitudinal creep forces and give the wheelset an angle-of-attack. This generates large lateral creep forces on the dry rail, which may cause the rail to roll over or result in flange climbing derailments. This is the reason for the AREMA guideline for a maximum difference in coefficient between the tops of opposite rails.

Tribometers are commonly used to measure the coefficient of friction on the surface of the rail. They are used routinely to check that lubricators are being effective, to monitor seasonal surface condition changes and to help with derailment investigations. Hand-operated and high-speed tribometers are available (Harrison *et al.*, 2002). The high-speed tribometer is usually propelled by a road-rail vehicle and can measure up to 30 miles/h (48 km/h).

25.6 Condition monitoring

US railroads make extensive use of wayside detectors to monitor car performance (Tournay, 2007). The detectors that have a bearing on wheel-rail interaction are:

- wheel impact load detectors;
- overload and imbalanced load detectors;
- truck performance detectors;
- truck hunting detectors.

Wheel impact load detectors are installed at approximately 86 locations in North America. They are the primary means of controlling large dynamic forces caused by out-of-round wheels. The limit for the dynamic force from a wheel is currently set at 90 000 lbs force (400 kN). A railroad is allowed to take a car out of service and replace the offending wheel when this limit is reached. Railroads are also notified when a wheel produces dynamic forces approaching the condemning limit. This allows them to plan to bring the car in for repair at a convenient time and results in less disruption to operations.

Overload and imbalanced load detectors help control the wheel–rail interface by avoiding car loading conditions that lead to higher stresses on the track and greater probabilities of derailment. These detectors use the quasi-static loads measured by wheel impact load detectors to determine end-to-end and side-to-side differences in loading as well as the total car weight. Research is underway to determine suitable limits for overloaded and imbalanced cars.

Truck performance detectors measure both vertical and lateral wheel–rail forces. They are installed in curves at 22 locations in North America. The typical installation is on a reverse curve. Measurements are made on each curve and on the tangent section of track between the curves. Performance indicators are calculated including the ratio of lateral to vertical load at the leading wheel, the total gauge spreading force from a truck and a truck warp index. Suitable limits for the performance indicators to warn car owners of poor performing cars are currently being developed.

Truck hunting detectors are a further development of the wheel impact load detector. They measure the lateral wheel–rail force over several cribs and detect when that load is oscillating. A hunting index has been developed and limits for acceptable hunting performance have been set (Tournay *et al.*, 2007a). Alternative designs of truck hunting detectors monitor the motion of wheelsets relative to the rail. These require their own definition of hunting index and their own performance limits (Tournay *et al.*, 2007b).

A significant advance in the management of data from wayside systems in North America has been the development of an integrated database to store the data and provide information. All the detector types described above are linked to the database, as is data from acoustic bearing detectors. The database system can send warnings directly to railroads and car owners when performance limits are exceeded. It provides performance reports and trending analysis.

25.7 Conclusions

The vast experience gained by the US railroads over years of operation is embodied in the standards and recommended practices published by the FRA, AAR and AREMA. They are recommended reading for anyone interested in learning more about US railroads. They are revised regularly, so the latest editions should be read.

US railroads are active in the International Heavy Haul Association (IHHA). The IHHA conference proceedings are a valuable source of technical information, and many papers are concerned with the wheel–rail interface. The IHHA published a very useful guide to best practices regarding wheel–rail interface issues (IHHA, 2001).

Many of the references quoted in this chapter are reports published by the AAR. Although the work behind these reports was funded by the US Class 1 railroads, the reports are available to anyone outside the industry. Periodically, the AAR publishes compilations of work and proceedings of symposia. The publication of the symposium on rail and wheel lubrication has a lot of information on that subject (Steele, 1987).

AREMA publishes the proceedings from its regular conferences on railway engineering. Reading these proceedings is a good way of keeping up with the

latest developments in the industry. AREMA's Practical Guide to Railway Engineering is also recommended reading (AREMA, 2003).

Although it is fairly dated, the work by Rajkumar and Stone (1987a, b) gives a comprehensive description of wheel failure mechanisms. Zarembski's book has a lot of useful information on managing the wheel–rail interface by rail grinding (Zarembski, 2005).

Finally, no description of the US railroads' wheel–rail interface would be complete without reference to Hay's book on railroad engineering (Hay, 1982). In addition to its historical perspective, this book gives the reader a fundamental understanding of the principles of railroad construction and maintenance in the USA and has several chapters concerned with the wheel–rail interface.

25.8 References

AAR (2007a), *Railroad Facts*, Washington DC, USA, Association of American Railroads Policy and Economics Department.

AAR (2007b), *Field Manual of the AAR Interchange Rules*, Washington DC, USA, Association of American Railroads Policy and Economics Department.

AAR (2007c), *Manual of Standards and Recommended Practices*, Washington, DC, USA, Association of American Railroads Policy and Economics Department.

APTA (2004), *Manual of Standards and Recommended Practices for Rail Transit Systems*, Washington DC, USA, American Public Transportation Association.

AREMA (2003), *Practical Guide to Railway Engineering*, Landover, MD, USA, American Railway Engineering and Maintenance-of-way Association.

AREMA (2007), *Manual for Railway Engineering*, Landover, M, USA, American Railway Engineering and Maintenance-of-way Association

Cummings S (2007), *Wheel Defect Prevention Research Consortium – Presentation to Non-members*, Pueblo, CO, USA, Transportation Technology Center, Inc.

FRA (2000a), *Code of Federal Regulations Title 49, Track Safety Standards Part 213, Subpart A to F*, US Department of Transportation Federal Railroad Administration – Office of Safety, June, Omaha, NE, USA, The Railway Educational Bureau

FRA (2000b), *Code of Federal Regulations Title 49, Track Safety Standards Part 213, Subpart G*, US Department of Transportation Federal Railroad Administration – Office of Safety, June, Omaha, NE, USA, The Railway Educational Bureau.

FRA (2000c), *Code of Federal Regulations Title 49, Freight Car Safety Standards Part 215*, US Department of Transportation Federal Railroad Administration – Office of Safety, June, Omaha, NE, USA, The Railway Educational Bureau.

Harrison H, McCanney T and Cotter J (2002), Recent developments in coefficient of friction measurements at the rail/wheel interface, *Wear*, **253**, 114–23.

Hay W W (1982), *Railroad Engineering*, Wiley, New York, USA.

Hou K, Thompson R, Lundberg W, Madrill B and Wu H (2006), *Wheel/Rail Contact Inspection System Development and Validation*, Report R-981, Pueblo, CO, USA, Transportation Technology Center, Inc.

IHHA (2001), *Guidelines to Best Practices for Heavy Haul Railway Operations: Wheel and Rail Interface Issues*, Virginia Beach, VA, USA, International Heavy Haul Association.

Kristan J (2002), Preliminary results of the AAR developed J6 bainitic rail tested in revenue service, *Technology Digest* TD-02-011, Pueblo, CO, USA, Transportation Technology Center, Inc.

Kristan J (2004), Preliminary results: rail performance evaluation at FAST, *Technology Digest* TD-04-010, Pueblo, CO, USA, Transportation Technology Center, Inc.

Leary J (1988), *Final Report on the Development of an Alternative AAR Interchange Wheel Profile*, Report R-708, Pueblo, CO, USA, Transportation Technology Center, Inc.

Leary J F and Gudiness T A (1992), *The Engineering and Economic Aspects of the AAR-1B Profile*, Report R-808, Pueblo, CO, USA, Transportation Technology Center, Inc.

Li D and Clark D (2007), Update of experiments at Western mega site, *Technology Digest TD-07-003*, Pueblo, CO, USA, Transportation Technology Center, Inc.

Li D and McDaniel R (2007), Update of experiments at Eastern mega site, *Technology Digest* TD-07-004, Pueblo, CO, USA Transportation Technology Center, Inc.

Lonsdale C P, Rusin T M, Dedmon S L and Pilch J M (2007), Wheel rim residual stress research using finite element analysis computer simulations, *Proceedings 15th International Wheelset Congress*, 23–27 September, Prague, Czech Republic, on CD.

LoPresti J and Kalay S (2000), FAST heavy axleload testing update, *Technology Digest* TD-00-024, Pueblo, CO, USA, Transportation Technology Center, Inc.

Rajkumar B and Stone D H (1987a), *Wheel Failure Mechanisms of Railroad Cars, Volume I – Final Summary Report*, Report R-679, Pueblo, CO, USA, Transportation Technology Center, Inc.

Rajkumar B and Stone D H (1987b), *Wheel Failure Mechanisms of Railroad Cars, Volume II – Technical Task Summaries*, Report R-680, Pueblo, CO, USA, Transportation Technology Center, Inc.

Reiff R and Gage S (1999), *Evaluation of Three Top of Rail Lubrication Systems*, Report R-936, Pueblo, CO, USA, Transportation Technology Center, Inc.

Reiff R, Gage S and Pasta C (1999), Evaluation of industry practices for wheel/rail friction control, *Technology Digest* TD-99-018, Pueblo, CO, USA, Transportation Technology Center, Inc.

Robles F C and LoPresti J (2007), Interim evaluation of premium rail steels at FAST, *Technology Digest TD-07-018*, Pueblo, CO, USA, Transportation Technology Center, Inc.

Sawley K (1998), Grinding trial results on Canadian National and Norfolk Southern railroads, *Technology Digest* TD-98-033, Pueblo, CO, USA, Transportation Technology Center, Inc.

Sawley K (1999), North American rail grinding on curves in track, *Technology Digest* TD-99-004, Pueblo, CO, USA, Transportation Technology Center, Inc.

Sawley K (2001), Wheel and rail profile maintenance, *Technology Digest* TD-01-019, Pueblo, CO, USA, Transportation Technology Center, Inc.

Sawley K and Jimenez R (2000), *The Comparative Wear Performance of Premium and Bainitic Rail Steels under Heavy Axle Loads*, Report R-941, Pueblo, CO, USA, Transportation Technology Center, Inc.

Sawley K and Pasta C (2002), Worldwide rail wear limit practices on freight railroads", *Technology Digest* TD-02-029, Pueblo, CO, USA, Transportation Technology Center, Inc.

Steele R (1987), *Rail & Wheel Lubrication*, Landover, MD, USA, American Railway Engineering and Maintenance-of-way Association

Stone D H, Sawley K, Kelly D and Schust W (1999), *Wheel/Rail Materials and Interaction:*

North American Heavy Haul Practices, Virginia Beach, VA, USA, International Heavy Haul Association.

Stone D H, Robles F C and Dahlman G (2007), Effects of microvoids, oxide inclusions and sulphide inclusions on the fatigue strength of wheel steels, *Technology Digest* TD-07-022, Pueblo, CO, USA, Transportation Technology Center, Inc.

Toma E, Cummings S and Reiff R (2002), *A NUCARS™ Study on the Effects of Top of Rail Lubrication on the Performance of Freight Cars*, Report R-957, Pueblo, CO, USA, Transportation Technology Center, Inc.

Tournay H (2007), *Use of Wayside Detection for Rolling Stock Performance Monitoring and Maintenance*, International Heavy Haul Association, Virginia Beach, VA, USA.

Tournay H, Wu H and Guins T (2004), *The Influence of Hollow-worn Wheels on the Incidence and Costs of Derailments*, Report R-965, Pueblo, CO, USA, Transportation Technology Center, Inc.

Tournay H, Chapman S and Walker R (2007a), Evaluation of cars registering salient hunting indices at or above 0.1, *Technology Digest* TD-07-012, Pueblo, CO, USA, Transportation Technology Center, Inc.

Tournay H, Chapman S, Keegan S and Blank R (2007b), Initial performance limits: three hunting detector types, *Technology Digest* TD-07-034, Pueblo, CO, USA, Transportation Technology Center, Inc.

Wu H, Madrill B and Kalay S (2006), New wheel profile design and preliminary service test results, *Technology Digest* TD-06-023, Pueblo, CO, USA, Transportation Technology Center, Inc.

Zaremski A M (2005), *The Art and Science of Rail Grinding*, Omaha, NE, USA, Simmons-Boardman Books Inc.

Copyrighted Materials
Copyright © 2009 Woodhead Publishing Retrieved from www.knovel.com

26

Managing the wheel–rail interface: the Japanese experience

M. ISHIDA, Railway Technical Research Institute, Japan

Abstract: Rolling contact fatigue (RCF) defects such as squats or surface shellings; gauge corner cracks or head checks; gauge face wear; rail corrugations which have four types classified from the viewpoints of mechanism and occurrence location; friction/adhesion which are related to wheel flange climb derailment in sharp curves, the stability of high-speed operation and braking performance in the case of wet wheel–rail interface conditions; lubrication to improve the steering performance of bogies to reduce wear and rail corrugations; and grinding to prevent RCF defects and recover wheel–rail contact configuration – there are very important wheel–rail interface issues in Japanese railways, and this chapter describes the current status of Japanese experience as related to them.

Key words: rolling contact fatigue (RCF) defects, wear, corrugation, friction/adhesion, lubrication, grinding.

26.1 Introduction

Rolling contact fatigue (RCF) defects such as squats or surface shellings; gauge corner cracks or head checks; gauge face wear; rail corrugations which have four types of long and very short pitch corrugations in tangent tracks, long pitch corrugations of high rail in gentle curves and relatively short pitch corrugations in sharp curves; friction/adhesion and lubrication – there are important wheel–rail interface issues for Japanese railways, and this chapter describes the current status of Japanese experience related to them.

The squat type of RCF defect gives rise to considerable costs for rail maintenance with regard to detection of cracks, instillation of reinforcement, rail renewal and so on. The initiation mechanism of squats has been experimentally and/or theoretically studied, and many findings have been obtained which help to understand it.[1–19] Also, statistical analysis has been carried out to identify influential factors in the initiation of squat in Japanese high-speed rails.[20,21] At first the crack is propagated horizontally. Second, after the crack has propagated to some extent, it may branch and begin to propagate transversely.[22–26] Predicting crack propagation is very important for maintaining rail integrity. However, because it is not easy to characterize the current status of rails with respect to both stress and externally applied wheel loads, it remains difficult to predict the progress of squat-type crack propagation. On the other hand, preventive grinding is currently the most

reliable and effective method of reducing the generation of RCF defects, and many Japanese railway operators are adopting it. The effect of preventive grinding on RCF defects has been studied both practically and experimentally, and number of preventive grinding methods have been proposed.[27–35] It is of great interest to obtain the optimal grinding period and grinding amount (the thickness of material removed from the rail surface) to control RCF defects. In addition, the effect of roughness of the ground surface of the rail on RCF and high-frequency noise emitted by wheels rolling on the fresh surface of the rail just after grinding is under discussion. Moreover, gauge corner cracks and/or head checks are a form of RCF defect commonly generated on high rails in curves.[36] Considering the initiation and propagation of such RCF defects, the balance or interaction between wear and RCF can be recognized as a key issue. In curved tracks, it is neccesary to focus on wear and RCF for optimization of track maintenance work.[37]

The flange wear of the wheel and the gauge face wear of the rail are important issues for wheel and rail maintenance cost. Wear is one of the typical and fundamental subjects in the field of tribology, so many studies have been carried out and significant findings have been obtained.[38–46] In this chapter, some features of wheel–rail wear caused by the wheel–rail contact condition in Japanese railways are described through investigation of actual wheel and rail performance since the late 1990s. In addition, the influence of some important parameters, such as stress, relative slip and hardness of material, on wear has been evaluated by laboratory simulation. On the other hand, focusing on the steering performance of the bogie, the influence on the wear of wheel flange and gauge face of lateral force and the angle of attack generated by railway vehicles negotiating sharp curves can be roughly estimated.[47–49] Rail corrugations are another wear issue, but these are related not only to wheel–rail tribology but also to vehicle–track dynamics. The main consequence of rail corrugations in Japan is noise and vibration which is an environmental issue.[50–62] However, the deterioration of vehicle and track structures caused by large interacting forces between rail and wheel excited by corrugations is also a problem.[55]

- very short pitch corrugations with wavelengths of 30–80 mm on straight tracks or on very gentle curves;
- relatively long pitch corrugations with wavelengths of 400–600 mm:
  - on high rails at sharp curves in narrow-gauge middle-speed tracks with a radius of curvature more than 600 m
  - on Shinkansen, high-speed tracks in straight tracks with radius of curvature around 1.2 m
- relatively short pitch corrugations with wavelengths of 80–150 mm on low rails at sharp curves with radius of curvature less than 600 m.[50]

Four types of rail corrugations have been recognised in Japan. Rail

corrugations have been much studied and many sophisticated findings have been obtained.[56,57] Grassie and Kalousek[58] carried out an extensive research review of rail corrugations, covering many railway systems in the world and providing helpful suggestions. However, the short pitch corrugations observed in the narrow-gauge tracks of Japanese railways are slightly different from their classification of corrugations. Many curves in the Japanese metro systems have radii of curvature less than 200 m. Almost all sharp curves suffer from short pitch corrugations on low rails. Matsumoto *et al.* systematically studied these short pitch corrugations.[59] In this chapter, the effect of the lateral force caused by vehicles negotiating sharp curves on the generation of low rail corrugations is investigated based on track site measurements and vehicle dynamics simulation.[60–62] In addition, the potential of lubrication to prevent low rail corrugations in sharp curves is described.[50,53]

In high-speed railways, adhesion between wheel and rail, particularly under wet conditions, is a very important function in maintaining safe running to keep and stable operation from the points of view of braking and driving performance. This chapter describes an experimental investigation and analytical study based on elastohydrodynamic lubrication (EHL) theory, with the focus on the effects of surface roughness and water temperature on the adhesion between wheel and rail under wet conditions.[63–78] The experiments were carried out using a twin-disc rolling contact machine under various conditions, such as three types of surface roughness and two water temperatures.[79] The maximum traction coefficient, which is commonly called adhesion coefficient in the railway industry, was estimated. As a result, the effect on the adhesion coefficient of water temperature and surface roughness of the wheel–rail interface was identified. Based on an understanding of the phenomenon of adhesion coefficient under wet conditions, adhesion-improving materials, such as abrasive block, which is installed in a bogie to improve the adhesion of wheel tread, or jetting devices for applying adhesion-improving materials in the form of ceramic particles to the contact between wheel and rail, have been developed.[80] In 2000, one terrible accident involving wheel flange climb derailment was caused by a combination of high coefficient of friction (COF) between wheel and rail and some other mechanical factors.[47] After that, COF between wheel and rail was measured and the effect of train operating conditions and meteorological conditions on the COF of the rail surface was identified.[80]

Railway vehicles negotiate sharp curves with a large lateral force interacting between the wheel–rail interface and the angle of attack of a leading axle in a bogie which depends on its steering performance. From the point of view of running stability and/or safety, this large lateral force is one of the main factors involved in wheel flange climb derailment. It is also a factor in material integrity issues, such as low rail corrugations, thin flange wear of the wheel and gauge face wear of the rail, and in the generation of rail squeal,

which is an environmental issue, particularly in urban railways.[47–50,53,54,58–62,80,81]

Currently, lubrication of the interface between the wheel tread and the rail top has been the focus of efforts to decrease these large lateral forces and to reduce wear of the wheel–rail interface, low rail corrugation and the resulting squealing noise. However, this lubrication gives rise to some risk of wheel sliding due to the resulting low traction coefficient, which means that the friction of such an interface must be appropriately controlled. The use of a friction modifier is considered to be one promising solution.[53,54,60,62] In this chapter, the influence of wheel–rail lubrication, as one of the main methods of moderating friction, on vehicle–track dynamic behaviour has been evaluated by vehicle running experiments in a test track. The experimental results were discussed from the perspective of interacting forces between wheel and rail with a vehicle dynamic simulator (VDS) developed by Railway Technical Research Institute (RTRI) in Tokyo. In addition, a kind of friction modifier, hereinafter called 'friction moderator (FM)', whose main component is carbon, has been developed, based on applying the materials to the wheel–rail interface along the whole length of curves, using a specially arranged nozzle installed on a bogie. Some performance tests have been carried out and their results discussed.

26.2 Rolling contact fatigue

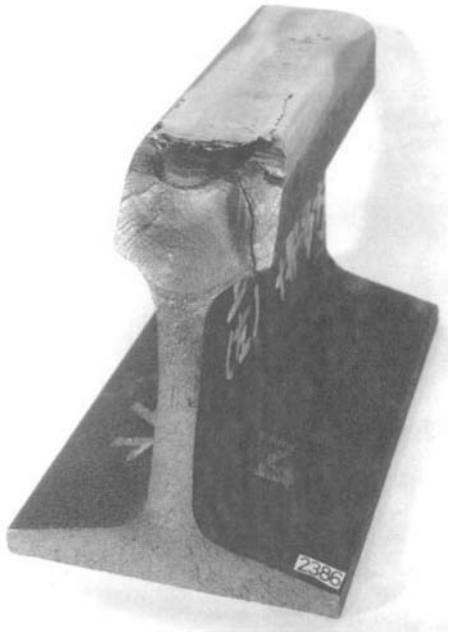
26.2.1 Influential factors

Figure 26.1 shows a typical photo of the squat type of RCF defect. Johnson applied 'shakedown theory' to the rolling contact between wheel and rail in order to understand the mechanism of RCF.[3] The squat type of RCF crack initiates in a very thin surface layer of the rail and develops inside the railhead. The plastic deformation caused by large contact stresses at the asperities in the wheel–rail interface can be considered to have a great influence on crack initiation. Focusing on the crack propagation of the squat type of RCF defects, the two stages of crack propagation, one of which is an early horizontal crack and the other of which is a transverse crack branched from the horizontal crack, are identified.[24] The development of the crack is facilitated by the presence of water.[8,10,22]

26.2.2 Crack propagation

Horizontal crack propagation

Fracture mechanics has been used by a number of researchers to study the development of RCF cracks.[4–6] Bower performed a theoretical study of horizontal crack propagation and developed a model.[8] A Hertzian contact was



26.1 Typical photo of the squat type of RCF defect.

considered, and the stress intensity factors at the crack tip were evaluated using a numerical procedure based on the distributed dislocations method. Akama and Mori[22] have developed a model of a two-dimensional RCF crack based on the boundary element (BE) method. As there is no discretisation of the interior data, the BE method is particularly suited to the problem of severe stress or strain gradients. The BE method is also suitable for contact problems that have to be solved by iteration with an incremental technique. When sliding occurs in the contact zone, the normal and tangential tractions have to be coupled with a friction parameter. Because the tractions are directly treated as nodal values in the BE method, the friction conditions are easily treated. An extension to three-dimensional problems, including residual stresses or thermal stresses, can be easily performed. BE analysis was adopted to estimate stress intensity factors at the inclined surface crack tip under a Hertzian contact. The analysis included an accurate model of friction between the crack faces and the effect of fluid in the crack. Stresses and opening displacement were evaluated at the seven nodes closest to the crack tip. Using a least-squares regression analysis of stresses and displacements and extrapolation to the crack tip, appropriate K-values were obtained. As a result of passing contact pressure, the crack tip is subjected to a complicated non-proportional cycle of Mode I (tensile) and Mode II

(shear) stress intensity K_I, K_{II}. The model accounts for (i) water in the crack lubricating the crack faces and reducing the crack face friction (and easing the Mode II crack propagation) and (ii) generating pressure at the crack tip (and easing the Mode I crack propagation). The effect of residual stresses can also be included. The method was used to study two cases. In one case there is no fluid pressure present, but the crack faces are considered to be lubricated to provide a friction coefficient μ_c of 0.1. In the other case, fluid entrapment is considered.

Akama and Mori's results[22] were compared with those of Bower[8] and were in good agreement with them.[22] They were also compared with the results of Kaneta *et al.*,[10] but a relatively large difference between Akama and Mori's results and those of Kaneta *et al.* was identified. Akama and Mori thought that the difference resulted from the fact that Kaneta's model was for a three-dimensional semi-circle crack without the effect of contact between the crack faces.[22] Akama and Suzuki[23] performed tests on wheel and rail steels to obtain the fatigue crack growth rate and the criterion for crack branching under mixed-mode loading. They simulated fatigue crack propagation in rolling by applying the complex cycle of Mode I and II stress intensity obtained in reference 22 to a cruciform specimen in an in-plane biaxial fatigue test rig. The effect of water trapped inside the cracks and the effect of heat-treatment of steel were studied. Akama and Suzuki proposed various models for crack growth rate. The main outcomes of the work are given below.

The co-planar crack growth rate focusing on the effective stress intensity factor can be fitted to the following equation:

$$da/dN = C[\Delta K_{Ieff}]^n \qquad [26.1]$$

where C and n are material constants. By means of a least square regression analysis, the constants for rail steel were obtained as follows:

$$C = 1.06 \times 10^{-10},\ n = 2.45 \qquad \text{(Rail steel)}$$

$$C = 8.19 \times 10^{-11},\ n = 2.64 \qquad \text{(Wheel steel)}$$

The fatigue crack growth rate under mixed-mode loading was assumed to be controlled by both the Mode I and Mode II mechanisms. The co-planar crack growth rate for rail steel is caused by both the effective stress intensity factors ΔK_{Ieff} and ΔK_{IIeff}. Wong *et al.*[24] proposed a model of growth rate expressed as:

$$da/dN = C\left[\Delta K_{IIeff} \times \{1 + (\Delta K_{Ieff}/\Delta K_{IIeff})^m\}\right]^n \qquad [26.2]$$

where C, m and n are material constants. By means of a least squares regression analysis, the constants for rail steel were obtained as follows:

$$C = 4.25 \times 10^{-11},\ m = 0.95,\ n = 2.23$$

The next approach was to assume that the crack growth rates were primarily controlled by the Mode I mechanism and modified by the Mode II. The model is expressed as:

$$da/dN = C\left[\Delta K_{\text{Ieff}} \times \{1 + (\Delta K_{\text{IIeff}}/\Delta K_{\text{Ieff}})^m\}\right]^n \quad [26.3]$$

with

$$C = 3.47 \times 10^{-11},\ m = 1.20,\ n = 2.17$$

It was found that the experiments correlate better with predictions based on the modified Mode I effective stress intensity factor range.

Transverse crack propagation

A typical squat has two cracks growing in opposite directions at about 10° to the top surface of the rail. As a train passes, the two squat cracks grow longitudinally. However, the crack growth is much faster in the direction of the train travel. These cracks are called horizontal cracks. After the crack has grown to a certain length, it may start to branch at about 60° to the surface. Such a crack is called a transverse crack. Transverse cracks may lead to rail breaks if they keep on growing. The crack is subjected to both tensile (Mode I) and shear (Mode II) stress intensities. However, the role of shear stress intensity in propagating the crack reduces after the crack has branched. Fractured surfaces of broken rails have revealed no evidence of shear mode. Striation patterns exist due to crack tip blunting and re-sharpening, indicating that cracks were propagating in Mode I. Kashiwaya modelled such a transverse crack as a circular crack inclined at an angle of 60° to the top surface and subjected to pure Mode I loading.[26] In the model, it was assumed that the transverse crack grows only by tensile stresses and the crack shape remains circular. Then crack growth calculations were carried out by taking into account stresses generated by train passage, residual stresses and thermal stresses.

The stress intensity factor K_I is given as:

$$K_I\, 2\sigma\sqrt{(a/\pi)} \quad [26.4]$$

where σ is the tensile stress and a is the radius of the circular crack. For this study, crack growth rates were obtained by using compact tension specimens of thickness 20 mm in order to ensure plane strain conditions under Mode I. The fatigue crack growth rate is strongly related to the range of stress intensity factor, and the crack growth rate can be adequately characterized by Paris law:

$$da/dN = C\ (\Delta K_I)^m \quad [26.5]$$

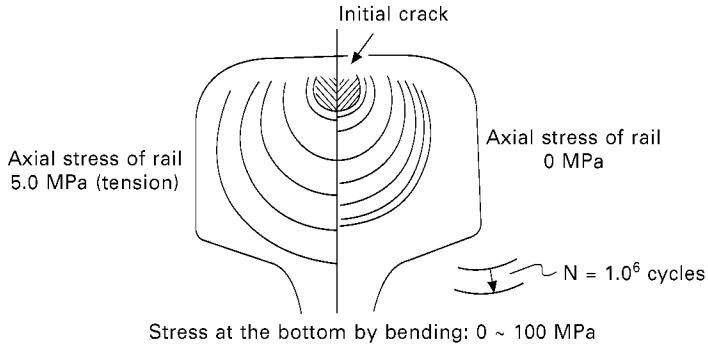
An important factor in estimating the crack growth rate is crack closure. The crack growth rate depends on the range of the stress intensity factor. If the crack is in compression, however, there will be no stress intensity at the crack tip. For stress ratios 0.6, 0.3 and 0.1, crack closures were investigated to clarify the relationship between the effective stress intensity factor range $\Delta K_{\mathrm{I\,eff}}$ and the nominal stress intensity factor range $\Delta K_{\mathrm{I\,nom}}$. In the current study, $\Delta K_{\mathrm{I\,eff}}$ was obtained by neglecting the compressive part of the cycle. According to the results of tension–compression tests with a standard test specimen made of JIS 60 kg rail,[82] the values of C and m were obtained as follows.[26]

$$C = 1.0 \times 10^{-11},\ m = 3$$

Examples of crack growth simulation of transverse cracks have been described, and the results are shown in Fig. 26.2. The right half of the railhead shows crack growth contours when there is no thermal stress and the left half shows crack growth contours when there is a tensile axial stress (thermal stress) of 50 MPa distributed uniformly over the railhead. The solid lines indicate crack fronts at increments of one million cycles corresponding to about ten million tons of traffic. It can be seen that the crack growth rate increases if tensile axial stress (thermal stress) is applied. This is because the average stress in the cycle changes, leading to a change in the cyclic stress intensity factor range. In other words, it is supposed that the crack growth rate of a transverse crack is mainly controlled by the thermal stress and that the growth rate will increase rapidly in winter when the tensile thermal stress becomes large.

26.2.3 Preventive grinding

It is important that the grinding interval (how frequently the grinding should be conducted) and grinding depth (how much surface material should be

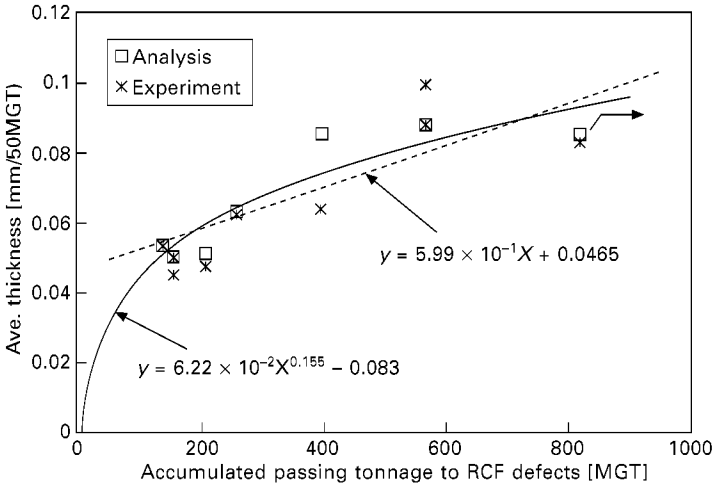


26.2 Crack growth simulation of transverse cracks.

removed) are determined scientifically. This knowledge will improve the efficiency of grinding work and decrease the maintenance cost.

Experimental study

The effect of preventive grinding on squats in Shinkansen rails was estimated experimentally using a large rolling disc machine in which the diameter of a rail disc was 350 mm and that of a wheel disc was 500 mm.[30] The materials of the test discs were the same as those of the wheels and rails used in the Shinkansen. Applied forces were arranged in the test rig to produce the same contact pressure, including dynamic effects, as in the Shinkansen wheels and rails. In the Shinkansen, the traction may vary significantly between each wheel, but an estimation based on the running resistance of trains produces an average traction coefficient of 0.01 because all the axles or wheels of the first-generation of Shinkansen vehicles are driven. Frankly, it is very difficult to estimate the variation of traction caused by every wheel because of the precision of the motor control system and the difference in wheel diameters and tribological conditions between wheels and rails. Spin creep was applied by having an appropriate conicity in the wheel disc. Lateral creep was applied by tilting the rail disc. Since squats rarely initiate in dry tunnels,[20,21] water was applied to the discs. Rail discs were ground every three million cycles, roughly corresponding to 50 MGT (million gross tonnage multiplied by the number of passing axles) to simulate preventive grinding. Tests were stopped after a certain amount of RCF developed on the rail disc. The corresponding MGT was plotted against the grinding depth in Fig. 26.3. The figure also

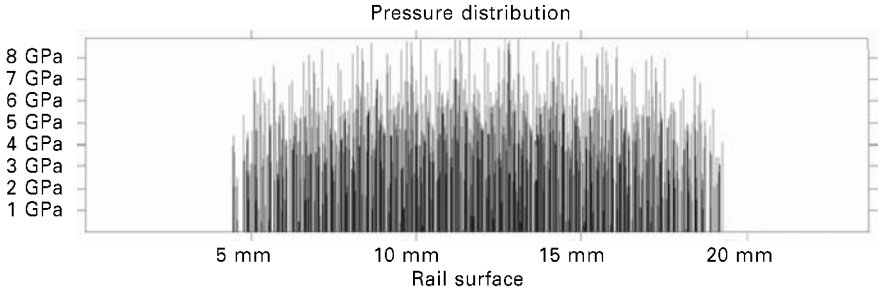


26.3 Effect of preventive grinding on restraining RCF defects.

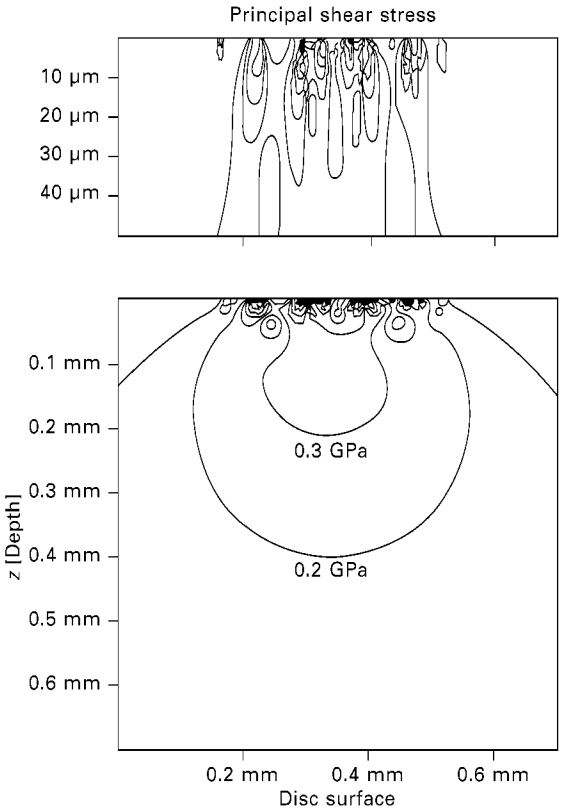
shows the results of a statistical analysis of these data. The relationship between grinding thickness and accumulated tonnage was analysed using the weighted probit analysis which can be adopted to treat small sample sizes and some suspended data.[83–85] In the analysis, assuming the distribution of grinding thickness at accumulated tonnage to be a normal distribution, a 50 % probability value for the distribution was estimated considering the coefficient of variation (standard deviation/mean value) of the distribution as a parameter of the weighted probit analysis. The non-linear regression curve fits the data better than the linear curve. Both curves give a grinding thickness of 0.06 mm at the accumulated tonnage of 200 MGT and 0.09 mm at 800 MGT. This means, for instance, that a grinding thickness of 0.1 mm will prevent the initiation of squats for up to 800 MGT with a fracture probability of 50 %.

Theoretical study

Kapoor *et al.* studied plastic deformation at asperities between wheel and rail and its effect on the shakedown process.[35] Here, their findings are summarised as follows. The maximum contact pressure which is carried purely elastically in the steady-state is known as the 'shakedown limit' and is the rational design criterion for tribological contacts such as ball bearings and railway rails and wheels. For frictionless rolling/sliding, this shakedown limit is four times the shear yield stress of the rail material. With increasing friction, the limit drops, initially gently and then rapidly at a friction coefficient of about 1/3. For Japanese rail materials, the shakedown limit is about 1200 MPa for frictionless sliding (corresponding to a yield stress in shear of about 300 MPa). The dynamic effects combined with the static axle loads causes the operating contact pressures for the Shinkansen to fluctuate in the range 900–1180 MPa. Since a dispersed power system is used in the Shinkanesen, the average traction coefficient is about 0.01. Because the shakedown limit is greater than the maximum contact pressure, the rail material is expected to deform elastically. However, substantial plastic deformation in a subsurface layer of thickness 15–20 μm is generally observed in cross-sections of rails. This plastic deformation is caused by asperities on the wheel and the rail surface. Figure 26.4 shows the contact pressure developed due to asperities between a typical Shinkansen rail and wheel (see details in reference 35). The nominal contact pressure estimated using a smooth Hertz contact approximation is 1 GPa. However, since asperities make contact over a smaller area, the asperity contact pressures can be seen to be about 8 GPa. The asperity contact area is generally of the order of a few tens of microns so the plastic deformation would occur at a depth comparable to this dimension. Figure 26.5 shows the resultant principal shear stress (τ_1) distribution due to the pressure distribution shown in Fig. 26.4. The effect of asperity stresses is confined to a very thin



26.4 Example of contact pressure developed due to asperities between a typical Shinkansen rail and wheel.



26.5 Resultant principal shear stress (τ_1) distribution.

surface layer of approximately 0.1 mm depth. Below this layer, the stresses agree with those estimated using a smooth Hertz approximation. It can be seen that plastic flow is confined to a thin surface layer where material fails, leading to initiation of squat cracks. Removing this layer by preventive

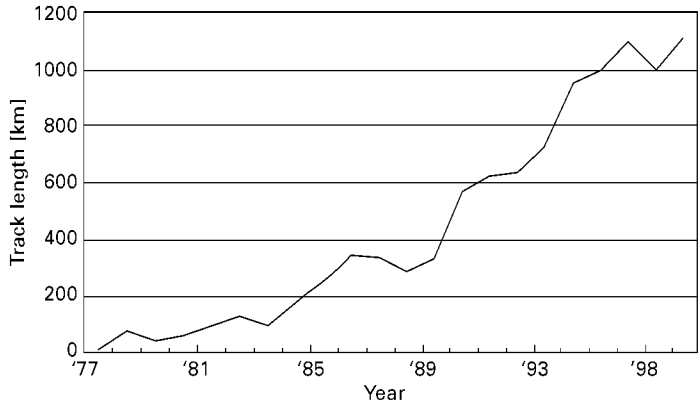
grinding reduces the chances of cracks initiating and developing into squats. Experiments by Ishida *et al.*[30] confirm this expectation.

Practice

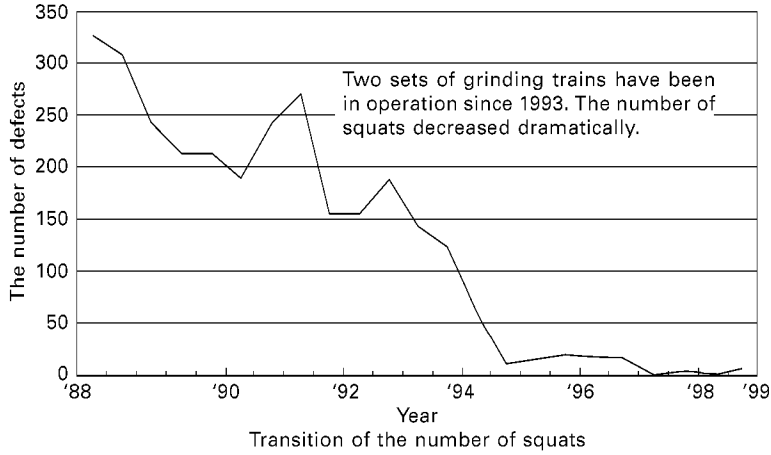
The Central Japan Railway Company adopted rail grinding to reduce rolling noise and to extend the rail life by preventing the initiation of squats in the Tokaido Shinkansen. Tada[33] reported the effect of the preventive grinding proposed in this chapter on reducing the number of squats. Figure 26.6 shows the history of rail grinding work. Grinding started in the late 1970s, and only about 100 km of affected track was ground every year. In 1995, two rail grinding cars were introduced. Since that time, rail grinding has been carried out on more than 1000 km of track per year. This covers almost the whole length of installed rails in the Tokaido Shinkansen. Figure 26.7 shows the effect of preventive grinding for one pass per year of a 48 stone grinding car in a special investigation section of 164 km of track. The figure shows that the number of squats has been steadily decreasing as a result of grinding. The target has been a grinding thickness of 0.08 mm/pass and a grinding interval of 40 MGT, which is similar to the 0.1 mm/50 MGT determined by Ishida *et al.*[30,31]

Surface roughness of rail formed by rail grinding

Preventive rail grinding is currently becoming very popular in Japan to remove the surface layer of RCF damage caused by repeated train loads. Also, curative rail grinding is carried out to remove longitudinal irregularities of the rail surface such as rail corrugations and rail welds. Such rail grinding is contributing to extending the service life of rails and to reducing rolling



26.6 History of rail grinding work.



26.7 Effect of preventive grinding for one pass per year of a 48-stone grinding car in a special investigation section of 164 km of track between the 436 and 518 km markers on the Tokaido Shinkansen.

noise and vibration. On the other hand, elastic–plastic stress analysis of asperity contact has shown that the contact stress of the wheel–rail interface considering surface roughness is very large. For instance, in the case where von Mises stress surpasses shear yield stress of rail materials, plastic deformation may take place and cause cracks due to ratcheting and so on. From the point of view of the efficiency of rail grinding work, an appropriate degree of rail roughness has some advantages with respect to grinding speed and ground thickness per cycle. However, an initially large roughness formed on the rail surface by rail grinding poses a problem with high-frequency rolling noise and has some potential to course RCF damage as mentioned above. On the other hand, such initial large roughness usually settles down to a normal level due to repeated train loads. This may give rise to two interesting questions. One is how long it takes or how much accumulated passing tonnage is needed for this initial large roughness to settle down. The other concerns the degree of RCF damage accumulated from the initial large roughness before normal roughness is established by passing train loads. Hence, it is very important to study the optimal initial roughness formed by rail grinding to avoid rolling noise and/or RCF and to obtain the greatest efficiency of grinding work.

Some experiments using a twin-disc machine were carried out to investigate the relationship between the capacity of accumulated passing tonnage to settle down initial roughness formed by grinding and the magnitude of that initial roughness. In addition, the test discs were studied, using an optical microscope to analyse the plastic flow of the surface layer and X-ray diffraction to measure the density of the crystal axis. The experimental results obtained

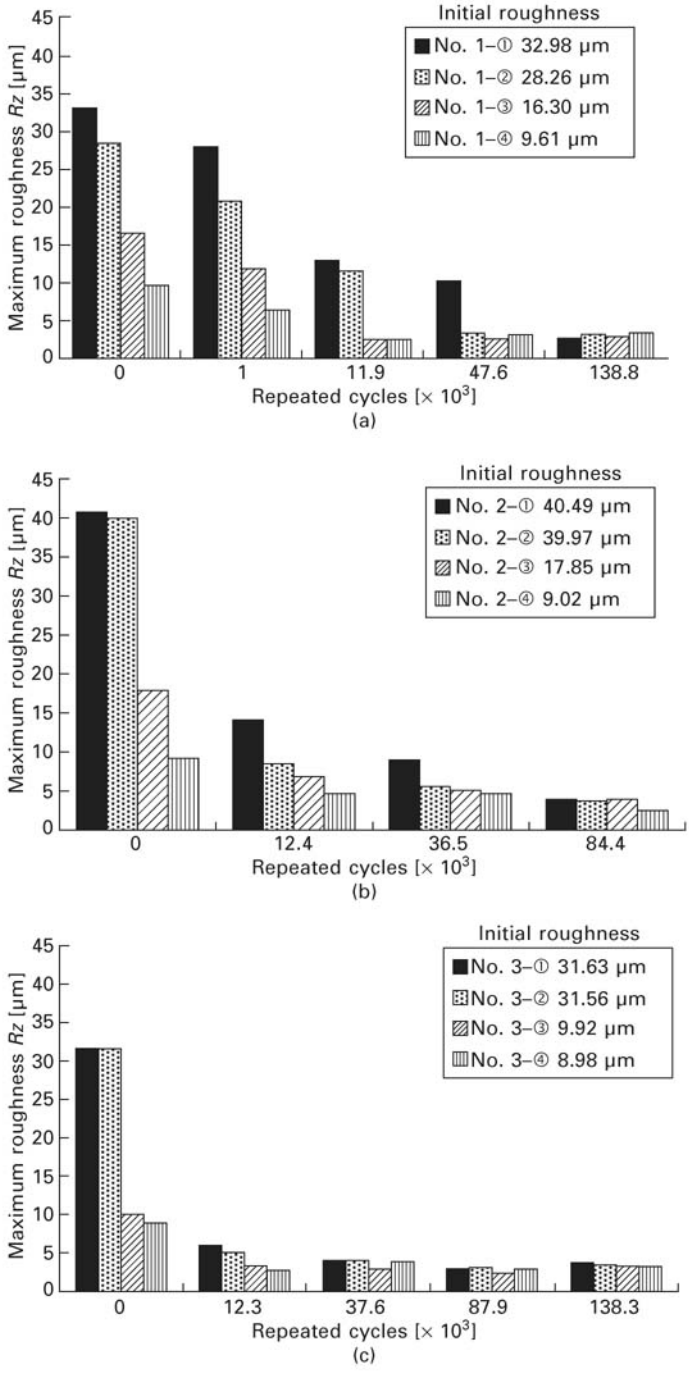
from rolling tests with the twin-disc machine and metallurgical analysis results of test discs are as follows.[86]

Variation of roughness with repeated cycles

Figure 26.8 shows the variation of roughness with repeated cycles in terms of the experimental arrangements given in Table 26.1. In this figure, *Rz* (arithmetic mean of surface roughness) decreased gradually in response to the increase of repeated cycles. Also, even the largest initial roughness disc settled down to almost the same level of roughness as other smaller initial roughness discs after 8×10^4 repeated cycles. Figure 26.9 shows the variation in roughness of test discs nos 1–3 with repeated cycles and the comparison among them. Test disc no. 3 with slip ratio 0.2 % showed the largest variation of roughness and settled down earlier than test discs with no slip. The tendency for initial roughness to settle down obtained in this study is roughly in agreement with the findings obtained so far in practical tests carried out on tracks.

Influence of initial roughness on crystal axis density

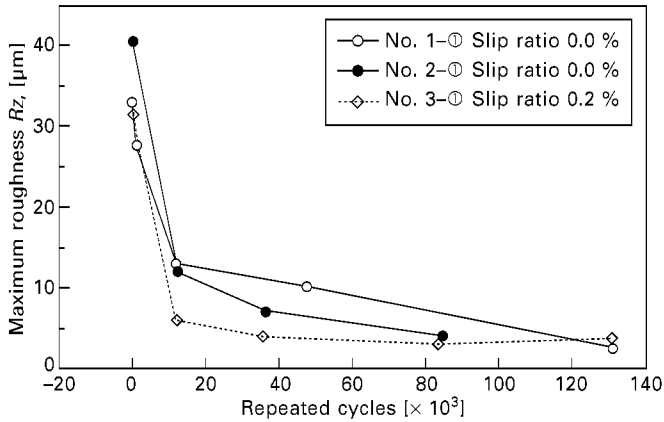
Crystal axis density was measured using the inverse pole figure method of X-ray diffraction[87] to investigate the influence of initial roughness on RCF damage. The inverse pole figure method of X-ray diffraction involves measuring degree of crystal order, whose factor is called the axis density. This factor is then expected to evaluate the degree of rolling fatigue.[88] In the measurements, M_0 is used as the origin of X-ray with 60 kV and 200 mA and the target crystal axis was 222.[88] At first, the inverse pole figure was measured at a depth of 10 microns from the contact surface. After that, the measurements were continued up to a depth of 2 mm from the contact surface using electro-solution polishing. Figure 26.10 shows the measured results for crystal axis 222 using test disc no. 1 without slip and test disc no. 3 with slip. In the figure, axis density indicates the degree of order of some particular crystallographic plane. This is because axis density at a depth sufficient that repeated loadings have almost no influence on it or that found using the new test discs is defined as 1. Also, in the figures, the axis density value in the very thin layer is larger than 1, which means the order of crystallographic plane is slightly recognized. However, the difference between the disc with slip and the one without slip was not clearly identified. On the other hand, in the measured results from rails actually ground in the track, the axis density of the thin layer of rail surface, where 26 MGT of accumulated passing tonnage was applied after grinding, was almost equal to 1, which suggests the grinding effect on removing RCF damage. From the findings obtained from actual rail samples and the results measured in this study, the repeated



26.8 Variation of roughness with repeated cycles on test discs: (a) test disc no. 1; (b) test disc no. 2; (c) test disc no. 3.

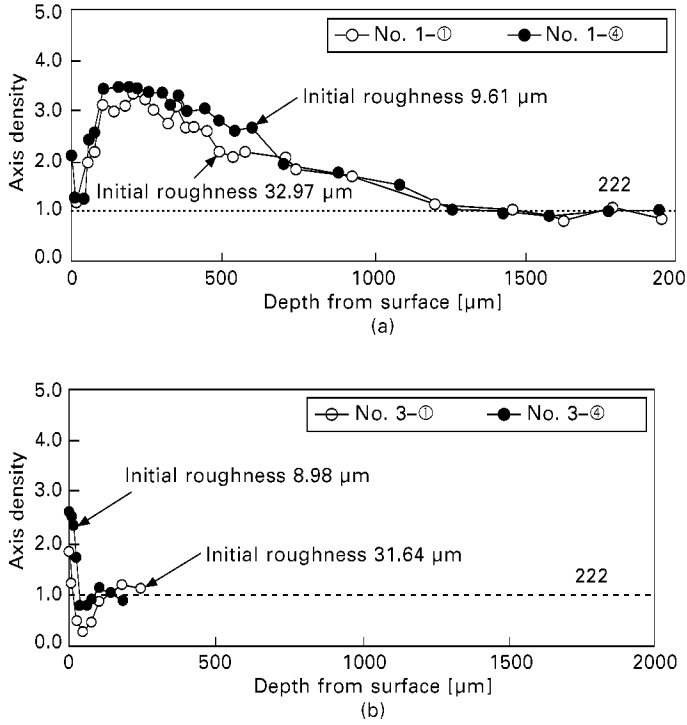
Table 26.1 Experimental arrangements of rolling tests focusing on decreasing initial surface roughness

Test discs	Roughness Rz [μm]	Contact pressure [MPa]	Rolling speed [km/h]	Slip ratio [%]	Repeated cycles	Passing tonnage (equivalent) [MGT]	Temperature & humidity [°C], [%]
No. 1– ①	32.97	750	30	0	136 731	1.37	24, < 30
– ②	28.25						
– ③	16.33						
– ④	9.61						
No. 2 – ①	35.46	750	30	0	84 463	0.84	25, < 25
– ②	35.56						
– ③	10.02						
– ④	7.54						
No. 3 – ①	31.64	750	30	0.2	138 335	1.38	23, < 32
– ②	31.56						
– ③	9.92						
– ④	8.98						



26.9 Comparison of variation of roughness on test discs.

loading cycles corresponded to 1.4 MGT of accumulated passing tonnage after grinding, which may suggest that the roughness of test discs decreases more due to repeated loadings, and a slightly damaged surface layer may be removed to the same level of actual rail samples. Further research on this issue is expected. However, the possibility that initial surface roughness, grinding mark, formed by current rail grinding work has a lesser degree of influence on RCF damage was identified from the results of this study and the achievements of an earlier study based an actual rail samples.



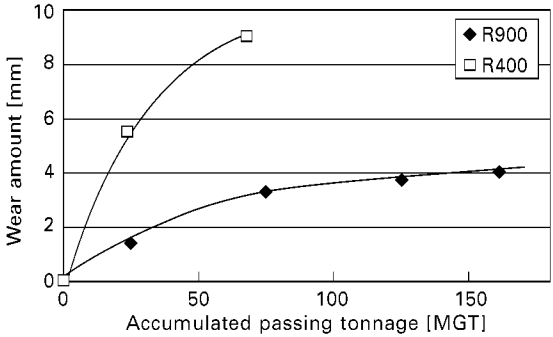
26.10 Variation of axis density in depth from surface with test discs no. 1 and no. 3: (a) comparison of no. 1 (1) and no. 1 (4); (b) comparison of no. 3 (1) and no. 3–(4).

26.3 Wear

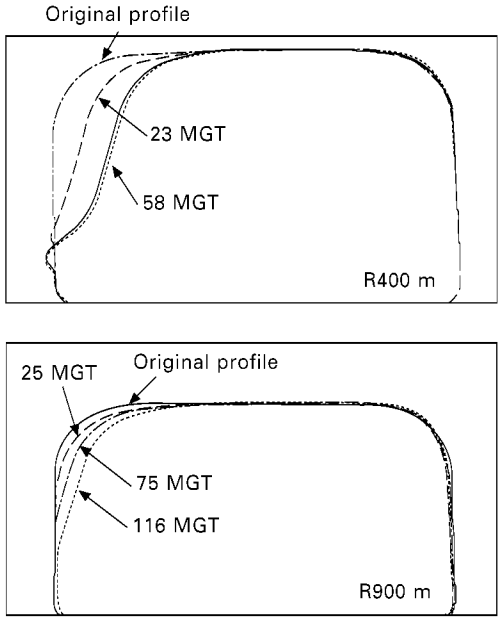
26.3.1 Worn profiles of wheel and rail

Gauge face wear of rail

The progress of gauge face wear was assessed at two Shinkansen track sites where the radii of the curved tracks were 400 m and 900 m. Figure 26.11 shows the progress of wear at the gauge face of rail installed in sharp curved tracks whose radii were 400 m and 900 m. In this figure, the wear progress, measured as the distance between a worn profile and an original profile at the gauge corner of 45 (shown in Fig. 26.12) track of 400 m radius curve (R400m), is roughly three times as much as that in 900 m radius curve (R900m). Figure 26.12 shows the variation of worn profiles at R400m and R900m compared with the original design profile of JIS 60 kg rail. Although wheel profiles must be considered, the trend of the progress of gauge face depending on the severity of curve negotiation can be identified. Also, Figs



26.11 Progress of wear at the gauge face of rail installed in sharp curved tracks.

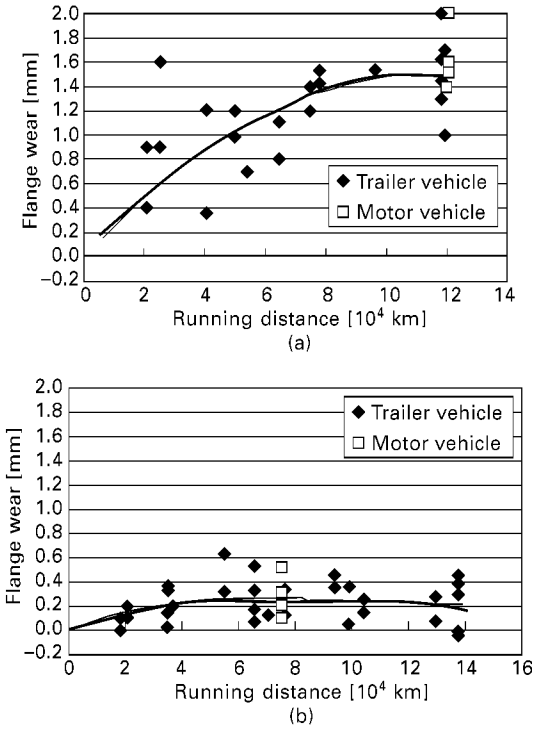


26.12 Variation of worn profiles at R400m and R900m in comparison with the original design profile of JIS 60 kg rail.

26.11 and 26.12 show that the wear progress rate at R900m curve typically comes down after 75 MGT.

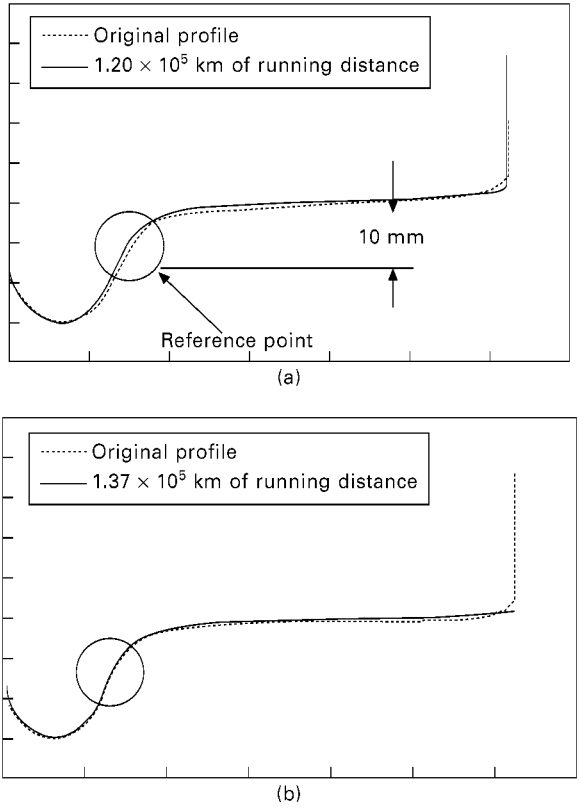
Thin flange wear of wheel

Figure 26.13 shows the progress of wear of wheel flange installed in narrow-gauge vehicles 89. In Fig. 26.13a, the amount of wheel flange wear of a conic

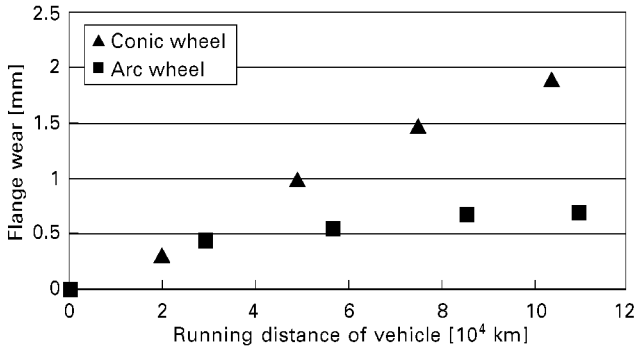


26.13 Progress of wear of wheel flange installed in narrow gauge vehicles: (a) conic wheel profile; (b) modified arc wheel profile.

profile wheel increases according to the increase in running distance reaching 1–2 mm depth from the original profile at a reference point located 10 mm down from the surface of the wheel tread which is shown in Fig. 26.14a at a running distance of 1.2×10^5 km. The wear amount of an arc profile wheel at the flange, reference point as shown in Fig. 26.14a, is smaller than that of the conic profile wheel. In this figure, the difference in the amount of wear between a trailer and a motor car is not clearly identified. Figure 26.14 shows examples of worn wheel profiles of Japanese narrow, gauge vehicles, one of each profile type shown in Fig. 26.13. The wear pattern and amount for both conic profile wheels and arc profile wheels are shown in this figure. Figure 26.15 shows the measured results of flange wear for Shinkansen wheels.[89] The reference point of measurements is the same as the point shown in Fig. 26.14a. The wear amount of the conic profile wheels is roughly three times as much as that of the arc profile wheels. Comparing with the wear amount shown in Fig. 26.13, it can be seen that the progress of flange wear of Shinkansen wheels has the same trend with respect to the difference between conic profile and arc profile as that of narrow-gauge vehicle wheels.



26.14 Examples of worn wheel profiles of Japanese narrow gauge vehicles: (a) conic wheel profile; (b) modified arc wheel profile.



26.15 Measured results of flange wear for Shinkansen wheels.

26.3.2 Laboratory simulation of the wear of wheel and rail

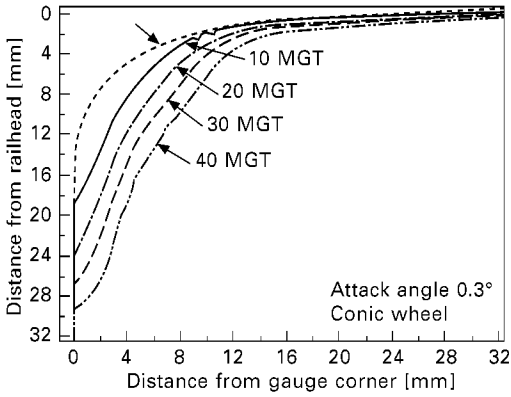
Wear simulation experiments

Laboratory simulation of flange wear and gauge face wear was carried out using a large twin-disc testing machine, which was developed to study RCF and wear phenomena caused by wheel–rail contact. Both wheel and rail specimens are made of the same materials as actual wheels and rails adopted in Japan. The initial contact geometry between rail and wheel specimens can be arranged to be the same as that of one of the typical Japanese railway systems. To simulate actual contact conditions between wheel and rail, a three-dimensional contact model has been developed for estimating contact stress of the actual wheel–rail interface. Linear elastic finite element (FE) analysis was carried out to estimate the contact stress between φ 860 mm conic and arc profile wheels and JIS 60 kg rail for Shinkansen, Japanese high-speed rail. In the laboratory simulation, the applied loads, a constant radial load and some thrust loads, were adopted based on the maximum compressive contact stress, which is almost equal to that of the actual wheel–rail contact in the Shinkansen. A constant radial load corresponding to the wheel load of 85 kN, which is the standard design load of track structures and components in the Shinkansen, was chosen. Also, thrust loads corresponding to lateral forces of 34 kN and 17 kN were estimated as the average of lateral forces measured at Shinkansen track sites of R400m and R900m (referring to Fig. 26.11), and those loads are equal to the design lateral forces similar to the wheel load of 85 kN. In addition, the attack angles of 0° and 0.3° were set to estimate the effect of slip ratio on wear at the contact point between wheel flange and rail gauge corner. The attack angle of 0.3° was also adopted as the average measured at Shinkansen track sites of R400m and R900m (referring to Fig. 26.12).

Results of laboratory simulation

The worn profiles and wear amount of rail and wheel

Accumulated passing tonnage is commonly used instead of the number of loading cycles as a parameter of the magnitude of repeated loading. The accumulated passing tonnage is calculated with a wheel load corresponding to a radial load established by experiments. Figure 26.16 shows an example of variation between a rail disc profile with which wear simulation was carried out together with an actual rail installed in Shinkansen track with the increase of accumulated passing tonnage. In this figure, the tendency of the worn rail disc profile looks similar to that of actual rail. In fact, the experimental arrangements may be similar to the actual contact conditions between wheel



26.16 Example of variation of a rail disc profile.

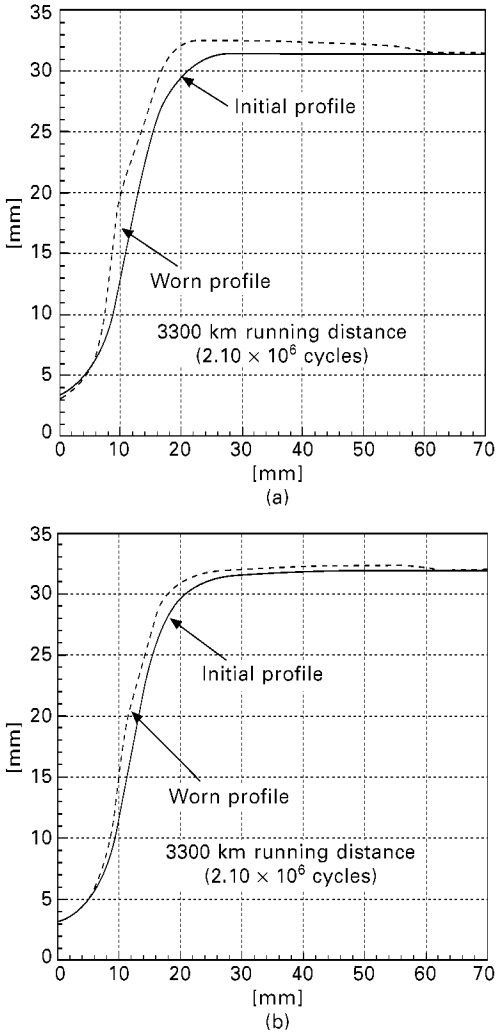
and rail in a certain curved track. Also, it is found that the wear at the initial stage mainly occurs just at the gauge corner, then gradually develops in the gauge face. Figure 26.17 shows the worn profiles of Shinkansen wheels at a running distance of 3300 km (2.10×10^6 cycles) and two different initial profiles of the conic and arc type. In this figure, the difference between worn profiles at the wheel flange depending on the design/initial profiles of wheels can be identified. As far as the test results are concerned, the arc profile wheel has better performance than the conic profile wheel with the focus placed on wear amount.

The effects of attack angle and lateral force on wear

Figure 26.18 shows the variation of wear of wheel disc and rail disc with some experimental arrangements. In this figure, the wear amount of the gauge face was measured at the gauge corner of 45° and that of the wheel flange was measured at the reference point shown in Fig. 26.13a. Considering the definition of wear amount for wheel flange and rail gauge face, comparing flange wear amount and gauge face wear amount makes no sense. Wear amount on both wheel disc and rail disc increases according to the increase of attack angle and lateral force, and the wear amount of the arc profile wheel is smaller on the whole than that of the conic profile wheel.

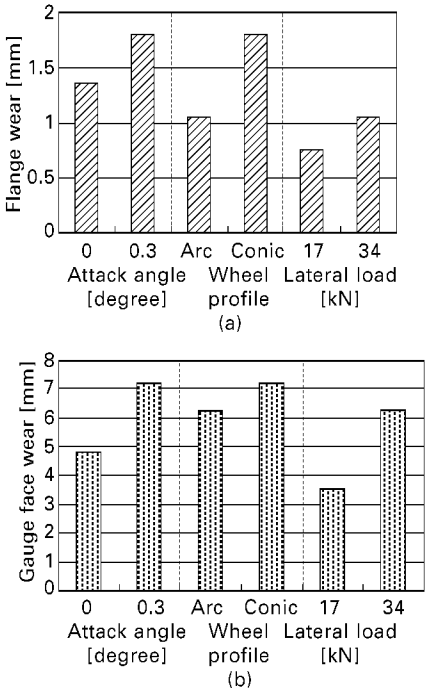
26.3.3 Influence of wear on vehicle–track interaction

The effect of dynamic vehicle–track interaction, with respect to lateral force, attack angle, etc., on the amount of wear is very important to understanding the mechanism of wear between wheel and rail. The dynamic measurements were carried out at three track sites. Two track sites were selected as



26.17 Worn profiles of Shinkansen wheels: (a) conic profile; (b) arc profile.

representatives of severe wear condition with no lubrication on Shinkansen sharp curved tracks whose radii of curve were 400 m and 900 m. The other track site was selected as a representative of sharp curved narrow-gauge tracks with lubrication. Although lubrication has been carried out on the high rail to prevent gauge face wear, its effect on vehicle–track interaction has not been widely reported. Some reports on the effect of low rail lubrication in preventing rail corrugations have only recently been published.[50,53,61,62] On the other hand, the effect of lubrication on flange climb derailment at low speed has recently been focused on.[47]



26.18 Variation of wear amount with some experimental arrangements of laboratory simulation (the profile of rail disc: JIS60 kg): (a) flange wear amount at 1.68×10^6 cycles corresponding to 2638 km of running distance; (b) gauge face wear at 2.4×10^6 cycles corresponding to 40 MGT.

Test arrangements

Dynamic measurements were carried out at two Shinkansen track sites of R400m and R900m.[49] The wheel load, lateral force and attack angle of a leading axle were measured. The cant deficiency at the R400m track measurement site is about 25 mm and that at the R900m track is about 20 mm, which means the cant deficiency of R400m is almost the same as that of R900m. The dynamic measurements in R400m track were performed just after rail renewal at an accumulated passing tonnage of 0.5 MGT as the initial stage of wear and then 23 MGT. Also, the dynamic measurements in R900m track were carried out just after rail renewal at a passing tonnage of 1.4 MGT as the initial stage and then that 25 MGT, 75 MGT and 161 MGT.

Test results

Profiles of worn rails at dynamic measurements

The progress of wear amount on the gauge face of rail installed in sharp curved tracks of R400m and R900m is shown in Fig. 26.11. Also, the

progress of worn profile of R400m track and R900m track compared with the original profile is shown in Fig. 26.12. These two figures are helpful in understanding the relationship between the timing of dynamic measurements and the worn profiles of the rail.

Lateral force at high rails

Figure 26.19 shows the lateral forces of initial and worn high rails caused by leading axles of some types of cars in R400m and R900m tracks. The lateral forces of the high rails in R400m track are almost twice as much as those in R900m track because of the severity of curve negotiation between the two curves. In this figure, the top, bottom and middle lines represent the maximum, minimum and average lines, respectively. The lateral forces of worn rails in R400m track tend to decrease as wear progresses. Also, the same tendency or lateral forces as in R400m track can be recognised in R900m track. However, lateral forces at the third measurements were larger than those at the second and fourth measurements in R900m track. The reason is that the coefficient of friction between wheel and rail at the third measurements may have increased more than that at the second and fourth measurements, depending on the contact surface of the wheel–rail interface.[49]

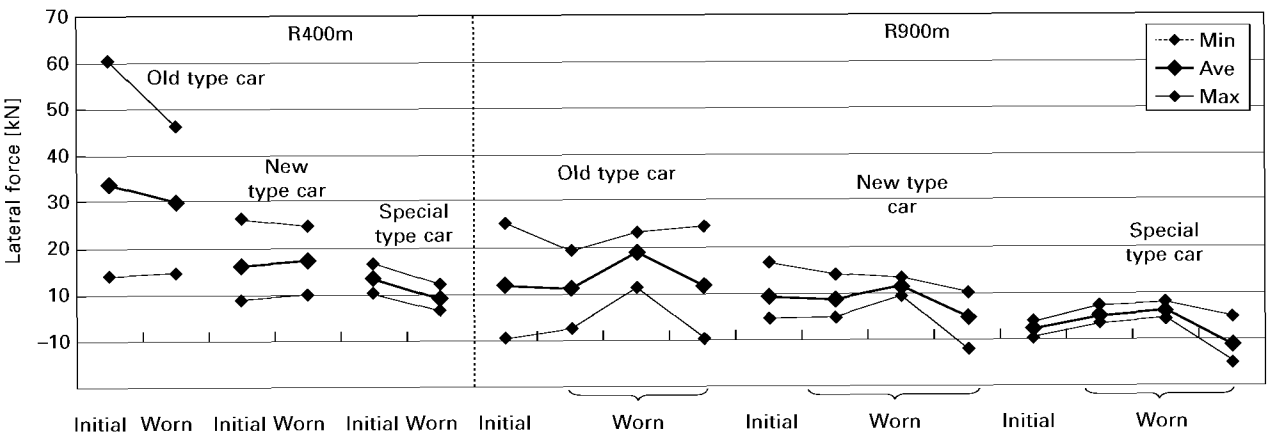
Attack angle of leading axles

Figure 26.20 shows the attack angles of initial and worn high rails caused by the leading axles at R400m and R900m tracks. The attack angle of worn rails at R900m track was measured at the initial stage and just after the accumulated passing tonnage of 75 MGT and 161 MGT. In this figure, the attack angle of worn rail is smaller than that of initial rails irrespective of vehicle types at both curved tracks. Also, the attack angles of R400m track were similar to those of R900m track. When the lateral forces shown in Fig. 26.19 were considered, the difference of high rail wear amount between R400m track and R900m track investigated here could have been caused mainly by lateral forces.

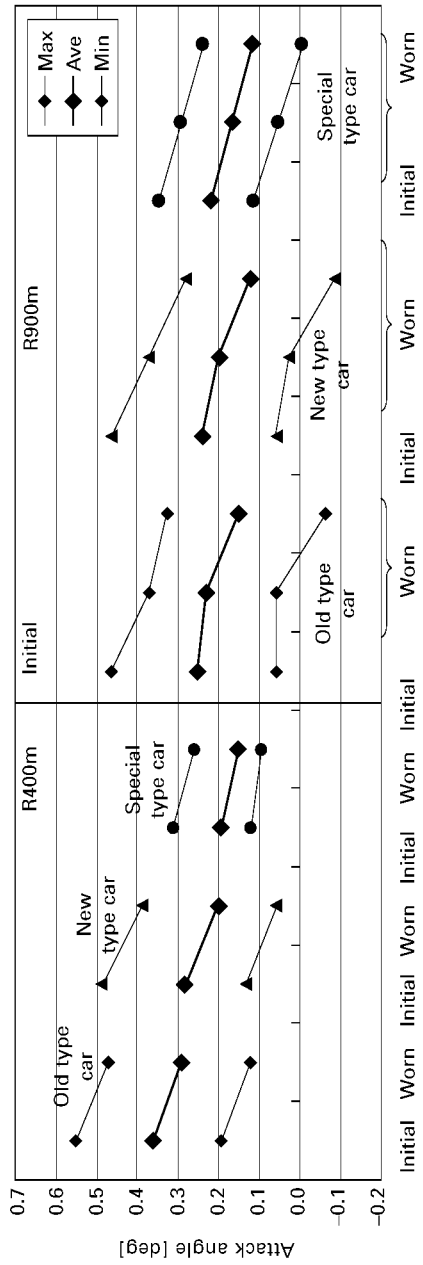
26.4 Corrugation

26.4.1 Short pitch corrugations in curved tracks

Short pitch corrugations on low rails shown in Fig. 26.21 pose acute problems of noise and track deterioration caused by vibrations excited by interaction between rail and wheel forces.



26.19 Lateral forces of initial and worn high rails caused by leading axles of some types of cars at R400m and R900m tracks.



26.20 Attack angles of initial and worn high rails caused by the leading axles at R400m and R900m tracks.

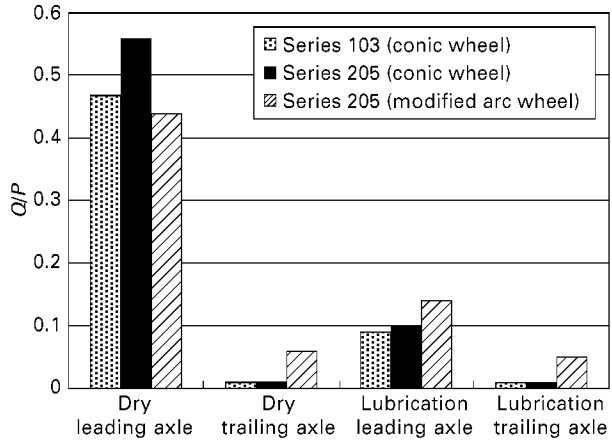
26.21 Typical short-pitch corrugations formed on low rail surface.

Influential factors in causing corrugations

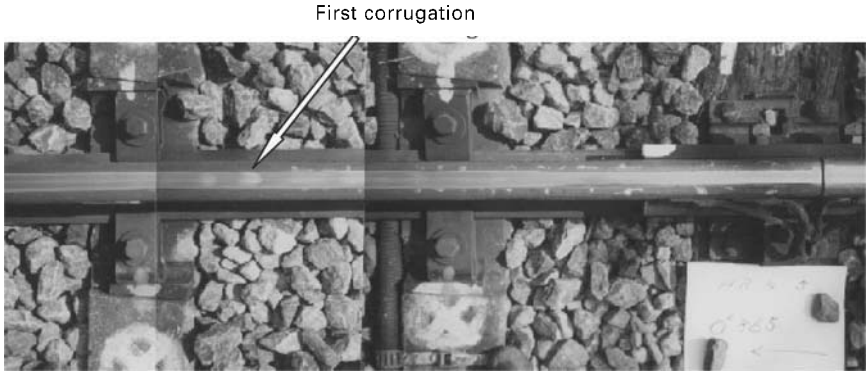
Tassily and Vincent[57] developed a theoretical model of corrugations formed on low rails and high rails at curved tracks in the Paris Rapid and Mass Transport Authority (RATP). They considered: quasi-static curving behaviour; dynamic interactions between the wheelset and the track; field tests. The frequency ranges of corrugations studied were 50–80 Hz and 250–400 Hz. They showed the importance of the relationship between transverse wear rate and roughness, and they found two sensitive frequencies at about 60 and 170 Hz. The bogie, vehicle and track structures considered in their analysis are not the same as those used in Japan, but the two sensitive frequencies of 60 Hz and 170 Hz are very similar to high rail corrugations and low rail corrugations formed in Japanese railways. This means that the transverse wear rate may play an important role in the formation of corrugations.

A question arose: which force, lateral creep force or longitudinal creep force, has more influence on short pitch corrugations? If the lateral creep force is more influential, then dynamic behaviour of a leading axle rather than a trailing axle of a bogie should be investigated. Vehicle running tests and dynamic measurements at a track site were carried out to understand the effect of lateral force and lateral vibration of rail on the formation of corrugations.[50] Also, the effect of lubrication on decreasing the lateral force

and preventing corrugation formation was studied. Figure 26.22 shows Q/P (lateral force/vertical force) of the low rail measured with strain gauges stuck the at rail base flange for lateral force Q and at the rail web for vertical force P during the vehicle running tests. In the figure, the Q/P of the leading axle was most significant and the effect of lubrication on reducing lateral forces was also important. In addition, Fig. 26.23 shows the first formation of corrugations on a new rail, which is located near the second tie from the rail joint in the direction of vehicle running. It is not clearly understood that the coupled vibration between rail joint and wheel gives rise first to corrugations roughly near the second tie from the rail joint, but almost the same phenomena were very often observed in investigations. It is surely clear from findings obtained so far that excitation at the rail joint is one of the major factors in the formation of corrugations.

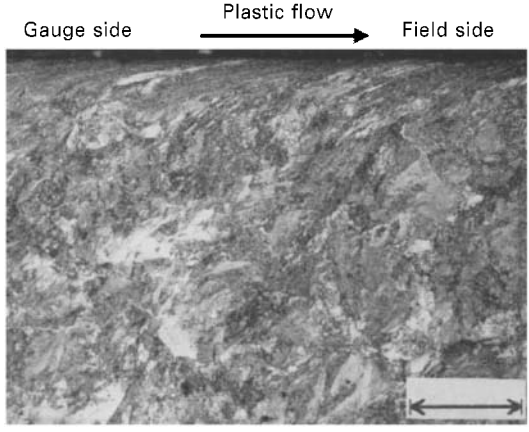


26.22 Q/P (lateral force/vertical force) of low rail measured at track site.

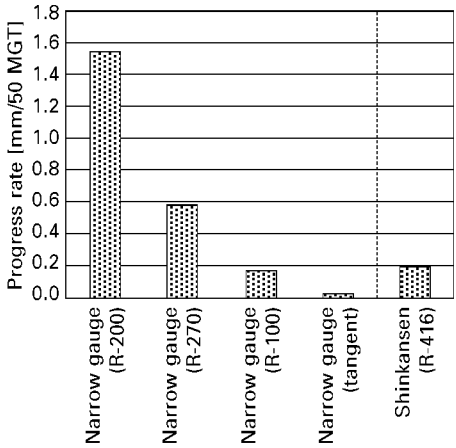


26.23 The first low rail corrugations formed on a new rail.

In light of the above discussions, investigation of longitudinal and lateral plastic flow on the surface of some corrugations was carried out.[50] There is little evidence of plastic flow in the longitudinal direction. Most of the plastic flow in Fig. 26.24 is in the lateral direction. Since plastic flow is greatly influenced by the friction force, plastic flow in the lateral direction indicates the importance of lateral creep forces. The direction of plastic flow (from gauge corner to field corner of the rail) corresponds to the direction of the lateral force on the rail applied by the wheel. Field investigation shows that lateral creep force and/or longitudinal creep force and the excitation at rail joints have a great influence on the formation of corrugations. Also, lateral creep force has greater influence than longitudinal creep force on plastic flow on the surface of corrugations. The excitation of vertical force between rail and wheel at rail joints causes the variation of friction force, estimated from the coefficient of friction multiplied by a resultant of longitudinal and lateral creep forces between wheel and rail. Then, the stick–slip phenomenon can be supposed to be induced between a resultant of mainly lateral creep force and longitudinal creep force, and the variation of friction force caused mainly by the vertical excitation at rail joints. This shows that the lateral creep force has a great influence on the formation of low rail corrugations, and it may indicate that the behaviour of the leading axle of the bogie should play a more important role than that of the trailing axle. Figure 26.25 shows the variation in progress rate of the corrugations with the radius of curvature. The progress rate is roughly inversely proportional to the radius of curvature. Considering that both lateral creep force caused by a leading axle and longitudinal creep force caused by a trailing axle increase together with the decrease of radius of curvatures, it is not clear which axle has more influence on the formation of corrugations. However, it is not impossible that lateral forces influence the development of corrugation.



26.24 Microstructure of rail surface layer in lateral direction.



26.25 Variation of progress rate of the corrugations with the radius of curvature.

Causes of corrugations

Experimental investigations and microscopic observation indicate that:

- lateral creep forces of a leading axle play an important role in low rail corrugations;
- rail joints can be the main trigger for exciting the fluctuation of Q/P;
- the fluctuation of Q/P may lead to stick–slip between wheel and rail, and wear takes place to form corrugations.

The fluctuation of Q/P based on the variation in vertical force P excited at rail joints and the resonance of rail tilting vibration may be the main factors in the formation of low rail corrugations, However, this proposition cannot be proved experimentally because the interval between strain gauges installed at a rail (usually 200 mm) is too short to allow the fluctuation in Q/P to be measured *in situ*. Moreover, because signals with frequencies of more than 100 Hz are filtered due to the resonance of the wheel structure, such measurement cannot be achieved using strain gauges installed at a wheel either. To understand theoretically the mechanism of low rail corrugations, further study which considers dynamic vehicle–track interaction, the curving performance of the vehicle and the lateral stiffness of the track is required.

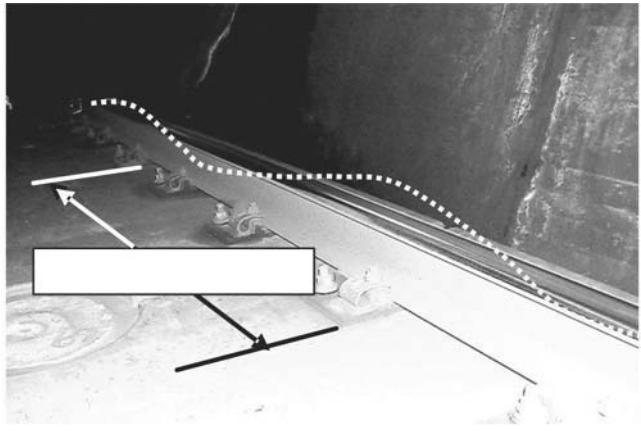
Lubrication

Lubrication between wheel and rail has been adopted to reduce (i) squeal, (ii) wheel flange vertical wear and (iii) rail gauge face wear. Lubrication is used

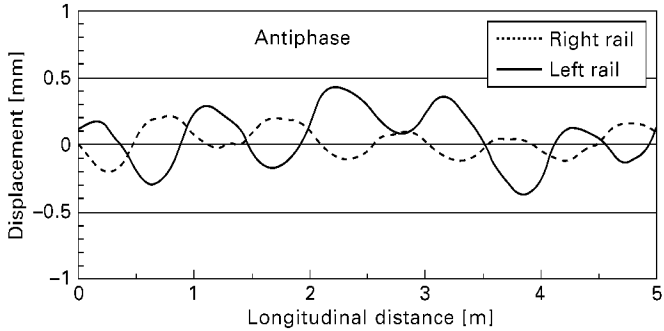
to prevent low rail corrugations at sharp curves. To investigate the effect of lubrication on preventing corrugations, the normal lubricant was applied on the crown of low rails at a sharp curve with a radius of curvature of about 300 m. Investigations were carried out up to about 200 MGT, and it was found that no corrugations formed.[50] It has been suggested that lubrication on the crown of low rails has a significant effect in reducing wheel flange climb derailment.[47] If the lateral creep force decreases due to lubrication, then the lateral force of the high rail, usually called flange reaction, will decrease as well. This means that wheel flange vertical wear and rail gauge face wear should also be reduced. On the other hand, lubrication of the rail crown carries some risk of wheel sliding. Such significant advantages and risks need to be balanced by choosing a lubricant with an appropriate friction coefficient.

26.4.2 Long pitch corrugation in the tangent track in a salty-environment tunnel

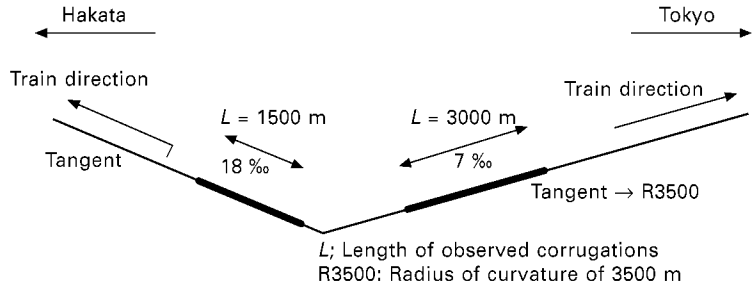
Figure 26.26 shows rail corrugations formed on slab track in the ascending slope of a salty, environment tunnel, called Shin-Kanmon Channel tunnel, of the San-yo Shinkansen line.[51,52] Since these corrugations are causing noise and large interacting forces between wheels and rails, a lot of maintenance work, such as frequent rail renewal and rail grinding, has been carried out to deal with those problems. Figure 26.27 shows a typical photo of the waveform and its phase lag for corrugations between left and right rails. In this figure, the wavelength is about 1.2 m and its frequency calculated by running speed is about 70 Hz, which is almost the same as the resonance



26.26 Rail corrugations formed in the ascending slope of Shin-Kanmon Channel tunnel (salt water environment tunnel).



26.27 Typical example of the waveforms and its phase lag of rail corrugations.



26.28 Location of rail corrugations observed in ascending slopes of Shin-Kanmon Channel tunnel.

frequency of the unsprung mass of a vehicle supported by track stiffness. Also, there are two possibilities for the phase lag between the right rail and left rail corrugations. One is the same phase of both rails and the other is the phase lag of 180° between right rail and left rail. Considering those facts, the resonance of the unsprung mass of the vehicle and the rolling frictional vibration of the wheelset may have a significant influence on the mechanism of rail corrugations.

Figure 26.28 shows the location of rail corrugations observed in ascending slopes of the tunnel. In this figure, corrugations are generated only in ascending slopes and not in descending slopes, and in almost tangent track, which suggests that the corrugations may be caused by driving wheels rather than following wheels. According to the findings obtained from Figs 26.27 and 26.28, the following possible mechanism for the formation of corrugations can be considered. The maximum traction force, or maximum traction coefficient, which is estimated by wheel load multiplied by adhesion coefficient, and is almost equal to the kinetic COF in normal dry conditions, changes in proportion to wheel load variation. On the other hand, wheels drive on rails with some rolling inertia. Rolling inertia does not always follow

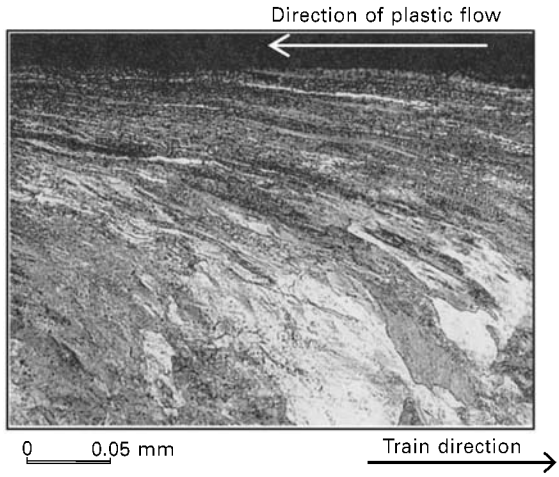
Previous Page
Copyrighted Materials
Copyright © 2009 Woodhead Publishing Retrieved from www.knovel.com

the variation in maximum traction coefficient, which basically depends on the surface conditions of wheel and which induce a longitudinal roll–slip phenomenon along rails. As a result, such a roll–slip phenomenon may cause rail corrugations. In fact, driving force can exceed the maximum traction force, depending on wheel load variation and kinetic COF, and wheels can slip on rails very slightly.

To verify the above-mentioned rail corrugation mechanism, it was desirable that the roll–slip phenomenon of the wheel be directly confirmed from the variation in wheel rotation; however, this was not easy at the time because the precision with which the rotational axle speed could be measured was not enough to identify the phenomenon of about 70 Hz. Then, investigating the possibility of roll–slip, the kinetic COF of the rail was measured, focusing on rust on the rail surface of due to the atmosphere in the tunnel under the sea, to as certain the appropriately low COF which may induce the roll–slip phenomenon.

Plastic flow on corrugated rails

To confirm longitudinal roll–slip caused by wheel load variation, the longitudinal plastic flow of the rail was investigated. Figure 26.29 shows longitudinal plastic flow on the rail surface layer. In this figure, longitudinal plastic flow is clearly observed. On the other hand, lateral plastic flow was not observed. This means that the large traction force or driving force necessary for a vehicle to ascend a slope may have a significant role in generating plastic flow and on the potential of the roll–slip phenomenon to cause such corrugations.



26.29 Microstructure of rail surface layer in longitudinal direction.

Factors influencing the maximum traction coefficient

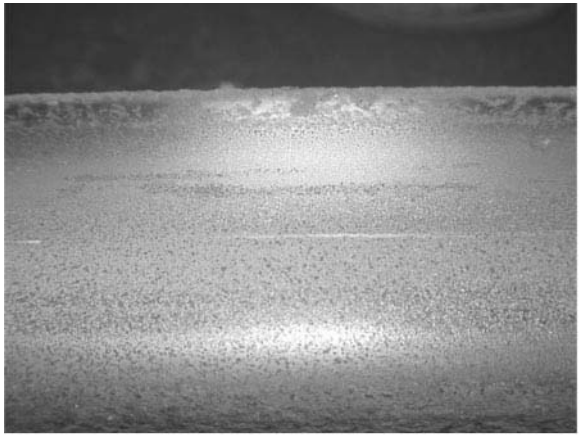
In order to evaluate the maximum traction coefficient, the kinetic COF on the top surface of the rail was measured. The measurements were carried out in both an open section and a tunnel section. Table 26.2 gives results of the measurements. Generally speaking, the measured data for kinetic COF in the tunnel are not too small. although it is not clear that magnitude of COF or maximum traction coefficient is necessary for Shinkansen trains to go up such ascending slopes, the adhesion problem of wheel sliding does not at least take place. Because the measured data should not, on the whole, be too small, even in the case of such ascending slopes, the potential for roll–slip depends on whether or not the COF of around 0.4 given in Table 26.2 is sufficient to give rise to the phenomenon.

Next, Fig. 26.30 shows the surface condition of the rail crown. In this figure, almost the whole area of the running surface of the rail looks very rusty. This surface may suggest that the COF cannot be evaluated properly because of rust, which means that the measuring apparatus for COF adopted in this study may not be suitable. Basically, COF depends on the measuring apparatus. In particular, the RTRI tribometer measuring apparatus for COF may not be suitable in this case (reference 52). The sliding resistance of the steel ball in the case of a normal surface may be smaller than that in the case of a thick, rusty surface or a smooth surface. Accordingly, the measured COF may not be a suitable parameter with which to evaluate the possibility of sliding. In view of this situation, the surface substance or oxidized layer of the rail crown was the focus of investigation into the potential for the roll–slip

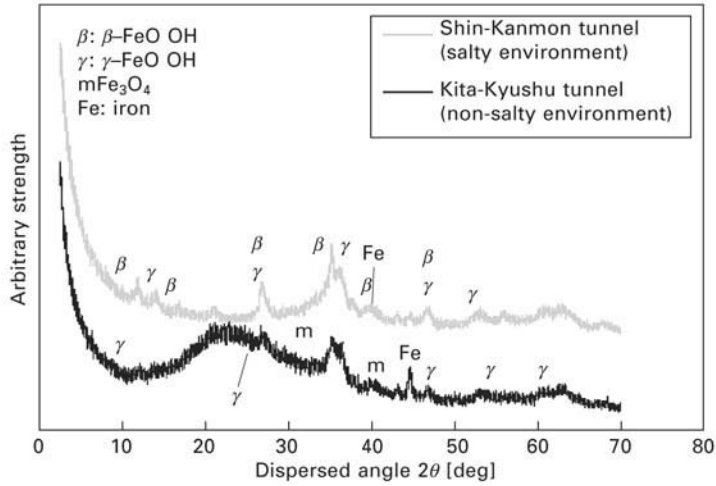
Table 26.2 Kinetic COF on top surface of rail

Place of measurements (distance from Tokyo)	Train direction of track	Kinetic COF		Atmospheric temperature [%]	Rail temperature [%]	Humidity [%]	Corrugation
		Left*[c]	Right*[d]				
1012k000m	Up[a]	0.44	0.48	19.4	18.3	86.4	No
1012k000m	Down[b]	0.39	0.4	19.4	18.5	86	No
998k600m	Up	0.39	0.38	23.1	22.5	82	No
998k600m	Down	0.4	0.4	23.2	22.1	83.1	Observed
997k000m	Up	0.44	0.46	22.5	21	79.2	No
997k000m	Down	0.35	0.49	21.4	21.1	78	No
995k700m	Up	0.42	0.4	20.5	19.3	85.6	Observed

[a] Up train direction bound for Tokyo.
[b] Down train direction bound for Hakata.
[c] Left: left rail is defined from the direction facing Hakata.
[d] Right: right rail is the other side of left rail.



26.30 State of top surface of rail.



26.31 Results of X-ray diffraction analysis.

phenomenon. At first, Fourier transform infrared absorption spectroscopy was revised to analyse the surface of rail installed in track *in situ*, but the analytical results achieved were not clear. Then sample rails were put beside track in a salty-environment tunnel and a dry mountain tunnel which was selected as a no-corrugation control to compare with the salty-environment tunnel. The sample rails were exposed in those two tunnels for six months to acquire enough rust for analysis. X-ray diffraction analysis was carried out to investigate the surface substances of the sample rails. Figure 26.31 compares the analytical results with X-ray diffraction for the two tunnels. One is Shin-Kanmon tunnel which is a channel tunnel and has a salty environment.

The other is Kita-Kyushu tunnel which is a mountain tunnel and has a non-salty environment. In this figure, some iron oxides and oxyhydroxides were identified, but β-FeOOH was found only in the case of the salty-environment tunnel. Since β-FeOOH is a material of low COF, the roll–slip phenomenon can be expected only in the salty-environment tunnel. In this study, *in situ* analysis has not been successful yet, but laboratory analysis, particularly X-ray diffraction, played a very important role in recognizing the potential for the roll–slip phenomenon to lead to the formation of corrugation.

Formation mechanism for rail corrugations in the Shin-Kanmon tunnel with salty-environment: considering the above-results, the factors involved in generating rail corrugations are considered to be as follows:

- A large traction force is generated by driving wheels to climb up the ascending slopes of the salty-environment tunnel.
- Wheel load variation is excited at the irregularity of rail welds. The phase of wheel load variation in the wheel on one side is almost the reverse of that in the wheel on the other.
- The maximum traction coefficient between wheel and rail is not large enough because of the small amount of β-FeOOH on the running surface of the rail only in the salty-environment tunnel.
- The roll–slip phenomenon may be induced based on the relationship between the driving force of the wheel and the maximum traction force, depending on the wheel load variation excited at the rail welds.
- Wheel load variation is excited more by the formation of corrugations, and the wear carried by the corrugation is progressed by the wheel load variation.

Roughly speaking, rail corrugations described here are considered to form under a combination of:

- insufficient maximum traction coefficient value caused by rust forming in the salty environment of the under sea tunnel, the variation in value depending on the wheel load variation excited at the rail welds; and
- the rolling inertia of driving wheels as they climb up the ascending slopes in tunnel.

26.5 Adhesion

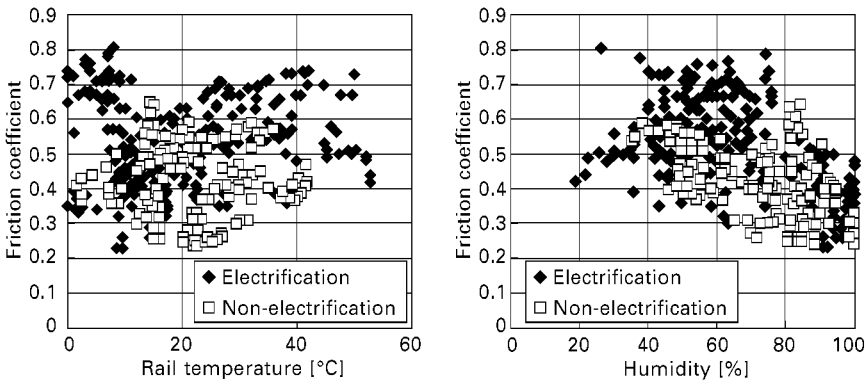
26.5.1 Friction coefficient of wheel–rail interface

Adhesion between rail and wheel is one of the most important parameters in a steel wheel on steel rail system. Basically, the purpose of studying adhesion has been so far to obtain an adhesion coefficient as high as possible for driving and braking, particularly in wet conditions. Some materials to improve the adhesion between wheel and rail have been developed in high

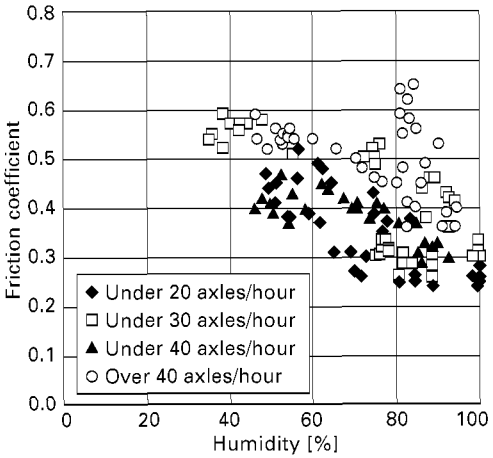
speed railways.[80] In contrast, lubrication is adopted to reduce friction power at the contact surface between wheel flange and rail gauge face in curves. Currently at lubrication top of low rail has been focused on reducing lateral forces and squeals excited by vehicles negotiating sharp curves. On the other hand, a terrible derailment and collision accident happened in an urban metro line, Tokyo, on the morning on 8 March, 2000.[47] Since then, much effort has been made to understand the mechanism of wheel flange climb derailment.[90] Many helpful findings have been obtained, but the friction coefficient between wheel and rail is considered to be very difficult to predict and the important question about what is the most influential factor on friction coefficient still remains unanswered.[80]

Measured results and discussions of friction coefficient

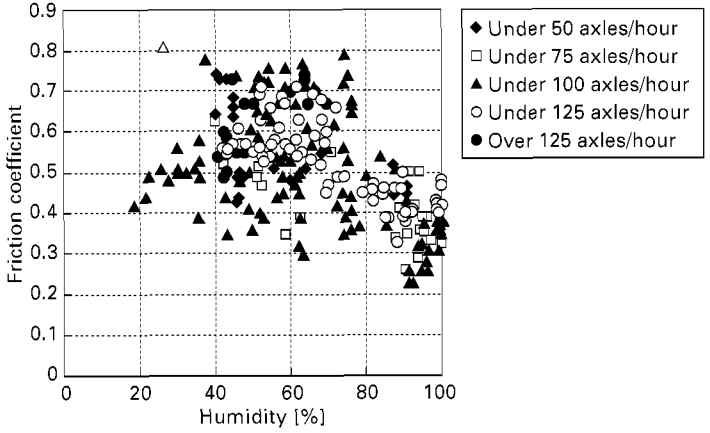
Figure 26.32 shows the measured results that include various data measured at 20 track sites in four seasons.[80] Then, the friction coefficients had large variation of 0.2–0.8 because the variation in rail temperature was 0–50 °C and that in humidity was 20–80 %. Also, the influence of electrification and non-electrification on the friction coefficient can be identified. It is roughly estimated that the data obtained in electrified lines are slightly larger than those in non-electrified lines. The reason is that oil mist or drop due to diesel engines gives on oily atmosphere or environment along the track. Also, the humidity may have better correlation than rail temperature with the friction coefficient. Focusing on the influence of operational frequency of trains, the data were divided into four ranks of number of passing axles per hour in Fig. 26.33 and five ranks in Fig. 26.34. The difference between four ranks looks significant in non-electrified lines in Fig. 26.33. However,



26.32 Influence of electrification and non-electrification on friction coefficient.



26.33 Influence of passing axle number on friction coefficient (non-electrified line).

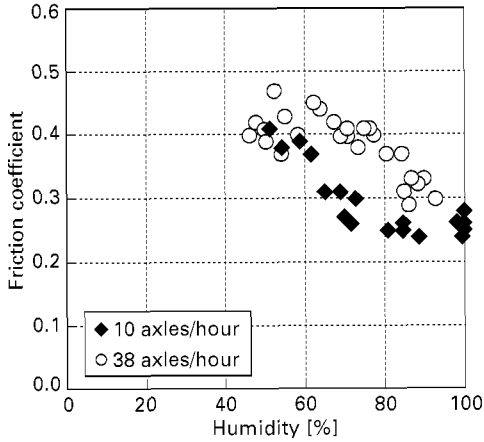


26.34 Influence of passing axle number on friction coefficient (electrified line).

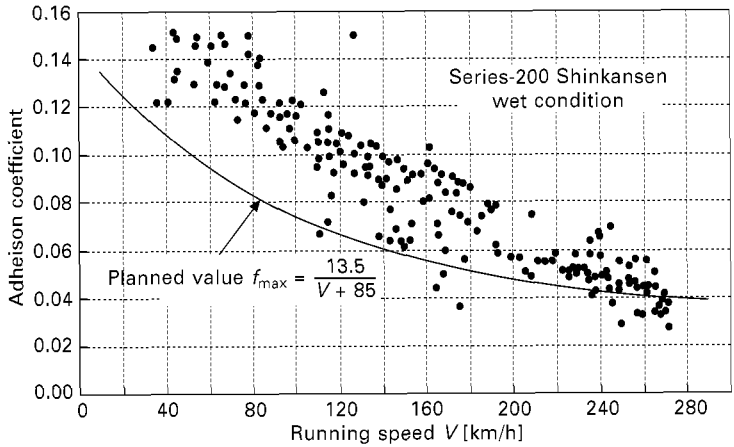
the influence of passing axle number per hour is not clearly identified in Fig. 26.34. Then, detailed analysis of the influence of passing axle frequency on friction coefficient was performed in the same revenue line. The difference between the two ranks is significant in Fig. 26.35.

26.5.2 Adhesion between wheel and rail under wet conditions

It is well-known that the adhesion coefficient between wheel and rail decreases with an increase of running speed under wet conditions caused



26.35 Influence of passing axle number on friction coefficient (selected non-electrified line).



26.36 Some measured data and the adhesion coefficient design curve adopted by Japanese railways.

by rain or snow, which has an influence on the running stability of vehicles, particularly the performance of driving and braking in high-speed railways. Figure 26.36 shows some measured data and the adhesion coefficient design curve adopted by Japanese railways.[74] In order to clarify the mechanism of the adhesion coefficient between wheel and rail under wet conditions, Ohyama performed laboratory tests with a high-speed rolling contact machine at a rolling speed up to 300 km/h with water sprayed onto the contact area between two rolling discs.[72–74] Chen *et al.* adopted a line contact (2-D) model and a point contact (3-D) model, which consist of EHL theory and the rough surface contact theory, to estimate the adhesion coefficient between

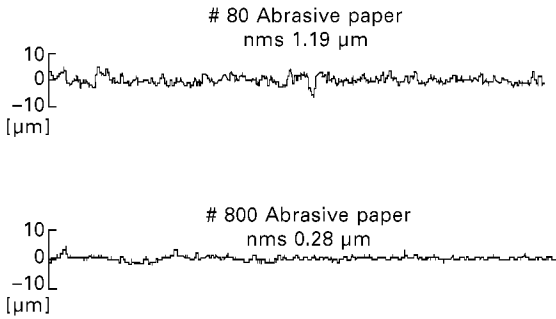
wheel and rail under wet conditions.[75,77,78] Then, some experiments were carried out focusing particularly on the water temperature and the surface roughness, using a twin-disc rolling contact machine to verify the adequacy of the theoretical analysis.

Experimental results and discussions

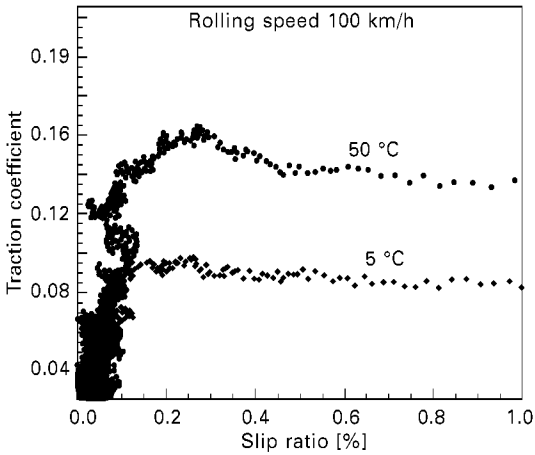
Effect of water temperature on the adhesion coefficient

The water temperatures were set at 5 °C (termed cold water) and 50 °C (termed hot water), based on an actual rail temperature in the winter and summer. The magnitude of the radial load was set by referring to the appropriate contact pressure between wheel and rail in Japanese Railways, with the focus placed on the Hertzian maximum pressure. Before the experiments, the surfaces of the wheel and rail discs were finished with abrasive paper # 80 and # 800, respectively. Figure 26.37 shows a sample of the surface roughness profiles of the rail disc finished by the above-mentioned abrasive papers. In this study, the surface roughness of a root mean square (rms), i.e. standard deviation of roughness, was adopted and measured by means of a stylus profiler along the axial direction of the rail disc.

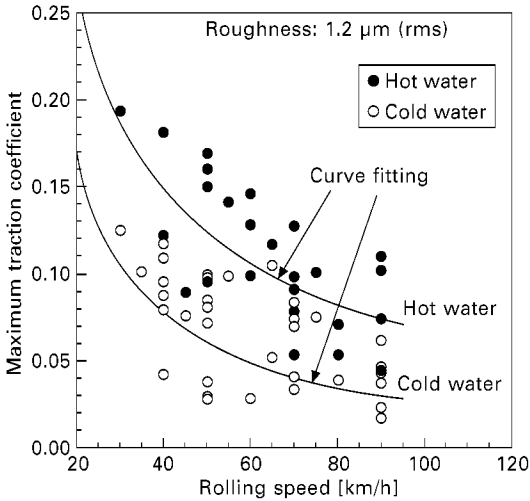
Figure 26.38 shows an example of the experimental results relating to the relationship between slip ratio and traction coefficient with the water temperature at 5 °C and 50 °C and a rolling speed of 100 km/h. It can be identified that the maximum traction coefficient (i.e. adhesion coefficient) at a high water temperature is larger than that at a low water temperature. Figure 26.39 shows the relationship between the maximum traction coefficient (i.e. adhesion coefficient) and rolling speed at the two different water temperatures of 5 °C and 50 °C. In the figure, black circles show all the values obtained in the experiments at 50 °C while white circles show all those at water 5 °C; moreover, solid lines show the curve fitted lines of the maximum traction coefficients at the two different water temperatures. Although the maximum



26.37 Surface roughness of the rail disc.



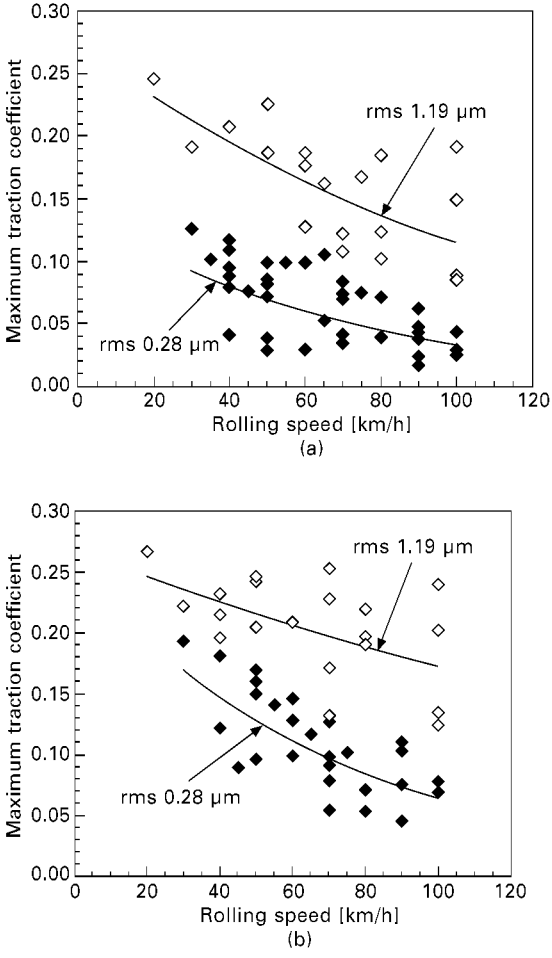
26.38 Relationship between traction coefficient and slip ratio at different water temperature.



26.39 Relationship between maximum traction coefficient and rolling speed at different water temperatures.

traction coefficients in the figure vary under the same contact conditions, such as rolling speed, axle load, surface roughness and water temperature, it is recognisable that the water temperature has a significant effect on the adhesion coefficient and a rise in the water temperature causes an increase in the adhesion coefficient. Figure 26.39 also shows that the rolling speed has a great effect on the maximum traction coefficient, regardless of the water temperature.

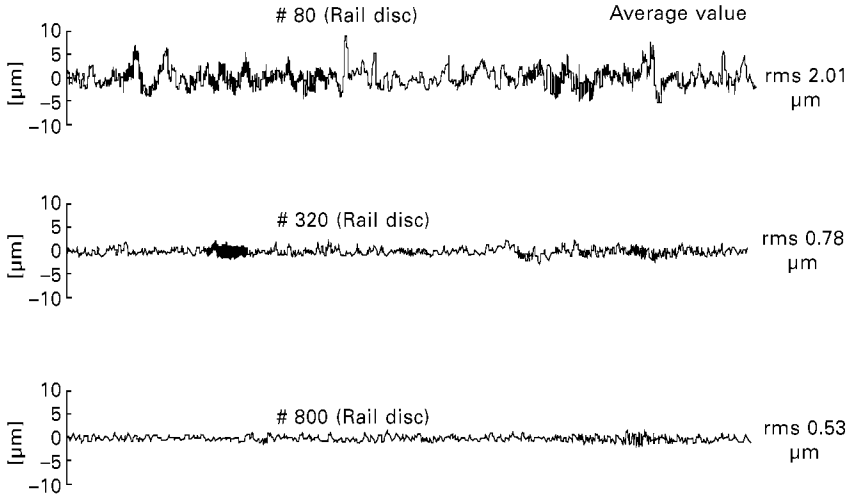
Figure 26.40 shows the relationships between the maximum traction coefficient (i.e. adhesion coefficient) and rolling speed for two kinds of surface roughness (rms.: 1.19 μm and 0.28 μm) under cold and hot water spray, respectively. It is clearly identified that the maximum traction coefficient for larger surface roughness is greater than that one for smaller surface roughness under the same contact conditions of rolling speed, axle load and water temperature, and the rolling speed has a great effect on the maximum traction coefficient in each surface roughness condition.



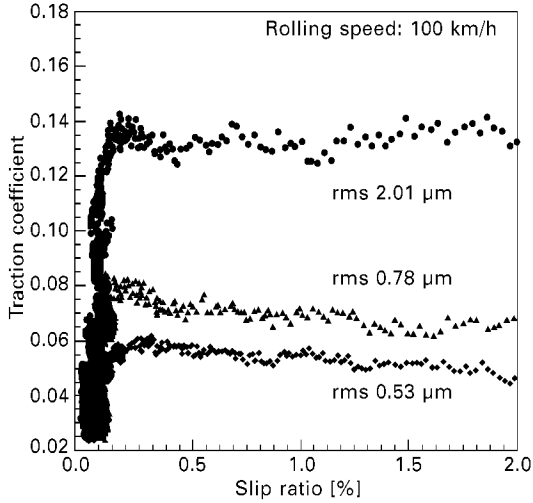
26.40 Relationship between the maximum traction coefficient and rolling speed for two kinds of surface roughness under cold and hot water conditions: (a) cold water (5 °C); (b) hot water (50 °C).

Effect of surface roughness on the adhesion coefficient (maximum traction coefficient)

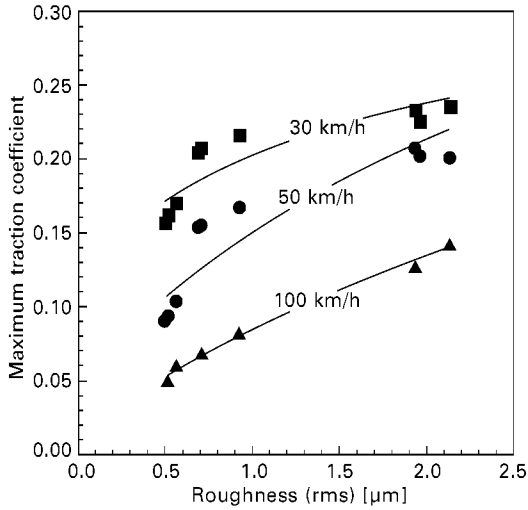
In the experiments, the surfaces of the wheel and rail discs were finished with abrasive paper # 80, # 320 and # 800, respectively, and the roughness configurations of the wheel and rail discs were arranged longitudinally, oriented to the rolling direction of the discs. Figure 26.41 shows a sample of the surface roughness profiles of the rail disc finished by the above-mentioned abrasive papers. Surface roughness (rms: standard deviation of asperity height) shown in the figure was the average of data obtained at six places on the rail disc along its axial direction using a stylus profiler. Figure 26.42 shows an example of the experimental results for the relationship between traction coefficient and slip ratio, under a rolling speed up to 100 km/h with three kinds of surface roughness profiles. It is clearly identified that the maximum traction coefficient in the case of the larger surface roughness is greater than that in the case of the smaller one. Figure 26.43 shows the maximum traction coefficients obtained under various rolling speed and surface roughness conditions in all the experiments. The maximum traction coefficient increases with an increase in surface roughness and, in contrast, decreases with an increase in rolling speed. Consequently, the effect of surface roughness on adhesion coefficient under wet conditions is significant.



26.41 Surface roughness profiles of the rail disc before the experiments.



26.42 Relationship between slip ratio and traction coefficient for three kinds of surface roughness.



26.43 Relationship between the maximum traction coefficient and surface roughness.

26.6 Lubrication

26.6.1 Effect of lubrication on vehicle–track interaction

Vehicle running tests focusing on the effect of lubrication on vehicle–track interaction were carried out on an RTRI test track whose radius of curvature

is 160 m. Wheel and rail dynamic behaviour excited by a vehicle negotiating a sharp curve was measured in the three lubrication arrangements detailed below. For the purposes of the measurements, a kind of grease which is commonly applied to the wheel–rail interface from a wayside system on Japanese railways was adopted.[62] Table 26.3 gives the technical details of track and lubrication arrangements. A lubricant was applied either to the top of the low rail or to the gauge face of the high rail for the distance of 55 m under 3.5 g/m. In addition, the kinetic COF was measured to monitor and evaluate the lubrication arrangements. The COF was measured on the top of the low rail and at the gauge face of the high rail under the lubrication arrangements set out in Table 26.3 just after the passing of the test vehicle. Table 26.4 shows the measured results for the COF on the top of the low rail and at the gauge face of the high rail depending on the lubrication arrangements given in Table 26.3. In this table, the effect of low and high rail lubrication in decreasing COF is roughly presented. Although the COF at the gauge face of the high rail is not influenced by the lubrication on the top of the low rail, the COF tends to depend to a large extent on changes in meteorological conditions and in the oxidized surface layer which is destroyed by large stresses due to contact loading when a vehicle passes. Therefore, it bears out the importance of identifying the COF on the top of the low rail and at the gauge face of the high rail in order to accurately evaluate vehicle–track interaction.

Vehicle dynamic simulation was carried out with a VDS. The COF in the simulation was set as 0.5 for no lubrication and 0.15 for lubrication based on the measured results given in Table 26.4. Moreover, an additional

Table 26.3 Technical details of track and lubrication arrangements

Radius	Rail	Cant		Lubrication arrangements
160 m	JIS 50N	90 mm	Case 1	No lubrication
			Case 2	Grease lubrication on the top of a low rail
			Case 3	Grease lubrication at the gauge face of a high rail
			Case 4[a]	Lubrication on both the gauge face and top of a high rail

[a] Lubrication arrangement not for vehicle running tests but only for vehicle dynamic simulation which is discussed later.

Table 26.4 Measured results of COF

Measurement point	Top of low rail	Gauge face of high rail
No lubrication	0.55	0.47
Top of low rail lubrication	0.20	0.60
Gauge face of high rail lubrication	0.43	0.15

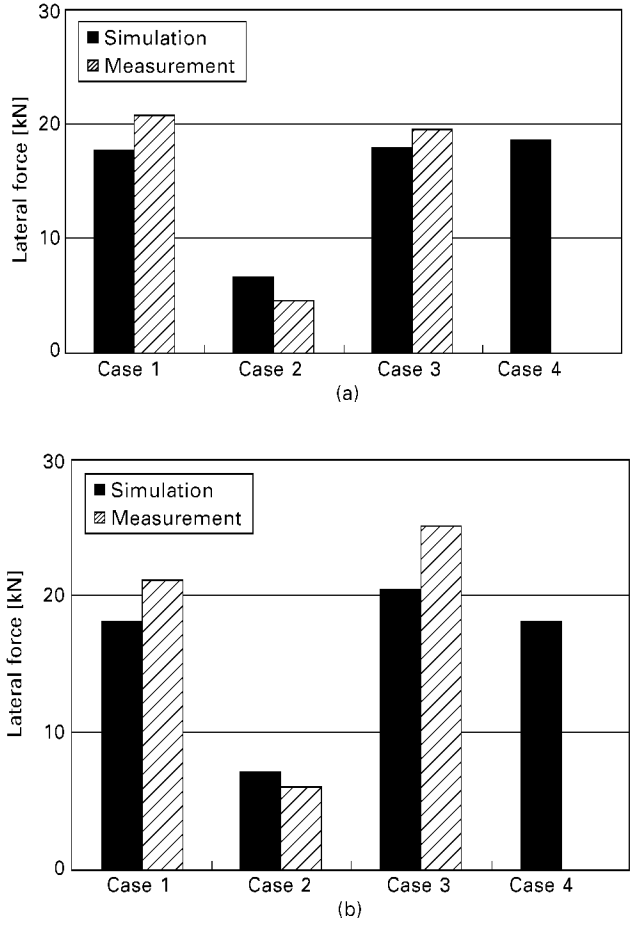
lubrication arrangement in which the top of the high rail was lubricated simultaneously with the gauge face of the high rail, hereinafter called top-gauge face 'high rail lubrication' for short, was analysed as case 4 given in Table 26.3. The reason why this top–gauge face high rail lubrication was considered in addition to the three lubrication arrangements for vehicle running tests is that the wheel of the leading axle normally makes contact with both the top and the gauge face of the high rail but that of the trailing axle normally makes contact only with the top of the high rail, which may cause the difference in vehicle dynamic behaviour between gauge face only high rail lubrication and top-gauge face high rail lubrication. It is expected that the effect of creep force of the trailing axle at the high rail side on the dynamic behaviour of a bogie will be identified.

Lateral force and the angle of attack of a leading axle

Figures 26.44 and 26.45 show the lateral force and the angle of attack of a leading axle analysed here together with measured results of vehicle running tests. There was some small difference in lateral force between the analytical results and the measured results in the three lubrication arrangements. However, the tendency and relation of the three lubrication arrangements between the simulations and the measurements are roughly in agreement. Also, the orders of magnitude of the angle of attack in the three lubrication arrangements obtained by the measurements are almost the same as those obtained by the simulations. In addition, comparing the analytical results in the case of gauge face only high rail lubrication with those of top-gauge face high rail lubrication, the lateral force of the low rail and the angle of attack in the case of high rail lubrication are slightly larger than those of the gauge face only high rail lubrication. Also, the lateral force of the leading axle in the case of top-gauge face high rail lubrication is about 2 kN smaller than that of gauge face only high rail lubrication.

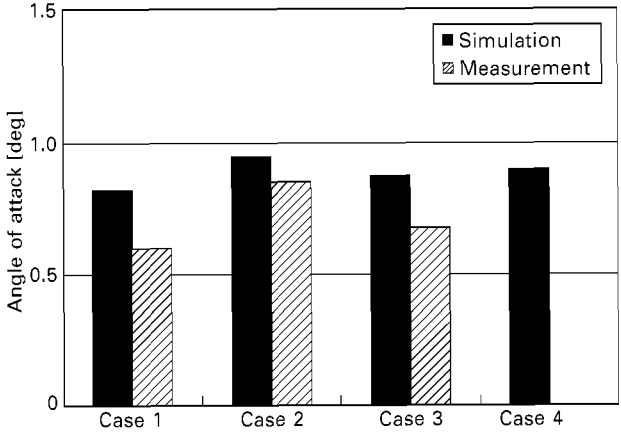
Interacting forces between wheel and rail when the vehicle is negotiating sharp curves

Figure 26.46 shows the four resultant forces of creep force and normal force analysed for four wheel–rail contacts of a leading bogie in the lateral and longitudinal directions except wheel loads. In the simulations, only the lateral force of the leading axle at the high rail side was the resultant of creep force between wheel tread and top of rail and the normal force between wheel flange and rail gauge face, which is usually called the flange reaction force; the other three lateral forces were creep forces which means no contact between wheel flange and the gauge face of the rail. In this figure, the lateral creep force at the low rail side in case 2 decreased so much that the normal

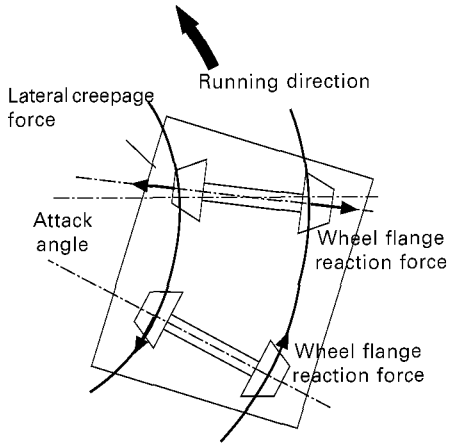


26.44 Lateral forces interacting between wheel and rail in a leading axle: (a) low rail side; (b) high rail side.

force between wheel flange and the gauge face of the high rail decreased in comparison with no lubrication. Then, as a result, the longitudinal creep force of the leading axle at the high rail side was also smaller than that in the case of no lubrication. Considering the balance of the interacting forces between the four wheel–rail interfaces in a bogie, the reason why the angles of attack of a leading axle in the three cases of lubrication were larger than in the care of no lubrication can be considered to be that the steering force based on the difference of longitudinal creep forces between the low rail and the high rail side in the cases of lubrication was smaller than that in the case of no lubrication.



26.45 Angle of attack of a leading axle.



26.46 Typical interacting forces from wheel to rail in a bogie negotiating sharp curves.

26.6.2 Friction moderating system

Friction moderator

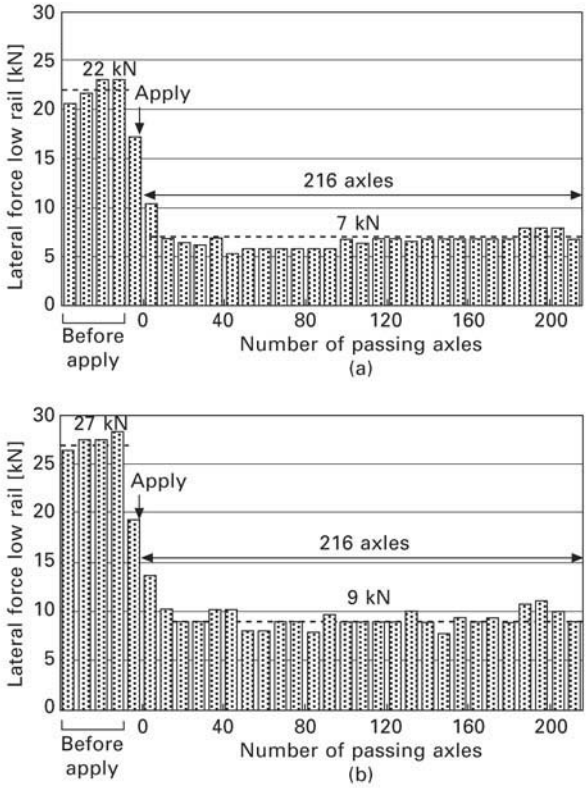
The effect of lubrication on wheel–rail dynamic behaviour was studied using vehicle running tests and vehicle dynamic simulations. As a result, low rail lubrication on the top of the rail was identified as having the effect of greatly reducing lateral force, which is one of the main factors in operating safety, material integrity and environmental issues. However, the traction characteristics of the lubricant adopted for low rail lubrication on the top of the rail are expected to have a positive slope at large slip ratio because

sliding between wheel and rail must be avoided. Then, friction modifier became the focus of research and was developed as the lubricant adopted for use at the top of the low rail in sharp curves.[53,54,62] On the other hand, RTRI developed the solid lubricant whose purpose is the same as friction modifier. The solid lubricant whose raw material is artificial graphite was crushed to a diameter of 0.2 mm and appropriately arranged about its graphitisation to obtain the required COF. The COF of graphite can be roughly controlled by the degree of graphitisation. Also, phenol resin was selected as the material coating the grain with a view to of preventing it from catching fire on braking because phenol resin is used for the composite brakeshoe. The material being developed is called friction moderator.

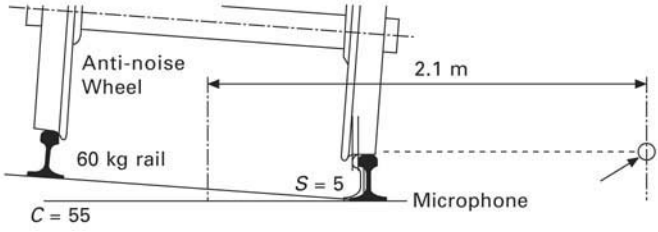
Effect of friction moderator on decreasing lateral force and squealing noise

Vehicle running tests were carried out to identify the effect of the friction moderator on the lateral forces on a sharp curve with a radius of 197 m (referred to below as curve (A)). Figure 26.47 shows the lateral forces of the low rail and high rail as measured at the track site. In the figure, approximately 22 kN of lateral force on the low rail before application of the friction moderator decreased to approximately 7 kN after application to the wheel–rail interface on the low rail side using the jetting device of the friction-moderating system that was installed on the test vehicle. The lateral force on the low rail was decreased by 68 % due to the effect of the friction moderator. Similarly, approximately 27 kN of lateral force on the high rail before application of the friction moderator decreased to approximately 9 kN after application, representing a reduction of approximately 66 % in the lateral force on the high rail. In addition, the effect of the friction moderator on reducing the lateral force continued up to at least 216 axles.

A microphone was set up at a location shown in Fig. 26.48 with trains passing in front of it to measure sound pressure levels. The microphone was placed on a sharp curve with a radius of 165 m (referred to below as curve (B)), and the measured results are shown in Fig. 26.49. In this figure, the sound pressure level of approximately 87 dB decreased to approximately 77 dB up to 168 passing axles from the first application of the friction moderator. After that, the sound pressure level gradually increased to more than 85 dB with each passing train up, but decreased again from the second application of the friction moderator. The effect of the second application on reducing the sound pressure level was identifiable as approximately 77 dB for 248 passing axles up to finishing the tests and starting the onboard lubrication system.



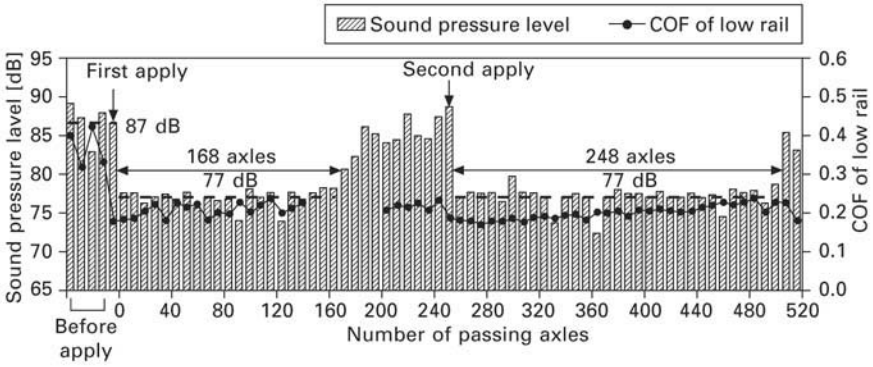
26.47 Effect of friction moderator on decreasing lateral force on curve (A): (a) low rail side; (b) high rail side.



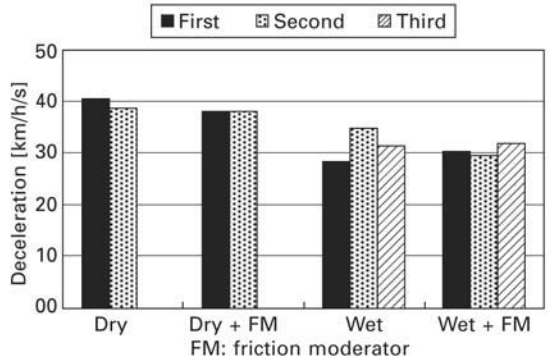
26.48 Location of microphone set up at the track site to measure sound power emitted by train running. [C = cant or super-elevation, s = slack or gauge widening]

Influence of friction moderator on braking distance

A commuter-type vehicle was installed with the jetting friction moderator for the wheel–rail interface on a commercial railway line. The test track was a straight section with an ascending slope of 12.5 %. The friction moderator was applied at 0.052 g/m to the wheel–rail interface of one rail side (not both

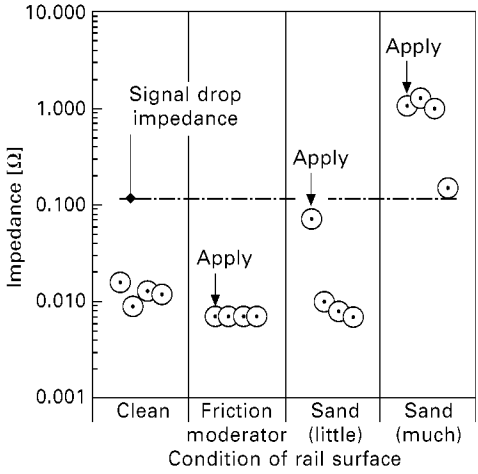


26.49 Effect of friction moderator on reduction of sound pressure level and COF of low rail on curve (B).



26.50 Results of brake test.

sides) from the location of initial braking to the estimated point of vehicle stop plus an additional distance equivalent to the length of a train set. In the tests, we applied the maximum braking force under normal operation from a speed of around 60 km/h until the train stopped. Figure 26.50 shows the braking test results under the test conditions of dry, dry with friction moderator, wet and wet with friction moderator. The test results were revised based on the effect of the slope on the braking force. In the figure, some variation of deceleration in each case is evident, but deceleration values (i.e. the velocity decrease rates) in the test cases with the friction moderator were almost the same as those without, so the influence of the friction moderator on the deceleration braking distance cannot be significant. In addition, since the COF of surfaces made wet by rain is commonly smaller than that of dry rail surfaces, the friction moderator is not applied to the wheel/rail interface in rains. An integrated automatic wiper control system means that when it rains most drivers will use the wiper installed at the front glass of the vehicle.



26.51 Shunting performance of friction moderator and sand.

Influence of friction moderator on track circuit

The track circuit which includes right/left rails, wheelsets shunting the two rails and insulation joints is usually adopted in Japanese railway systems. Shunting malfunction is therefore a very important issue in railway safety, and the materials present on the wheel–rail interface have great influence here. For the friction moderator coated with phenol resin, the raw material (carbon) shows conductivity, but the phenol resin merely shows insulation. Figure 26.51 shows the test results for shunting performance [62]. A large amount of sand showed large impedance, which would represent a problem with shunting. In addition, even a small amount of sand showed relatively large impedance just after being placed on the rail surface, after which its impedance decreased. Friction moderator showed almost the same level of impedance as a clean, dry rail surface, which suggests that the risk of shunting problem due to friction moderator will be the same as with a clean and/or dry surface rail.

26.7 References

1. Hertz, H. (1896), On the contact of elastic solids, in Jones, D.E. and Scholt, G.A. (eds), *Miscellaneous Papers by H. Hertz*, Macmillan, London, UK.
2. Way, S. (1935), Pitting due to rolling contact, *ASME Journal of Applied Mechanics*, **2**, A49–A58.
3. Johnson, K.L. (1985), *Contact Mechanics*, Cambridge University Press, Cambridge, UK.
4. Keer, L.M. and Bryant, M.D. (1983), A pitting model for rolling contact fatigue, *Journal of Lubrication Technology, Trans ASME*, **105**, 198–205.

5. Kaneta, M., Yatsuzuka, H. and Murakami, Y. (1985), Mechanism of crack growth in lubricated rolling/sliding contact, *ASLE Transactions*, **28**(3), 407–14.
6. Hearle, A.D. and Johnson, K.L. (1985), Mode II stress intensities for a crack parallel to the surface of an elastic half-space subjected to a moving point load, *Journal of the Mechanics and Physics of Solids*, **33**, 61–81.
7. Nishida, S., Sugino, I.K., Urashima, C. and Masumoto, H. (1985), Study on contact rolling fatigue of rails, *Bulletin JSME*, **28–243**, 1819–24.
8. Bower, A.F. (1988), The influence of crack face friction and trapped fluid on surface initiated rolling contact fatigue cracks, *Journal of Tribology ASME*, **110**, 704–11.
9. Inoue, Y., Satoh, Y. and Kashiwaya, K. (1991), Estimation of growth in rolling contact fatigue damage to rail steel with development of texture, *Quarterly Report of Railway Technical Research Institute*, Tokyo, Japan, **32**(1), 35–41.
10. Kaneta, M. and Murakami, Y. (1991), Propagation of semi-elliptical surface cracks in lubricated rolling/sliding elliptical contacts, *Journal of Tribology ASME*, **113**, 270–75.
11. Bold, P.E., Brown, M. W. and Allen, R.J. (1991), Shear crack growth and rolling contact fatigue, *Wear*, **144**, 307–17.
12. Murakami, Y., Sakae, C. and Ichimaru, K. (1994), Three-dimensional fracture mechanics analysis of pit formation mechanism under lubricated rolling-sliding contact loading, *STLE Tribology Transactions*, **37**, 445–54.
13. Beynon, J.H., Kapoor, A. and Tyfour, W. R. (1996), Deterioration of rolling contact fatigue life of pearlitic rail steel due to dry–wet rolling-sliding line contact, *Wear*, **197**, 255–65.
14. Bogdanski, S., Olzak, M. and Stupnicki, J. (1996), The effects of face friction and tractive force on propagation of 3D 'squat' type of rolling contact fatigue crack, in Zobory Z (ed.), *Proceedings 2nd Mini Conference on Contact Mechanics and Wear of Rail/Wheel Systems*, Budapest University of Technology Budapest, Hungary 164–73.
15. Brown, M.W., Hemsworth, S., Wong, S.L. and Allen, R.J. (1996), Rolling contact fatigue crack growth in rail steel, in Zobory, I. (ed.) *Proceedings 2nd Mini Conference on Contact Mechanics and Wear of Rail/Wheel Systems*, Budapest University of Technology Budapest, Hungary, 144–53.
16. Clayton, P. (1996), Tribological aspects of wheel–rail contact: a review of recent experimental research, *Wear*, **191**, 170–83.
17. Grassie, S.L. and Kalousek, J. (1997), Rolling contact fatigue of rails: characteristics, causes and treatments, *Proceedings 6th International Heavy Haul Conference*, Cape Town, South Africa, 7–11 April, 381–404.
18. Murakami, Y., Sakae, C. and Hamada, S. (1999), Mechanism of rolling contact fatigue and measurement of K_{th} for steels, in Beynon, J.H. and Brown, M.W. *et al.* (eds), Engineering Against Fatigue, Taylor and Francis, Rotterdam, the Netherlands, 473–85.
19. Fletcher, D.I. and Beynon, J.H. (1999), A simple method of stress intensity factor calculation for inclined fluid-filled surface-breaking cracks under contact loading, *Proc. IMechE, Part Journal of Engineering Tribology*, **213**, 299–304.
20. Ishida, M. (1989), Statistical analysis of rail shelling on Shinkansen, *Proceedings 4th International Heavy Haul Conference*, Brisbane, Qld, Australia, 11–15 September, 205–9.
21. Ishida, M. (1990), Relationship between rail shellings and track surroundings, *Quarterly Report of RTRI*, 31–1, 22–8.

22. Akama, M. and Mori, T. (2000), Boundary element analysis of surface initiated rolling contact fatigue cracks in wheel/rail contact systems, *Proceedings 5th International Conference on Contact Mechanics and Wear of Rail/Wheel Systems*, Tokyo, Japan, 25–27, July, 161–6.
23. Akama, M. and Susuki, I. (2000), Fatigue crack growth simulation and estimation for rolling contact, *Proceedings 4th International Conference on Fracture and Strength of Solids*, 1035–41.
24. Wong, S.L., Bold, P.E., Brown, M.W. and Allen, R.J. (1996), A branch criterion for shallow angled rolling contact fatigue cracks in rails, *Wear*, **191**, 45–53.
25. Bogdanski, S. and Brown, M.W. (2002), Modelling the three-dimensional behaviour of shallow rolling contact fatigue cracks in rails, *Wear*, **253**, 17–25.
26. Kasiwaya, K. and Ishida, M. (2003), The prediction model of growth rate of rail transverse crack, *Proceedings of Railway Mechanics*, **7**, JSCE, 79–84 (in Japanese).
27. Kalousek, J., Sroba, P. and Hegelund, C. (1989), Analysis of rail grinding tests and implications for corrective and preventative grinding, *Proceedings 4th International Heavy Haul Conference*, Brisbane, Qld, Australia, 11–15 September, 193–204.
28. Linn, S., Abell, D., Kalousek, J. and Sroba, P. (1993), Planning of production rail grinding on the Burlington Northern Railroad, *Proceedings 5th International Heavy Haul Conference*, Beijing, China, June, 316–23.
29. Grassie, S.L. (1997), Requirements for transverse railhead profile and railhead roughness following grinding, *Proceedings 6th International Heavy Haul Conference*, Cape Town, South Africa, 7–11 April, 549–564.
30. Ishida, M. and Abe, N. (1997), Experimental study on the effect of preventive grinding for Shinkansen rails, *Proceedings 6th International Heavy Haul Conference*, Cape Town, South Africa, 7–11 April, 565–75.
31. Ishida, M., Abe, N. and Moto, T. (1998), The effect of preventive grinding on rail surface shellings, *Quarterly Report of RTRI*, 39–3, 136–41.
32. Grassie, S.L. (1999), Appropriate specification of grinding requirements, *Proceedings International Heavy Haul Association Specialist Technical Session on Wheel/Rail Interface*, Moscow, Russia, 14–17 June, 503–10.
33. Tada, Y. (1999), Rail grinding method in Shinkansen, *Sinsenro*, **53** (8), 4–7 (in Japanese).
34. Sawley, K. J. (1999), *North American Rail Grinding: Practices and Effectiveness*, Report R-928 Association of American Railroads, Washington DC, USA.
35. Kapoor, A., Frank, F.J., Wong, S.K. and Ishida, M. (2002), Surface roughness and plastic flow in rail wheel contact, *Wear*, **253**, 257–64.
36. Jin, Y., Ishida, M. and Aoki, F. (2006), Investigation and analysis on rail head checks, RTRI Report, 20–11, 29–34 (in Japanese).
37. Burstow, M.C., Benyon, J., Watson, A.S., Beagle, A.E. and Beagles, M. (2003), Current developments in the whole life rail model to predict rolling contact fatigue in rails, *Proceedings World Congress on Railway Research*, Edinburgh, UK, 28–30 September, 446–51.
38. Steele, R.K. and Devine, T.J. (1982), Wear of rail/wheel systems, in Kalousek, J., Dukkipati, R.V. and Gladwell, G.M.L. (eds), *Proceedings of the Conference on Contact Mechanics and Wear of Rail/wheel Systems*, Vancouver, BC, Canada, July 6–9, University of Waterloo Press, Waterloo, ON, Canada, 293–315.
39. Ohno, K. and Ogawa, Y. (1983), The influence of rust on the adhesion between wheel and rail, *RTRI report* , A-83-70 (in Japanese).

40. Bolton, P.J. and Clayton, P. (1984), Rolling-sliding wear damage in rail and tyre steels, *Wear*, **93**, 145–65.
41. Steel, R. (1990), Overview of the fast HTL/HAL rail performance tests, Proceedings IHHA Workshop on Heavy Axle Loads, Pueblo, CO, USA, 14–17 October.
42. Tyfour, W.R. and Beynon, J.H. (1994), The effect of rolling direction reversal on the wear rate and wear mechanism of pearlitic rail steel, *Tribology International*, **27**(4), 273–82.
43. Clayton, P. (1996), Tribological aspects of wheel–rail contact: review of recent experimental research, *Wear*, **19**, 170–83.
44. Zakharov, S., Komarovsky, I. and Zharov, I. (1998), Wheel flange/rail head simulation, *Wear*, **215**, 18–24.
45. IHHA (2001), *Guidelines to Best Practices for Heavy Haul Operations, Wheel and Rail Interface Issues*, International Heavy Haul Association, Virginia Beach, VA, USA, 3-33–3-40.
46. Komarovsky, I. and Zharov, I. (2001), Influence of hardness on wear resistance of wheel and rail steels for different loading conditions, *Journal of Friction and Wear*, **22**(2), 134–39.
47. Ishida, M. and Nakahara, T. (2001), Derailment accident in Hibiya Line and tribology, *Journal of Japanese Society of Tribologists*, **46**(7), 46–53 (in Japanese).
48. Ishida, M., Takikawa, M. and Jin, Y. (2001), Gauge face wear caused with vehicle/track interaction, *Proceedings World Congress on Railway Research*, Koln, Germany, 25–29 November, on CD.
49. Ishida, M., Jin, Y., Aoki, F. and Takikawa, T. (2004), Influential factors on wear of wheel flange and rail gauge face, Proceedings 14th International Wheelset Congress, Orlando, FL, USA,17–21 October, on CD.
50. Ishida, M., Moto, T. and Takikawa, T. (2002), The effect of lateral creepage force on rail corrugation on low rail at sharp curves, *Wear*, **253**, 172–77.
51. Ishida, M., Aoki, F., Sone, Y., Ban, T. and Shirouzu, K. (2005), Rail corrugations caused by low coefficient of friction in a submarine railway tunnel, *Proceedings of World Tribology Congress*, Washington, DC, USA, 12–16 September, WTC2005-64346.
52. Ishida, M., Aoki, F., Sone, Y., Ban, T. and Shirouzu, K. (2005), Study on the cause of rail corrugations formed by salty environment, *RTRI Report*, 19-9, 11–16 (in Japanese).
53. Eadie, D.T., Kalousek, J. and Chiddick, K.C. (2002), The role of high positive friction (HPF) modifier in the control of short pitch corrugations and related phenomena, *Wear*, **253**, 185–92.
54. Eadie, D.T., Santoro, M. and Kalousek, J. (2003), Railway noise and the effect of top of rail liquid friction modifiers: changes in sound and vibration spectral distributions, *Proceedings 6th International Conference on Contact Mechanics and Wear of Rail/Wheel Systems*, Gothenburg, Sweden, 10–13 June, 503–10.
55. Ishida, M. (2005), Effect of rail grinding and/or ballast tamping on track deterioration caused by rail corrugations, *Proceedings Railway Engineering 2005*, London, UK, 29–30 June, on CD.
56. Frederick, C.O. (1986), A rail corrugation theory, *Proceedings 2nd International Conference on Contact Mechanics and Wear of Rail/Wheel Systems*, Kingston, RI, USA, 8–11 July, 181–211.
57. Tassily, E. and Vincent, N. (1991), Rail corrugations: analytical model and field tests, *Wear*, **144**, 163–78.

58. Grassie, S.L. and Kalousek, J. (1993), Rail corrugations: characteristics, causes and treatments, *Proceedings of IMechE, Part F: Journal of Rail and Rapid Transit*, **207**, 57–68.
59. Matsumoto, A., Sato, Y., Ono, H., Tanimoto, M., Oka, Y. and Miyauchi, E. (2002), Formation mechanism and countermeasures of rail corrugation on curved track, *Wear*, **253**, 178–84.
60. Eadie, D.T., Oldknow, K.D., Maglalang, L., Makowsky, T., Reiff, R., Sroba, P. and Powell, W. (2006), Implementation of wayside top of rail friction control on North American heavy haul freight railways, *Proceedings World Congress on Railway Research*, Montreal, QC, Canada, 4–8 June, on CD.
61. Oldknow, K.D., Reiff, R.P., Vidler, B. and Elvidge, D. (2006), Verifying top of rail friction control through dynamic rail deflection monitoring, *Proceedings World Congress on Railway Research*, Montreal, QC, Canada, 4–8 June, on CD.
62. Ishida, M., Ban, T., Iida, K., Ishida, H. and Aoki, F. (2006), Effect of moderating friction of wheel/rail interface on vehicle/track dynamic behaviour, *Proceedings 7th International Conference on Contact Mechanics and Wear of Rail/Wheel Systems*, Brisbane, QLD, Australia, 24–27 September, 227–33.
63. Johnson, K.L. and Cameron, R. (1967), Shear behavior of elasto-hydrodynamic oil films at high rolling contact pressures, *Proc. IMechE*, **182**, part I, **14**, 307.
64. Johnson, K.L. (1970), Regimes of elasto-hydrodynamic lubrication, *Journal of Mechanical Engineering Science*, **12**(1), 9–16.
65. Johnson, K.L. and Roberts, A.D. (1974), Observations of viscoelastic behavior of an elasto-hydrodynamic oil film, *Proceedings of Royal Society*, Series A, **337–1609**, 217.
66. Johnson, K.L. and Tevaarwerk, J.L. (1977), Shear behavior of elasto-hydrodynamic oil film, *Proceedings of Royal Society*, Series A, **356–1685**, 215.
67. Patir, N. and Cheng, H.S. (1978), An average flow model for determining effects of three-dimensional roughness on partial hydrodynamic lubrication, *Transactions of ASME*, **100**, 12–17.
68. Conry, T.F., Johnson, K.L. and Owen, S. (1979), Viscosity in the thermal regime of traction, in Dowson, D. (ed.), *Thermal Effects in Tribology: Proceedings of the 6th Leeds-Lyon Symposium on Tribology Held in the Institut National Des Sciences Appliquees De Lyon, France, September 18–21, 1979*, Mechanical Engineering Publications for the Institute of Tribology, Leeds University, Leeds, UK and the Institut national des sciences appliquees Lyon, Lyon, France, Paper VIII (i).
69. Bugarcic, H. and Lipinsky, K. (1982), Mechanical and tribological research on two newly developed rolling friction test rigs at the Technical University of Berlin, in Kalousek, J., Dukkipati, R.V. and Gladwell, G.M.L. (eds), *Proceedings of the Conference on Contact Mechanics and Wear of Rail/wheel Systems*, Vancouver, BC, Canada, July 6–9, 1982, University of Waterloo Press, Waterloo, ON, Canada, 445–62.
70. Tevaarwerk, J.L. (1982), Traction in lubricated contacts, in Kalousek, J., Dukkipati, R.V. and Gladwell, G.M.L. (eds), *Proceedings of the Conference on Contact Mechanics and Wear of Rail/wheel Systems*, Vancouver, BC, Canada, July 6–9, 1982, University of Waterloo Press, Waterloo, ON, Canada, 121–32.
71. Krause, H. and Poll, G. (1982), The influence of real material and system properties on the traction/creep relationships in rolling contact, in Kalousek, J., Dukkipati, R.V. and Gladwell, G.M.L. (eds), *Proceedings of the Conference on Contact Mechanics and Wear of Rail/wheel Systems*, Vancouver, BC, Canada, July 6–9, 1982, University of Waterloo Press, Waterloo, ON, Canada, 353–72.

72. Ohyama, T. and Maruyama, H. (1982), Traction and slip at higher rolling speeds: some experiments under dry friction and water lubrication, in Kalousek, J., Dukkipati, R.V. and Gladwell, G.M.L. (eds), *Proceedings of the Conference on Contact Mechanics and Wear of Rail/wheel Systems*, Vancouver, BC, Canada, July 6–9, 1982, University of Waterloo Press, Waterloo, ON, Canada, 395–418.
73. Ohyama, T. (1987), Study on influence of contact condition between wheel and rail on adhesion force and improvement at higher speeds, *RTRI Report*, 1–2 (in Japanese).
74. Ohyama, T. (1991), Tribological studies on adhesion phenomena between wheel and rail at high speeds, *Wear*, **144**, 263–75.
75. Chen, H., Yoshimura, A. and Ohyama, T. (1998), Numerical analysis for the influence of water film on adhesion between rail and wheel, *Proc IMechE, Part Journal of Engineering Tribology*, **212**, 359–68.
76. Harrison, H., McCanney, T. and Cotter, J. (2002), Recent developments in coefficient of friction measurements at the rail/wheel interface, *Wear*, **253**, 114–23.
77. Chen, H., Ban, T., Ishida, M. and Nakahara, T. (2002), Adhesion between rail/wheel under water-lubricated contact, *Wear*, **253**, 75–81.
78. Chen, H., Ishida, M. and Nakahara, T. (2005), Analysis of adhesion under wet conditions for three-dimensional contact considering surface roughness, *Wear*, **258**, 1209–16.
79. Ohno, K., Ban. T. and Obara, T. (1996), Improvement of adhesion between wheel and rail by ceramics particle injection, *Journal of Japanese Society of Tribologist*, **41**(12), 124–9 (in Japanese).
80. Ishida, M., Ban, T., Takikawa, M. and Aoki, F. (2003), Influential factors on rail/wheel friction coefficient, *Proceedings 6th International Conference on Contact Mechanics and Wear of Rail/Wheel Systems*, Gothenburg, Sweden, 10–13 June, 23–27.
81. Matsumoto, A., Sato, Y., Nakata, M., Tanimoto, M. and Kang, Q. (1996), Wheel-rail contact mechanics at full scale on the test stand, *Wear*, **191**, 101–6.
82. JIS (Japanese Industrial Standard) E1101-1993 (Rails), JISC, Tokyo, Japan.
83. Nishijima S. (1980), Statistical analysis of fatigue test data, *Materials*, **29**(316), 24–9 (in Japanese).
84. Nishijima S. (1980), Statistical analysis of small number of fatigue data, *Proceedings JSME(A)*, **46**(412), 1303–13 (in Japanese).
85. Johnson L.G. (1964), *The Statistical Treatment of Fatigue Experiments*, Elsevier, New York, USA.
86. Chen, H. and Ishida, M. (2006), Influence of surface roughness of rail formed by rail grinding on rolling contact fatigue, *Proceedings of World Congress on Railway Research 2006*, Montreal, QC, Canada, 4–8 June, on CD.
87. Nagashima, S. (1984), Texture structure, *Maruzen*, Tokyo, Japan.
88. Satho, Y. and Iwafutci, K. (2005), Crystal orientation analysis of running surface of rail damaged by rolling contact, *Wear*, **258**, 1126–34.
89. Fujimoto, H., Jin, Y., Takikawa, M., Asano, K., Ishida, H. and Ishida, M. (2002), Wear and its prevention of wheel flange and rail gauge corner, *Proc J-Rail 2002*, Kobe, Japan 27–29 November, 505–508 (in Japanese).
90. Report on Study Group Investigating Prevention of Flange Climb Derailment Caused by Vehicle Negotiating Sharp Curves at Low speed, 2004, March, Ministry of Land, Infrastructure, Transport and Tourism, Bureau of Railways & Railway Technical Research Institute, Tokyo, Japan (in Japanese).

Copyrighted Materials
Copyright © 2009 Woodhead Publishing Retrieved from www.knovel.com

27

Managing the wheel–rail interface: the Australian experience

S. MARICH, Marich Consulting Services, Australia

Abstract: Since the 1980s, a very large number of technical papers has been written on the wide range of factors that are influenced by the wheel–rail contact characteristics. The primary aim of this chapter is to discuss the actual implementation of the technologies that has occurred within the various Australian Railway Systems, including difficulties and areas that may still require practical/realistic explanations and improvements.

Specific reference will be made to the following main aspects:

- rail–wheel wear and lubrication;
- rail corrugations;
- rolling contact fatigue and traction defects;
- control of wheel–rail interaction through profiling;
- rail grinding;
- friction management, to reduce noise and applied forces;
- rail and wheel materials.

Key words: rails, wheels, deterioration, wear, lubrication, defects, interaction, materials, grinding, corrugations, friction management.

27.1 Introduction

Since the 1980s, considerable research and a very large number of technical papers have been written on the wide range of factors that are influenced by the wheel–rail contact characteristics.

The papers have ranged from the very detailed and theoretical analysis of the wheel–rail contact patch, with all its intricacies, to the more practical applications of some of the theoretical findings.

There is no doubt that the work conducted in this area has led to some major economic, operational and safety improvements, particularly in terms of:

- increased rail and wheel lives;
- increased operating speeds and axle loadings;
- decreased deterioration particularly of track components;
- reduced risk of rail and wheel failures, and hence potential derailments or at least reduced costs of maintenance.

The heavy-haul systems in Australia were among the first to conduct research/development on a wide range of wheel–rail interface issues (Marich,

1984; Marich and Walker, 1993). These systems serve as good examples of the major benefits gained from the successful interpretation and implementation of at least some of the findings, which in the 1980s and 1990s led to more than doubling of the rail and wheel lives while at the same time allowing the operating nominal axle loads to be increased from 30 tonnes up to 40 tonnes (Marich and Maass, 1986; Marich *et al.*, 2001a).

The primary aim of this chapter is to discuss the actual implementation of the technologies that has occurred on mainline tracks within at least some of the Australian Railway Systems, including the difficulties encountered and the areas that may still require practical/realistic explanations and improvements. The chapter does not cover the wheel/rail improvements that have occurred in turnouts.

The following sections describe a range of major wheel–rail interaction aspects, all of which have a direct influence on the potential lives of rails and wheels, including;

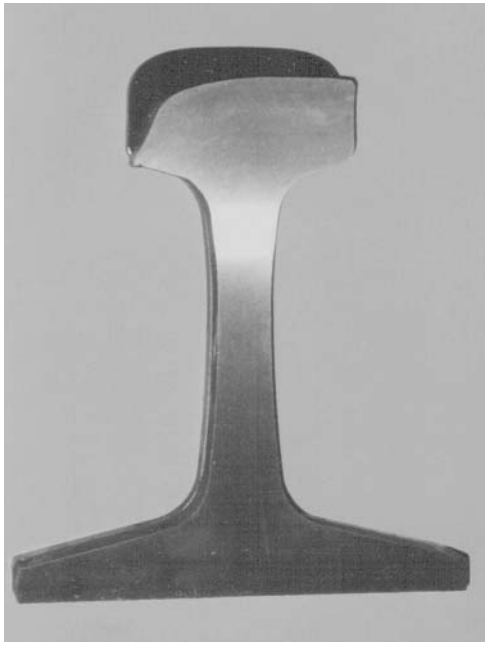
- rail/wheel wear and lubrication;
- rail corrugations;
- rolling contact fatigue and thermal/traction defects;
- control of wheel–rail interaction through profiling;
- rail grinding;
- friction management;
- rail and wheel materials.

27.2 Rail–wheel wear and lubrication

It is obvious that rail (and wheel) wear is one of the major deterioration modes. Wear generally occurs over a relatively long period of time. Technically, this is an advantage since the time allows potential beneficial changes to be made. However, unfortunately the same characteristic often reduces the sense of urgency on the part of the track owners and operators to become interested in implementing changes. On the other hand, at least in some systems, the vehicle owners and operators have provided the driving force because of the adverse economic consequences associated with wheel flange wear.

As illustrated in Fig. 27.1, rail wear occurs predominantly on the gauge face of the high rails in the tighter radius curves (generally less than 500–800 m), due to the higher lateral loads and creepages applied by the wheels. However, as also illustrated in Fig. 27.1, wear can also occur on the running surface of both high and low rails, either due to normal wheel–rail interaction and/or rail maintenance activities such as rail grinding. The equivalent wheel wear modes are of course flange and tread wear.

The various rail (and wheel) wear regimes that can occur at the various wheel–rail contact stress levels and creepage/slippage levels have been well



27.1 Rail wear – high rail in sharp curve.

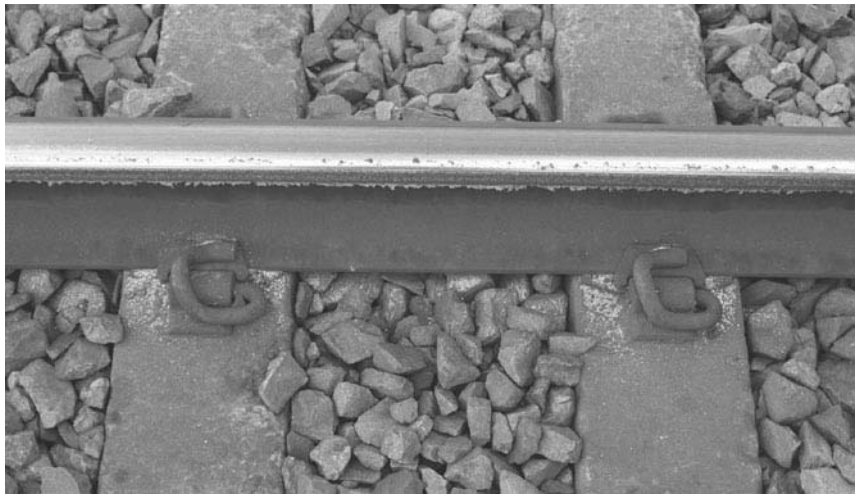
documented (Bolton and Clayton, 1984; Clayton, 1996; and Zakharov, 2001). Furthermore, it has been clearly shown (Mutton *et al.*, 1986) that within any one wear regime the energy at the wheel–rail interface, and hence the resultant wear, is proportional to:

- the force applied normal to the wheel–rail interface;
- the creepage level;
- the friction level.

Hence, the majority of the work associated with wear has concentrated on reducing the influencing factors and ensuring that a severe wear regime, which is typified by very coarse wear particles or debris and very rough contact surfaces, as illustrated in Fig. 27.2 (note the wear debris near the foot of the rail), is not obtained in track.

The main factors that influence the lateral forces, creepage and friction levels are:

- Curve radius, which is generally of course a fixed parameter.
- Wheel–rail interaction characteristics, and in particular the steering forces that are established when particular profiles are implemented (to be discussed in a later section).
- Bogie and track stiffness/damping characteristics – however, in most cases these tend to be also fixed parameters, at least in the short term.



27.2 Large wear debris indicative of a severe wear regime.

- Bogie geometry characteristics and, in particular, inter-axle yaw, bogie lozenging, wheel diameter mismatch and bogie stiffness. All of these aspects have a direct influence on the bogie and wheel maintenance practices.
- Track geometry characteristics and, in particular, track gauge and super-elevation (and the associated ruling speed). For these, a very important aspect occurs in mixed passenger and freight lines, where the design has to balance the often-conflicting requirements of passenger comfort and track damage.
- Lubrication characteristics at the wheel–rail interface.

Another major factor that has a direct influence on the wear is, of course, the hardness of the rails and wheels. This aspect will also be discussed in a later section.

In the shorter term, generally the main reliance has been based on establishing effective wheel flange–rail gauge face lubrication and friction conditions, particularly in the sharper curves, to ensure that a severe wear regime does not occur. This is of particular importance in passenger operations, where generally the wheel materials used are relatively soft and hence more prone to severe wear.

Effective lubrication can also have marked beneficial benefits, particularly in terms of:

- reducing the flanging wear noise associated with wheel–rail interaction, which is very important in built up areas;
- reducing the energy (fuel) consumption associated with wheel–rail interaction.

However, it should always be kept in mind that the benefits of lubrication have to be balanced by the potential adverse effects which ineffective lubrication may have, particularly in relation to:

- the possible contamination of the running surface of the rails, which can lead to the loss of traction by the locomotive wheels and, consequently, the potential development of wheel skids (burns) on the rails and the need for additional tractive effort;
- the enhancement of contact fatigue growth from cracking (checking) on the gauge corner and running surface of the rails (and wheels);
- the adverse influence on the bogie steering characteristics and the lateral wheelset forces, which occurs when the running surface friction of one rail in a curve becomes very different to that of the other rail;
- the general contamination of the track, which generally occurs near track mounted lubricators.

The best way of controlling the application of lubricant to the rails (and hence wheels) is by using vehicle-mounted systems. However, for various practical reasons, the majority of Australian Railways Systems still rely on track-mounted applicators to provide the required lubrication. In this regard, the fieldwork conducted mainly in Australia has clearly shown that the lubrication effectiveness can be markedly improved by paying particular attention to:

- the type of lubricant applied;
- the setting up and maintenance of lubricators;
- the locations at which the lubricators are placed in track.

Some of the more important guidelines that have been established and applied by at least some systems include (Marich *et al.*, 2000, 2001b):

- the need to locate the lubricators in moderate radius (400–1000 m) **feeder curves** ahead of the sharper curves, which are the main target, as long as there are some indications of very minor wheel flanging and no indication of heavy wheel flanging;
- the need to avoid locating the lubricators where steady-state wheel flanging does not occur, such as in tangent track, the low rails of curves and close to the tangent point of curves;
- the need to avoid having lubricators servicing the up and down rails too close to each other (within about 300–500 m).

Some of the important wheel–rail lubrication aspects that still require further investigation include:

- The valid assessment of lubricant types in the laboratory, thus avoiding the need for extensive field tests. The field tests should concentrate

only on the lubricant types that have a high probability of delivering substantial potential advantages.

- The ability to have systems capable of detecting the track locations where lubrication is becoming less efficient, rather than the current procedure of relying on observations made by track personnel which often occur too late. In this regard, the work being conducted in Australia on using a vehicle-mounted noise-based system for assessing the wear/lubrication regime is well underway in achieving this capability (Dwight and Jiang, 2006).
- The assessment of differential steady-state rail wear within severe curves, which at some locations has been shown to vary by up to 50–100 %, from the start to the end of the curves. This is not only a lubrication 'problem' but is likely to involve both wheel–rail and bogie curving characteristics. The ability to reduce such variability even towards the middle of the range would lead to considerable economic benefits in terms of rail replacement.

27.3 Rail corrugations

The development and potential treatments of rail corrugations have been discussed in a large number of technical papers (for example: Grassie and Kalousek, 1993; Sato and Matsumpto, 2000; Grassie, 2005).

Field observations and analytical studies have clearly shown that the corrugations can be of two main types:

- Short-pitch (Fig. 27.3a) – which generally exhibit wavelengths in the range 30–90 mm and occur mainly in lower axle-loads (< 20 tonnes) passenger systems, often on the low rails in the sharper curves. These corrugations are thought to develop from the differential wear (not plastic flow) caused by a repetitious longitudinal sliding action of the wheel on the rail, whether through acceleration, braking or lateral motion across the rail. The longitudinal oscillations can develop due to the excitation of the torsional resonance of the wheelset, which is currently regarded as the main wavelength-fixing mechanism. The energy required for the excitation of the wheelset is enhanced by the stick–slip phenomenon that may occur at the wheel–rail contact patch in sharp curves, because of the differential wheel diameters in a solid wheelset, i.e. the need for one wheel to catch up to the other. The phenomenon is primarily influenced by the:
  - Geometry of the wheelset and, in particular, the difference in the rolling radius differential between wheels on the same wheelset, relative to the curve radius. Steering wheel–rail profiles would exacerbate the process by increasing the difference in contact wheel diameters and

(a) (b)

27.3 (a) Short-pitch rail corrugations. (b) Long-pitch rail corrugations.

creep forces. Hence, there is a conflict between achieving reduced wear and development of corrugations.

— The running surface friction on the high and low rails, with the stick–slip process being exacerbated by higher friction values (which are necessary for wheel traction) or considerable differences in the running surface friction values of the high and low rails. This re-emphasises the need to have an effective lubrication regime.

The main concerns associated with the short-pitch corrugations are the increase in wheel–rail noise and, to a lesser extent, vibration levels that are obtained in corrugated track.

- Long-pitch (Fig. 27.3b) – which generally exhibit wavelengths in the range 150 mm up to about 450–500 mm (or greater) and occur mainly in higher axle loads (> 20 tonnes) mixed freight or unit train operations, on either the high or low rails in the sharper curves. The extensive theoretical and fieldwork conducted since the 1970s has clearly shown that these corrugations develop because of excessive wheel–rail contact stresses causing plastic flow of the rail material. The actual shape of the corrugations is associated with the combined vertical resonance of the vehicles' unsprung mass and the track stiffness (often referred to as the P2 resonance). The phenomenon is therefore exacerbated by all of those factors which lead to higher dynamic loadings and hence contact stresses, and plastic flow of the rail material, including:

  — higher nominal wheel loads;

- higher vehicle speeds, which increase the dynamic loads;
- larger vertical dips at welds/joints, which increase the impact loads;
- higher track stiffness (concrete sleepers are much stiffer than timber sleepers), which increases the dynamic loads;
- higher railpad stiffnesses, which increase the dynamic loads;
- higher bogie suspension stiffnesses, which increase the dynamic loads;
- smaller wheel radii, which increase the wheel–rail contact stresses;
- higher friction/creep at the wheel–rail contact;
- softer rails, which increase the propensity for plastic flow of the material;
- poor matching of wheel and rail profiles, which leads to narrow wheel–rail contact and hence higher contact stresses.

The longer-pitch corrugations are of concern because they increase the dynamic wheel loads (and vibration), and therefore the rate of deterioration and potential failure of various track and vehicle components, such as:

- rails and defects (such as shelling);
- welds and boltholes at insulated rail joints;
- railpads and sleepers, particularly in the rail seat region;
- ballast, which tends to powder and become rounded;
- rail clips (which also become loose);
- track geometry (mainly because of the vibration and the ballast deterioration) – indeed, corrugations can lead to the skewing of sleepers and, in extreme cases, the loss of lateral track geometry;
- wheels and defects;
- bearings;
- bridges and abutments, particularly when the track is not ballasted.

The higher dynamic loads also increase the subsequent rate of corrugation development and the rate of rail profile deterioration. The rails therefore require more maintenance effort (grinding) at shorter intervals.

Analytical models ranging from the relatively simple to the very complex have been recently developed, based on time-domain and frequency-domain approaches, which attempt to predict corrugation initiation and growth rates. Some, albeit limited, success has been achieved in validating the results with actual field data. This is not surprising considering the complexity and variability of the wheel–rail and vehicle–track systems.

Consequently, even although the rate of corrugation development is a very important parameter to know for maintenance scheduling purposes,

it is still not certain whether such models can be applied successfully to actual railway systems consisting of a wide range of vehicle track and wheel–rail interaction characteristics.

To date, in existing systems the control of the longer-pitch corrugations has been achieved primarily by:

— the use of higher strength rail steels, particularly in the sharper curves, to reduce the propensity for plastic flow of the material;
— the application of improved wheel/rail profiles, to reduce the resultant wheel–rail contact stresses;
— the application of appropriate rail maintenance procedures, such as rail grinding, primarily to:
  - reduce the possible dipping at welds that can be a major driver for corrugation development,
  - allow the rails to work-harden in a controlled manner, and thus become more resistant to corrugation development, and
  - ensure that the corrugation severity remains within a controllable regime.

Some of the important aspects of rail corrugation development that still require further investigation include:

- the development of relatively simple, realistic and applicable guidelines on creepage/slippage and friction conditions required to minimise the development of the shorter-pitch corrugations;
- the development of technical/cost functions that quantify the overall track damage associated with corrugation development, so that appropriate control procedures may be justified;
- The development of relatively simple models that can be applied by the track owners and operators to determine general guidelines on corrugation development and growth;
- the establishment of the most appropriate rail grinding procedures required to minimise the redevelopment and subsequent growth of corrugations, in particular the metal removal requirements in terms of the initial corrugation severity.

27.4 Rail contact fatigue and thermal/traction defects

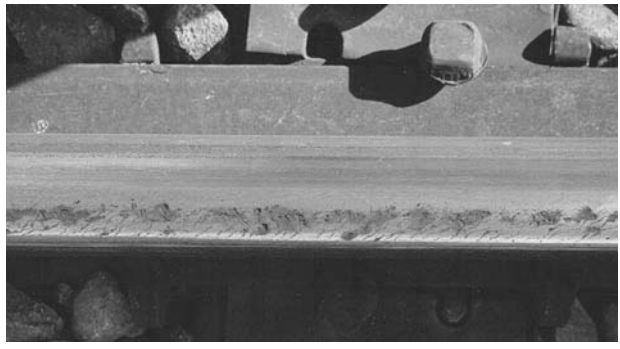
The term rolling contact fatigue (RCF) is generic in nature and used to describe a range of defects, which are due basically to the development of excessive cyclic shear stresses at or close to the wheel–rail contact interface, causing plastic deformation and exceedance of the shear limit (ductility) of the rail material (Grassie and Kalousek, 1997).

RCF defects are mainly of the following types:

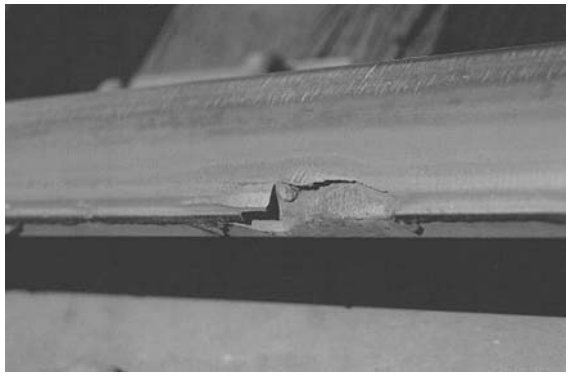
- ***Gauge corner checking*** – a surface condition that occurs mainly in the gauge corner region of the high rails in sharper curves. The gauge corner checking cracks are initiated at or very close to the rail surface, generally occur at about 2–5 mm intervals along the rail, and can grow to the same depth, at a downward angle of about 10–30° to the rail surface, gradually spreading across the railhead. Once this occurs, they usually break out as small 'wedges or spalls', as illustrated in Fig. 27.4.

  Gauge corner checking can also occur in shallow curves and tangent track, where wheelsets/bogies/vehicles tend to exhibit lateral dynamic or hunting behaviour, and where there is negligible gauge face wear of the rails.
- ***Shelling*** – internal defects that initiate at a depth of 2–8 mm below the gauge corner of generally the high rails in curved track. Shelling defects do not form as regularly along the rail as gauge corner checking defects. Shelling cracks develop on a horizontal or longitudinal plane consistent with the shape of the rail on the gauge corner. The cracks can continue to grow in a longitudinal direction on that plane for some distance at an angle of about 10–30° to the rail surface, and then either spall out into a shell, as illustrated in Fig. 27.5, or turn down and form *transverse defects*, as illustrated in Fig. 27.6, which can continue to grow on a transverse plane and eventually lead to rail failure, if not detected in time.

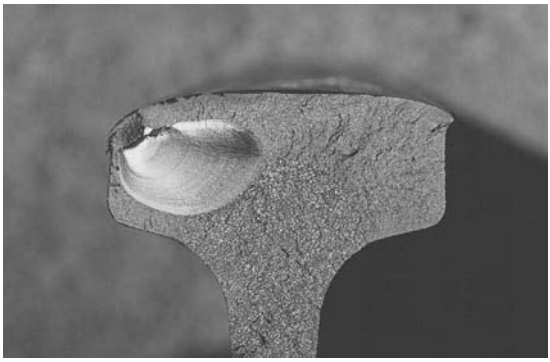
  It should be noted that sometimes, and in particular in older rails, the initiation of both shelling and transverse defects generally occurs at irregularities in the steel (non-metallic inclusions), which can greatly enhance the initiation process. Because of their internal nature, transverse defects cannot be visually detected in their early growth stages, and hence their detection must rely on regular ultrasonic rail inspection.
- ***Flaking*, or *running surface checking*** – also a surface condition that



27.4 Severe gauge corner checking defects in high rail (covering more than 20 mm of the running surface).



27.5 Severe shelling defect in high rail.



27.6 Transverse defect.

occurs on the running surface of the low and/or high rails. Initially, the defects appear as a mosaic or snakeskin-like pattern on the railhead. In the latter stages of development, the cracks produce 'spalls', that can be up to about 10–15 mm wide, up to 3 mm deep, and can be continuous along the rail length. Examples of *flaking* and the associated spalling are shown in Fig. 27.7.

Past work has clearly shown that development of RCF defects is a function of the various factors that influence the resultant wheel–rail contact shear stresses and the ability of the material to withstand such stresses, including:

- the nominal, dynamic and impact wheel loadings, and the range of factors that influence such loadings, including track geometry, bogie characteristics, wheel and rail vertical irregularities, track super-elevation, etc;

27.7 Severe running surface defects in low rail.

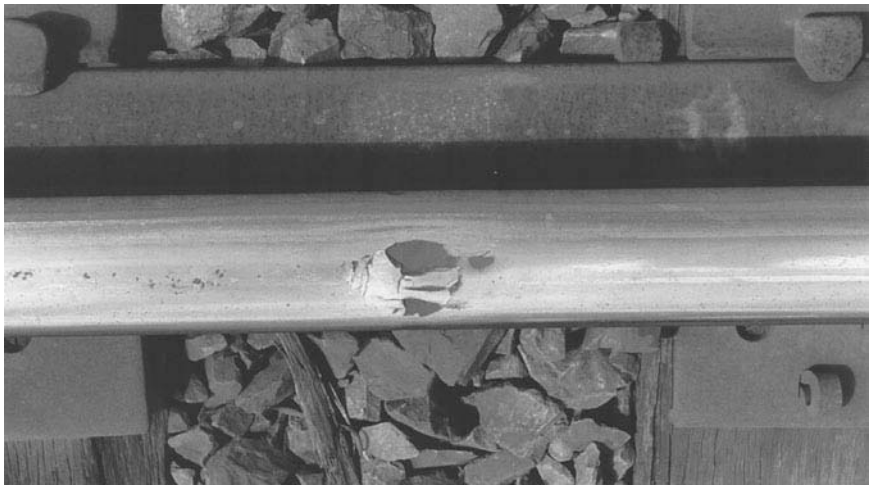
- the respective radii of the wheels and rails at their contact region, and consequently the wheel and rail profile characteristics;
- the radius of the wheels (smaller radii result in higher stresses);
- the resultant traction/creep forces at the wheel–rail contact zone;
- the cleanliness of the rail (and wheel) materials, in the case of the shelling and transverse defects;
- the strength (yield) and ductility (in shear) of the rail (and wheel) materials.

The effect of traction/creep forces on the resultant shear stress distribution, in combination with the relatively random occurrence of inclusions within the rail-head, also explains the differences between the checking and shelling defects.

Thermal/traction defects, on the other hand, are associated with a combination of thermal and traction/creep/slippage effects that can develop at the wheel–rail interface in any track section, and under all types of operating conditions (ranging from passenger to heavy-haul). Thermal/traction defects are mainly of two types:

- ***Wheel or engine burns*** – defects that form on the running surface of the rails (Fig. 27.8), and generally occur in pairs directly opposite to each other on the two rails. Wheel burn defects are caused by the continuous slipping of the locomotive or traction wheels on the rails, which occurs when the longitudinal creepage between wheels and rails reaches saturation.

  The slipping action of the wheels increases the temperature near the surface of the rails to very high values (above transformation of the steel). The subsequent fast cooling causes the rail material to transform to a hard and brittle martensite phase, which in severe cases can extend to depths of 4–6 mm from the running surface.



27.8 Severe wheel burn with spalling.

The main factors that enhance wheel slip are:

— excessive track grades;
— poor train driving procedures, such as rapid acceleration;
— inadequate locomotive/traction power;
— contamination of the running surface of the rails, which can reduce the friction to undesirable levels (less than about 0.30–0.35) – this may occur when the rail lubrication is not applied efficiently, particularly close to the track-mounted lubricators.

However, adverse running surface contamination and loss of traction may also occur when some light rain follows a long hot and dry spell. Under these conditions, it is suggested that the running surfaces of the rails become contaminated with a range of deposits, including oils and pollens. The subsequent light rain can then form a lubricating film, which may reduce the friction on the running surface of both rails to unsatisfactory levels even in the absence of wheel–rail lubricant. Heavier rain would clean the running surface and lead to higher acceptable friction values.

- ***Squat defects*** – surface or near-surface initiated defects that generally develop on the running surface of the rail head (Fig. 27.9) in shallower curves and tangent track. Squats can occur as discrete defects or as closely spaced multiple defects. Generally, if allowed to develop the defects contain subsurface cracking, which occurs very quickly (in less than 5–10 MGT of traffic) and is typically on a horizontal plane, approximately 3–5 mm *below* the rail surface.

  Each squat defect consists of two main subsurface cracks, a leading one that propagates in the direction of train travel, and a trailing one

27.9 Multiple squat defects.

that propagates in the opposite direction. The leading crack is usually several times longer than the trailing crack and contains one main crack with a number of secondary or minor cracks branching off that crack.

Historically, squats have been classified as part of the RCF defects family. However, more recent work (Marich and Mackie, 2002; Kerr *et al.*, 2008) has shown that the majority of defects are actually initiated from a 'white etching', hard and brittle layer, which is most commonly found on infrequently ground rail, is 5–60 μm deep (0.005–0.060 mm), and can have a hardness of up to 750–780 HB.

The 'white etching' layer can form on the rail surface because of adiabatic (low-temperature/high-strain rate) shear between the rail and wheel surfaces, caused by the microslip of the locomotive/traction wheels that are under traction. Hence, special care is required, particularly with vehicles that can sustain higher applied traction levels or locomotives that have not been well maintained.

In other words, the initial development of squats is very similar in nature, but not in degree, to the development of wheel burns, which are of course associated with the much more severe final stage of the slip mechanism. This leads to much higher temperatures, and much greater depths of transformation and hardening (up to 4–6 mm, rather than the 0.06 mm observed with the squats). Furthermore, the other characterising difference between squats and wheel burns is that the former generally develop only on one rail, whilst the other rail remains relatively unaffected. Once the squat defects are initiated, it is most likely that their subsequent growth does occur by a RCF mechanism.

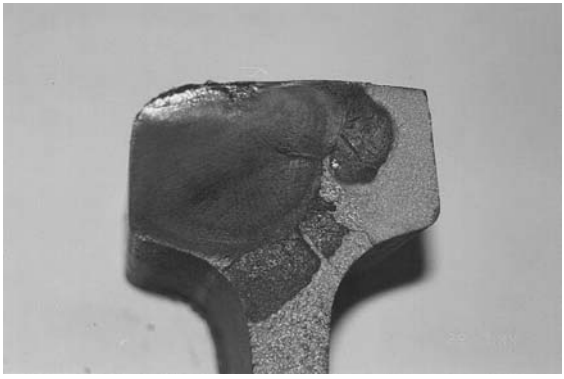
The above discussion applies to squat defects that develop on the running surface of the rails. However, it should be noted that squat defects can also develop from the gauge corner checking cracks present in the high rails (Kerr *et al.*, 2008).

RCF and thermal/traction defects are of particular concern for of two main reasons:

- They may lead to rail failures if not detected in time, particularly in the case of the transverse defects that **may** develop from the gauge corner checking, the shelling, the wheel burns and the squats defects, as illustrated in Fig. 27.10 for the case of a wheel burn. Transverse defects associated with large wheel burns (and possibly large squats) are of most concern because of the potential for fast crack growth associated with the impact wheel loadings that are produced.
- They can mask the ultrasonic signal during routine inspection and hence prevent the detection of larger and deeper defects that may be present within the railhead, including any such defects that may have developed from the shallower initial cracks.

Currently, the main procedures adopted by systems to control the development of RCF and thermal/traction defects have included:

- ***Adoption of higher strength rail steels in the more critical track locations***, to increase the allowable shear stress limits. Higher strength head-hardened or heat-treated rails have been particularly successful in reducing the development of checking, shelling and transverse defects in well-maintained systems. Recent steel making and quality control procedures adopted by all of the major rail manufacturers have also led to much cleaner steels. This aspect has been of particular relevance to shelling and transverse defects development, rather than to gauge



27.10 Large transverse defect developed from a severe wheel burn defect.

corner and running surface checking defects. However, higher strength and cleaner steels have not been of any benefit in terms of the thermal/traction defects.

- ***Improvements in the wheel–rail lubrication procedures***, to reduce the risk of rail contamination and hence the enhancement of fatigue crack growth in the case of the RCF defects, and defects development in the case of the thermal/traction defects.
- ***Application of appropriate rail maintenance strategies***, to provide control over the development of RCF defects, in particular, the grinding of rails at suitable intervals to remove the damaged surface layer. This aspect has been recognised as the most essential control procedure, particularly for gauge corner and running surface checking defects.
- ***Improvements in the wheel–rail interaction characteristics*** by the implementation of preferred wheel and rail profiles, to reduce the wheel–rail contact stresses, improve the wheelset steering characteristics and hence reduce the lateral traction at the rail surface and, in some cases, provide transitional relief from wheel contacts in the affected regions.
- ***Application of suitable ultrasonic testing procedures***, to ensure that the fatigue cracks do not reach their critical sizes and hence lead to rail failures.
- ***Improvements in the rail field stressing procedures***, to reduce the risk of fatigue crack growth in the transverse plane.

It is evident that most of the above procedures have concentrated on the RCF defects rather than the thermal/traction defects.

Some of the important aspects of RCF and thermal/traction defects development that still require further investigation include:

- The development of realistic models that are capable of explaining the reasons why only a very small proportion of the checking and thermal/traction defects that form in the rails actually develop into transverse defects. This would require an improved knowledge of the characteristics associated with the work-hardened layer that develops below the wheel–rail contact zone over time due to controlled plastic deformation of the material.
- The more thorough explanation of why the squat defects only develop on one rail, and what particular operating/track conditions enhance their development.

27.5 Control of wheel–rail interaction through profiling

Most/all major systems, ranging from high axle load heavy-haul to lower axle loads passenger, have now accepted that the control of the wheel–rail interface

can have a major impact on the cost-effectiveness of the operations, and are implementing special rail and wheel profiles with the aim of improving the wheel–rail interaction characteristics and hence:

- the loadings that are applied to the rails by the wheels, which have a direct influence on the wear rate and also the noise levels;
- the wheel–rail contact stresses, which have a direct influence on the development of corrugations and RCF defects;
- the rail stresses, and hence the allowable railhead wear limits;
- the wheelset steering in curves, and hence the rail gauge face and wheel flange wear.

Profiles are usually designed around their naturally worn shape, so that the profiles are stable over time and hence future maintenance required to restore profiles is minimised.

Rail and wheel profile design also needs to consider the mix of traffic, track configuration, track geometry and geographical differences. Control of the wheel–rail interface via rail and wheel profiles has been recognised as a powerful tool for directing rail and wheel behaviour and consequently optimising system performance, through providing just the right amount of wheel–rail contact, balancing stresses and steering in curved track, and avoiding hazardous contact situations such as might cause hunting in tangent track. Appropriate rail and wheel management has also made possible the introduction of higher axle loads than would otherwise be possible.

The rail and wheel profiles have to be system-specific, depending on the major rail deterioration modes that need to be improved or at least controlled.

Taking as an example the standard gauge coal, freight and passenger lines in Australia, the work on wheel–rail interaction conducted in the late 1990's (Kerr and Marich, 2001) led to the development and subsequent implementation of:

- One wheel profile to suit all wheels. This profile features a worn wheel shape, moderate steering capability on curves and increased run-off to reduce the effect of tread hollowing.
- Two sets of rail profiles for curved track (with radii less than 800–100 m), namely:
  - one set applicable for mainly passenger or light axle load lines, which provides increased steering capability on sharp curves by concentrating the conformal wheel–rail contact in the gauge corner of the high rails;
  - one set for higher axle load operations with less steering capability, but with more resistance to gauge corner damage by spreading the conformal contact load on the high rails, and additional field side

relief on the low rails to reduce the adverse effects of tread hollow wheels.

- One rail profile for application to tangent track and shallow curves (with radii generally greater than 1000 m), which minimises gauge corner contact in the rails and hence reduces the potential for wheelset/bogie hunting behaviour.
- The low and tangent rail profiles also provide conformal contact towards the centre of the running surface to reduce the contact stresses, and hence the potential for corrugation development, and the rail bending stresses, hence increasing the allowable rail wear limits.

The various rail and wheel profile combinations adopted are illustrated in Fig. 27.11.

More recently, special rail profiles have been designed and implemented in some systems with the aim of adopting a transitional rail maintenance (grinding) strategy and hence reducing the cyclic maintenance costs. The profiles have been applied in several situations, including:

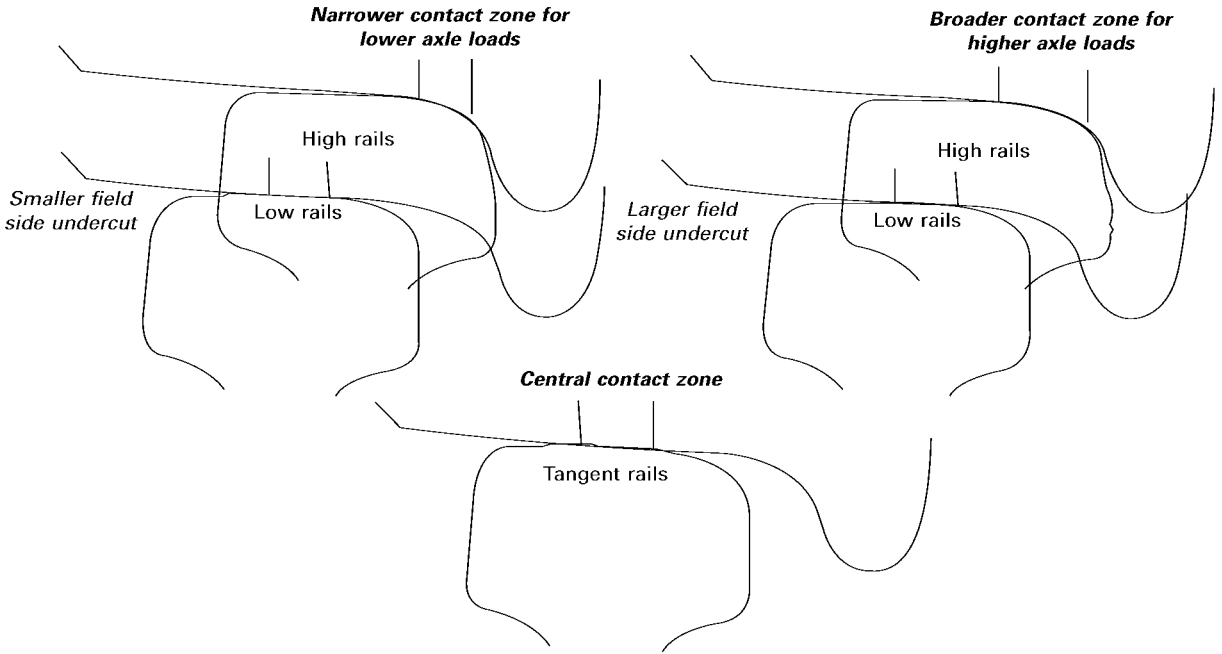
- rails exhibiting severe gauge corner and running surface checking defects;
- rails exhibiting very broad wheel–rail contact bands, which are due to wheels with excessive tread hollowing levels.

The other considerable improvement that has occurred has been the redesign of the contact region in some of the new rails, so that when these are installed they require minimum maintenance to establish the desired operating profiles. It is believed that in the future the design of rail (and wheel) profiles will become even *smarter*, to provide the required balance between wear, fatigue and deformation. The main challenge will be for the various systems to implement the established technologies. One major aspect that will need to be better defined and implemented in the future is the technical/economic consequences associated with deviations from the preferred profiles, which would then provide realistic and acceptable operating tolerances for both the track and vehicle owners.

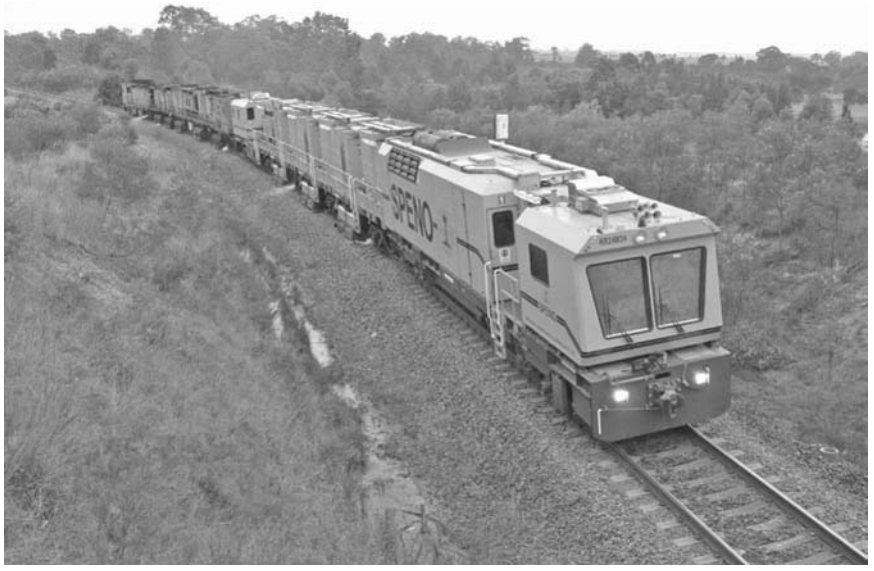
27.6 Rail grinding

Since the 1980s, rail grinding, as illustrated in Fig. 27.12, has become an accepted rail maintenance practice in railway systems in Australia (and throughout the world), ranging from higher axle load heavy-haul, to general freight to lower axle load passenger (Marich, 2005a).

Each system has its own particular set of objectives for the grinding of their rails. These objectives may range from the very broad, for example, control of the wheel–rail interaction characteristics, to the very narrow, for example: reduction in wheel–rail noise levels or removal of rail corrugations.



27.11 Modified wheel and rail profiles for mainline track.



27.12 Rail grinder.

Generally, rail grinding has been applied to achieve one or more of the following technical/economic objectives:

- implement modified rail profiles, to improve the wheel–rail interaction characteristics, and hence reduce rail–wheel wear and contact stresses;
- correct/control rail corrugations and dipped welds and joints, which increase the rate of deterioration of both track and vehicle components, increase the noise levels and can sometimes lead to speed restrictions;
- correct/control RCF defects, which increase the risk of rail failures and sometimes reduce the effectiveness of the in-line ultrasonic rail testing;
- correct/control other rail defects (such as wheel burns, squats, vertical and horizontal split heads), which also increase the risk of rail failures;
- increase the allowable rail wear limits, by providing improved wheel–rail contact conditions;
- reduce the adverse influence of 'rogue' wheels and bogies, which can exacerbate rail wear and defect development;
- reduce noise and vibration, again by reducing the localised vertical irregularities at welds and joints and controlling the rail corrugations;
- moderate the adverse influence of higher axle loads, by providing improved wheel–rail contact conditions;
- reduce the sensitivity to wheelset–bogie lateral instability (hunting), again by providing improved wheel–rail contact conditions and interaction characteristics.

Rail grinding can successfully achieve all of the above technical objectives. However, it has also be recognised that to ensure that rail grinding is cost-effective, certain steps need to be taken by the personnel representing both the rail asset owner and the rail grinding contractor. These include:

- **Establishing clear objectives** – The primary objectives associated with the rail grinding need to be well defined by the rail asset owner, explained and accepted by all of the personnel involved in the rail grinding process, since the rail grinding procedures and strategies that are adopted depend to a major extent on the set objectives.

  If the objectives become confused, it is very likely that the rail grinding will not be cost-effective and indeed, in some cases, will even lead to negative benefits.
- **Establishing standards** – A serious rail grinding campaign **should not** begin without the rail asset owner having established detailed, reasonable and practical Technical Standards. Such standards need to cover a wide range of factors that are regarded as important, ranging from the allowable profile tolerances, to the metal removal requirements, to the surface finish requirements, to the quality assurance procedures, to the health and safety issues, to the competencies required by the personnel involved in the grinding operations.
- **Establishing grinding strategies** – A serious rail grinding campaign **should not** begin without the rail asset owner having established a longer term cost-effective grinding strategy, which depends largely on the reasons for grinding and the potential longer term commitments that are made.
- **Improving the grinding efficiency** – There are a wide range of actions that can be taken by both the rail asset owner and the rail grinding contractor to improve the rail grinding efficiency and hence its cost-effectiveness, including:
  - periodic review and fine-tuning of grinding cycles and procedures for specific regions and track sections – this requires inspection of the rail condition not only immediately after grinding but also between grinding cycles;
  - adoption of longer term *planned strategies* for transition from corrective/ defect to preventative/cyclic maintenance;
  - improved and more consistent grinding operations, which means planned maintenance strategies for the rail grinding equipment and hence reduced risk of break-downs;
  - improved track conditions suitable for grinding, including the prior removal of potential obstacles, such as lubricators, high ballast and crossings, and improved track access, so that the grinding operations can be monitored and controlled;

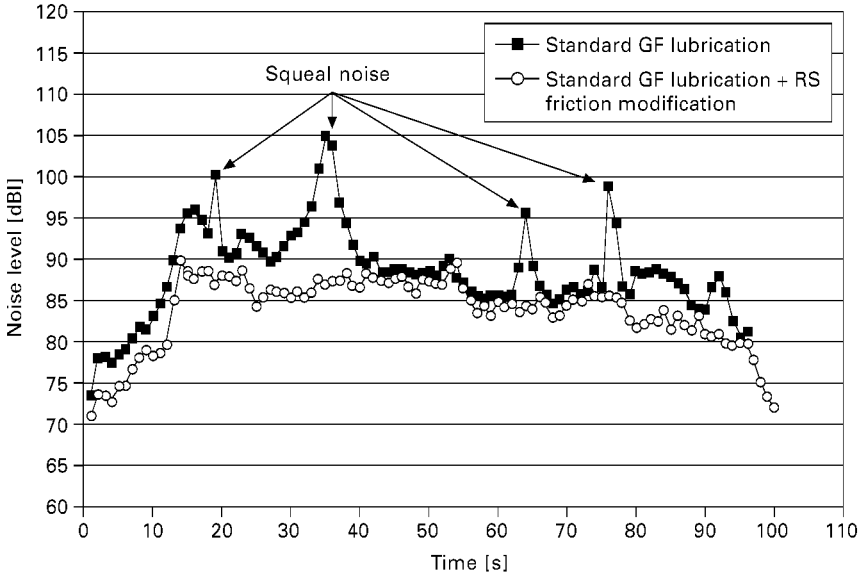
- implementation of auxiliary procedures that reduce the rail grinding effort required, for example, weld straightening or repair of larger surface defects.
- prior rail/track inspection by appropriate personnel representing both the rail asset owner and the grinding contractor, essential in determining the actual grinding requirements and potential obstacles;
- improved understanding and co-ordinated planning/management of the grinding operations, including operational matters and the assessment of fire risk and also minimising the travel requirements of the rail grinding unit between work sites;
- full discussion and disclosure of the grinding objectives and strategies;
- implementation and acceptance of onboard rail profile and other rail surface condition monitoring systems, that can be used to determine both the grinding procedures/requirements and for quality assurance/control purposes; establishment of appropriate data collection and analysis procedures, which are essential for rational decision making;
- improved understanding of the range of factors influencing metal removal and surface condition, including the relationship between the metal removal rates, the grinding stone wear rates and the operating parameters;
- scheduling of the rail grinding operation preferably after track geometry rectification rather than immediately before.

The potential future improvements that can be made to the maintenance of rails by grinding are mainly associated with the acceptance and implementation by the track owners of all of the above factors.

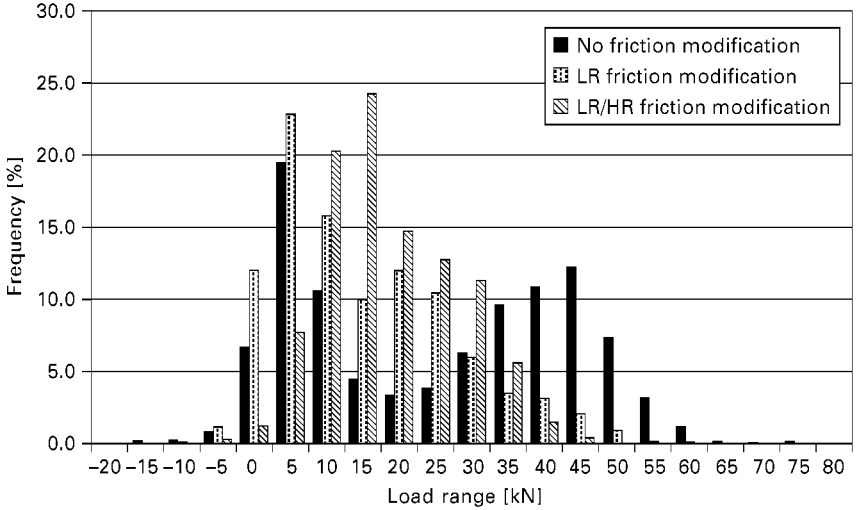
27.7 Friction management

The control of wheel flange/rail gauge face friction through having effective lubrication has already been discussed in Section 27.2. The other major development that has occurred since around 2000 has been in the area of friction modification at the rail running surface–wheel tread interface. It is now well established (Eadie *et al.*, 2003; Reiff Davies, 2003; Marich and Mackie, 2004) that the application of appropriate friction modifiers to both rails can lead to:

- a marked attenuation of the wheel squeal noise associated with certain wheelsets/bogies, as illustrated in Fig. 27.13;
- an overall reduction in the wheel–rail rolling noise, as also illustrated in Fig. 27.13;



27.13 Reduction in wheel–rail noise levels due to friction modification.



27.14 Reduction in lateral forces in high rails due to friction modification.

- a considerable reduction in the lateral forces applied by the wheelsets on both the high and low rails in sharp curves – for example, as illustrated in Fig. 27.14, the reductions in the maximum force levels can be as high as 40–45 %.

Consequently, appropriate friction modification offers the potential for major benefits in terms of reduced applied lateral loadings and hence component deterioration, including of course rail and wheel wear. Additional potential future benefits of friction modification could include:

- the ability to control the overall tractive effort required in certain track sections;
- the ability to provide control over the wheel–rail creep–slip mechanisms and hence the development of certain surface defects, such as short-pitch corrugations and squats.

The actual effect of friction control at the running surface on the applied load levels has been successfully analysed and modelled. At least some analytical work, although not to the same extent, has also been conducted recently on the wheel–rail noise issues, and in particular the squealing noise that is generated by some wheelsets/bogies (Anderson *et al.*, 2008). Having established the considerable benefits associated with friction modification, the main challenge in the future will be to develop and apply the most appropriate and cost-effective application systems, which can be either track-or vehicle-mounted.

27.8 Rail and wheel materials

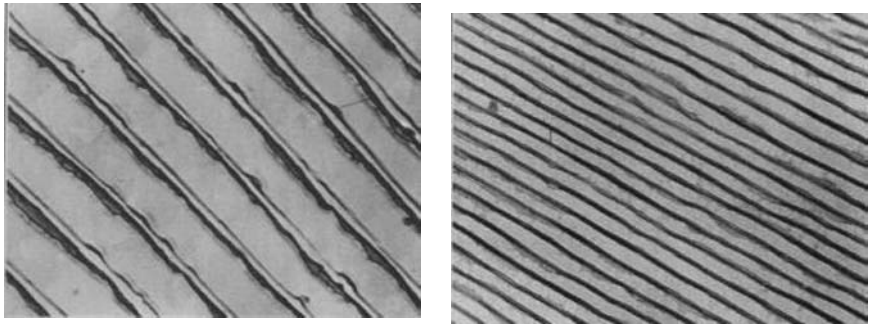
As discussed previously (Marich, 2005b) very extensive work has been conducted since the 1970s on wheel and rail materials, particularly in the areas of:

- improvements in the mechanical properties and material types;
- improvements in the designs, particularly of wheels.

27.8.1 Mechanical properties and material types

From the early 1950s, major improvements have been made to chemistry and heat-treatment procedures associated with fully pearlitic rail steels, which have led to considerable increases in the strength, hardness and fatigue characteristics of the materials. Thus the strength and hardness of the plain carbon steels, which are around 900–1020 MPa and 260–290 HB, respectively, have been increased to 1100–1300 MPa and 350–400 HB first by the addition of suitable alloying elements and then by the development of special accelerated cooling procedures and the production of the heat-treated, fully pearlitic higher carbon steels. These improvements have been due mainly to the refinement of the microstructure achieved by the processes, as illustrated in Fig. 27.15.

Since the 1980s, the heat-treated rails have become almost a standard



(a) (b)

27.15 Pearlitic microstructure in (a) plain carbon and (b) heat treated rails (magnification × 25 000).

product, particularly in high axle load operations or in systems with a large proportion of very sharp curves. Nevertheless, it should also be noted that in Australia probably more than 80 % of the rails in use are still of the plain carbon type.

The popularity of the heat-treated rail steels, at the expense of the alloyed rail steels, stemmed from various reasons, including:

- ease of manufacture;
- reduced costs, even allowing for the need of special heat-treatment facilities;
- improved overall mechanical properties, and in particular ductility and fracture resistance; and
- improved welding characteristics, although the rails still require air quenching after welding to increase the hardness in the fusion zone and hence avoid dipping at the welds particularly under high axle load operations.

It has been clearly shown that the implementation of heat treated rails has led to marked reductions in rail wear, RCF defects and rail corrugation (long-pitch) development. Indeed, it has been a major factor in the ability to implement and sustain higher nominal axle loads.

Three heat-treated rail types have become commercially available, namely:

- head-hardened (HH);
- deep head-hardened (DHH) or low-alloy heat-treated (LAHT);
- fully heat-treated (FHT).

As their name implies, the main difference between these rails relates to proportion of the rail cross-section that has been heat-treated and hardened. Thus in the HH rail the heat-treatment extends about 30 mm from the running

surface, in the DHH/LAHT rail the heat-treated zone is over 45 mm deep and in the FHT rail the whole section is heat-treated.

The DHH/LAHT and FHT rails exhibit a more gradual hardness reduction from the rail surface than the HH rails, although all product types show similar hardness values up to depths of at least 5–10 mm. However, generally the field performance of the various heat-treated rail types has been found to be similar. This is not surprising considering that the major modes of rail deterioration (wear, corrugations and RCF defects) are all associated with the material characteristics and behaviour close to the contact surface (i.e. within 5–10 mm). Furthermore, the gradual development of a work-hardened layer establishes a hardened zone which remains essentially constant as wear occurs, and is relatively independent of the original material properties at depths greater than 10–15 mm.

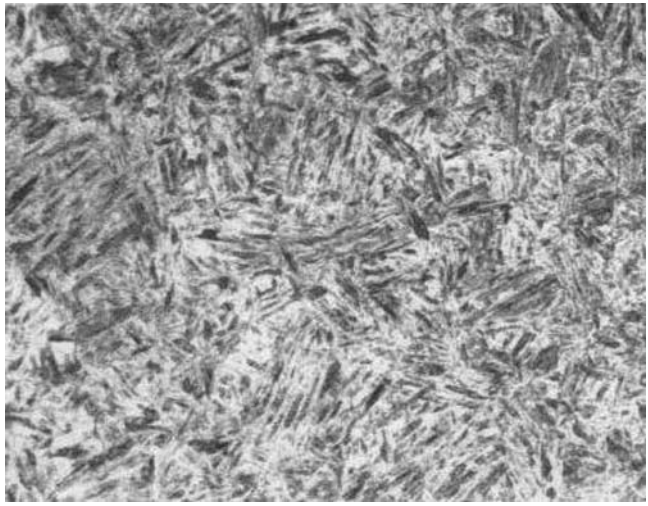
The main difference that would be expected between the HH and FHT rails is in terms of the residual stress distribution present in the rail section. The heat-treatment of the full section produces a more balanced stress distribution. An improvement in the residual stresses would improve the rail's behaviour, particularly in terms of resistance to fatigue initiation and growth in the rail web (i.e. horizontal split web failures which may initiate in welds but develop into the rails).

Further improvements in performance have been obtained with the development of micro-alloyed heat-treated rails, primarily through their improved work-hardening behaviour when subjected to compressive loadings, and improved welding characteristics, to the extent that air quenching is no longer required. However, these materials have not been pursued, mainly due to some manufacturing difficulties.

Since around 2000, two quite different rail steel types have also been developed and produced, namely:

- Hyper-eutectoid steels, in which the carbon levels have been increased further up to about 0.9 %, with the aim of increasing the thickness of the cementite lamellae within the pearlitic microstructure and hence the resultant hardness levels (up to 400 HB). In these steels, acceptable (albeit lower) ductilities and a microstructure that is free of pro-eutectoid cementite are achieved by controlled rolling and in-line accelerated cooling of the rails. Initial field trials have indicated some improvements, particularly in terms of the wear and plastic flow behaviour.
- Lower carbon, bainitic (and martensitic) steels, which exhibit totally different types of microstructures, as illustrated in Fig. 27.16.

Both of these steel types exhibit considerably higher strength, hardness and ductility (in the case of the bainitic/martensitic steels) than even the heat-treated rails, while the low-carbon bainitic/martensitic steels also exhibit better impact and work-hardening characteristics, with the fatigue crack



27.16 Example of microstructure in lower carbon bainitic/martensitic rails (magnification × 500).

initiation and growth characteristics being equivalent to those of the head-hardened rails. Their weld performance is also very acceptable, particularly in the case of the lower carbon bainitic/martensitic materials.

However, as yet neither of the new materials has been implemented in large quantities. The main reasons could include the very conservative nature of the railway industry and the acceptable performance of the standard heat-treated pearlitic steels, particularly in combination with appropriate management procedures.

Similar developments have occurred with the wheel materials, such that at present a wide range of wheel material classes is available and has been adopted by the industry, namely:

- AAR Class A, with a hardness in the range 255–321 HB;
- AAR Class B, with a hardness in the range 277–341 HB;
- AAR Class B micro-alloyed, with a hardness in the range 321–363 HB;
- AAR Class C, with a hardness in the range 321–363 HB;
- AAR Class C micro-alloyed, with a hardness in the range 362–401 HB.

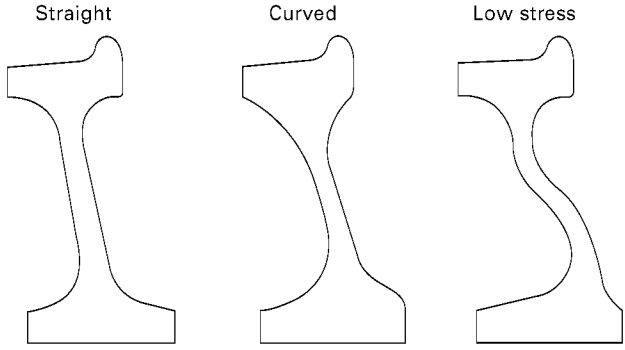
The choice of the most appropriate wheel material depends on the balance between resistance to wear, RCF defect development and tread hollowing against resistance to thermal fatigue damage. Although some work has been done on the development of lower carbon bainitic/ martensitic wheel materials, as for rails as yet these new materials have attracted little interest by the industry.

27.8.2 Designs

To reduce the susceptibility to thermal damage associated with abnormal braking conditions, the wheel plate design has been modified considerably since the 1970s, firstly to a curved section and then to a low stress (S-plate) section, as illustrated in Fig. 27.17. The extensive analytical, laboratory dynamometer and field testing clearly showed that the new design allowed much greater flexibility of the wheel plate, and hence reduced the adverse changes in residual stress levels on overheating. The new plate designs have now been fully accepted and implemented by the industry.
On the other hand, with the exception of increasing the rail size, negligible changes have occurred in the basic design of flat bottom rails.

The characteristics of the rail sections generally used in higher axle load, heavy-haul operations, or heavy-duty passenger operations are summarised in Table 27.1 and Fig. 27.18. It can be seen that the major differences are as follows:

- the generally smaller dimensions of the 60 kg/m section, although it is of interest to note that the head height is similar to the 68 kg/m section;
- the increased head height of the newer 71 kg/m section (by more that 5 mm) relative to the 68 kg/m section;

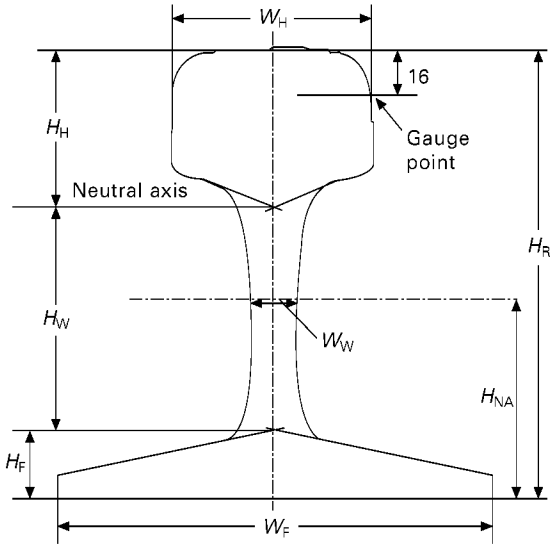


27.17 Plate profiles for wheels.

Table 27.1 Main dimensions of some heavy duty rail sections (mm)

	W_H	W_W	W_F	H_H	H_W	H_F	H_R	H_{NA}
60 Kg/m	70.0	16.5	146.0	49.0	93.0	28.0	170.0	80.0
68 Kg/m	74.6	17.5	152.4	49.2	106.4	30.2	185.7	85.0
71 Kg/m	74.4	17.5	152.4	54.7	104.0	30.2	188.9	88.0

W_H = head width, W_W = web width, W_P = foot width, H_H = head height, H_W = head width, H_P = foot height, H_R = rail height, H_{NA} = neutral axis height



27.18 Rail sectional dimensions.

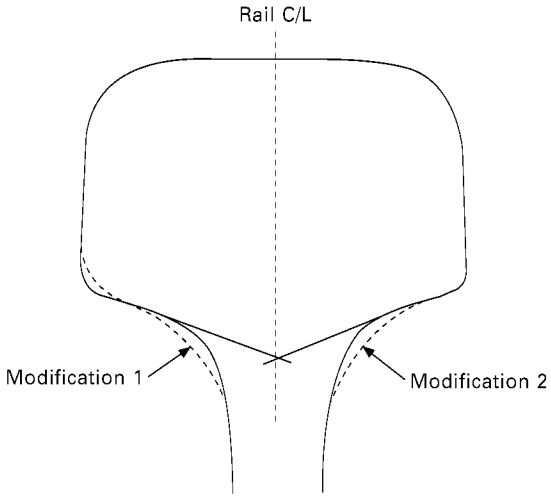
- the increased height of the neutral axis, with increased rail section size.

The new, heavier section therefore appears to have been designed to provide extra material in the railhead, presumably to increase the wear life associated with running surface wear, rather than gauge face wear. Since most of the running surface wear is due to rail maintenance activities such as rail grinding, rather than wheel–rail interaction, the new rail section would therefore allow a greater overall metal removal by grinding before reaching the limiting stress condition.

However, it is also of interest to note that finite element modelling analysis of the 60 kg/m and 68 kg/m rail sections has indicated that the smaller section actually exhibits lower stresses on the underside of the railhead, which is one of the most critical locations in determining the allowable railhead wear limits, particularly as the level of head wear increases. This is not surprising considering the larger rotational bending moments associated with the larger and taller sections.

These results therefore would indicate that the smaller rail section would be preferable in terms of allowable rail head loss, and hence rail life. Hence bigger may not always be better. It should also be noted that a considerable proportion of rail sections in the Australian mainline tracks are still less than 60 kg/m.

Previous finite element modelling analysis (Tew *et al.*, 1990) has also shown that relative minor variations in the design of the rail fishing surface/



27.19 Modified rail sections.

web transition (refer to Fig. 27.19) can reduce substantially the stresses obtained in this critical area. The results obtained clearly indicated that the current rail sections are far from optimum, particularly when assessed in terms of the critical stresses that determine the economic rail life based on the allowable railhead wear limits. Unfortunately, very little if any further activities appear to have been pursued in this area.

In the next 5–10 years, it is felt that some of the main challenges within the railway industry in terms of wheels and rails materials and designs include:

- establishing the most appropriate information to allow the most cost-effective decisions to be made on the wheel material type(s) most suitable for particular operations, to provide the required balance between wear/fatigue/deformation resistance and thermal fatigue resistance;
- critically assessing the track locations in which the higher strength rail steels are most cost-effective;
- modifications to the rail sections, with the aim of reducing the most critical rail stresses.

27.9 Conclusions

In this chapter, an attempt has been made to review the vast amount of work and the considerable advances that have occurred in the area of wheel–rail contact technologies, mainly since the 1980s. The main areas covered have been:

- rail–wheel wear and lubrication;
- rail corrugations;

- RCF and traction defects;
- control of wheel–rail interaction through profiling;
- rail grinding;
- friction management;
- rail and wheel materials.

Particular emphasis has been placed on the practicalities of these technologies and their implementation within the Australian Railway Industry.

Some of the advances, but in particular their combinations, have already led to major increases in the lives of both rails and wheels, even at the very high axle loads currently applied in some systems, and in the very sharp curves that are often present in other systems. Indeed, these days rail lives in excess of one gigatonne and wheel lives in excess of two million kilometres are becoming increasingly common. The risk of rail and wheel failures has also reduced substantially, not only because of the advances in wheel–rail technologies but also because of the improvements made to non-destructive (mainly ultrasonic) testing procedures.

Some more specific activities have been identified as still requiring practical and realistic explanations and improvements, within the areas discussed. However, it is felt that the next 5–10 years should be a time for review, consolidation and implementation, with the aim of avoiding or at least minimising the duplication of research activities. In the foreseeable future, a much greater emphasis will need to be placed on:

- the development and application of overall rail and wheel management procedures and strategies, which allow informed decisions on the most cost-effective combinations of the technologies that have already been developed;
- the appropriate training of railway engineers and support staff, whose functions include not only the making of day-to-day decisions but also the implementation, maintenance and assessment of the new technologies.

As usual, further developments will require the active interest and support of the railway operators, track owners and suppliers. It is also essential that these critical parties have the knowledge and patience required to provide guidelines and longer term plans for future research and development activities.

27.10 References

Anderson D, Wheatley N, Fogarty B, Jiang J, Howie A and Potter W (2008), Mitigation of curve squeal noise in Queensland, New South Wales and South Australia, *Proceedings RTSA CORE2008 Conference on Railway Engineering*, Perth, WA, Australia, 7–10 September, 625.

Bolton P J and Clayton P (1984), Rolling-sliding wear damage in rail and tyre steels, *Wear*, **93**, 145–65.

Clayton P (1996), Tribological aspects of wheel–rail contact: a review of recent experimental research, *Wear*, **191**, 170–83.

Dwight R and Jiang J (2006), On-board wheel-rail noise monitoring for the direction of track maintenance, *Proceedings RTSA CORE2006 Conference on Railway Engineering*, Melbourne, VIC, Australia, 30 April–3 May, 521.

Eadie D T, Vidler B, Hooper N E and Makowsky T W, Top of rail friction control: lateral force and rail wear reduction in a freight application, *Proceedings International Heavy Haul Association Specialist Technical Session on Implementation of Heavy Haul Technology for Network Efficiency*, Dallas, TX, USA, 5–9 May, 5.73–5.80.

Grassie S L and Kalousek J (1993), Rail corrugation: characteristics, causes and treatment, *Proceedings of the Institution of Mechanical Engineers*, **287**, 57–68.

Grassie S L (2005), Rail corrugation: advances in measurement, understanding and treatment, *Wear*, **258**, 1224–34.

Grassie S L and Kalousek J (1997), Rolling contact fatigue of rails: characteristics, causes and treatments, *Proceedings 6th International Heavy Haul Conference*, Cape Town, South Africa, 7–11 April, 38–404.

IHHA (2001) *Guidelines To Best Practices For Heavy Haul Railway Operations: Wheel and Rail Interface Issues*, International Heavy Haul Association, Virginia Beach, VA, USA.

Kerr M and Marich S (2001), At last, compatible wheel and rail profiles for standard gauge lines, *Proceedings 13th International Rail Track Conference*, Canberra, ACT, Australia, 11–14 November, Rail Track Association, Cherrybrook, NSW, Australia.

Kerr M, Wilson A and Marich S (2008), The epidemiology of squats and related rail defects, *Proceedings RTSA CORE2008 Conference on Railway Engineering*, Perth, WA, Australia, 7–10 September, 83.

Marich S (1984), Heavy haul railway research – an overview, Proceedings Institution of Engineers, Australia Annual Engineering Conference, Brisbane, Qld, Australia April, 35.

Marich S and Maass U (1986), Higher axle loads are feasible – economics and technology agree, *Proceedings 3rd International Heavy Haul Conference*, Vancouver, BC, Canada, 13–17 October, Paper IA-1.

Marich S and Walker W (1993), The evolution of railway research and development at a heavy haul railway', *Proceedings of the Institution of Mechanical Engineers*, **207**, 89.

Marich S, Mackie S and Fogarty R (2000), The optimisation of rail/wheel lubrication practice in the Hunter Valley, *Proceedings RTSA CORE2000 Conference on Railway Engineering*, Adelaide, SA, Australia, 21–23 May, 4.1.

Marich S, Cowin A and Moynan M (2001a), Managing the wheel/rail interface under very high axle loads and tonnages at bhp iron ore, Australia, *Proceedings 13th International Wheelset Congress*, Rome, Italy, 17–21 September 2001.

Marich S, Kerr M and Fogarty R (2001), Optimising trackside lubrication for the wheel/rail system, *Proceedings 13th International Rail Track Conference*, Canberra, ACT, Australia, 11–14 November, Rail Track Association, Cherrybrook, NSW Australia.

Marich S and Mackie S (2002), The development of squat defects under high axle load operations, *Proceedings RTSA CORE2002 Conference on Railway Engineering*, Wollongong, NSW Australia, 10–13 November 17.

Marich S and Mackie S (2004), The control of wheel/rail noise and forces through friction modification, *Proceedings RTSA CORE2004 Conference on Railway Engineering*, Darwin, NT, Australia, 70–24 June 01.1.

Marich S (2005a), Some facts and myths about rail grinding – the Australian experience, *Proceedings 8th International Heavy Haul Conference*, Rio de Janeiro, Brazil, 14–16 June, 457.

Marich S (2005b), Major advances in rail technologies achieved in the past 10–20 years – the Australian experience, *Proceedings 8th International Heavy Haul Conference*, Rio de Janeiro, Brazil, 14–16 June, 747.

Mutton P J, Boelen R and Marich S (1986), Requirements for wheel and rail materials, *Proceedings 6th International Rail Track Conference*, Melbourne, VIC, Australia, 17–19, March, Rail Track Association, Cherrybrook, NSW, Australia.

Reiff R and Davies K (2003), Implementing locomotive based, top of rail friction control for improving network efficiency, *Proceedings International Heavy Haul Association Specialist Technical Session on Implementation of Heavy Haul Technology for Network Efficiency*, Dallas, TX, USA, 5–9 May, 5.87–5.95.

Sato Y and Matsumpto A (2000), Review of rail corrugation studies, Proceedings 5th International Conference on Contact Mechanics and Wear of Rail/Wheel Systems, Tokyo, Japan 25–27 July, 74.

Tew G P, Soeleiman S, Choo P and Chitty G (1990), Rail profile modifications for improved performance, *Proceedings 8th International Rail Track Conference*, Sydney, NSW, Australia, 22–25 October, Rail Track Association, Cherrybrook, NSW, Australia.

Copyrighted Materials
Copyright © 2009 Woodhead Publishing Retrieved from www.knovel.com

28

Managing the wheel–rail interface: the Dutch experience

A. ZOETEMAN, R. DOLLEVOET, R. FISCHER and J.-W. LAMMERS, ProRail EIM and Delft University of Technology, The Netherlands

Abstract: This chapter discusses ProRail's practical approach and projects to manage wheel–rail interfaces together with its partners, the transport operators and maintenance contractors. Policy and projects from both the rail and the wheel maintenance perspective are discussed. Phenomena such as railway noise levels and rolling contact fatigue have been reasons to intensify contacts with partners in the industry and the users of the railway network since 2000. This has led to strategic research programmes, together with the academic world (see e.g. Chapter 13 for research on squats and Chapter 12 for research on track irregularity). This chapter provides a general overview of the state of play and is not intended to include all details of wheel–rail interface management. A list of references is included in the final section for further reading on Dutch wheel–rail interface research.

Key words: infrastructure management, rolling contact fatigue (RCF), railway noise, squeal, wheel maintenance, weigh-in-motion systems.

28.1 Introduction

ProRail is the Dutch Infrastructure Manager for the railway network under a Management Concession until 2015, which comprises 6500 km of track (of which 4875 km is considered as main tracks), including 4500 bridges and tunnels and 2000 level crossings. The network is used daily by 5400 passenger trains (servicing 1.2 million passengers) and 300 freight trains (transporting 100 000 tonnes of cargo). After Switzerland and Japan it is the most intensively used railway network in the world, regarding passenger transport. ProRail manages the infrastructure processes, including design and construction, maintenance and renewal, traffic management and capacity planning and passenger information services. Each year on average 1 billion Euros is spent in new construction and 1 billion Euros in upkeep of the existing network, which has a replacement value of more than 30 billion Euros.

The number of operating companies operating on the Dutch rail infrastructure, after the restructuring of the railways, has increased steadily to 35, with many smaller operators for freight and for regional passenger traffic. The largest operator is still the incumbent railway, Netherlands

Railways (NS), which works under a transport concession until 2015 and realises more than 80 % of the train kilometres. New operating companies are able to enter the network after rolling stock acceptance, managed by the Transport and Water Management Inspectorate, and the allocation of slots according to non-discriminatory procedures. Infrastructure use is charged on a variable cost basis, including parameters such as train and tonne kilometres, number of stops at stations, and energy used (kWh and diesel oil).

The railway restructuring has caused important changes in the management of the wheel–rail interface (WRI), requiring the involvement of many stakeholders. Moreover, there have been developments which have changed the physical parameters in the WRI. Rail traffic has increased at a rate of 5% per year for some years and continued growth is forecast for the period until 2020, as long as the railway sector and the Ministry of Transport are able to remove physical limits quickly enough through new logistic models and investment. Finally, the regulation for WRI is more and more prescribed by Brussels, through technical specifications for interoperability (TSIs), in order to open markets further and to improve cross-border operations with more, different types of rolling stock. These developments will demand continuous research and (re-)optimisation of the wheel–rail interface.

This chapter provides an overview of ProRail's practical approach and ProRail-funded research projects to manage the WRI in an optimum manner. First, Section 28.2 describes the perspective from the track side of the interface, in particular the battle against rolling contact fatigue (RCF). Section 28.3 describes the developments at the vehicle side of the interface. Section 28.4 introduces two special WRI aspects that depend on the quality of interface, riding comfort and noise. Section 28.4 will deal with the management opportunities that are provided by the weigh-in-motion system, Quo Vadis, which is installed all along the Dutch rail network. Section 28.5 concludes the chapter and recommends areas for further research.

28.2 Optimising rail maintenance

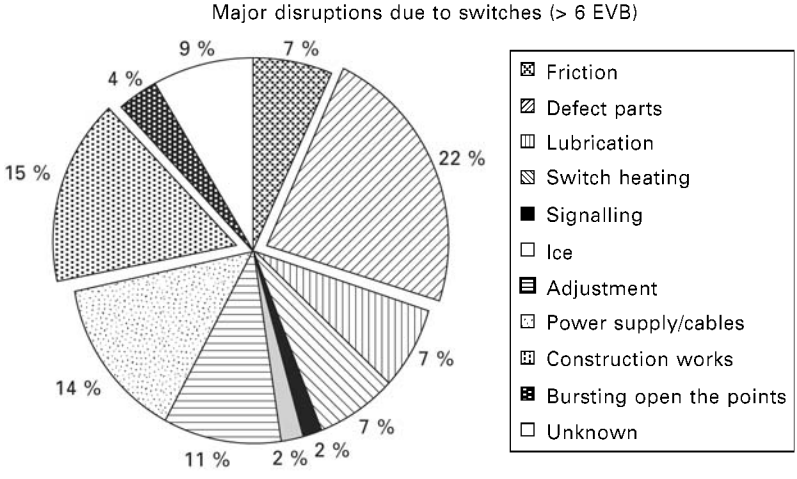
Tracks, turnouts and crossings belong to the crucial assets of each Infrastructure Manager. Failing tracks (e.g. rail breaks and defect insulation joints) and turnouts make train traffic impossible or require significantly reduced speeds (e.g. 40 km/h). Problems in the stability of the track are not uncommon in the Dutch (sub)soil conditions, particularly in the western part of the country. Moreover, the assets have a replacement value of more than 8 billion Euros and consume annually around 60 % of the total maintenance costs and 75% of renewal costs, due to their usage-based, relatively rapid deterioration patterns. Details of the Dutch track and rail system are given below. This section focuses on the strategy to manage rail sections with RCF in an optimum manner. RCF is caused by wheel–rail interface problems. The total

yearly costs of all maintenance activities related to the control of RCF are more than 30 million Euros (ProRail, 2005).

The standard track system consists of 54E1 rails on concrete monoblock sleepers and a ballast bed of 30 cm. The new Betuweroute and High Speed Line South are equipped with 60E1 rails. Duoblock concrete sleepers have also been widely applied and are still performing well on mainlines. NP46, an older, less heavy rail type, is still in service on lines with lower traffic loads. Around 5000 km of tracks now have continuously welded rails (CWR), which is standard technology for all mainlines and has been rolled out rapidly in railway yards in recent years in order to reduce noise levels. Turnouts are relatively complex, failure-prone systems, which is why standardisation is being sought. Standard angles for turnouts are 1:9, 1:12, 1:15, 1:18.5 and 1:34.7. Often, special engineering is required on railway yards where limited space is available (e.g. double-slip turnouts).

Degradation and traffic-disrupting failures are a result of a variety of events. Figure 28.1 shows a representative distribution of failure causes for turnouts; some can be classed as to technical defects (e.g. due to wear or track settlement), others are caused by process errors (e.g. inadequate or wrongly carried out maintenance). This shows that improving performance and lowering maintenance cost requires several ingredients. This chapter discusses only the technical approach and projects, not other possible measures to improve WRI management, such as education of operators and service level agreements with suppliers.

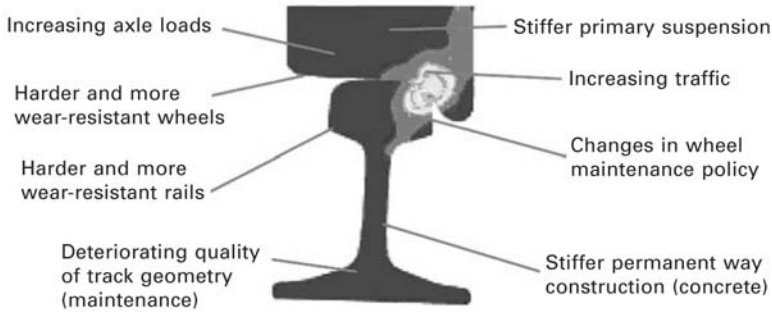
The biggest change in wheel-rail contact since the mid-1990s has been the rise of the RCF phenomenon, beginning with the appearance of head



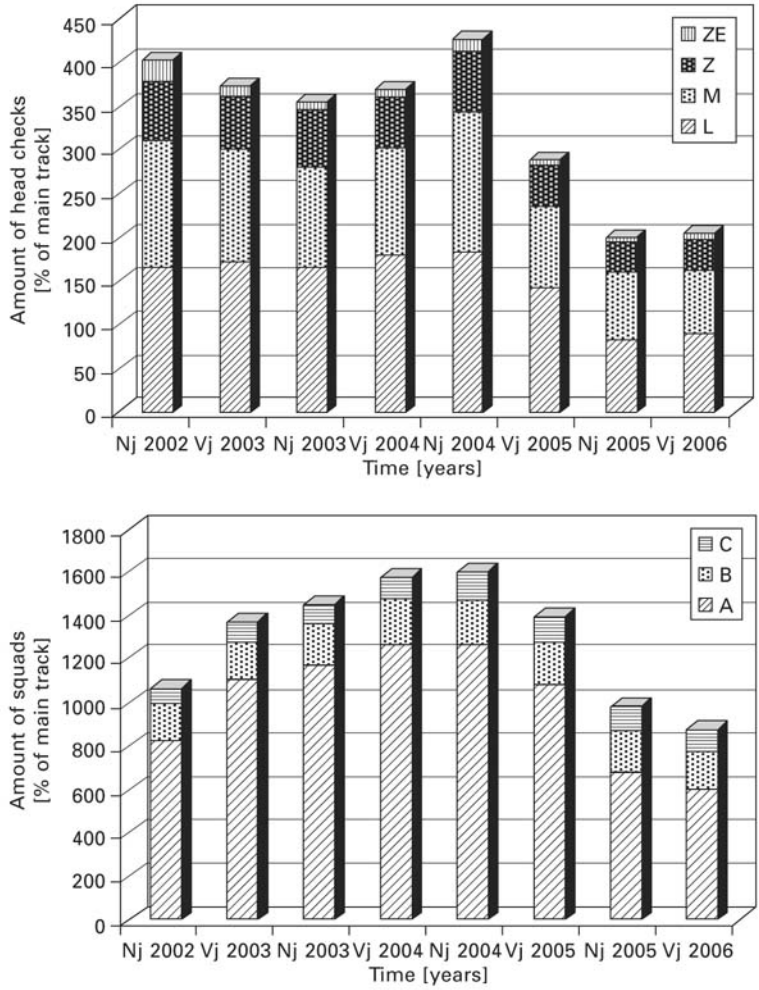
28.1 Typical distribution of causes of traffic disrupting failures for turnouts.

checks, followed some years later by squats. Since 2001, ProRail has been actively investigating the causes of and developing strategies to reduce the RCF problem (Smulders and Hiensch, 2003). It has become clear that the causes for the sudden rise in RCF are multiple. Figure 28.2 shows changes that have occurred on both wheel and rail side, e.g. the introduction of trains with more seating capacity and an 8–10-fold increase in yaw stiffness (e.g. double-deck intercity trains) and the change from rail type NP46 to UIC54.

After the RCF problem had been sufficiently recognised, and the often referred to Hatfield accident in the UK played a role in this, ProRail developed its first strategy and procedures for dealing with RCF. The first type of strategies implemented were aimed at improving detection of RCF problems through increased inspection (e.g. improvements to ultrasonic inspection, the latest development being the introduction of eddy current measurements), in combination with rail parts renewals. Next came the introduction of a cyclic, gradual grinding regime in 2004 (pilot phase) and 2005 (national rollout), with different grinding cycles for different loads and curvatures – directly adopted from Canadian heavy-haul practice (Kalousek and Magel, 1999). To begin with, a number of years was needed to bring the entire network up to a required base condition, after which more limited grinding cycles and depths could be achieved, and this was the situation in 2008. The grinding strategy has been fully employed. After the first two years of grinding, serious discussions took place regarding its effectiveness; doubts were caused by a serious delay in the reduction of head checks. Grinding does not remove the more severe head checks, due to their depth, and parts replacement remains important. Nevertheless, the grinding strategy went ahead as planned, and this led to a 75 % reduction in the most severe category of head checks, as seen from the statistics below (Fig. 28.3): from 25 km infected with severe headchecks in 2002, to 5 km in 2007. Note that in the Netherlands four defect severity classes for head checks and three for squats are distinguished. Table 28.1 gives an overview of the defect severity classes used in visual



28.2 The typical changes that the wheel–rail interface experiences in the Netherlands.



28.3 Statistics of measured head checks and squats 2000–2008. [Nj = Autumn; V_j = Spring] See Table 28.1 for explanation of abbreviation.

inspection. Improvements in head check detection will play a role in the absolute numbers, which makes the success even move marked.

The business case for preventive gradual grinding was strong from the start. More than 3 Euros return is estimated for every euro invested in grinding, considering only the direct RCF-maintenance. However, moving more from the 'stress regime' that causes RCF towards a so-called 'wear regime', where rail life is determined by wear, is not only advantageous for RCF. It also reduces the occurrence of corrugations and increases the lifespans of the rail through better wheel–rail contact and reduced dynamic forces. A fundamental research programme on RCF was put in place Quite

Table 28.1 Overview of applied defect severity classes for RCF in The Netherlands[a]

Type of RCF	Defect severity	Abbreviation	Description
Head checks	Light	L	Surface crack < 10 mm
	Medium	M	Surface crack 10–19 mm
	Severe	Z	Surface crack 20–29 mm
	Extreme	ZE	Surface crack > 20 mm
Squats	Light	A	A simple imprint combined with a black spot
	Medium	B	A 'V-shaped' crack with a black spot on both sides
	Severe	C	Cracks with a length between 20 and 29 mm, railhead is dented, big black spot

[a]If RCF is visually detected a 50 m fragment of rail is considered 'RCF-infected', independent of the amount of RCF in the fragment.

rapidly and has led to the introduction of an Anti-Head Check rail type (Fig. 28.4). The 54E1 Anti-Headcheck Profile is applied in curves ($R < 3000$ m) and, since 2006, in turnouts. Steel qualities are differentiated depending on location, while studies of new types of steel, Bainite, and fundamental research on squats initiation are ongoing (Hiensch *et al.*, 2006; Li, Zhao *et al.*, 2007).

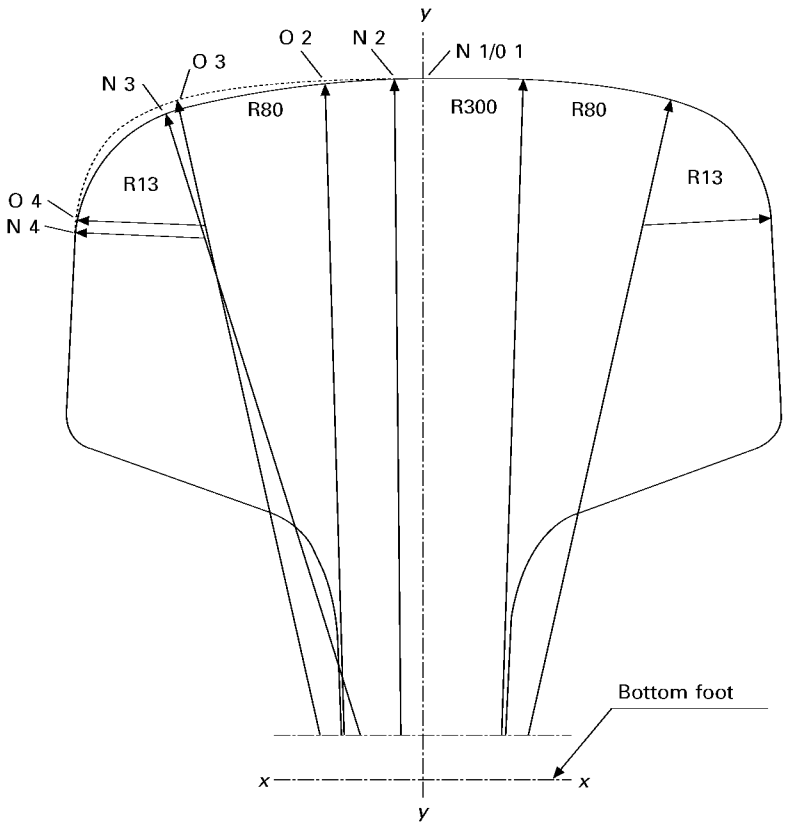
In 2007 and 2008 more steps were taken to 'attack' the WRI head checks problems at source, by avoiding unnecessary loads and frictions on the track. This is where the vehicle side of the interface comes into play, and an application of that is described in Section 28.1.3. The following subsections provide an insight into the composition and optimisation of RCF maintenance plans for the rail infrastructure, managed by ProRail.

28.2.1 Development of tailored RCF maintenance plans

The first pilot scheme for better co-ordination of maintenance RCF activities began in 2004, when about 40 % of the total main track was affected by RCF. Fragmented development of the RCF-maintenance plans; proved to be a significant problem the planning of small, and large-scale maintenance tasks are separate activities within ProRail, while the execution is also realised by different maintenance contractors.

The increase in RCF did not allow such division by responsibilities. ProRail had to re-engineer its RCF-treatment process and define clear roles and responsibilities. Also, decision support systems (DSS) had to be developed in order to deal with RCF in a more cost-effective manner. This means the right maintenance activity at the right time (grinding when possible and renewal when necessary). (The costs of grinding are approximately 5 % of

Anti-head check profile 54E1 (UIC 54)



28.4 The Dutch anti-head check profile 54 E1.

Co-ordinate route rail head and radius

nr	*x*	*y*
N 1	0	159.0
N 2	–3	159.0
N 3	–26.0	155.4
N 4	–35.1	143.6

O 1	0	159.0
O 2	–10.0	158.8
O 3	–24.9	157.0
O 4	–35.0	144.9

N R13	–22.1	143.0
N R80	–2.2	79.0
N R300	0	–141.0

O R13	–22.1	144.2
O R80	–7.3	78.9
O R300	0	–141.0

N is anti-head check profile
O is normal profile

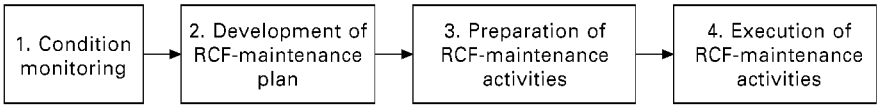
the costs of rail renewal.) The core of the process is the development of tailored RCF-maintenance plans based on general rules of thumb. The next section will elaborate on the re-engineered process.

28.2.2 Process of RCF treatment re-engineered

The main goal of the process is to generate cost-effective maintenance activities in order to mitigate the RCF problem in the Netherlands.[1] Figure 28.5 depicts a high-level process flow diagram for RCF treatment (phase 1). The diagram shows the sequence of steps and their relationships in the process. The process flow starts when the condition of the rails is monitored. In order to collect the necessary data, two types of measurements are carried out. A visual assessment is carried out in order to record visual detectable rail defects. Besides that, an ultrasonic measurement train collects all defects (also non-visual detectable defects). The elements of the track that are not measurable by train (such as points, trackside, etc.) are measured by hand with manual ultrasonic devices. Both types of measurements are carried out at least once a year and up to four times a year, depending on the track classification, according to UIC Leaflet 714 (UIC, 1989).

The main goal of the next process step (phase 2) is to convert the collected data to an RCF-maintenance plan comprising an overview of maintenance activities (grinding versus renewal), bandwidth of execution date and a life-cycle analysis for each line section.[2] This process step is elaborated in the next section.

After the plans are developed, the maintenance activities need to be prepared (phase 3). The maintenance activities are combined to enable a public tender, since ProRail has outsourced all the maintenance activities to contractors. Furthermore, train operating companies are consulted to negotiate feasible maintenance slots. Finally, the maintenance activities are carried out (phase 4). Deviations from the original plan and the intended product quality are assessed. The next section will describe the development of RCF-maintenance plans (phase 2).



28.5 Operational high-level process flow diagram for RCF treatment.

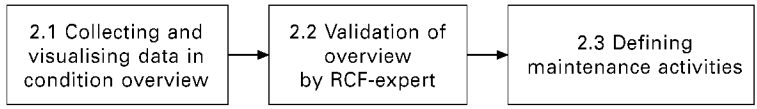
[1]Cost-effectiveness is defined as the relationship between the expenditures (costs) and relative outcomes (effects) of these expenditures.

[2]A *line section* is defined as track between to junctions (i.e. railway stations) or a junction itself.

28.2.3 RCF-maintenance plans

This section focuses on the development of tailored RCF-maintenance plans (phase 2 in Fig. 28.6).[3] Figure 28.6 gives an overview of sub-processes of phase 2 (development of the RCF-maintenance plan). Three main phases are distinguished. In order to support the process of phase 2, we developed a DSS to process the condition data automatically to standardised maintenance activities. The DSS uses a set of 'rules of thumb' for RCF planning, depicted in Table 28.2.

Figure 28.7 gives an illustrative example of the developed form. As shown in Fig. 28.7, the type and severity of the rail fault lead to a specific maintenance activity. For instance, the left rail of the track contains grindable squats (1) in the last months of 2007. The DSS automatically generates a

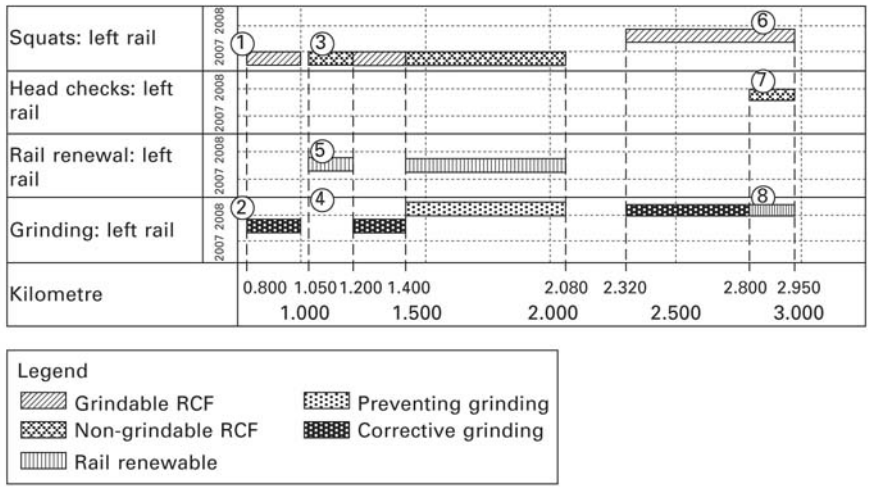


28.6 Distinguished subprocesses of phase 2, development of RCF maintenance plan.

Table 28.2 Rules of thumb for the basis of the developed DSS

Activity	Condition	Realisation period
Corrective grinding	Fault depth of < 2.0 mm	Within six months after detection
Preventative grinding	To remove the top layer of the rail	Within one year after rail renewal
Cyclic grinding	Is based on passed tonnage: – Curved track with a radius < 3000 m → After 15 MGT – curved track with a radius > 3000 and < 9000 m → after 30 MGT – Non-curved track → after 45 MGT	Within three months after passed tonnage
Removing rail corrugation (at least 0.5 mm)		Within six months after detecting corrugated rail
Rail renewal	Fault depth of ≥ 2.0 mm. The challenge is to combine rail faults in big renewal projects in order to reduce cost and the number of welds.	Depends on severity class. A fault depth of > 25 mm needs to be replaced immediatel.

[3]Tailored in this context means that all maintenance activities are based on recent inspection results and not only on generic maintenance concepts.



28.7 Illustrative example of a tailored RCF-maintenance plan.

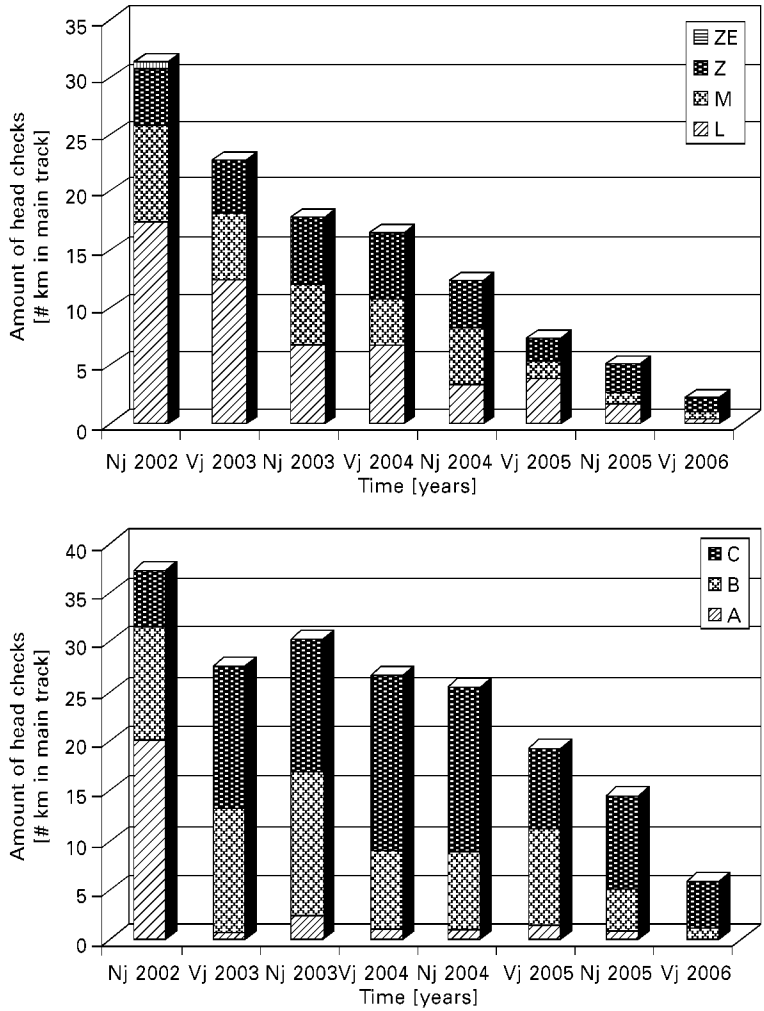
maintenance activity 'corrective grinding' (2) in the second half of 2008. Furthermore, 'non-grindable RCF' (3) leads to 'rail renewal' (5) and 'preventive grinding' (4) within six months after the rail renewal. Finally, a combination of 'grindable RCF' (6) and 'non-grindable RCF' (7) leads to 'corrective grinding' combined with 'rail renewal' (8).

Next, these automatically generated overviews of maintenance activities are validated by means of visual inspections by RCF experts in order to reduce the chances of flaws in the condition monitoring and data processing process. Finally, maintenance activities are defined and grouped in the most cost-effective manner. Subsequently, those packages of maintenance activities are placed in a 'backlog', waiting to be prepared in phase 3.

The result of the implemented re-engineered process in the pilot areas is shown in Fig. 28.8. The figure clearly shows a desirable trend in the amount of head checks and squats in the pilot area (a decrease of 93 and 86 % respectively in three years). It has also led to a reduction in maintenance costs due to the smarter, more timely co-ordination between parts replacement, grinding and overall rail renewal. The re-engineered process will be implemented on a national scale in the coming years.

In summary, we can state that the management and reduction of head checks in the Netherlands is partly under control, thanks to a combination of measures:

- implementing several types of grinding regimes;
- combining required maintenance of rail and (larger) rail replacement projects, i.e. the optimisation of RCF-related maintenance and renewal works;



28.8 Amount of head checks and squats in the pilot area, being the maintenance contract areas Leiden, Rotterdam and Den Haag.

- increased rail quality and an adjusted rail profile; and last but not least;
- a through understanding of its causes through scientific research.

28.3 Optimising wheel maintenance

The last few years have seen a renewed interest on the part of the Dutch Railway industry in the dose relationship between wheel and rail. There is of course a relationship in the obvious sense, the wheel being in constant contact

with the rail, but it is also recognised that similar problems are found in both the rail and the wheel, e.g. RCF. Consequently, changes to the wheel shape, material or maintenance have a significant influence on the rail. Based on this train of thought, the wheel–rail interface has been optimised from both perspectives. The following subsections discuss optimisation examples for wheel maintenance, which contribute both to longer wheel life and a more 'track-friendly' design and maintenance of vehicles. It is based on Vermeij *et al.*, (2008), in which more details can be found.

28.3.1 Alternative wheel profiles

Until 2005, the UIC-ORE s1002 profile was used for all passenger trains in the Netherlands. This profile was developed in the 1970s and, since then, the railway has changed significantly. Heavier axle loads and stiffer bogies have resulted in new problems like RCF. One way to reduce the occurrence of RCF is optimisation of wheel and/or rail profiles. A logical approach would be to optimise the *combination* of wheel and rail profiles. This is such a complex problem, though, that the rail profile for the Dutch track was first optimised separately, based on the principle that the rail shoulder should be relieved. Subsequently, the wheel profile has been optimised based on this Anti Head Checks rail profile.

The goal for the optimised wheel profile was to reduce slip forces and increase spread in contact position. This resulted in a dedicated profile for intercity trains, the so-called HIT-profile. Slow trains are still re-profiled with the s1002 profile, because these trains have lower axle loads and lower bogie yaw stiffness. For the development of the new wheel profile, an efficient combination of practice and theory has been used. It is important to start with a theoretical analysis in order to make sure that only feasible profiles (e.g. stability guaranteed) are tested in practice. Practical analysis is inevitable, because it is impossible to simulate every aspect of the outside world.

The optimised wheel profile, discussed here, has resulted in an increase in wheelset life of up to 30 %. This has mainly been achieved by reducing RCF cracking, but also by reducing flange wear. Because the additional costs of turning a different profile are minimal, the cost reduction is of the same magnitude. The impact of this alternative wheel profile on the reduction of RCF in the railhead is not exactly known. One of the aims of the alternative wheel profile was to reduce the slip forces; this means that what is beneficial for the wheel will also be beneficial for the rails.

28.3.2 Condition-based wheelset maintenance

Traditionally, wheelset maintenance was driven by visual inspection of the wheels during short-term (periodic) maintenance. Since 2002 a major reduction

in wheel defects has been achieved by implementing a wheel impact load detector system: Gotcha®/Quo Vadis, a wheel impact load detector (WILD) system. (See www.gotchamonitoringsystems.com for more information on this system.)

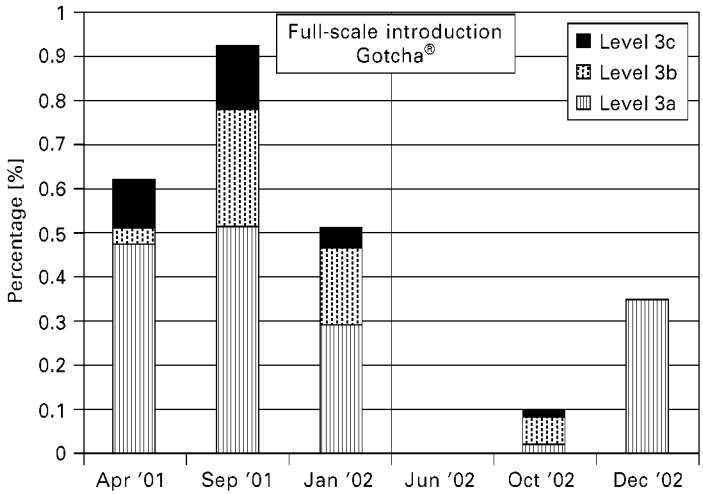
One of the functions of Gotcha®/Quo Vadis is to measure the wheel quality of the trains, at least once per day during normal operation. If the wheel quality falls below a certain level, wheelset maintenance will be planned. Within a predefined time, the train will be sent to the wheel lathe. The maintenance is condition-based and may take place during short-term maintenance (planned) or when wheel defect levels are exceeded (unplanned).

The impact of this approach to wheelset maintenance on the rail is huge. Since the introduction of Gotcha®/Quo Vadis the number of heavy impacts has been reduced significantly, resulting in less excessive loads on the track. In turn, this results in a higher track life. A sample of the reduction of heavy impacts is shown in Fig. 28.9. Other functions of the system are discussed in Section 28.5.

28.3.3 Changing re-profiling intervals

The (condition-based) maintenance of wheelsets in the Netherlands was aimed to maximise re-profiling intervals. However, this does not necessarily result in the most cost-effective maintenance regime, nor in an optimum situation for the rail.

Based on a life-cycle cost analysis for wheelset maintenance the *scraping principle* has been introduced. The scraping principle belongs to the preventive maintenance category. During each short-term maintenance, all wheels are



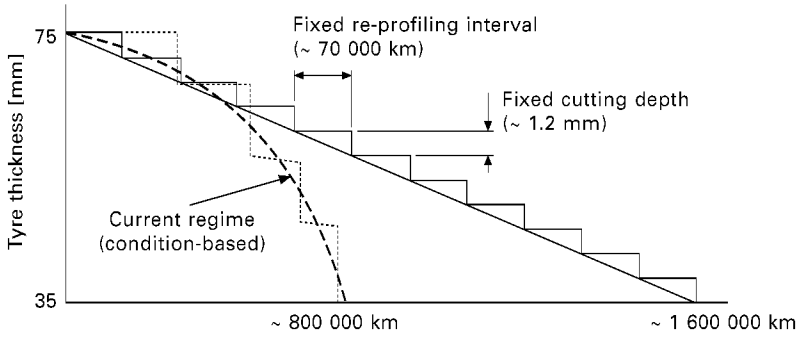
28.9 Impact of WILD driven wheelset maintenance on high impacts.

re-profiled, with a cutting depth of around 1 mm. The result is that wheels remain round and, consequently, the dynamic load during their life-cycle is lower. Defect initiations like small cracks and pitting are removed at an early stage. Relatively large cutting depths due to damage accumulation are prevented. Using a small cutting depth, the work-hardened layer is not removed, and this results in a slower development-of-out of roundness. The scraping principle is shown in Fig. 28.10.

In the first quarter of 2006, a field test was begun with the scraping regime. Two VIRM double-deck intercity EMUs (24 axles, average axle load of 14.3 tonnes) and four ICM single-deck intercity EMUs (16 axles, average axle load of 12.3 tonnes) were maintained according to the scraping principle. Both vehicle types are shown in Fig. 28.11.

In Fig. 28.12 the defect value growth for a single wheel, as measured by Gotcha, is shown. In February 2006, this wheel was introduced into the scraping regime. With condition-based maintenance, the defect value increases in time, while after introduction of scraping the defect value continues to be at a low level.

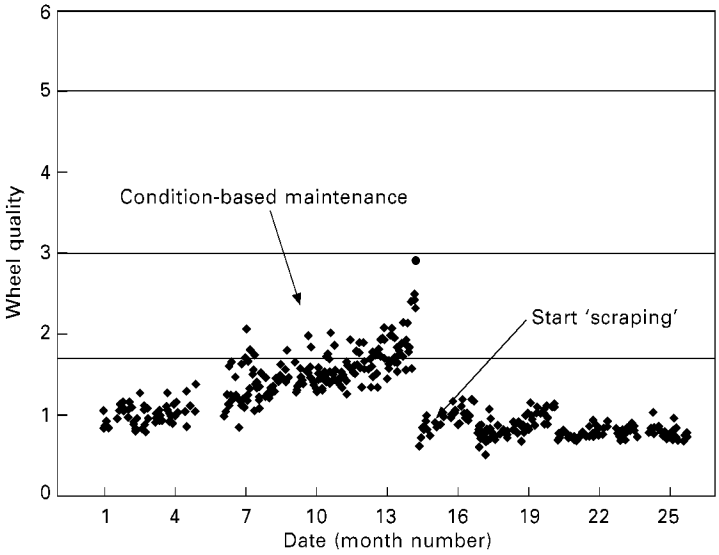
This scraping regime is now being used successfully for several train types and the feasibility of its introduction for other train types is being



28.10 Scraping principle (dashed line represents current regime).

(a) (b)

28.11 VIRM double deck intercity (a) and ICM single deck intercity (b).



28.12 Defect level growth for single wheel before and after introduction of scraping.

investigated. The scraping principle has one huge advantage for the rail: the dynamic load decreases significantly due to the increased smoothness of the wheel surface. This, together with the use of WILD systems, is keeping impacts due to wheel defects to a minimum.

28.4 Special aspects in optimising wheel–rail interface: riding comfort and noise

Two specific aspects are discussed in this section, the management of riding comfort and noise. It shows two examples where sophisticated knowledge of vehicle–track interaction and friction management are needed to develop predictive models.

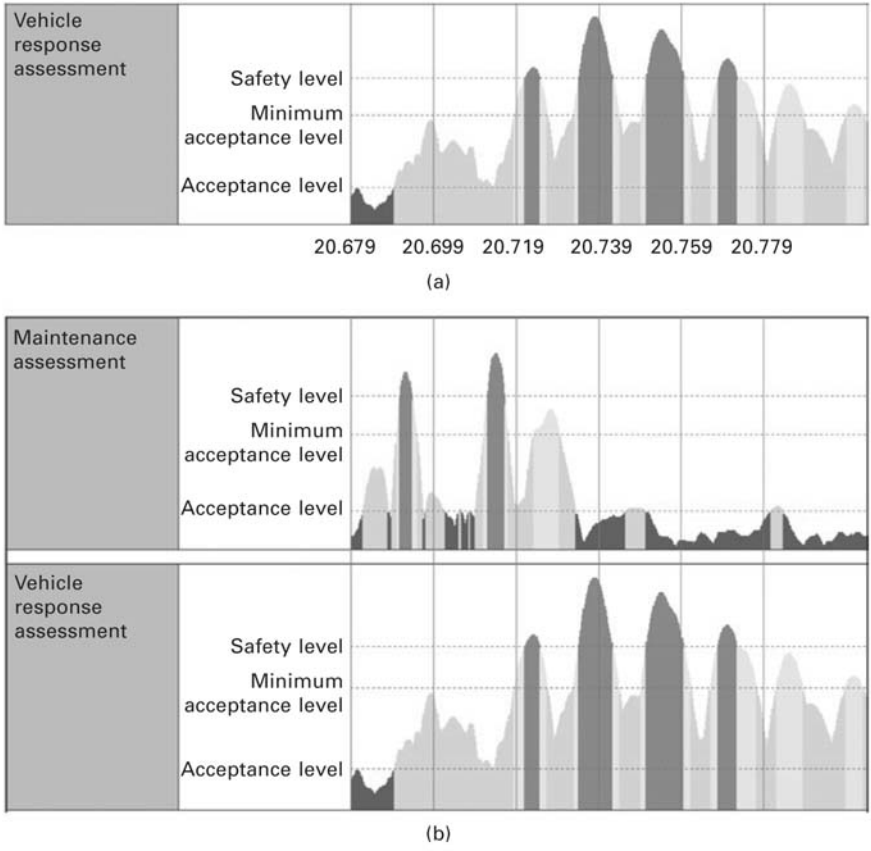
28.4.1 Management of riding comfort through IRIS/pupil

For some years, ProRail has been developing and using the IRIS system, provided by Erdmann Consulting and fed by data from the Eurailscout inspection coaches, to manage its geometry for the purpose of safety. However, results from vehicle response analysis (VRA) could never be related directly to track geometry. This was added around 2005 through a new system linked to IRIS, called PUPIL. The development has been done in accordance with the latest European CEN and UIC regulations. A new

set of track geometry standards has been derived from UIC regulations regarding vehicle response.

PUPIL makes it possible to assess, in accordance with the European standards (CEN and UIC), the forces and accelerations which result from combined and harmonic deviations. The assessment method is based on recently developed calculation models (using parameters of standard vehicle types) and presents the results in a user-friendly manner.

A practical example will show the analysis that is made possible by PUPIL. In Fig. 28.13a and b, a horizontal axis represents a part of a train journey over the track. The broken horizontal lines indicate the various quality levels; that is, the comfort and safety of the vehicle as expressed in its response data. At the beginning of the journey, the response is within the permitted quality standards. However, the graph also shows that – at this



28.13 An example of PUPIL output: (a) vehicle response assessment; (b) comparison between vehicle response assessment and maintenance assessment

specific speed – the safety level is reached and even exceeded at several points later in the journey.

If we now compare the vehicle–response assessment with the maintenance assessment of the track geometry, we see that the latter peaks just before the former (see Fig. 28.13b). It is thus clear that the minimum acceptable safety levels are being exceeded and that this affects the vehicle response. Following analysis of the associated geometrical parameters – track height, cant difference, lateral track displacement and so on – an overall assessment can be reached based on all factors.

This method enables us to make objective, quantified statements about the relationship between the quality of the track geometry and the vehicle's responses at various speeds. Consequently, it has also become possible to verify track geometry standards objectively and quantitatively, based on the analysis of vehicle responses. More information on track geometry management in general at ProRail and the management of riding comfort by PUPIL can be found in Ginkel (2005).

28.4.2 Railway noise

Railway noise is another special aspect. While trains are an environmentally-friendly mode of transport, they produce a significant amount of noise. In those areas most affected, railway noise can be a big problem. In a densely populated country such as the Netherlands, noise is an important aspect to manage in capacity management and infrastructure projects. Strict noise legislation results in capacity limits, or even bans on specific types of railway traffic on specific lines. The interaction between wheel and rail is one of the primary sources of noise.

For this reason, government-funded innovation and investment programmes have been established. This has led to several new products used in the track for noise reduction, such as 'silent bridges' with embedded rails, prefab rail-attached dampening devices and trackside lubrication devices on yards, all being rolled out in the last five years. Research and investments related to new applications are still ongoing, and one example related to railway squeal is discussed in the next sections.

To be able to facilitate the expected and desired growth of traffic on the Dutch tracks while remaining within noise limits, ProRail has co-ordinated the Dutch Innovation Program Noise (IPG). The IPG was executed between 2001 and 2007, and was aimed at testing and implementing new source-related and cost-effective measures. In cooperation with passenger and freight operators, different solutions have been tested to reduce the noise emission of the railway system. Implementing these new solutions is the goal over the coming years. For this purpose, ProRail has started the Railway Noise Knowledge Centre to collect and disseminate information on noise

aspects, also in international collaboration with parties as UIC, DB and SNCF.

Another instrument is the noise-dependent track access charging system that has been developed within the IPG. Operators can receive a reduction in the track access charge if they use less noisy rolling stock. After a few years, a penalty for the use of noisy rolling stock will be implemented. The system was introduced by ProRail in 2008.

The noise emission level of conventional railway traffic is dependent to a large extent on the roughness of both the wheels and tracks. A major issue in the IPG was the objective to reduce the wheel roughness of cast iron braked rolling stock. This aspect has been identified throughout Europe as the most effective way to reduce noise emission. Much effort has been put into testing K- and LL-brake blocks with promising results.

With respect to infrastructure, the IPG has produced several projects. For instance, the introduction of rail dampers has been instigated. At the time of writing (summer 2008), around 50–100 km of track has been fitted with two different types of rail dampers. Another project involves the introduction of a rail roughness monitoring system. With this system, measures to reduce noise can be implemented, after which ProRail is allowed to increase the noise capacity of railway lines and allow more trains on the infrastructure.

28.4.3 Squeal

In addition to rolling noise, curve squeal (a high-pitched 'screeching' noise produced by trains when negotiating narrow-radius curves and switches) is a major source of local noise nuisance caused by railways. The origin of curve squeal has been reported in several papers (Transportation Research Board, 1997; Kalousek and Magel, 1999; Beer *et al.*, 2003). Squeal is generated from lateral slip between wheel and low rail. Further, it is understood that squeal occurrence is dependent on the friction behaviour in combination with the size and direction of the forces acting in the contact patch; more specifically, the presence of the so-called 'stick–slip loop. Since high levels of lateral slip mainly occur in narrow-radius curves, squeal is a phenomenon related to curves in this category. For most railways, narrow curves typically are defined to have a radius below 500 m in metro and below 100 m in tramway.

Curve squeal is characterised by extremely high noise levels (up to 120 dB), resulting in a major discomfort to local residents, passengers and train operators. With train traffic increasing, it is becoming more and more difficult to comply with national noise legislation and licence, leading to bottlenecks in track capacity distribution resulting a decline in the efficiency of the rail business.

From the UIC final report on combating squeal noise (Oertli *et al.*, 2005), reporting on a friction modifier benchmark investigation, it is understood

that to date there is no optimal solution for the problem of curve squeal. In this report, asymmetric rail profiles were recognised as a promising effective mitigation measure, but this concept was not pursued.

Squeal involves a so-called 'on-off' system; it is either there or not. The chance of squeal occurring can be reduced for a given site by taking preventive measures. The higher the 'instability' of a given site, the more frequently a situation will arise in which squeal occurs. The purpose of preventive measures is to make the instability of a site as low as possible, so that squeal noise can be expected only in extremely unfavourable conditions.

The occurrence of squeal is of a discontinuous nature. This is caused by the variation in contact positions, forces, friction, profiles, track gauge, etc. Weather conditions (air humidity) also have a considerable influence on whether or not curve squeal occurs.

28.4.4 The Anti-Squeal Rail

Early in 2006 a project started within the national Dutch noise reduction programme 'IPG' (*Innovatie Programma Geluid*) aiming to develop and introduce a sustainable solution to prevent curve squeal within the Dutch rail infrastructure. From this research, it was understood that the occurrence of railway squeal noise can be prevented in a sustainable and maintenance, free way through the optimisation of wheel–rail friction behaviour and wheel–rail contact position. This has been achieved by introducing a new rail surface design resulting in a 'Anti-Squeal Rail'.

The technique of particle impregnation is applied to achieve this innovation. In this process, material particles, selected to optimise the interface to reduce squeal noise and wear, are impregnated directly into the rail running surface resulting in a sustainable solution. The anti-squeal rail is introduced as the 'low rail' in the curve, the high rail retaining a standard profile. The introduction of new interface materials provides the unique opportunity to combine surface friction design with a special anti-squeal rail profile design. Squeal noise prevention combining surface friction and profile optimisation can produce a quiet and stable system. In addition to the paper presented in Istanbul in 2007 (Hiensch *et al.*, 2007), a further paper was presented at the World Congress on Railway Research (WCRR) in May 2008 in South Korea (Hiensch and Lammers, 2008).

28.4.5 Rail friction optimisation

To manage friction behaviour in a sustainable way, we need to move away from the traditional steel–steel interface. Surface optimisation can take place through particle impregnation, resulting in a PI-rail®. In this impregnation process, material particles selected to optimise the interface to reduce squeal

noise and wear are impregnated directly, without mechanical pre-treatment, into the rail running surface resulting in a sustainable solution. Hard wear-resistant 'islands' which are load-bearing in the contact surface with the wheel are produced, dominating the friction behaviour. The rail impregnation development process has been performed in close cooperation with Swedish company Duroc AB.

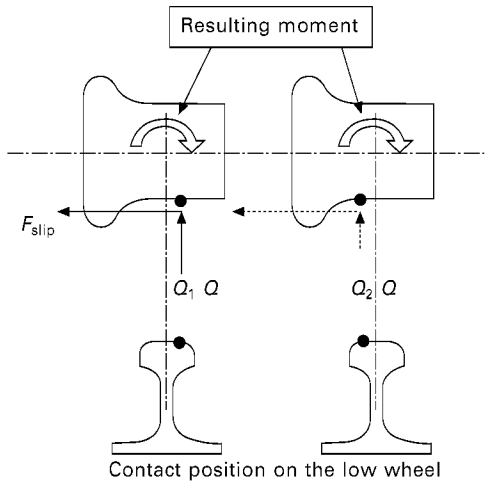
Materials have been selected not only from a functional but also from a production technology point of view. Titanium carbides (TiC) and tungsten carbide (WC) have been selected for their ability to provide a load-bearing and wear-resistant surface. Boron nitride (BN) has been selected to further enhance friction behaviour. BN is often referred to as 'white graphite'. It is a lubricious material with excellent thermal stability. After the introduction of unleaded fuel, BN was introduced into the automotive industry as a valve seat insert material to take over the lubrication functionality of lead. It is also used in brake linings of Formula I racing cars to reduce squeal noise during braking.

28.4.6 Rail profile optimisation

From earlier research and publications in the area of profile optimisation, it was recognised that by using asymmetrical profiles a reduction can be obtained in the frequency of occurrence of curve noise. To decrease squeal instability, rail profiles should reduce lateral slip, position the contact of the wheel on the low rail towards the flange and stimulate longitudinal slip.

The efficiency of this asymmetrical 'anti-squeal rail profile' is also affected by the contacting wheel profile and the running properties of the rolling stock. Design and effectiveness of an anti-squeal rail profile therefore partly depend on the local track and rolling stock characteristics. The development of the squeal-optimised rail profile is therefore supported by train–track simulations (calculation of wheel–rail contact positions, occurring slip, stress levels). The model requires both the track characteristics (such as curve radius, cant of the track, installation gradient, permissible speed) and material data that affect the contact position between wheel and rail (such as the wheel profile and yaw stiffness of the wheel set) as input. The design principles of the anti-squeal rail profile have been validated using a twin-disc test rig. This is presented in Hiensch *et al.* (2007).

Twin-disc tests performed by DeltaRail BV confirmed that the position of contact on the wheel at the low rail strongly influences squeal instability. The generation of squeal could be controlled by moving the contact position (see Fig. 28.14). Situation 1 (solid arrow) is generating squeal noise during twin-disc testing. Due to the lateral slip, E_{slip}, a resulting moment is enforced on the wheel. When the contact between wheel and rail is positioned on the outside of the wheel (situation 1–) the resulting moment will contribute



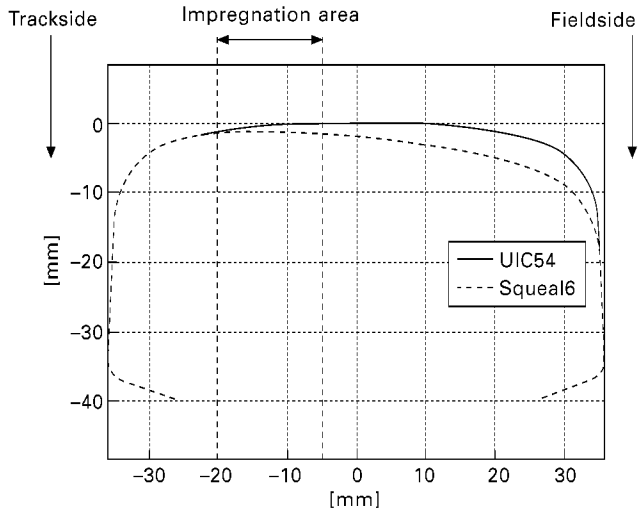
1 : Increase F_{slip} → Increase Q → Increase in squeal instability
2 : Increase F_{slip} → Decrease Q → Increased squeal instability

29.14 Influence of the contact position between wheel and low rail on squeal instability (the squeal test rig generated squeal only for situation 1. For situation 2 squeal remained absent).

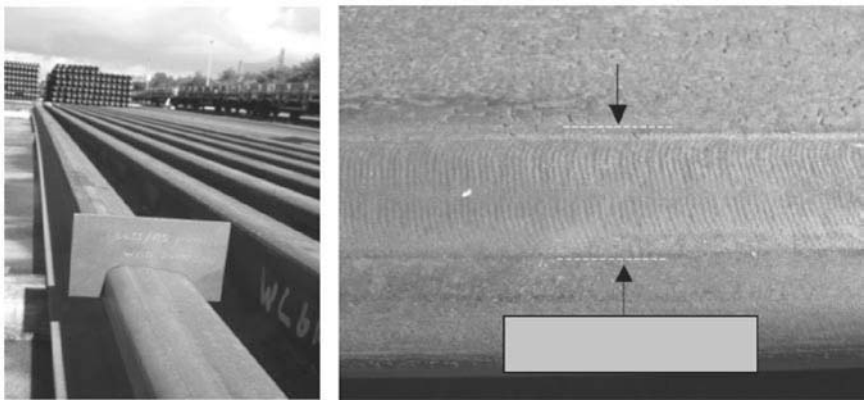
to the loading of the contact increasing the vertical force ($Q1 > Q$). This will result in an aggravating and unstable process of increased lateral slip values followed by an increased resulting moment. Wheel vibration will occur followed by squeal generation. For situation 2 (dashed arrow), the contact position is towards flange side of the wheel. The resulting moment is relieving the contact ($Q_2 < Q$), avoiding an unstable process and decreasing squeal instability. These validated insights were the starting point for further simulation work, resulting in the design of the anti-squeal profile as shown in Figs 28.15 and 28.16. This work has been published in detail in Dirks and Wiersma (2007).

28.5 Monitoring traffic movement: Gotcha/Quo Vadis

Gotcha®, the NS system for inspecting wheel defects, was introduced in Section 28.1.3. It is automatically linked with the Quo Vadis system by ProRail which, since 2004, has measured all details on the static and dynamic axle-loads of passing trains. It is installed on about 40 locations on the network, in such a way that 80 % of traffic movements and 96 % of tonne kilometres are measured. (Each measuring location is automatically calibrated, using known 'standardised weights' from electric locomotives travelling over the network.) This section highlights some of the functions that such a so-called



28.15 The proposed anti-squeal rail profile for the modelled situation – low rail: the area of wheel – rail contact to be impregnated prior to field testing is indicated.



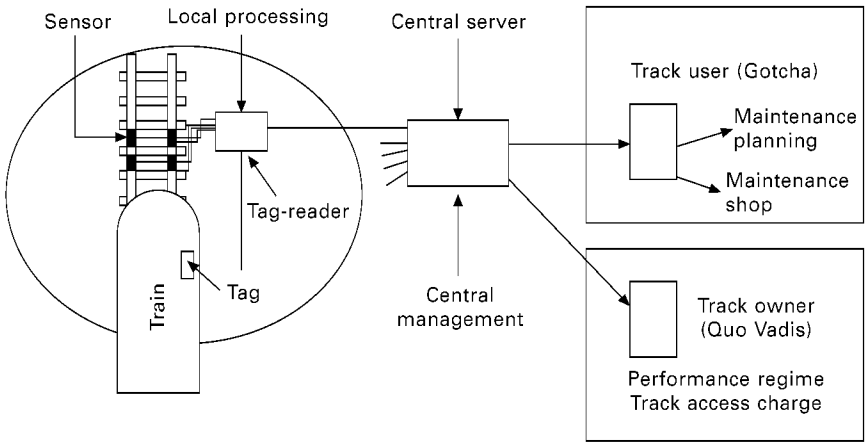
28.16 Developed prototype ready for installation.

'weigh-in-motion' system can fulfil for WRI management and immediate action in wheel-rail interface problems. These functions were described more in detail in a publication in the *European Railway Review* (Zoeteman and Buurman, 2006).

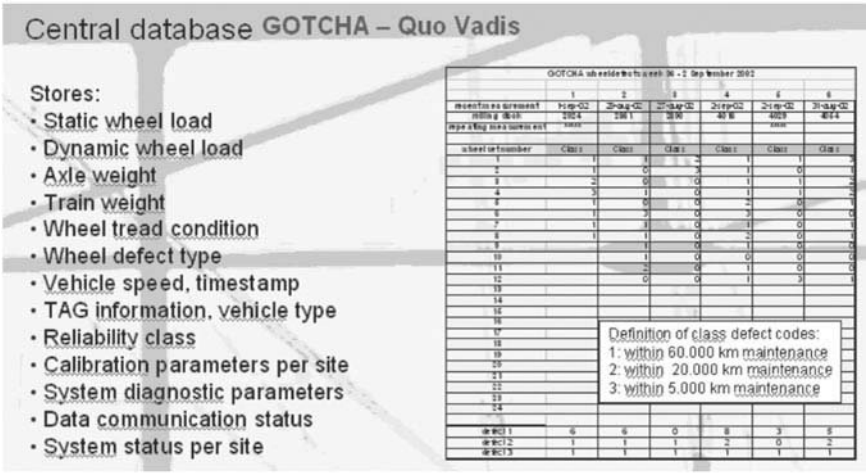
Quo Vadis's sensors are mounted under the rail and connected with a reader in a trackside box. The reader generates an optical signal that is transmitted to the sensor over a fibre optic cable. The sensor converts the (minute) vertical deflection of the rail as a result of the passing wheel into a change in the optical signal. This signal is converted in turn by the reader

into a precise electrical signal that is available for further processing (see Figs 28.17 and 28.18). Figure 28.17 shows that the raw data are collected locally and transferred to a central ProRail server after each train passage using GSM-R/GPRS. Next, they can be used in the specific applications. Quo Vadis's measurements are also available on the Internet, where operators can easily track their haulage and access charging.

Quo Vadis has proved its use in infrastructure charging and in wheel maintenance planning. First, ProRail gained a much better and more up-to-date insight into the traffic performance taking place on the network. The data from Quo Vadis together with the information from the train log files



28.17 Overview of Quo Vadis/Gotcha.



28.18 Data available in a central database.

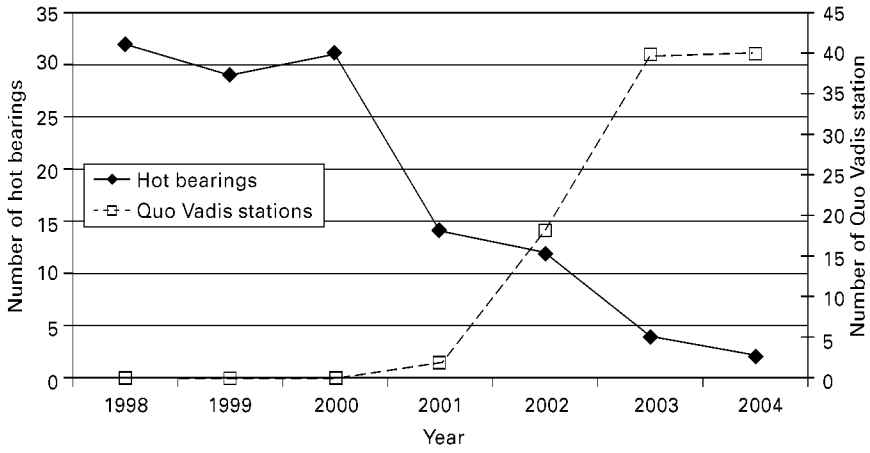
provide information about the actual use of all infrastructure elements. It is now known for each line section how often, and with what load, trains run over it. This is also known for each set of automatic turnouts, even for each direction and position.

Several applications and benefits were also found in infrastructure maintenance and rationalisation of infrastructure layouts, using the aggregated traffic data from Quo Vadis. First of all, information derived from the measurements was directly input to more refined usage, and condition-based inspection and maintenance strategies (e.g. resulting in a 10 % reduction in costs of visual inspections). We will not illustrate this here extensively, as this has been included in Zoeteman and Buurman (2006).

Thanks to the Quo Vadis/Gotcha® measurements, wheel quality has increased noticeably in recent years, while repair costs have fallen. This has been demonstrated by data collected by NS and ProRail. Measuring degradation of wheels early is beneficial for several reasons. First, it allows prioritisation in repairs and proactive scheduling (thus optimising the operation of the underfloor wheel lathes). Second, it reduces the use of material while repairing wheelflats. Finally, it improves wheel quality and (predictability of) rolling stock availability. This is particularly important in autumn, when slippery rails cause problems on the railway network. The most important result has been a drastic 90 % reduction in the number of broken springs and hot axles, largely thanks to Quo Vadis (Fig. 28.19).

28.6 Conclusions

Looking back, it can be stated that changes since 2000 have led to a new situation on the Dutch railway track, requiring a renewed interest in wheel–rail



28.19 Progress in hot bearings reduction (NS) and Quo Vadis roll-out.

interface management. ProRail's *rails* are subjected today to a 'stress regime' according to the definition of the International Heavy Haul Association, with high friction and loads from the operation. This results in cracks and premature loss of rail life due to RCF (IHHA, 2001). Once damage occurs, it will accelerate track degeneration. This can be avoided by grinding to introduce artificial wear and moving more towards a 'wear regime', where initial head checks do not 'survive' and do not have the potential to grow into deep defects. This is the reason why the Anti-Head Check profile and the cyclic grinding regime have been highly successful so far in improving wheel–rail contact, reducing RCF and, at the same time, corrugations and extending rail life.

The interaction of *wheels* and rail initiates the forces and thus the stresses. Although infrastructure management and operations management have been separated in the last decade, in order to liberalise the railway market, it is becoming more and more accepted that on the vehicle-side too design and maintenance should be 'track-friendly', which is beneficial for both wheel and rail life. In the end, it is much more effective than any measures taken at the infrastructure side, when the first signs of RCF appear.

The discussed successes in head check reduction, and wheel profile optimisation, discussed and the management of the specific interaction aspects of *noise and riding comfort* clearly show the potential of WRI research in contributing to day-to-day improvements. However, overall maintenance and renewal costs for tracks remain hard to manage and tend to increase, not least because of ever more stringent labour safety and noise regulation. Solutions do not remain stable, as the railway is an open system, e.g. new types of vehicle will enter after the European TSIs will come into force. ProRail acknowledges therefore the requirement of continuous research for optimal WRI management.

ProRail's research programmes, together with Dutch universities and partners in the rail sector, will focus in the coming years on defining optimum quality levels for track components, short wave irregularity research, inspection technology (e.g. automatic inspection of squats) and the management of rail friction, i.e. control of slippery rails in the autumn. This project, called AdRem, has not been included in this chapter; more information is available in Arias-Cuevas *et al.* (2009).

Weigh-in-motion technology such as Gotcha/Quo Vadis provides vital data collection for implementing the above strategies and should become a standard element in the toolkit of any infrastructure manager. It allows the development of accurate 'footprints' of rolling stock quality. Of course, all of these developments should be part of a total asset management approach, which aims to improve infrastructure performance and reduce life-cycle costs. More information on ProRail's asset management approach can be found in Zoeteman (2006, 2007).

28.7 References

Arias-Cuevas O., Z. Li, R. Lewis and E. A. Gallardo-Hernandez (2009), Rolling-sliding laboratory tests of friction modifiers in leaf contaminated wheel–rail contacts, *Tribology Letters*, **33**, 97–109.

Beer de F. G., Janssens M. H. A. and Kooijman P. P. (2003), Squeal noise of rail-bound vehicles influenced by lateral contact position, *Journal of Sound and Vibration*, **267**, 497–507.

DeltaRail (2006), *RCF Maandrapportage Oktober 2006*, Utrecht, the Netherlands.

Dirks B. and P. Wiersma (2007), Asymmetric rail profile to prevent railway squeal noise (ref. 618155), Proceedings InterNoise, Istanbul, Turkey, 28–31 August, paper in -07-272.

Ginkel W. van (2005), Innovative solutions to old problems, *European Railway Review*, Issue 4, Russell Publishing, Brasted, UK.

Hiensch E. J. M., Dinks B., Horst J. and van der Stelt J. (2007), Rail head optimisation to reach a sustainable solution preventing railway squeal noise, *Proceedings InterNoise*, Istanbul, Turkey, 28–31 August, paper in-07-271.

Hiensch E. J. M., Horst J. J., Wiersma P. K., Van der Linden F. L. J., Lapidaire B. and Dollevoet R. P. B. J. (2006), Relationship between track geometry and the development of rolling contact fatigue damage, *Proceedings 7th World Congress on Railway Research*, Montreal, Canada, 4–8 June, on CD.

Hiensch E. J. M. and J. W. Lammers (2008), Preventing railway squeal noise through railhead optimisation, *Proceedings 8th World Congress on Railway Research*, Seoul, South Korea, 18–22 May, on CD.

IHHA (2001) *Guidelines to Best Practices for Heavy Haul Railway Operations: Wheel and Rail Interface Issues* International Heavy Haul Association, IHHA, Virginia Beach, VA, USA.

Kalousek J. and E. Magel (1999), Noise control at the interface of rail and wheel, *Railway Track and Structures*, 1 August.

Li. Z., X. Zhao, C. Esveld and R. Dollevoet (2007), Rail stresses, strains and fatigue under dynamic wheel-rail interaction, *Proceedings International Heavy Haul Association Specialist Technical Session*, Kiruna, Sweden, 11–13 June, 389–96.

Oertli J. *et al.* (2005), *Combating Curve Squeal, Phase II*, Final Report, August International Union of Railways, Paris, France.

ProRail (2005), '*Beheerplan Prorail 2006*', Utrecht, the Netherlands.

Smulders J. and M. Hiensch (2003) RCF management and research program in the Netherlands: approach and solutions to control the wheel–rail interface to reduce RCF damage, *Proceedings 6th World Congress on Railway Research*, Edinburgh, UK, 28–30 September, on CD.

Transportation Research Board (1997), *Wheel/Rail Noise Control Manual*, TCRP Report 23, Transportation Research Board, NRC, National Academy Press, Washington DC, USA, 167–71.

UIC (1989), *UIC Code 714: Classification of Lines for the Purpose of Track Maintenance*, International Union of Railways, Paris, France.

Vermeij I., Liefting en, G. and Bontekoe T. (2008), Optimisation of wheelset life time with a focus on increasing wheel tyre life, *Proceedings 6th World Congress on Railway Research*, Seoul, South Korea, 18–22 May, 28–30 September, on CD.

Zoeteman A. and G. den Buurman (2006), Quo Vadis: a vital instrument in asset management, *European Railway Review*, Issue 3, Russell Publishing, Brasted, UK.

Zoeteman A. (2006), Asset maintenance management: state of the art in the European railways, *International Journal of Critical Infrastructures*, **2**, (2/3), 171–86.
Zoeteman A. (2007), ProRail's management of tracks and turnouts, *European Railway Review*, Issue 3, Russell Publishing, Brasted, UK.

Copyrighted Materials
Copyright © 2009 Woodhead Publishing Retrieved from www.knovel.com

INDEX

Index Terms	Links		
A			
AAR 1:20	690		
AARlB profile	690		
AAR1B wheels	675	682	
Acid Bessemer rail steel	365	366	
acoustic efficiency	480		
acoustic grinding	598		
procedures	489		
ADAMS/Rail	87		
adhesion	53	514	737
adhesion coefficient as function of velocity	517		
creep curves during testing with applied contaminants	55		
friction coefficient of wheel-rail interface	737		
measured results and discussions of friction coefficient	738		
and friction modification	510		
friction modifiers	56		
influence of electrification and non-electrification on friction coefficient	738		

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links
adhesion (*cont.*)	
influence of passing axle number on friction coefficient	
electrified line	739
non-electrified line	739
selected non-electrified line	740
maximum traction coefficient and rolling speed	
at different water temperatures	742
for two kinds of surface roughness under cold and hot water conditions	743
maximum traction coefficient and surface roughness	745
measured data and adhesion coefficient design curve in Japanese railways	740
rail disc surface roughness	741
slip ratio and traction coefficient	
at different water temperature	742
for three kinds of surface roughness	745
surface roughness profiles of rail disc before experiments	744
traction and creep in wheel	54
wheel and rail under wet conditions	739

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links		
adhesion (*cont.*)			
surface roughness effect	744		
water temperature effect	741		
AdRem	815		
Adtranz *see* Bombardier Transportation			
aerodynamic noise	478		
air-borne particles, from wheel-rail contact	550		
alarm limits	259		
low-level limits in different countries	260		
trains exceeding load magnitude	261		
aluminothermic welding	356	379	396
Alvsjo test track	520		
Alzheimer's disease	566		
American Public Transportation Association	691		
American Railway Engineering and Maintenace-of-way Association	672	687	
American Railway Engineering Association	687		
Amtrak	691		
anisotropy	225		
Anti-Head check profile	797	803	815
Anti-lock Braking System	410	416	
Anti-Squeal Rail	810	811	
Archard equation	189		

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links			
Archard's Wear Law	98	116	122	272
see also Holm's Wear Law				
Arlanda Airport	554	557	558	559
	568	571		
Arlanda Central	552	557	558	560
	568	571		
Arlanda South	557	558	559	560
	561	562	568	
ash	563			
asset condition monitoring train	654			
Association of American Railroads	293	675	687	
ASW-MK2	682			
asymmetric grinding	672			
austenitic steel	143			
Australian Railway Systems	760	763	789	
automatic train control	642			
Automatic Train Operation	653			
automatic transformers system	649			
Automatic Truck Identification	641			
axis density	714			

B

Bainite	797			
bainitic steel	151	157	163	784
comparative properties	158			
experimental steel microstructures	155			
non-UK specifications	161			
TTT curve for ferrite-pearlite	154			
UK specifications	159			

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links			
ball-on-disc test device	513			
Banverket	608	621	635	649
bar gauge	589			
Barber-Sheffel bogies	676			
barium sulphate	556			
barytes	556			
Bessmer process	5			
BoBo vehicle	638			
body force method	295			
bogie-on-roller rig	515			
Bombardier Transportation	636			
booted sleeper corrugation see trackform-specific corrugation				
boron nitride	811			
Botniabanan line	642			
boundary element method	297	480	705	
boundary lubrication	98	448		
Boussinesq-Cerruti formulae	192	193		
brakeshoe metal pickup	680			
braking dust	556			
breaking distance	517			
breaking weight	517			
Brinell hardness	671	688		
British Rail	366	445		
British Standard BS4490	320			
British Steel Corporation	131			
BU roller rig	176			
BV50	642	651		

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links	
C		
Canada’s freight railroads, wheel shelling	680	
NRC anti-shelling wheel profile vs standard AAR1B wheel shape	683	
relation to brakeshoe metal pickup	681	
Canadian National Railway	669	682
Canadian Northern Railway	669	
Canadian Pacific Railway	669	670
AAR common wheel grades	671	
degree curvature vs radius in metres	671	
family of eight rail profiles for wheel/rail interaction improvement	673	
management methods to control rail and wheel wear	672	
frame-braced bogies	676	
friction management	673	
harder and cleaner rail steels	672	
preventive rail grinding	672	
track materials	674	
wheel profiles	675	
North American rail classes	671	
rail wear for different high rail lubrication strategies	674	

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links	
Canadian Pacific Railway (*cont.*)		
seasonality of shelled wheel removals	681	
top-of-rail friction management program effect on lateral forces on low rail	675	
Canudas de Wit *et al.* friction model	107	
carbon manganese		
carbon content on tensile strength and ductility	140	
pearlite fraction on yield stress	139	
carcinogen	567	
carrier frequency	389	
Carter's creep coefficient	78	
Cartier Railway Company	676	682
cellulose	513	
cementite	342	
Central Japan Railway Company	712	
chemical mass balance	558	
cobalt-based alloys	523	
Comité Européen de Normalisation	22	
computer-aided tomography	333	
computer simulation	84	
contact conditions and creep forces	84	
patch size and shape	85	
current computer packages	87	
vehicle dynamics	86	

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links
constant damping loss factor	503
CONTACT	37 79 81 191 421
contact analysis	74
creep forces	77
traction coefficient	77
fundamental wheel-rail contact quantities	75
numerical approaches comparison	78
comparison of contact area during curving	84
contact area vs wheelset lateral positioning	82
contact pressure comparison	82 84
curving cases -profile pairing	83
Kalker	78
overview of approximate methods	80
three principally different approaches	81
contact filter effect	482 483
contact modelling	62
boundary conditions	64
contact between two bodies	63
normally loaded contacts	65
contacting bodies and elliptical contact area	66

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links
contact modelling (*cont.*)	
elliptical contacts correction factors	69
Hertz equation	68
Hertzian theory	65
line and circular contact	67
maximum shear stress magnitude and location	69
numerical methods	68
pressure cell division	71
pressure distribution	65
tangentially loaded contacts	71
analytical solutions for line and circular contacts	71
numerical methods	73
sphere and plane contact	72
contact patch	60
contact position plot	61
S 1002/UIC60 non-elliptical contact patches	62
contaminants	
definition	437
discussion	450
wear and traction coefficient for various contaminants	452
wear data and test conditions	452
effect on wear, fatigue and traction	437

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links
contaminants (*cont.*)	
effect on wheel-rail isolation	536
friction modifiers	448
effects of crack truncation	450
effects on creep curve	449
isolation caused by leaves between contact load and voltage across the contact	440
leaves	438
oil or grease	447
effects on rolling contact fatigue	448
lubricating film formation	447
solid particulates	440
abrasive wear by particles	444
brittle particles entrainment and succession fracture	441
crushed granite particle	443
crushed sand	443
effect of sand on traction coefficients in twin-disk contact	444
ejection of sand particles from twin-disc contact	441
granite particles distribution before and after testing	442
isolation by sand particles	445
particle entrainment into the wheel-rail contact	440

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links			
contaminants (*cont.*)				
rail and wheel images				
after 6000 cycles of operation				
granite particle contamination	442			
traction sand and effect on				
friction	443			
wheel and rail wear rates				
under conditions of sand				
Contamination	445			
wheel-rail conductance variation				
with sand concentration in				
contact	446			
water	445			
influence of humidity on dry and				
leaf contaminated rail	446			
liquids and rolling contact				
fatigue	447			
reduced traction	445			
continuously welded rail	377	409	619	794
corrective grinding	693	801		
corrugation	231	646	725	728
	815			
see also rail corrugation				
Dang Van equivalent stress	233			
factors in rail corrugations				
generation	737			
first low rail corrugations formed on				
new rail	729			

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links	
corrugation (*cont.*)		
kinetic COF on top surface of rail	735	
lateral force/vertical force of low rail measured at track site	729	
long pitch corrugation in tangent track in salty-environment tunnel	732	
factors influencing the maximum traction coefficient	735	
plastic flow on corrugated rails	734	
longitudinal plastic flow on rail surface layer	734	
progress rate variation of corrugations with radius of curvature	731	
rail surface layer microstructure in lateral direction	730	
in rails	646	
Shin-Kanmon Channel tunnel ascending slope		
formed corrugation	732	
observed rail corrugations location	733	
short pitch corrugations in curved attacks	725	728
causes of corrugations	731	
influential factors in causing corrugations	728	
lubrication	731	
state of top surface of rail	736	

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links	
corrugation (*cont.*)		
typical short-pitch corrugation formed on low rail surface	728	
waveforms and its phase lag of rail corrugations	733	
X-ray diffraction analysis results	736	
Corrugation Analysis Trolley	599	
Corus	165	
Coulomb friction model	100	
and viscous friction model	102	104
Stribeck friction model	104	
Coulomb's law of friction	74	193
coupled boundary element model	501	
crack growth laws	298	
biaxial specimen	299	
crack growth mechanisms	285	
fluid-assisted	286	
fluid entrapment	287	
hydraulic pressure transmission	287	
shear driven crack growth	286	
squeeze film fluid action	287	
phase II	288	
rail bending	289	
residual stress	288	
thermal stress	288	
crack truncation	450	

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links	
creep	514	573
creep force	462	495
creepage	462	495
critical sand flow rate	539	
critical speed	462	500
curve squeal	477	494
reducing squeal noise	496	
squeal noise generation		
mechanism	494	
creep force-creepage		
relationship	495	
forces acting on wheels of bogie in		
a curve	494	
cyclic flattening	662	
cyclic grinding regime	815	
D		
Dahl friction model	107	
Dankowicz friction model	106	
Canudas de Wit et al. friction		
model	107	
combined Coulomb, Stribeck, viscous		
and Dankowicz friction		
model	108	
Dahl friction model	107	
DARTS-NL package	387	
DeCAyS *see* Deterioration Cost		
Associated with the Railway		
Superstructure		

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links	
decibels	478	
decision support system	797	
DeltaRail BV	811	
Derby wear index	199	
Deterioration Cost Associated with the Railway Superstructure	624	
computational tool	617	
track data	618	
data entered	618	
reference information	619	
track geometry data on Swedish national railway network	618	
vehicle and operational data	619	
DIFF	271	
DIFF3D	271	
digital computed radiography	333	
discrete point reacting springs model	482	
distance to go radio signalling system	654	
DR-2 bogies		676
dry friction *see* Coulomb friction model		
dry wear tests	165	
ductility	138	
Duroc	165	
Duroc AB	811	
Dutch HSL-South	398	399
Dutch Innovation Program Noise	808	

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links			
Dutch Rail Infra Manager ProRail	400			
dye penetrant	579			
dynamic absorber	486			
dynamometer	786			
E				
eddy current	433			
measurements	795			
Edmonton Transit, wheel life and ride quality	679			
elastic shakedown limit	334	336		
elastohydrodynamic lubrication	447	450	703	740
elastoplastic rolling contact mechanics	4			
electro-solution polishing	714			
electrodiscrete element method	548			
electromagnetic brakes	568			
electron microscopy	556	562		
EN 14 363	615			
energy dissipation mechanisms	405			
equivalent conicity	458	459	461	462
	463	464		
ER7	8			
ERRAC *see* European Rail Research Advisory Council				
Euler method	117			
European Rail Research Advisory Council	5			
European Rail Research Institute	412			

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links
European Railway Agency	24
European Standard EN 13306:2001	649
European Standard EN 13450	648
European Standard EN 50126	649
European Traffic Management System	642
Excel	617
F	
Facility for Accelerated Service Testing	293 363 688
false flange	586
FAST *see* Facility for Accelerated Service Testing	
FASTSIM	37 79 193 199 273
fatigue	49
interaction with fatigue	52
rail and wheel	49
crack growth phases	50
crack opening from fluid pressurisation	51
head check cracks at gauge corner	50
surface fatigue damage	51
fatigue impact	234
Federal Railroad Administration	687

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links			
ferrite	342			
fibre-optic sensors	259			
field tests	254			
contact force history time	258			
measured sleeper bending moment at rail seat	256			
strain gauge-based wheel impact load detector for wheel-rail contact force measurement	254			
vertical wheel-rail contact force measurement	255			
finger gauge	593			
finite element and dislocation based methods	296			
finite element method	483			
finite element model	335	390	501	788
finite element simulation	387			
flaking	342	768		
flange back wear	235			
flange lubricant	438			
flange reaction	732			
flange reaction force	747			
flangeway clearance	588			
flanging noise	496			
flowability	695			
fluid crack pressurisation	447			
flush-butt welds	356			

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links
focused ion beam milling	332
Fourier analysis	478
Fourier transform infrared absorption spectroscopy	736
fretting fatigue	216
friction	93
and adhesion modification	510
future trends	524
coefficient of friction, adhesion and braking distance	513
breaking distance as function of adhesion coefficient and varying slope	519
examples of rail-wheel adhesion coefficients	516
friction coefficients measured with salient system tribometer	516
relationship between tangential force and creep at the wheel-rail contact	515
schematic figure with forces acting on braking vehicle	518
schematic of a wheel	514
definition	513
leaf residue and blackish layer on top of rail	511

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links	
friction (*cont.*)		
measured coefficient of friction vs sliding distance for leaf-lubricated and unlubricated tests	512	
modification	519	
friction modifier behaviour	521	
high-pressure washing of rails	522	
ideal friction coefficients in wheel-rail contact	520	
possible models for low friction at wheel-rail contact	523	
proposed multi-layer low-adhesion model on steel surface	524	
friction models	99	104
Coulomb and viscous friction model	102	104
Stribeck friction model	104	
Dankowicz friction model	106	
Canudas de Wit et al. friction model	107	
combined Coulomb, Stribeck, viscous and Dankowicz friction model	108	
Dahl friction model	107	
pure sliding and oscillating contacts	100	
1 DOF system	100	
Coulomb and viscous friction		

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links
friction models (*cont.*)	
forces as function of sliding speed	101
Coulomb friction model	100
random process	108
small displacements	105
viscous friction model	102
step response simulation of 1 DOF system	103
friction moderator	704 749
friction modifier	7 56 201 363 438 448 449 497 519 532 679 695 780
application methods	
mobile applicators	532
onboard applicators	532
wayside applicators	532
effects of application	780
reduction in lateral forces in high rails	781
reduction in whee/rail noise levels	781
high coefficient modifier	521
low coefficient modifier	520
potential future benefits	782
solid, effect on wheel-rail isolation	541

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links
friction modifier (*cont.*)	
disc arrangement and friction modifier applicator	543
friction and contact impedance during 3 % slip test	543
measured static impedance values for wheel-rail discs	544
very high coefficient modifier	521

G

galling	43	488			
Gallivare ore	637				
gauge comer cracking	363	768			
general elastic contact mechanics	4				
GENSYS	81	87	89	200	617
Global System for Mobile Communications-Railways	642				
Gotcha	805				
GotchdQua Vadis	804	812			
Grand Trunk Pacific Railway	669				
graphite	750				
Green’s functions	297				
grinding depth	708				
grinding interval	708				
grinding pattern	594				
ground-borne noise	497	505			
from trains in tunnels	501				

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links		
ground-borne noise (*cont.*)			
effect of tunnel structure	502		
force transmissibility of hysteretically damped single degree of freedom system	502		
insertion loss at 25 m for different resilient track types	504		
vibration field around the unlined tunnel at 100 Hz	501		
ground vibration	497		
from surface railways	498	500	
spectra of trains measured at house 50 m from the track	500		
vertical velocity level at Ledsgård for train speed of 70 kmi/h	499		
vertical velocity level at Ledsgård for train speed of 200 kmi/h	499		
vibration phenomena overview	497		
GSM-WGPRS	814		
GWM 550 grinding train	399		
H			
Hadfield manganese steel	143		
Hall-Petch type equation	138	140	148
hand-held tribometer	522		
hand-pushed rail tribometer	515		
Harsco Track Technologies	693		
Hatfield accident in UK	795		

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links			
head checks	49	282	284	421
	644			
internal head check geometry	285			
in rails	644			
visual appearance	284			
heat-affected zone	382			
heavy-haul corrugation	353	354		
cause	355			
characteristics	353			
on low rail of curve on mixed traffic line	354			
treatment	356			
hematite	566			
Hertz analysis	37	711		
Hertz contact	4	704	705	
Hertz equation	68			
Hertz pressure distribution	336			
Hertz spring	423			
Hertz theory	61	65	190	
non-elliptical contact patch and multi-Hertzian approximation	191			
Hertzian deflection model	493			
Hertzian maximum pressure	741			
Hertzian stress analysis	291			
high-carbon, hyper-eutectoid steel rail	133	164		
high-cycle fatigue	334			
high-density polyethylene	523			
high positive friction	542	679		

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links		
high-pressure water-jets	522		
HIT-profile	803		
hollow wear	586		
Holm’s Wear Law	98		
hunting	462	579	778
hyper-eutectoid steels	784		
hysteretic damping model	503		
I			
impact noise	493		
independently rotating wheels	458		
INFRASTAR	18		
infrastructure costs models, for wheel-rail interface	608		
computational tools and input data	616		
computational tool DeCAyS	617		
lateral wheel-rail forces and wear numbers simulation	617		
track data	618		
vehicle and operational data	619		
conclusions and future trends	624		
examples of results	622		
deterioration cost for different axle loads	623		
deterioration per gross ton-km expressed as cost elasticity	624		
model calibration and costs	620		
calibration procedures	621		
principles	620		

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links		
infrastructure costs models, for wheel-rail interface (*cont.*)			
notation	626		
purposes	609		
track deterioration model	609		
vertical dynamic wheel load model	628		
Ingot casting	5		
inhalable particles *see* air-borne particles			
INNOTRACK	14	19	21
work process	20		
Innovative Programma Geluid	810		
INTERFACE	99		
International Heavy Haul Association	697	815	
International Union of Railways	488		
IORE	636		
IRG transition diagram	8	98	
IRIS system	806		
iron ash	563		
iron oxides	737		
iron oxyhydroxides	737		
ISO 2631	497	500	
ISO 3095	598		
isotropic hardening	335		

J

Jenkins formula	391		
Jernbaneverket	635		

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links		
Jiang fatigue parameter	238		
JIS 60 kg rail	708	717	721
K			
K-brake block	488	809	
K-Industrier	636		
Kalker's creep coefficients	79		
Kalker's theory	421		
kinematic hardening	335		
kinetic friction force	513		
Kita-Kyushu tunnel	737		
klingel wavelength	663		
KTH Machine Design	95		
L			
laser triangulation technique	590		
lateral creep force	728	730	
leaves			
effects on wheel-rail isolation	541		
leaf thickness against break-down load	542		
voltage reading against load for single dead leaf layer	542		
light-rail corrugation	356		
cause	357		
characteristics	356		
measurement	357		
treatment	357		

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links		
lignin	513		
linear elastic fracture mechanics	225	294	
lipping	339		
Liseberg station	552	558	
LKAB *see* Luossavaara-Kiirunavaara Aktiebolag			
LL-brake block	26	488	809
load cycle	215		
London Underground Ltd	655		
London Underground Victoria Line			
Europe Metro experience in managing wheel-rail interface	653		
identified wheel-rail interface problems	662		
rail damage caused by concrete contamination	665		
transverse profile cyclically flattened low rail	663		
introduction to and historic wheel-rail interface issues	653		
lubrication management	659		
lubrication monitoring system front page	661		
peak equivalent rail roughness over time	662		
upgrade	654		
wheel-rail interface monitoring	654		

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links	
London Underground Victoria Line (*cont.*)		
monitoring diagram	656	
regular track measurements	655	
TRV trace from scope and benefit review report	657	
variance in wheel profile distribution over time	658	
longitudinal creep force	728	730
Loram Inc	693	
low-alloy hyper-eutectoid rail steels	672	
low coefficient of friction	679	
low-cycle fatigue	334	
lubricant	449	450
lubrication	731	745
adverse effects	763	
angle of attack of leading axle	749	
benefits	762	
brake test results	752	
effect of friction moderator		
on decreasing lateral force on curve	751	
on reduction of sound pressure level and COF of low rail on curve	752	
effect on vehicle-track interaction	745	
interacting forces between wheel and rail when the vehicle is negotiating sharp curves	747	

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links
lubrication (*cont.*)	
lateral force and angle of attack of leading axle	747
friction moderating system	749
effect on decreasing lateral force and squealing noise	750
friction moderator	749
influence on braking distance	750
influence on track circuit	753
and friction modification	695
important aspects that require further investigation	763
lateral forces interacting between wheel and rail in leading axle	748
location of microphone set up at track site to measure sound power emitted by train running	751
measured results of COF	746
shunting performance of friction moderator and sand	753
technical details of track and lubrication arrangements	746
typical interacting forces from wheel to rail in bogie negotiating sharp curves	749
and wear of rail/wheel	760

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links
Luossavaara-Kiirunavaara Aktiebolag	633
four-axle ore wagon with three-piece bogies	639
iron ore operation, split between the North and South circulation	638
new locomotive IORE	638
objectives	635
shipping harbours	635
M	
MacMeter ride quality measurements	680
magnetic particle inspection	579
magnetite	566
Malmbanan (Swedish ore line)	
car properties	639
electrical power system	649
general description	634
infrastructure configuration	642
locomotives and cars	636
maintenance, wheel-rail interaction	651
total maintenance cost in percent for 132 km of Malmbanan	652
track maintenance practices	649
track structure	648
train control system	642

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links		
Malmtrafik i Kiruna AB	635		
Malmtrafikk AS	635		
Manual for Railway Engineering	687		
martensite	153		
Martin process	5		
massive bainite *see* martensite			
maximum Hertzian strain subsurface	318		
MEDYNA	88		
Metro-Cammel 67 Stock trains	653		
mild wear	94	110	
milling	594		
Miner's rule	336		
MiniProf	59	301	
instrument	582	590	593
measured profiles	60		
rail and wheel profile	60		
mixed lubrication *see* partial elastohydrodynamic lubrication			
mobile flash-butt rail welding	379		
mobile lubricators	202		
mode shapes	483		
moving irregularity model	486		
multi-pass grinding	582		
multiple sclerosis	566		

N

National Research Council Canada	676		
negative damping	495		
Netherlands Railways	792		

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links	
noise	480	
see also specific type of noise		
non-destructive testing	650	
non-linear Hertz contact model	388	
non-steady-state contact mechanics	366	
NRC-ASW	676	682
NUCARS	87	688

O

one-third octave bands	479	
Open Hearth rail steel	365	
order spectrum	251	
radial deviation from mean wheel radius vs circumferential coordinate	253	
ORE model	610	
ORES1002 profiles	460	
ORES1002/UIC60	194	
out-of-round railway wheels	245	
see also wheel out-of-roundness		
classification and quantification	246	
discreet wheel tread defects	246	
instrument for measurement	250	
measurement	250	
order spectrum	251	
rolling contact fatigue	247	
roughness level spectrum	251	

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links	
out-of-round railway wheels (*cont.*)		
tread braking	249	
wheel flats	246	
wheel polygonalisation	248	
discrete wheel tread defects	252	
alarm limits	259	
field tests	254	
wheel impact load detectors	257	
simulation of consequences	267	
feedback loop	273	
long-term irregular wheel wear	271	
mathematical models and computer programs	267	
spectral density of normal contact force vs wave number κ	269	
train speed on maximum wheel-rail contact force	272	
wheel roughness from tread braking	261	
rolling noise	264	
roughness reduction strategies	266	
thermoelastic instability and tread braking	262	
ovalisation	215	216
Oxygen rail steel	365	

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links			
P				
P1 force	389	390		
P2 force	355	358	389	390
P2-force	615			
P2 resonance	765			
Paris law	226	707		
Parkinsonism	566	567		
partial elastohydrodynamic lubrication	98			
partial least square method	558			
pearlitic rail	131	136		
brittle cracks	317			
chemical compositions	316			
disc material properties for SUROS tests	323			
effect of surface micro-roughness on pressure distribution in twin-disc contact	341			
evolution and failure of pearlitic microstructure	311			
FIB milling and SEM-imaging through single initiating micro-cracks	332			
higher strength grades	141			
lamellae structure for 220, 260 & 350HT grade rail steel	141			
material response to cyclic loading	334			

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links		
pearlitic rail (*cont.*)			
modelling	333		
material failure after ratcheting	336		
materials response to cyclic loading	334		
microstructural evolution by ratcheting	337		
rolling contact fatigue crack initiation	340		
shakedown limits	335		
wear by ratcheting	339		
wear-fatigue interaction	342		
nano-hardness results for R84 and RN samples	330		
nomenclature	344		
observations of microstructural evolution and failure	312		
experimental observations	317	322	327
key microstructural features summary	329	331	
observations from rail	312	315	317
ongoing metallurgical work	331		
periodic unit cube of Voronoi-generated microstructure	338		
plain carbon and low-alloy structures from air-cooling	132		
carbon and manganese content influence on C-Mn steel			

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links	
plain carbon and low-alloy structures (*cont.*)		
strength	139	
carbon content influence on tensile strength and ductility	140	
iron-carbon phase diagram	133	
isothermal ‘time-temperature-transformation’ diagram	137	
PE ferrite and pearlite railhead microstructures	135	
rail production	131	
tache ovale type fatigue fracture	131	
ratcheting wear simulation	343	
representation of pearlitic microstructure	339	
surface and subsurface micro-hardness results	319	
two head check cracks following the angle of plastically sheared rail steel	340	
work-hardening of structures	142	
pectin	513	
phenol resin	750	753
phosphorus	523	
PI-rail	810	
piezoviscosity	447	
pin-on-disc machine	512	515
pin-on-disc test	113	

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links			
pinned-pinned mode	485			
pinned-pinned resonance corrugation see roaring rails/pinned-pinned resonance corrugation				
plain contact fatigue	216			
planned visual inspection	650			
plastic bending	351	371		
plastic flow	351	371	646	
plastic ratcheting	334	337		
plastic shakedown limit	334			
$PM_{2.5}$ see air-borne particles				
PM_{10} *see* air-borne particles				
pneumonia	567			
Poisson's ratio	365			
polycrystals	337			
polytetrafluoroethylene	142	523		
positive insertion loss	503			
preventive grinding	593	673	677	693
	701	801		
preventive maintenance	649			
ProRail	792			
Dutch experience in managing the wheel-rail interface	792			
monitoring traffic movement	812			
data available in central database	814			
overview of Quo Vadis/Gotcha	814			
progress in hot bearings reduction and Quo Vadis roll-out	815			

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links	
ProRail (*cont.*)		
optimising rail maintenance	793	
amount of head checks and squats in pilot area	802	
changes that the wheel-rail interface experience in The Netherlands	795	
development of tailored RCF maintenance plans	797	799
distribution of failure causes for turnouts	794	
Dutch anti-head check profile	54	
E1	798	
high-level process flow diagram for RCF treatment	799	
overview of applied defect severity classes for RCF in The Netherlands	797	
RCF-maintenance plans	800	
RCF treatment re-engineered process	799	
rules of thumb for the basis of developed DSS	800	
statistics of measured head checks and squats 2000-2008	796	
sub-processes of RCF maintenance plan development	800	
tailored RCF-maintenance plan	801	

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links
ProRail (*cont.*)	
optimising wheel maintenance	802
alternative wheel profiles	803
changing re-profiling intervals	804
condition-based wheelset maintenance	803
defect level growth for single wheel	806
impact of WILD driven wheelset maintenance on high impacts	804
scraping principle	805
VIRM double deck intercity and ICM single deck intercity	805
special aspects in optimising wheel-rail interface	806
Anti-Squeal Rail	810
developed prototype ready for installation	813
influence of contact position between wheel and low rail on squeal instability	812
proposed anti-squeal rail profile for modelled situation	813
PUPIL output	807
rail friction optimisation	810
rail profile optimisation	811

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links	
ProRail (*cont.*)		
railway noise	808	
riding comfort through IRIS/pupil management	806	
squeal	809	
Prud'homme limit	465	
PUPIL	806	
Q		
Qhull	337	
Quebec Cartier Mine-Heumann	677	
Quo Vadis	793	
R		
R260	9	
radiation efficiency	480	
radiation ratio	480	
rail bending	289	
schematic representation	290	
rail corrugation	349	
characteristics, causes and treatments	352	
conclusions and recommendations	370	
corrugation analysis trolley	373	
corrugation classification	350	
effect on vehicle dynamics	467	
further investigation	767	
heavy-haul corrugation	353	354
cause	355	

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links
rail corrugation (*cont.*)	
characteristics	353
on low rail of curve on a mixed traffic line	354
treatment	356
light-rail corrugation	356
cause	357
characteristics	356
treatment	357
main types	764
long-pitch	765
short-pitch	764
mechanism	351
damage mechanism	351
wavelength-fixing mechanism	350
other P2 resonance corrugation	358
causes	360
characteristics	358
treatments	360
P2 resonance and pinned-pinned resonance corrugation on metro system	358
P2 resonance corrugation on a tram system	359
roaring rails/pinned-pinned resonance corrugation	364

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links	
roaring rails/pinned-pinned resonance (*cont.*)		
cause	365	
characteristics	364	
treatment	366	
rutting	360	
cause	362	
characteristics	360	
treatment	363	
short-pitch and long-pitch rail corrugations	765	
short wavelength corrugation from two sites on a metro system	359	
trackfom-specific corrugation	368	
cause	369	
characteristics	368	
treatment	370	
tunnel wall acceleration level before and after rail grinding	469	
Rail Corrugation Analyser	599	
rail grinding	776	778
rail grinder	778	
steps to take to ensure cost effectiveness	779	
establishing clear objectives	779	
establishing grinding strategies	779	

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links		
rail grinding (*cont.*)			
establishing standards	779		
improving grinding			
efficiency	779		
technical/economic objectives	778		
rail rectification	650		
rail renewal	801		
rail replacement	650		
Rail Safety and Standards Board	512		
rail steel metallurgy	125	782	
austenitic rails for switches and			
crossings	143		
bainitic rail	151	157	163
comparative properties	158		
experimental steel			
microstructures	155		
non-UK specifications	161		
TTT curve for ferrite-pearlite	154		
UK specifications	159		
designs	786		
evolution and failure of pearlitic			
microstructure	311		
heat-treated rail types	783		
historical review	125		
lower carbon bainitic/martensitic			
rails	785		
main dimensions of heavy			
duty rail sections	786		

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links	
rail steel metallurgy (*cont.*)		
mechanical properties and material types	782	
modified rail sections	788	
pearlitic microstructure in plain carbon and heat treated rails	783	
pearlitic rails	131	136
higher strength grades	141	
plain carbon and low-alloy structures from air-cooling	132	
rail production	131	
work-hardening of structures	142	
rail development	126	131
non-UK rail steel specifications	129	
UK rail steel specifications	127	
rail material developments	164	
high carbon, hyper-eutectoid rail steel	164	
two-material rail	165	
rail sectional dimensions	787	
rail steel types	784	
bainitic (and martensitic) steels	784	
hyper-eutectoid steels	784	
wear and rolling contact fatigue of pearlitic rail	145	147

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links	
rail steel metallurgy (*cont.*)		
BS11 rail maximum strain distortion	148	
gauge side of a worn BS11 rail	149	
pearlitic rail lamella structures and the microstructural alterations	150	
RCF cracks in low-carbon BS11 rail and high-carbon 260 grade rail	146	
void formation in 260 grade rail	153	
white etching constituent formation	152	
welding rail	143	
alumino-thermic welding	144	
flash-butt welding	144	
manual metal arc welding	145	
Rail Surface Damage model	610	611
rail surface fatigue	280	
crack branching predictions	299	
crack growth rate calculation	293	
fracture mechanics	294	
Mode I, II, III	294	
stress intensity factor calculation methods	295	
experimental investigations	290	
field and full-scale test track		

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links	
rail surface fatigue (*cont.*)		
investigations	292	
laboratory investigations	290	
twin-disc system	291	
phases in crack growth	281	
rail wear	300	
wear-fatigue interaction	302	
rolling contact fatigue	282	
fatigue damage types	283	
head checks	282	284
squats	282	
tongue lipping	283	
underlying mechanisms	285	
rail wear	300	
see also wear		
wear-fatigue interaction	302	
crack growth predictions	304	
crack growth rate plots	303	
rail cross section of phase II crack growth	302	
rail welds	377	
and damage formation	380	
deteriorating geometry from non-uniform shakedown and work hardening along rail weld in HAZ	385	
qualitative behaviour of steel hardness distribution along rail surface	384	

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links
rail welds (*cont.*)	
squat initiation on flash-butt rail welds or in HAZ	387
squat initiation on thermite rail welds or in HAZ	386
weld and parent material and the HAZ after thermite welding and grinding	383
dynamic wheel-rail contact force peak value as function of the vertical peak deviation and maximum absolute inclination	395
at rail welds dependence on train velocity	391
at welds dependence on maximum absolute inclination	396
geometry assessment, Dutch rail welding regulations	395
acceptance levels according to traditional method vs QI method	401
cumulative distributions of maximum absolute first derivatives on Dutch HSL-South	399
ground rail weld and geometry assessment using digital straightedge	403

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links
ground rail weld and geometry (*cont.*)	
screen output of digital measuring device	402
speed-dependent norm values determination in QI definition	400
vertical geometry along in-situ made rail weld	397
weld with indentations due to non-uniform shrinkage after welding and grinding	398
weld with pronounced step due to bad rail end alignment	398
weld with smooth surface but with negative height coordinates	398
inclination norm values adopted by ProRail for QI determination	401
and the resulting 2 m wave amplitude	401
irregularities, energy considerations and deterioration	402
maxima of power input into the track due to P1 for a ramp with a basis of 25 mm	406
power input into track per train wheel	404

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links		
welding irregularities and dynamic effects in wheel-rail interface	384	386	
computed time histories of contact force for two measured welds	394		
dynamic peak forces for rail joints and welds	392		
relationship between maximum dynamic contact forces and velocity	393		
types of deterioration of track system	389		
wheel-rail system dynamic response to artificial rail weld with irregularities	388		
welding processes	378		
aluminothermic welding	379		
flash butt and aluminothermic rail welding	378		
manual rail weld grinding and grinding train	381		
mobile flash-butt rail welding	379		
railway noise	808		
Railway Noise Knowledge Centre	808		
Railway Safety and Standards Board	611		
Railway Technical Research Institute	704	745	750
railway wheel wear	172		
contact mechanics influence	189		
normal problem	190		

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links	
railway wheel wear (*cont.*)		
tangential problem	193	
wheel and rail profiles	194	
state-of-the-art uniform modelling	195	
ingredient for wear modelling	195	
research development	199	
tribological issues	175	178
wear mapping	183	
wear modelling	188	
wear regimes	178	
wear testing	175	178
wear transitions	180	
uniform wear reduction	201	
friction modifiers	201	
vehicle design	203	
vehicle dynamics active control	204	
wear-resistant wear profiles	202	
wear process	173	
schematic representation	174	
wear regions	174	
Rapid and Mass Transport Authority	359	728
ratcheting failure	336	
ratcheting process	150	
ratcheting threshold	334	
Rayleigh-Timoshenko beam	268	
RCF *see* rolling contact fatigue		
re-railing	650	

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links						
reference profile	584						
reinvestments	649						
residual stress	288						
longitudinal stress distribution	289						
resonance	483						
retentivity	695						
R350HT	9						
R350LHT	9						
roaring raildpinned-pinned resonance							
corrugation	364	371					
cause	365						
characteristics	364						
pinned-pinned resonance corrugation							
on mainline railway	364						
treatment	366						
rolling contact fatigue	4	11	126	215			
	216	247	311	353			
	363	409	447	448			
	450	471	510	533			
	579	651	682	701			
	704	793					
contact pressure developed from							
asperities between typical							
Shinkansen rail and wheel	711						
crack growth simulation of transverse							
cracks	708						
crack initiations	340						
crack propagation	704						

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links
rolling contact fatigue (*cont.*)	
horizontal crack propagation	704
transverse crack propagation	707
crack zones just above the gauge corner compared to the crown of the rail	315
cracks in worn high rails from curves	313
early crack propagation investigation methods	332
digital computed radiography	333
focused ion beam milling and SEM-imaging	332
multiple polishing and microscopic examination	332
experimental arrangements of rolling tests	716
high matrix strain at the RCF crack location	314
history of rail grinding work	712
important aspects that need further investigation	773
influential factors	704
main procedures to control development	773
main types	768
flaking, or running surface checking	768

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links	
rolling contact fatigue (*cont.*)		
gauge corner checking	768	
shelling	768	
overview of largest crack and detail of crack propagation	327	
preventive grinding	708	
effect of one pass per year of stone grinding car	48 713	
effect on refraining RCF defects	709	
experimental study	709	
influence of initial roughness on crystal axis density	714	716
practice	712	
roughness variation with repeated cycles	714	
surface roughness of rail formed by rail grinding	712	
theoretical study	710	
resultant principal shear stress distribution	711	
roughness variation comparison on test discs	716	
severe gauge corner checking defects in high rail	768	
severe running surface defects in low rail	770	
severe shelling defect in high rail	769	

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links			
rolling contact fatigue (*cont.*)				
squat type of RCF defect	705			
subsurface-initiated	217	223		
crack formation and shear deformation	218			
wheel fracture	218			
surface-initiated	219	221	226	229
material responses in uniaxial loading	228			
shakedown map	228			
typical appearance of severe RCF	220			
test disc at 10 and 100 % RCF life	328			
and thermal/traction defects	767			
transverse defect	769			
variation of axis density depth from surface with test discs	717			
variation of roughness with repeated cycles on test discs	715			
X-ray, CAT scan of small RCF crack on worn rail	333			
rolling noise	481			
contact filter effect	483			
effect of railpad stiffness on predicted components of rolling noise	491			
generation model	481			

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links		
rolling noise (*cont.*)			
measured decay rate of vertical vibration along the track for three different rail pads	485		
measured noise reduction achieved in field tests	492		
noise radiation	486		
predicted noise components from wheels, rails and sleepers	487		
reduction	487		
optimised track design	490		
optimised wheel design	490		
shielding measures	492		
smooth rails	488		
smooth wheels	488		
surface roughness	481		
TSI limits for rail roughness spectrum	489		
vertical mobilities of wheel-rail system	484		
wheel and track dynamics	482		
rolling radius difference	576		
rolling stock	636	637	638
roll-slip phenomenon	734		
root mean square	478		
rough surface contact theory	740		
roughness	482		
roughness level spectrum	251		

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links
roughness level spectrum (*cont.*)	
X2 trailer wheels and 14 freight wagon wheels	252
Royal Institute of Technology	608
RSD *see* Rail Surface Damage model	
RTRI tribometer	735
running surface checking	768
rutting	360
cause	362
characteristics	360
corrugation on a metro	361
corrugation on mainline railway	361
treatment	363
wheelset vibration torsional modes	362
S	
San-yo Shinkansen line	732
sand	
average conductance against sand flow rate for wet and dry tests	540
conductance against amount of sand in model wheel-rail contact	540
critical concentration for test and rail conditions	541
particle entrainment in to dry and wet disc contact	539

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links			
sand (*cont.*)				
RMS voltage for test run dry with dry sand	537			
RMS voltage plots for test run dry with dry sand	538			
RMS voltage plots for test run with water and sand	539			
twin-disc contact application	537			
wheel-rail isolation	536			
Sanding	531			
Sandite	522	531		
scanning electron microscopy	312	332	441	
scraping principle	804	805		
scrubbing	579			
scuffing	43			
second-order polynomial	405			
severe wear	94	110		
shadow zone	501			
shakedown limit	710			
shakedown theory	704			
shear yield stress	335			
shelling	49	50	644	768
gauge comer shelling in rails	645			
Shin-Kanmon Channel tunnel	732	736		
ascending slope				
formed corrugation	732			
observed rail corrugations				
location	733			

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links	
Shinkansen	4	
Shinkansen rail	709	721
Shinkansen trains	735	
Shinkansen wheel	89	719
shuffle block grinder	598	
silent bridges	808	
Silent Track damper	491	
SIMPACK	87	
single-axle running gear	614	
single-degree-of-freedom system	502	
single-point observation method	95	
rolling and sliding contact	96	
wear simulation	114	
slippage *see* spin creepage		
smooth rail	598	
sound	478	
generation mechanisms		
radiation by structural		
vibration	479	
sound produced by unsteady		
aerodynamic flow	479	
sound power level	479	
sound pressure level	478	
spalling	645	
Speno grinding train	399	
Speno International	693	
spin creepage	145	
spreadability	695	

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links		
squat defects	771		
squats	50	282	
characteristics	418		
severe squat with network of surface and subsurface cracks	420		
squat-like head check	420		
correlation with track parameters	413		
definition of half sleeper span centred on sleepers and the other half between sleepers	413		
friction	418		
hardness and geometry variations at welds	417		
rail surface irregularities	415		
short pitch corrugation and corrugation-like wave pattern after squat	414		
track structure parameters	413		
welds and materials	415		
counter measures	433		
detection	433		
developed from wheel burn	418		
development stages	410		
at the end of a fishplate	427		
further research	434		
growth process	427	429	431

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links	
squats (*cont.*)		
initiation due to differential wear and differential plastic deformation	423	
differential wear and deformation cause increased dynamic force	423	430
differential wear and deformation occurrence	423	
FE modelling of a wheel rolling over fish-plated rails	428	
vertical and longitudinal contact forces under different fishplate preloads	429	
past research review	411	
initiation and growth mechanisms	412	
metallurgical research	411	
stress and crack growth research	412	
on railway rails	409	
three-dimensional dynamic rolling contact solutions in elastoplasticity	421	
comparison of plastic strain under different friction conditions	422	
forces and surface stresses distribution of dynamic contact	423	

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links	
squats (*cont.*)		
friction effect	421	
material strength effect	421	
normal pressure distributions with effect of high-frequency waves visible	425	
tangential surface stress distribution	426	
time histories of effective strain is yield stress	422	
time histories of vertical contact forces and maximum von-Mises stress	424	
unsprung mass effect	422	
at welds	416	
squeal	809	
squeeze film	287	
SRRA *see* Strategic Rail Research Agenda 2020		
SSAB steel plant	637	
static friction force	513	
stellites	523	
stick-slip loop	809	
stick-slip phenomenon	730	764
Stockholm Central Station	553	
32-stone machine	582	
strain gauge load circuits	258	

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links
Strategic Rail Research Agenda 2020	5
stress field	214
stress intensity factor	295
stress regime	796
Stribeck friction model	104
STRIPES	81
subcritical Hopf bifurcation	462
sulphur	523
supercritical Hopf bifurcation	462
SUROS disc	
crack details and respective test conditions and traction coefficients	324
fatigue results and effect of higher maximum contact stress	323
microstructures prior to testing	321
RCF crack length-depth ratio statistics and comparative rail data	326
SUROS twin-disc testing	338
Swedish National Rail Administration	511 513 515
Swedish National Road and Transport Research Institute	621
Swedish railway infrastructure manager see Banverket	
swing-nose crossing	493
switches and crossings	423

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links			
T				
T-gamma	611			
tache ovale	49	51	131	151
	283	288		
fatigue fracture in old rail	131			
taping line	587			
technical specifications for interoperability	23	477	793	
TGV	4			
thermal damage	221			
brick pattern	222			
wheel flats formation	223			
thermal loading	230			
thermal stress	288			
thermal/traction defects				
further investigation	773			
large transverse defect developed from a severe wheel bum defect	773			
main procedures to control development	773			
main types	770			
squat defects	771			
wheel or engine burns	770			
multiple squat defects	772			
and RCF	767			
severe wheel burn with spalling	771			
squats vs wheel burns	772			

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links			
thermoelastic instability	262	488		
thin flange	585			
Thomas process	637			
three-body abrasion process	444			
titanium carbides	811			
Tokaido Shinkansen	712			
tolerances	396			
tongue lipping	283	646	647	
TRACK	271			
track deterioration model	609			
basic models used	610			
general considerations and assumptions	612			
mechanisms	609			
model of vertical dynamic wheel load	615			
model simplifications and additions	614			
modified RSD model	611			
principal mathematical model	613			
vertical wheels loads - test data vs model calculations	616			
track lubrication	623			
track recording vehicle	655	657	663	667
trackform-specific corrugation	368			
cause	369			
characteristics	368			
measurements showing difference in corrugation on different trackfoms	369			

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links			
trackform-specific corrugation (*cont.*)				
track with resiliently-mounted baseplate	368			
track with sleepers in resilient boots	368			
treatment	370			
trackside lubricators	202			
Track-Wheel Interaction Noise Software	264	271	490	493
traction	514			
traction friction	362			
traction noise	478			
traction ratio	362			
traffic control centres	642			
traffic-dependent costs	608	620	621	
transition diagrams *see* wear maps				
transmissibility	502			
Transport and Water Management Inspectorate	793			
Transportation Research Board	809			
Transportation Technology Centre	688			
transversal profile wear effect	457			
flange climb derailment	466			
guidance and stability	460			
conicity graph for UIC60 1:20 rail coupled with new and worn ORES1002 wheel profiles	461			
rolling radius variation vs wheel set lateral displacement	460			
hollow-worn and asymmetric wheel profiles	466			

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links
transversal profile wear effect (*cont.*)	
turnout passing	467
vehicle dynamics	457
vehicle steering	463
attitude of railway bogie during curve negotiation	465
wear regions of wheel and rail profiles	458
regions of wear on rail	459
regions of wear on wheel and rail	458
tread braking	261
rolling noise	264
roughness spectra on vertical wheel-rail contact force	265
tread vs disc brakes	264
strategies for roughness reduction	266
thermoelastic instability	262
hot spots during brake experiments	262
tread datum position	587
tread rollover	232
tribology	15 175 178
wear mapping	183
Class D tyre material vs BS11 rail material wear coefficient maps	185
GENSYS simulation of R8T wheel material wear	187

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links	
tribology (*cont.*)		
R8T wheel material vs		
UIC60 900A rail steel wear		
coefficient map	186	
wheel material wear data in terms		
of contact conditions	187	
wear modelling	188	
wear regimes	178	
abrasive score and oxide film		
delamination	179	
full-scale test rig	180	
wear rates & regimes for R8T		
wheel steel vs UIC60 900A rail		
steel	178	
wheel disc run at low $T\gamma/A$	181	
wear testing	175	178
BU300 full-scale roller rig	177	
SUROS twin disc test		
machine	176	
wear transitions	180	
friction vs slip in twin-disc		
contact	182	
plastic deformation depth	183	
twin-disc contact temperatures and		
wear coefficients for UIC60		
900A rail materials vs R8T		
wheel materials	184	
wear features & regimes	182	

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links				
tribology (*cont.*)					
wheel-rail contact	34				
systems approach to interface management and research	35				
tribometers	696				
truck performance detectors	684				
TRV *see* track recording vehicle					
TSI *see* technical specifications for interoperability					
tungsten carbide	811				
twin-disc machine	513	713			
twin-disc system	291				
twin-disc test	440	444	451	453	
	544	811			
set-up	534				
twin-disc tribometer	516				
TWINS *see* Track-Wheel Interaction Noise Software					

U

UIC860 *900A*	163				
UIC code	518	615			
UIC Leaflet	519	461			
UIC Leaflet	714	799			
UIC/ORE	624				
UIC/ORE model	610				
UIC-ORE s1002 profile	803				
UIC60 rail profile	642				

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links
UIC60 rails	460
ultrasonography	433
underfloor wheel lathes	587
US railroad industry	
condition monitoring	696
overload and imbalanced load detectors	696
track performance detectors	697
truck hunting detectors	697
wheel impact load detectors	696
whole impact load detectors	696
experience in managing wheel-rail interface	686
lubrication and friction modification	695
properties of wheels in interchange service	689
railroad size and revenue in 2006	686
weight of rail in place	692
wheel and rail materials	688
wheel and rail profiles	690
wheel and rail surface damage	692
V	
VAMPIRE	87
vanadium micro-alloyed	142
Vancouver's Skytrain	

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links	
Vancouver's Skytrain (*cont.*)		
spin creepage dominance in wheel-rail contact conditions	679	
wheel-rail problems	678	
vehicle design	203	
vehicle dynamic simulator	704	746
vehicle dynamics	204	
active primary suspension	205	
active wheel steering	205	
effect of damage	456	
classification	456	
localised damages on wheel and rail profiles	471	
rail corrugation	467	
simulated time history of vertical wheel-rail contact force	471	473
transversal profile wear	457	
wheel out-of-roundness	469	
independently driven wheels	205	
secondary suspension actuation	205	
vertical forces	259	
VIA	271	
vibration and noise		
basics	478	
conclusions and future trends	503	505
curve squeal	494	
reducing squeal noise	496	
squeal noise generation mechanism	494	

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links	
vibration and noise (*cont.*)		
ground vibration and ground-borne noise	497	
ground-borne noise from trains in tunnels	501	
ground-borne vibration from surface railways	498	500
vibration phenomena overview	497	
impact noise	493	
reduction of rolling noise	487	
optimised track design	490	
optimised wheel design	490	
shielding measures	492	
smooth rails	488	
smooth wheels	488	
rolling noise	481	
noise radiation	486	
surface roughness	481	
wheel and track dynamics	482	
from the wheel-rail interface	477	
Vickers hardness	382	
VICT	271	
VMA *see* vanadium micro-alloyed		
VOCO	88	
VOCOLIN	88	
Volpe Centre	688	

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links				
von Mises stress	421	423	713		
Voronoi method	337				
W					
wavelength fixing mechanism	468				
wear	40	94	351	717	
see also specific type of wear					
attack angles of initial and worn high rails caused by leading axles	727				
influence of wear on vehicle-track interaction	722				
lateral force at high rails	725				
leading axles attack angle	725				
test arrangements	724				
worn rails profiles at dynamic measurements	724				
Japanese narrow gauge vehicles worn wheel profiles	720				
laboratory simulation of wear of wheel and rail	721				
attack angle and lateral force on wear	722				
wear simulation experiments	721				
worn profiles and wear amount of rail and wheel	721				
large wear debris indicative of severe wear regime	762				

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links
wear (*cont.*)	
lateral forces of initial and worn high rails caused by leading axles	726
and lubrication of rail/wheel	760
main factors that influence the lateral forces, creepage and friction levels	761
mechanisms	43
abrasive scratching	44
abrasive wear mechanisms	44
adhesive wear mechanisms	43
oxidative wear	46
particle hardness value	45
modelling and mapping	47
rail wear rate vs average daily precipitation	49
UIC60 900A rail steel wear map	48
wheel and rail form change	48
rail disc profile variation	722
rail wear on high rail in sharp curve	761
rates and transitions	45
mild and severe wear regimens	46
R8T wheel wear rates	47
Shinkansen wheels	
measured results of flange wear	720

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links	
wear (*cont.*)		
worn profiles	723	
situations	41	
cracking and material removal	42	
steel-on-steel wear sliding wear map	41	
variation of wear amount with some experimental arrangements of laboratory simulation	724	
wear progress at gauge face of rail installed in sharp curved tracks	718	
wheel flange installed in narrow gauge vehicles wear progress	719	
worn profiles at R400m and R900m vs JIS 60 kg rail original design profile	718	
worn profiles of wheel and rail	717	
rail gauge face wear	717	
wheel thin flange wear	718	
wear maps	96	
effect of surface roughness on transition load of a lubricated sliding steel	99	
Lewis and Olofsson	97	
wear number	611	617

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links		
wear regime	796		
wear resistant	202		
wear simulation	109		
classic wear models	111		
concluding remarks	122		
initial-value process	116		
interacting rollers wear	117		
contact point and co-ordinate system-sliced contact	120		
cylindrical rollers	118		
ellipsoidal roller 1 simulated wear	121		
modified and cylindrical rollers	119		
numerical integration in rolling and sliding contact	117		
pressure distribution	121		
single-point observation method	114		
sliding distances determination	115		
sliding wear in a rolling and sliding contact	112		
sliding distance determination	114		
wear curve from a pin-on-disc test	113		
wear mechanisms	110		
weigh-in-motion system	793	813	815
weld batter	382		

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links		
welding	143		
alumino-thermic welding			
rail failure due to excessive porosity	144		
flash-butt welding	144		
manual metal arc welding	145		
wheel burns	770		
wheel fatigue	211		
appearance and mechanisms	213	217	221
classification	214	217	
fatigue components	213		
subsurface-initiated rolling contact fatigue	217		
surface-initiated rolling contact fatigue	219	221	
thermal damage	221		
conclusions	239		
future trends	241		
fatigue interactions	230		
corrugation and out-of-round wheels	231		
material characteristics	237		
single rail irregularities	234		
track geometry	230		
wheel wear	235		
prediction of fatigue	223	229	
subsurface-initiated rolling contact fatigue	223		

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links	
wheel fatigue (*cont.*)		
surface-initiated rolling contact fatigue	226	229
thermal fatigue	230	
wheel flats	246	598
wheel impact load detector	804	806
wheel materials		
available wheel material classes	785	
designs	786	
mechanical properties and material types	782	
plate profiles for wheels	786	
wheel out-of-roundness		
classification		
periodic	469	
stochastic	469	
effect on vehicle dynamics	469	
peak impact load value corresponding to different values of traveling speed and of polygonalisation depth	471	
wheel polygonalisation	248	469
out-of-roundness order sample	248	
wheel seat fretting fatigue	381	
wheel tread corrugation	469	
wheel wear	235	
hollow wear	236	
wheel-rail contact	189	
air-borne particles	550	

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links
wheel-rail contact (*cont.*)	
contact mechanics	58
computer simulation tools for railway vehicle dynamics	84
contact analysis	74
contact patch	60
future trends	89
general contact modelling	62
geometry	59
normal problem	190
tangential problem	193
tribology	34
adhesion	53
contact mechanics	35
contact position and stress variation	36
contact pressure maps	39
contact zones	36
fatigue	49
FE, CONTACT, & Hertz analysis	38
shakedown map	40
systems approach to interface management and research	35
wear	40
wheel and rail profiles	194
contact patches	195

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links	
wheel-rail contact (*cont.*)		
frictional power density distribution	197	
normal pressure distribution	196	
wheel-rail creep	4	
wheel-rail interface	3	
applications	16	
extra surface layer on railhead	19	
high axle loads	16	
high train speeds	17	
metro traffic	17	
two-layer materials	18	
Australian experience in managing	759	
control of wheel-rail interaction through profiling	774	
friction management	780	
modified wheel and rail profiles for mainline track	777	
rail and wheel materials	782	
rail contact fatigue and thermal/traction defects	767	
rail corrugations	764	
rail grinding	776	778
rail/wheel wear and lubrication	760	
Canadian experience in managing	669	
Canadian National		

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links	
wheel-rail contact (*cont.*)		
Railway	682	
Canadian Pacific Railway	670	
Cartier Railway Company	676	
summary and conclusions	684	
wheel life and ride quality on Edmonton Transit	679	
wheel shelling on Canada's freight railroads	680	
wheel-rail problems on Vancouver's Skytrain	678	
corrective maintenance	593	600
change of transverse profile or initial rail inclination	602	
irregular welds or joints	601	
tight gauge change	602	
wheelburns	601	
correlation of measured profile with the profile to which it should have been ground	585	
corrugation analysis trolley	600	
definition of reference points	585	
Dutch experience in managing	792	
monitoring traffic movement	812	
optimising rail maintenance	793	
optimising wheel maintenance	802	
special aspects in optimising wheel-rail interface	806	

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links	
wheel-rail contact (*cont.*)		
end view of grinding train	580	
Europe Metro experience on London Underground Victoria Line	653	
identified wheel-rail interface problems	662	
introduction to the Victoria Line and historic wheel-rail interface issues	653	
lubrication management	659	
ongoing work and future plans	666	
Victoria Line upgrade	654	
wheel-rail interface monitoring	654	
false flange or hollow wear measurement	589	
flange height and thickness measurement	589	
friction and wear simulation	93	
friction models	99	104
single-point observation method	95	
wear maps and transition diagrams	96	
wear simulation	109	
future trends	26	
active control	27	
data collection and condition monitoring systems	27	

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links
wheel-rail contact (*cont.*)	
integrated computer models	27
legislation and standardisation	27
new materials	27
particle emissions	28
geometrical features	457
grinding module showing typical annular grinding stone	580
ground rail after carrying a day of traffic	581
Hatfield derailment	577
history and present situation	3
current set-up	5
material development	5
railway operation	3
wheel-rail contacts theory	4
hollow wear of wheel treads	590
Japanese experience in managing	701
adhesion	737
corrugation	725
lubrication	745
rolling contact fatigue	704
wear	717
maintenance	576
future trends	603
importance of friction	578
routine maintenance of longitudinal/circumferential profile	594

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links	
wheel-rail contact (*cont.*)		
routine metal removal and maintenance of transverse profile	579	
MiniProf instrument for measuring transverse profile	591	
MiniProf measurements superposed to show metal removal	583	
models for infrastructure costs	608	
calibration of model and cost	620	
computational tools and input data	616	
conclusions and future trends	624	
examples of results	622	
track deterioration model	609	
noise and vibration	477	
ongoing research, development and standardisation efforts	18	21
European standards	22	
INNOTRACK	19	21
interoperability technical specifications	23	
particle emissions	21	
profile optimisation	22	
phenomena	6	
deterioration mechanisms	101	

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links
wheel-rail contact (*cont.*)	
environmental issues	12
management aspects and costs	14
operational control, condition monitoring and maintenance	13
safety issues	11
train-track interaction	9
typical interface	6
undesired phenomena	7
wear and fatigue mechanisms of curve radius	12
wheel profile contact in inclined rail	8
wheelset and track forces and relative motion	10
pre- and post-grind corrugation measurements over 200m grinding site	597
profile of wheels removed from service for thin flange	588
rail corrugation analyser	601
rail grinding train of typical size used for mainline railway work in Europe	581
rail profile that should be ground in order to widen gauge	602
rail surface that has been ground for adhesion improvement in leaf-fall season	603

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links
wheel-rail contact (*cont.*)	
requirements following grinding	
for acceptable longitudinal	
profile	596
for acceptable transverse	
profile	584
research fields	14
Swedish experience in managing	
electrical power system	649
general description of	
Malmbanan	634
infrastructure configuration	642
locomotives and cars	636
maintenance, wheel-rail	
interaction	651
railway infrastructure maintenance	
in severe environment	633
track maintenance	
practices	649
train control system	642
system aspects and optimisation	24
LL break blocks	26
matching bogie type to track	
curvature	25
RCF in high-speed	25
transverse profile	
measurement over 7-month	
period	592

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links
wheel-rail contact (*cont.*)	
of partially ground rail	586
rail measurement using bar gauge holding template of desired profile	591
US experience in managing	
condition monitoring	696
lubrication and friction modification	695
wheel and rail materials	688
wheel and rail profiles	690
wheel and rail surface damage	692
wheel-rail isolation	528
effects of contaminants	536
leaves	541
sand	536
solid friction modifier	541
modelling approaches	545
conductance against resistivity for varying layer thickness	546
disc contact showing metal to metal contact bridges	547
electrodiscrete element method	548
number of contacts against conductance for varying contact radius	547

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links
wheel-rail isolation (*cont.*)	
wheel-rail discs separated by sand	545
wheel-rail surfaces in partial contact	546
wheel-rail surfaces separated by sand	545
problems	529
testing for isolation	532
dynamic test method	533
impedance measurements	533
simplified TI21 electrical circuit used for voltage determination	535
static test apparatus	534
static test method	534
track circuit simulations	535
twin disc test set-up	534
third bodies in wheel-rail contact	530
sanding apparatus	531
types	530
track circuits	528
contaminant film breakdown voltages	530
schematic	529
types	530
white etching	772

This page has been reformatted by Knovel to provide easier navigation.

Index Terms	Links			
white etching constituent	151	317		
white etching layer	342			
Widmanstatten ferrite	137			
WILD *see* wheel impact load detector				
Winkler model	289			
Winkler surface model	120	122		
Wöhler S-N curve	213			
X				
X-ray diffraction	713	714	736	737

This page has been reformatted by Knovel to provide easier navigation.