Rolling Contact Fatigue: A Comprehensive Review

Office of Railroad Policy and Development
Washington, DC 20590
NOTICE
This document is disseminated under the sponsorship of the Department of Transportation in the interest of information exchange. The United States Government assumes no liability for its contents or use thereof. Any opinions, findings and conclusions, or recommendations expressed in this material do not necessarily reflect the views or policies of the United States Government, nor does mention of trade names, commercial products, or organizations imply endorsement by the United States Government. The United States Government assumes no liability for the content or use of the material contained in this document.

NOTICE
The United States Government does not endorse products or manufacturers. Trade or manufacturers’ names appear herein solely because they are considered essential to the objective of this report.
# Rolling Contact Fatigue: A Comprehensive Review

Rolling contact fatigue (RCF) is a pervasive and insidious problem on all types of railway systems. Although it is a dominant cause of maintenance and replacements on heavy haul rail lines, it is also a significant economic and safety challenge for commuter and metro lines. It is a subject of intense research around the globe, with strong academic research being undertaken in Europe particularly, with more practical work being performed in Australia, South Africa, and North America. The safety implications of RCF include being responsible for about 100 FRA reportable derailments annually in North America. The poster child for hazardous RCF is the Hatfield derailment in the United Kingdom, which incurred four deaths, 39 injuries, and economic fallout easily exceeding GBP1 billion, dismemberment of the railway authority and manslaughter charges against several railway officials. The economic implications of RCF to the North America railway industry for rail replacement alone amounts to over USD300 million annually, with costs of inspection and derailments, as well as damage to track and rolling stock, and derailment costs further increasing that number. Of the USD100+ million dollars spent annually on rail grinding in North America, at least ≥30 percent can be attributed to RCF. A review of the types or RCF defects on wheels and rails, causal mechanisms and monitoring and maintenance practices has been undertaken for the purpose of identifying gaps and the most pressing areas for research and development.

### Subject Terms
- Rolling contact fatigue, RCF, rail gauge corner cracking, wheel tread cracking, monitoring RCF, RCF damage model, friction management, rail grinding
### METRIC/ENGLISH CONVERSION FACTORS

#### ENGLISH TO METRIC

<table>
<thead>
<tr>
<th>Length (Approximate)</th>
<th>Metric to English</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 inch (in) = 2.5 centimeters (cm)</td>
<td>1 millimeter (mm) = 0.4 inch (in)</td>
</tr>
<tr>
<td>1 foot (ft) = 30 centimeters (cm)</td>
<td>1 centimeter (cm) = 0.4 inch (in)</td>
</tr>
<tr>
<td>1 yard (yd) = 0.9 meter (m)</td>
<td>1 meter (m) = 3.3 feet (ft)</td>
</tr>
<tr>
<td>1 mile (mi) = 1.6 kilometers (km)</td>
<td>1 meter (m) = 1.1 yards (yd)</td>
</tr>
<tr>
<td>1 kilometer (km) = 0.6 mile (mi)</td>
<td>1 yard (yd) = 0.9 meter (m)</td>
</tr>
</tbody>
</table>

#### AREA (Approximate)

| 1 square inch (sq in, in²) = 6.5 square centimeters (cm²) | 1 square centimeter (cm²) = 0.16 square inch (sq in, in²) |
| 1 square foot (sq ft, ft²) = 0.09 square meter (m²) | 1 square meter (m²) = 1.2 square yards (sq yd, yd²) |
| 1 square yard (sq yd, yd²) = 0.8 square meter (m²) | 1 square kilometer (km²) = 0.4 square mile (sq mi, mi²) |
| 1 square mile (sq mi, mi²) = 2.6 square kilometers (km²) | 10,000 square meters (m²) = 1 hectare (ha) = 2.5 acres |
| 1 acre = 0.4 hectare (ha) = 4,000 square meters (m²) | 1 acre = 0.4 hectare (ha) = 4,000 square meters (m²) |

#### MASS - WEIGHT (Approximate)

| 1 ounce (oz) = 28 grams (gm) | 1 gram (gm) = 0.036 ounce (oz) |
| 1 pound (lb) = 0.45 kilogram (kg) | 1 kilogram (kg) = 2.2 pounds (lb) |
| 1 short ton = 2,000 pounds (lb) | 1 tonne (t) = 1,000 kilograms (kg) |
| 1 tonne (t) = 1.1 short tons | 1 tonne (t) = 1.1 short tons |

#### VOLUME (Approximate)

| 1 teaspoon (tsp) = 5 milliliters (ml) | 1 milliliter (ml) = 0.03 fluid ounce (fl oz) |
| 1 tablespoon (tbsp) = 15 milliliters (ml) | 1 fluid ounce (fl oz) = 30 milliliters (ml) |
| 1 fluid ounce (fl oz) = 30 milliliters (ml) | 1 liter (l) = 2.1 pints (pt) |
| 1 cup (c) = 0.24 liter (l) | 1 liter (l) = 1.06 quarts (qt) |
| 1 pint (pt) = 0.47 liter (l) | 1 liter (l) = 0.26 gallon (gal) |
| 1 quart (qt) = 0.96 liter (l) | 1 cubic foot (cu ft, ft³) = 0.03 cubic meter (m³) |
| 1 gallon (gal) = 3.8 liters (l) | 1 cubic yard (cu yd, yd³) = 0.76 cubic meter (m³) |
| 1 cubic foot (cu ft, ft³) = 0.03 cubic meter (m³) | 1 cubic meter (m³) = 36 cubic feet (cu ft, ft³) |
| 1 cubic yard (cu yd, yd³) = 0.76 cubic meter (m³) | 1 cubic meter (m³) = 1.3 cubic yards (cu yd, yd³) |

#### TEMPERATURE (Exact)

<table>
<thead>
<tr>
<th>°F</th>
<th>°C</th>
</tr>
</thead>
<tbody>
<tr>
<td>[(x-32)(5/9)]°F = y °C</td>
<td>[(9/5)y + 32]°C = x °F</td>
</tr>
</tbody>
</table>

For more exact and or other conversion factors, see NIST Miscellaneous Publication 286, Units of Weights and Measures. Price $2.50 SD Catalog No. C13 10286

Updated 6/17/98
Acknowledgments

The author is extremely grateful to the U.S. Federal Railroad Administration for its financial support of this review through grant DTFR-53-05-H-00203-mod-4.

The author would next like to recognize the assistance of many colleagues during the production of this report who supplied technical information, photographs, various supplementary information and cross references. Thanks are due to the following individuals, listed in alphabetical order:

Chris Barkan (University of Illinois, Urbana-Champaign, IL), Bill Bell (CSX, Jacksonville, FL), Li Cheng (Salient Systems, Austin, TX), Mathew Dick (ENSCO, Falls Church, VA), Marty Gearhart (Union Pacific, Omaha, NE), Stuart Grassie (Loram, Germany), Martin Hiensch (DeltaRail, The Netherlands), Cory Hogan (Transportation Technology Center, Inc. (TTCI), Pueblo, CO), Jude Igwemezie (Applied Rail Research Technologies, Toronto, Canada), Makoto Ishida (Railway Technical Research Institute, Tokyo, Japan), Joe Kalousek ([retired] National Research Council Canada), Tyler Kerr (Canadian Pacific Railway, Calgary, Alberta, Canada), Jean Kneale (NRC Canada), Yan Liu (NRC Canada), Francisco Robles-Hernandez (University of Houston, Houston, TX), Michael Roney (Canadian Pacific Railway), Ken Rownd (TTCI), Wolfgang Schoech (Speno, Geneva, Switzerland), Peter Sroba (Sroba Rail Services, Australia), Daniel Szablewski (TTCI), Harry Tournay (TTCI), and Allan Zarembski (Zeta-Tech, Cherry Hill, NJ).

Special thanks are extended to Daniel Stone of Hunter Holiday Consulting (Pueblo, CO) and Kevin Sawley of Interfleet (Derbyshire, U.K.), who, more than any others, entertained numerous queries and graciously accepted and responded to or passed along questions related to rail and wheel metallurgy and performance.
Contents

Executive Summary .................................................................................................................. 1

1 RCF of Wheels and Rails ..................................................................................................... 5
  1.1 The International Scope ............................................................................................... 5
  1.2 Safety Implications of RCF ......................................................................................... 8
    1.2.1 (Rail) Ultrasonic No-Tests ................................................................................. 8
  1.3 Types of RCF Defects ................................................................................................. 9
    1.3.1 Rail Gauge Corner Cracking .............................................................................. 9
    1.3.2 Wheel Tread Cracking ....................................................................................... 12
    1.3.3 Rail and Wheel Shelling .................................................................................... 12
    1.3.4 Rail Transverse Defects ..................................................................................... 16
    1.3.5 (Rail) Squats ......................................................................................................... 18
    1.3.6 (Rail) Crushed Heads ......................................................................................... 19
    1.3.7 (Wheel) Spalling .................................................................................................. 20
    1.3.8 (Wheel) Shattered Rim ....................................................................................... 21
    1.3.9 (Wheel) Vertical Spit Rim .................................................................................... 22
  1.4 Economic Implications of RCF ..................................................................................... 23

2 Formation of RCF .............................................................................................................. 25
  2.1 Crack Initiation ............................................................................................................. 26
    2.1.1 Surface Crack Initiation in Rails ......................................................................... 26
    2.1.2 Subsurface Crack Initiation ............................................................................... 29
    2.1.3 Contact Stress ..................................................................................................... 32
    2.1.4 Traction ................................................................................................................ 33
    2.1.5 Material Strength ................................................................................................. 34
    2.1.6 Thermal Softening ............................................................................................... 35
    2.1.7 Material Embrittlement ....................................................................................... 36
    2.1.8 Models of Crack Initiation ................................................................................. 37
  2.2 Crack Propagation ......................................................................................................... 38
    2.2.1 Shear Strength ..................................................................................................... 40
    2.2.2 Strain Rate Dependence ..................................................................................... 41
    2.2.3 Contact Stresses ................................................................................................. 42
    2.2.4 Moisture and Grease ......................................................................................... 45
    2.2.5 Fracture Toughness ............................................................................................. 46
  2.3 Rail (and Wheel) Fatigue Failure ................................................................................... 47

3 Detection and Monitoring of RCF ...................................................................................... 49
  3.1 Existing Systems .......................................................................................................... 49
    3.1.1 Eddy Current ......................................................................................................... 51
3.1.2 Ultrasonic Measurements ................................................................. 56
3.1.3 Vision Technologies ........................................................................ 57
3.1.4 Acceleration and Force-Based Measuring Systems ......................... 58
3.2 Current Gaps ...................................................................................... 59
3.2.1 Surface Defects ............................................................................... 59
3.3 Internal Defects .................................................................................. 59
4 The Influence of Operating Parameters ..................................................... 61
4.1 Metallurgy ......................................................................................... 61
4.1.1 Hardness ....................................................................................... 63
4.1.2 Cleanliness .................................................................................... 64
4.1.3 Steel Type .................................................................................... 65
4.1.4 Layered Steels ............................................................................... 68
4.2 Traction Coefficient .......................................................................... 68
4.3 Wheel Loads ...................................................................................... 72
4.4 Truck/Bogie Characteristics ................................................................ 73
4.5 Track Geometry ................................................................................. 74
4.6 Wheel-Rail Profiles ........................................................................... 75
4.7 Wear – Material Attrition ................................................................... 75
5 Management of RCF ............................................................................ 77
5.1 RCF Damage Model .......................................................................... 77
5.2 Managing Wheel-Rail Profiles ............................................................. 78
5.2.1 Wheel Profiles ................................................................................ 78
5.2.2 Rail Profiles .................................................................................. 80
5.2.3 Rail Profile Shapes Developed from Practical Field Experience .......... 80
5.2.4 Computer Modeling for Rail Profile Optimization ............................ 82
5.2.5 Managing Wheel Profiles ............................................................... 85
5.3 Managing Forces ............................................................................... 85
5.3.1 Wheel Impact Load Detector ......................................................... 86
5.3.2 Skewed Truck Detector .................................................................. 86
5.3.3 Truck Performance Detector ......................................................... 86
5.3.4 Instrumented Wheelset ................................................................... 86
5.3.5 Acceleration-Based Systems .......................................................... 87
5.3.6 Vehicle-Dynamics Simulations ....................................................... 87
5.4 Nondestructive Rail Inspection ............................................................ 87
5.5 Friction Management ......................................................................... 88
5.5.1 Lubrication .................................................................................... 90
5.5.2 Top-of-Rail/Wheel-Tread Friction Control ....................................... 90
5.5.3 Improved Performance Trucks/Bogies ........................................... 92
Figures

Figure 1. Well-Defined Gauge Corner Cracking from High-Speed Passenger Operation .......... 9
Figure 2. Cracking of the High and Low Rails, Dye Penetrant Enhanced .......................... 10
Figure 3. Examples of Typical High and Low Rail Checking and Shelling .......................... 11
Figure 4. A) Cracks at the Field Side of the Wheel Are Typically Most Severe; B) Extensive Cracking at the Flange Root, Like That Seen Here, Is Much Less Common ....................... 12
Figure 5. Shelling of a Rail; Shown Are Longitudinal Sections Machined Vertically (parallel to the rail web) at Distances Ranging from 4 to 28 mm from the Gauge Side of the Rail; The Steel Is a 1984 CC CFI Rail Steel .......................................................... 13
Figure 6. Progression of Wheel-Tread Surface Fatigue from Microcracks to Fully Shelled [171] ........................................................................................................ 14
Figure 7. Gauge Corner Collapse Can Lead to Deep-Seated Shells with Very Large Chunks of Metal Spalling from the Rail ............................................................. 15
Figure 8. Examples of Gauge Corner Collapse ................................................................... 16
Figure 9. Gauge Corner Fatigue on a Light Rail System Because of Flattening of Rail (with a reciprocating stone grinder) and Overloading of the Rail Gauge Corner ....................... 16
Figure 10. Transverse Defects Initiated by Deep-Seated Shells, Probably from Gauge Corner Collapse ........................................................................................................ 17
Figure 11. Vertical Split Head Defect .................................................................................. 18
Figure 12. A) The Squat Appears as a Dark Spot on the Top of the Rail, B) In Cross Section, One Can See Two Primary Cracks—a Short Leading One and a Longer Trailing Crack; Photographs Are Courtesy of Makoto Ishida, RTRI, Japan .................................................. 18
Figure 13. The Crushed Head Defect Is Associated with Collapse and Deformation of the Underside of the Rail Head as a Result of Cracks That Initiate at the Mid-Gauge and Propagate to the Center of the Rail ........................................................................ 19
Figure 14. Examples of Wheel Surface Martensite ............................................................. 20
Figure 15. A) Thin Streaks of Martensite on the Rail Surface Because of High Creep Rates from Wheels under High Traction; B) a Spinning Wheel Can Burn a Deep Patch of Martensite into the Rail ........................................................................................................ 21
Figure 16. A Shattered Rim Failure [262] ............................................................................. 22
Figure 17. A Typical Vertical Split Rim [260] ...................................................................... 23
Figure 18. Hydraulic Mechanism Is No Doubt Contributing to Shelling on This High Rail; Cracking Is Enhanced by Dye Penetrant ................................................................. 24
Figure 19. Ratcheting in Rail Steels Associated with Contact Fatigue ............................... 25
Figure 20. Cyclic Stresses Associated with a Rolling Contact, Nondimensionalized with Respect to the Width of the Line Contact (2a) and the Peak Normal Stress (P_o) .......... 26
Figure 21. Plot of the Stress Trajectory (σ_{zz,σ_{xx}}) / 2P_o versus τ_{xy}/P_o for One Complete Rolling Cycle at Different Depths z/a below the Surface .................................................. 27
Figure 22. Different Models for Yielding of Materials Include Elastic Plastic, Linear Hardening, and Kinematic Hardening ............................................................................. 28
Freight Wheel Profile against the RE136, 141AB, and NRC-H2 Rail Grinding Template (top of rail friction is 0.45, and the gauge face value is 0.3) .................................... 70

Figure 44. Traction Forces Measured with Instrumented Wheelsets Running on the Amtrak Northeast Corridor ................................................................. 71

Figure 45. Surface Traction and Subsurface Plastic Flow – Longitudinal and Lateral Traction Directions for RCF Crack Initiation at Different Locations on the Wheel Tread; One or Several Distinct Bands of Cracks May Appear as a Result of Contact with Tangent Track (TT) or the High Rail (HR) or Low Rail (LR) of Sharp Curves (SC) and Mild Curves (MC) .................................................................................. 72

Figure 46. Track Geometry Errors Contribute Directly to Many Instances of RCF: A) Gauge Tightening over a 6-Foot Length of Track Otherwise at Standard Gauge Lead to a Localized Band of Gauge Corner RCF; B) A Poor Ballast Shoulder Led to a Crosslevel Error and the Formation of a Deep-Seated Shell on Tangent Rail ........................................ 75

Figure 47. Controlling Wear at “The Magic Wear Rate” Maximizes Rail Life .................. 76

Figure 48. The RCF Damage Function (1/Nf the number of cycles to failure) as a Function of the Wear Number T [15] .................................................................................. 78

Figure 49. A Dynamic Shakedown Plot Summarizes Contact Conditions for Three Different Wheel Profiles against the Same Rail Shape As-Is Negotiates a 1,500-Meter, High Cant Deficiency Curve .................................................................................. 79

Figure 50. The Typical Rail Profile Changes that Arise because of Common Wheel and Rail Contact Conditions ............................................................................. 80

Figure 51. NRC’s Complete Set of Eight 200-Millimeter (8 inches) Profiles: TT Is the Tangent Rail Template, whereas a) H1 to H4 Are High Rail Templates with Progressively More Gauge-Corner Relief and b) L1-L3 Are Low Rail Templates with Progressively More Field Side Relief .................................................................................. 82

Figure 52. The Worn Rail Profile Is an Envelope of the Worn Wheels that Run over It ........ 84

Figure 53. The Improved CP-TF2 Rail Profile Spreads Wear across the Rail Head, Minimizing Contact at the Gauge Side where cracks Initiate, Placing Most of the Load through the Center of the Rail, but also Allowing the Field Side Metal to Bear a Portion of Stress; This Contact Distribution Is Believed to Be the Best Compromise between Controlling Initiation while Managing Propagation on Existing Defects .................................................. 85

Figure 54. The Coefficient of Friction at the Unmanaged Wheel-Rail Interface Can Vary Widely, from 0.05 to 0.7 .................................................................................. 89

Figure 55. Friction Management Involves Two Different Interfaces (top-of-rail/tread-of-wheel and gage face/wheel flange) with Different Requirements for Each .............................................. 89

Figure 56. The Effect of Friction Management of Wear and RCF – Full-Scale Rig Test Results; Left to Right: New 60E1 Profile, “Dry” after 100k Wheel Passes, FM 100k Wheel Passes, and FM 400k Wheel Passes (from Reference 257) .................................................................................. 91

Figure 57. Solid Stick and Fluid Friction Modifier Products Developed by Kelsan Technologies for Control of Wheel-Rail Friction .................................................................................. 92

Figure 58. A Spray Nozzle Used to Dispense Thin Films of the Kelsan Friction Modifiers to the Top of Rail, Usually from a Locomotive; A Same Approach Is Taken by Timken to Apply Their Trackglide Friction Modifier from a Locomotive to the Top-of-Rail ........................................ 92
Figure 59. Crack Initiation and Fatigue Model Presented by Kapoor [139] ............................... 96
Figure 60. The Friction Transformation Process “Heals” Cracks in a Laboratory Test............. 98
Tables

Table 1. Listing of Those Organizations Publishing Papers Related to Studies of RCF .......... 7
Table 2. Rail Defect Classifications ......................................................................................... 17
Table 3. Demands on Rail (from Reference 124) ..................................................................... 34
Table 4. Mechanical Properties of High-Strength Rails (Reference 108, Table 3.3) .............. 35
Table 5. Comparing High Cycle Fatigue and Low Cycle Fatigue Approaches to Modeling of
     Crack Growth ....................................................................................................................... 39
Table 6. Increasing Hardness Corresponds to Increasing Resistance to the Initiation of RCF;
     Standard North American Classifications; Typical Values of Shear Strength (k) Calculated
     at One-Sixth of the Vicker’s Hardness [170] .................................................................... 41
Table 7. Impact Resistance of Rail Steels [Reference 108, Table 3-3] ................................... 47
Table 8. Example of a Defect Classification and Required Action ......................................... 48
Table 9. Technologies Applied to Rail Testing [Table 1 of Reference 200] ............................ 50
Table 10. Comparison of the Three Common Detection Technologies Applied to Rail
     Inspection [extension of Reference 200, Table 1] ............................................................. 51
Table 11. General Properties of Wheel and Rail Steels .......................................................... 62
Table 12. Prescribed Properties of Rail Steels According to UIC 860V 8 [291] ..................... 63
Table 13. Standard European Rail Steels and Their Hardnesses [291] ................................. 63
### List of Symbols

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>a, b</td>
<td>The half length of the wheel/rail elliptical contact patch in the longitudinal (in direction of travel) and lateral (across the rail) directions, respectively (in metres or inches)</td>
</tr>
<tr>
<td>z/a</td>
<td>Distance beneath surface – nondimensionalized</td>
</tr>
<tr>
<td>x/a</td>
<td>Distance in front of the center of the contact patch, nondimensionalized</td>
</tr>
<tr>
<td>K</td>
<td>Shear strength (Pa, psi)</td>
</tr>
<tr>
<td>Hv</td>
<td>Vickers Hardness uses a pyramid shaped indenter (kg/mm$^2$)</td>
</tr>
<tr>
<td>HB</td>
<td>Brinnell Hardness uses a spherical indenter (kg/mm$^2$)</td>
</tr>
<tr>
<td>N, V</td>
<td>Normal (or vertical) force loading two bodies together (N, lbf)</td>
</tr>
<tr>
<td>$P_o$</td>
<td>Maximum contact stress (Pa, psi)</td>
</tr>
<tr>
<td>T</td>
<td>Total traction or shear force (lbf, Newtons)</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>The nondimensionalized slip (slip rate/velocity)</td>
</tr>
<tr>
<td>$\mu$</td>
<td>Friction coefficient</td>
</tr>
<tr>
<td>$\tau$</td>
<td>Shear stress (Pa, psi)</td>
</tr>
<tr>
<td>$\tau_{zx}$, $\tau_{xy}$</td>
<td>Shear stress in the zx and xy planes, respectively</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>Normal stress (Pa, psi)</td>
</tr>
<tr>
<td>$\sigma_{zx}$, $\sigma_{xy}$</td>
<td>Normal stress in the zx and xy planes, respectively</td>
</tr>
</tbody>
</table>
Executive Summary

Rolling contact fatigue (RCF) is a pervasive and insidious problem on all types of railway systems. Although it is a dominant cause of maintenance and replacements on heavy-haul rail lines, it is also a significant economic and safety challenge for commuter and metro lines. It is a subject of intense research around the globe, with strong academic research being undertaken in Europe particularly, with more practical work being performed in Australia, South Africa, and North America.

The safety implications of RCF include being responsible for approximately 100 Federal Railroad Administration (FRA) derailments annually in North America. The poster child for hazardous RCF is the Hatfield derailment in the United Kingdom, which resulted in 4 deaths and 39 injuries, economic fallout easily exceeding £1 billion, dismemberment of the railway authority, and manslaughter charges against several railway officials. The economic implications of RCF to the North American railway industry for rail replacement alone amounts to over USD300 million annually, with costs of inspection and derailments, as well as damage to track and rolling stock further increasing that number. Of the more than USD100 million spent annually on rail grinding in North America, at least 30 percent or more can be attributed to RCF.

A review of the types or RCF defects on wheels and rails, causal mechanisms, and monitoring and maintenance practices has been undertaken for the purpose of identifying gaps and the most pressing areas for research and development.

The extensive cross-referencing, combined with a specific listing of leading organizations currently involved in research and development that are included in this report should prove valuable in developing possible future collaborative projects aimed at eliminating gaps in knowledge or practices.

Causal Mechanisms

Crack initiation

- Two prominent models used in rolling contacts for understanding crack initiation are the Shakedown Limit and the Dang Van Criterion. Their implementation is relatively straightforward, but opportunities exist for better characterizing the material properties upon which they depend. This includes the development of repeatable testing methodologies that mimic the true state of stress and the short loading duration of the wheel-rail system. The latter is relevant in that the wheel-rail contact loading is typically completed in a very short time (e.g., 0.5 milliseconds (ms)) and high strain rates occur (e.g., 1.0 s⁻¹). Currently, materials characterization tests are typically performed under nearly static conditions.

- Although detailed models of the wheel-rail contact and vehicle-track interaction are being developed, the input of realistic operating conditions is sometimes overlooked. As an example, the effect of friction and traction coefficient is well appreciated, but to date, nearly all work assumes a simplified traction-creepage relationship. Laboratory and field evidence suggests that the relationship differs from that of Kalker. Further field characterization of friction characteristics and implementation in vehicle track interaction (VTI) models is required. Similarly, the distribution of vehicle characteristics, wheel-rail profiles, and material properties is needed.
Crack propagation
- The rolling contact involves a complex multiaxial state of stress dominated by a combination of Mode II shear and compressive Mode I stress. Grease and water in the interface dramatically affect the stress state. Reduced friction at the crack face considerably increases the impact of shear stresses on crack propagation. Water that is drawn into the crack by surface tension can be pressurized toward the crack tip and generate large tensile stresses that drastically accelerate crack growth.
- At high contact stress values, RCF life decreases with increasing stress, but below a threshold value, RCF life approaches infinity. A much stronger relationship is seen with the creep ratio in which minimum life occurs at approximately 0.3 percent creep.
- High-strength materials better resist crack propagation. Technologies for improving shear strength (hardness) and fracture toughness include minimizing inclusions, ensuring favourable residual stresses, and alloying (most notably with molybdenum).

Monitoring Technologies
- It is understood that current ultrasonic measurement systems are able to collect a signature from the rail surface with information relevant to the severity of RCF. No literature was found to suggest that this has been investigated seriously as a technique for assessing surface damage.
- Vision systems offer the potential to identify surface cracks and support a fatigue management system, but the capabilities of current instruments is limited. Improved hardware, combined with appropriate interpretive algorithms, could make this approach a reality.
- Eddy current systems provide a return signal that is proportional to the length of the surface breaking crack. Such measurements would be extremely valuable for a field study of the relationship between RCF and operating conditions, for safety monitoring of track, for maintenance planning, and for the practice of rail grinding. Work is needed to a) validate available measurement systems; b) demonstrate the reliability in field conditions; c) study the relationship between crack depth and crack length under several environmental and loading conditions; and d) study how the rate of crack growth varies with a multitude of operational factors.
- Acceleration-based systems, including instrumented wheelsets and axle-box-mounted accelerometers, are able to detect cases of deep RCF surface defects but are not yet able to support a preventive maintenance program.
- Inspection of wheel steels: small, nonmetallic inclusions are believed to be the cause of some wheel failures such as shattered and broken rims. The current statistical sampling and analysis approach is inadequate for finding random inclusions, and current inspection technologies (e.g., ultrasonic) are unable to find a 1-millimeter (mm) defect. Improved tools are required.

Management of RCF
Among the tools available for managing RCF, the most powerful ones continue to be development and adoption of improved rail steels, optimized wheel-rail profiles, friction management, and rail grinding.
Work should continue toward developing rail steels that maximize cleanliness and minimize pro-eutectoid ferrite. Bainitic steels, especially lower bainite, appear to offer improved resistance to RCF. Cleaner and cleaner steels may continue to improve resistance to subsurface failure in both rails and wheels.

Although the significant contribution of wheel-rail profiles to minimizing RCF and techniques for their design is well understood, practical approaches for managing profiles are not well developed. Management tools exist, but their analytical capabilities remain simplistic, and their penetration into the rail industry poor. Tolerances on wheel and rail profiles remain to be developed. Although various wheel-rail interaction indices suitable for evaluating risk and maintenance requirements have been developed, more can still be done to develop improved criteria and support their adoption by the industry.

Friction management – the ability of friction management to reduce RCF has been demonstrated recently and the theory is sound, but the practical applications and economic and safety cases still need to be convincingly demonstrated for a range of systems.

Rail grinding is a well established method of removing surface defects and profiling rail. The optimal rail grinding interval and metal removal as a function of specific operating conditions (e.g., axle loads, metallurgy, track curvature, and rolling stock characteristics) is not known, except by generalized “best practice” principles. The integration of VTI, rail defect, and rail wear data into rail grinding decisions remains an area for development. The mechatronic rail grinder—one that measures profile and cracks at the lead end of the train, computes the required speed and patterns, and measures the results on the tail end—remains a theoretical and practical challenge.

Control of Track Geometry Defects

Track geometry perturbations are often associated with local clusters of RCF. These perturbations can be identified by train-based acceleration measurements with the probability of clusters arising readily predicted by dynamic models. The severity of RCF depends not only on the characteristics of the track geometry defect but also on the profiles, friction conditions, truck characteristics, track curvature, and train speed. Although operating limits for geometry errors can thus be established to minimize fatigue, it would be difficult to do so in a general sense for the railroad industry because of the system-specific dependencies.

Improved Suspensions

In the United Kingdom, it has been noted that as train speeds increase on the network, the primary yaw stiffness has also increased to meet the need for improved ride quality. This in turn has generally compromised curving performance and increased the RCF damage severity, especially in shallower, high cant deficiency curves. However, it was further noted that careful suspension geometry design may mitigate the adverse effects of higher primary yaw stiffness.

In North America, the contribution of advanced (e.g., M976 compatible) trucks to reduced RCF has not been well quantified, but there is some evidence of improved curving performance resulting in longer wheel life, including with respect to shelling. Because the truck designs are intended and tested to provide reduced curving resistance, and thereby “lower the stress state,” it is reasonable to expect that reduced rail RCF will follow also. Yet outside of computer models, this has been difficult to quantify.
VTI Monitoring Systems

Accelerometers and strain-gauged components (including instrumented wheelsets) onboard rolling stock have proven effective in identifying high dynamic and damaging forces as well as unsafe conditions such as elevated risk of wheel climb. Once regions of high dynamic forces were identified, there is reasonably a correlation with localized incidents of RCF. Under quasi-static curving, high forces can be correlated to increased rates of RCF development. However, a full quantification of high RCF areas would require measurement of not only force but also angle of attack and longitudinal creepage as well as the precise location of the contact patch. This is beyond the capabilities of current onboard or wayside instrumentation and partly explains the widespread use of dynamics models to explain the problem.
1. RCF of Wheels and Rails

RCF has been a subject of significant scientific scrutiny for several decades, but as tools and practices for controlling corrugation and wear have become increasingly effective, RCF has emerged as a governing reason for rail replacement and maintenance and for rail failure and safety concerns. Since the mid-1990s especially, much research has been undertaken on all continents to understand the fundamental causes of RCF, approaches to modeling, and development of maintenance approaches. The intent of this review is to identify the areas where improved knowledge and practices need development before a fully effective strategy for managing RCF can be realized.

1.1 The International Scope

RCF is a problem that knows no borders. It is the subject of research in every railroading country in the world including the following:

- United States: Rail grinding practices [228, 254] and steel developments [213, 214, 263] are the primary approaches taken by U.S. railroads to deal with RCF. The rail steels have progressively increased in hardness, with 400 BHN rail now commonly used, especially but not exclusively in curved track.

- Canada: RCF is a major initiator of broken rails and is the main reason for rail grinding and loss of rail life. Work in that Nation has focused on optimal rail grinding practices and profile design [47] and proving friction management [52, 218, 250] for controlling RCF.

- United Kingdom: As a result of the Hatfield derailment in 2000 because of a broken rail initiated by surface RCF, there has been intense study into the causes of and remedies for gauge-corner cracking [21, 44, 88]. Strong contributions have been made in the past several years related to the modeling of crack initiation [76], understanding the effect of track curvature and track geometry perturbations [15] and developing systems for managing RCF [272]. In addition, there has been a tremendous effort to develop comprehensive models of the interaction between vehicle and track and the impact on RCF [15, 217].

- Sweden: On the electrified heavy iron ore lines, 25-ton axle loads cars run of UIC 60 rail with UIC 1100 and UIC 900 metallurgy. The rail suffers severe head checking and spalling, primarily in switches and the high rail of curves, which contributes to broken rails. Corrugation and gage face wear are significant concerns also [89]. Rail grinding has proven very effective in mitigating the costs of RCF [194]. Sweden has a very active railway research program that includes significant research effort into RCF [e.g., 56, 58, 60, 210, also Table 1].

- France: The focus is on RCF appearing on the rails of its high-speed lines, especially gauge corner cracking [43] and squats [40].

- Japan: The surface defect of greatest concern is squats [112] with the main treatment being preventive rail grinding [221]. In the mid-1990s, 70 percent of the 220 kilometers (km) of rail replaced annually on the Tokaido Shinkansen was for RCF [148]. Head checking and shelling have also been experienced on the high-speed lines and on lines throughout the Tokyo urban area [270].
- China: With the dramatic increase in rail traffic, the tonnage on some Chinese lines has escalated dramatically. RCF is becoming increasingly problematic [27], with shelling of the high rail gauge corner and on the top of low rail being most common. The approach to treatment appears to focus on the adoption of improved rail steels [305].

- Australia: Leading the world in terms of axle load, this country has been at the forefront of developing rail grinding strategies [153] and advanced rail and wheel steels [178, 280] to deal with problems of wear, fatigue, and plastic flow. A significant finding reported in 1987 is that high strength rails, and in particular the head-hardened varieties, showed a strong improvement in fatigue performance [179].

- South Africa: Pioneered steerable bogies [82] for the reduction of shear stress, wear, and fatigue. More recently, the development of integrated systems for managing RCF [80] have been developed. As in other heavy haul countries, improved rail grinding practices [231] continue to form a key part of the RCF management strategy.

- Germany: German railways (DB AG) has suffered from various forms of surface fatigue for many years, with head checks developing most rapidly on newly installed high rails in shallow curves [91, 193]. Rail grinding and improved rail profiles [91] as well as high-strength rail steels [99] are the current approaches taken to treat the problem.

- Russia: RCF defects on rails and wheels continue to be a challenge for the Russian railways [304]. Improved steels [234] and rail/wheel profiles [303] appear to be the prime focus of researchers in Russia.

- Brazil: RCF on wheels and rails in Brazil has focused mainly on its iron ore mining railroads [70]. Recent projects to increase axle load and dramatically increase annual tonnage shipped from those same mines have only exacerbated the problem. Improved steering trucks [283] and rail grinding [203, 236] are the primary tools for its treatment, with friction management also gaining consideration [125].

An analysis was conducted of RCF-related publications between 1999 and 2010 as listed in the Scopus database. It should be noted that the Scopus database does not capture the numerous papers that are being presented on the subject in railway-related forums such as the American Railway Engineering and Maintenance-of-Way Association, the International Wheelset Congress, the International Heavy Haul Association, and the World Congress on Railway Research. The numbers of publications counted by Scopus will naturally be skewed toward academic institutions that place a strong emphasis on generating peer-reviewed publications. Also, in Scopus, where there were multiple participants in a publication, each one is given credit for the publication, which helps to explain why some commercial companies, which sponsor research at universities, show relatively high number counts. Despite these limitations, the summary in Table 1 is believed to a fair list of the major organizations conducting research into RCF.
Table 1. Listing of Those Organizations Publishing Papers Related to Studies of RCF

<table>
<thead>
<tr>
<th>Organization</th>
<th>No. of pubs.</th>
<th>Country/Headquarters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chalmers Tekniska Högskola (27)</td>
<td>27</td>
<td>Sweden</td>
</tr>
<tr>
<td>University of Sheffield (16)</td>
<td>16</td>
<td>U.K.</td>
</tr>
<tr>
<td>RTRI, Railway Technical Research Institute (7) + Vehicle Strength (1) + Track</td>
<td>11</td>
<td>Japan</td>
</tr>
<tr>
<td>Dynamics Laboratory (1) + Frictional Materials Laboratory (1) + Vehicle and Bogie Parts Strength (1) + Transportation Technology Center, Inc. (9) + TTCI U.K. Ltd. (1)</td>
<td>10</td>
<td>U.S./U.K.</td>
</tr>
<tr>
<td>University of Birmingham (9)</td>
<td>9</td>
<td>U.K.</td>
</tr>
<tr>
<td>Voest-Alpine AG (7) + Voest-Alpine Schienen GmbH (2)</td>
<td>9</td>
<td>International</td>
</tr>
<tr>
<td>AEA Technology Rail bv (4) + AEA Technology (4)</td>
<td>8</td>
<td>Netherlands/U.K.</td>
</tr>
<tr>
<td>Newcastle University United Kingdom (7) + Newcastle University (1)</td>
<td>8</td>
<td>U.K.</td>
</tr>
<tr>
<td>Semcon AB (6) (CARAN)</td>
<td>6</td>
<td>Sweden</td>
</tr>
<tr>
<td>Southwest Jiaotong University (6)</td>
<td>6</td>
<td>China</td>
</tr>
<tr>
<td>Imperial College London (5)</td>
<td>5</td>
<td>U.K.</td>
</tr>
<tr>
<td>Politecnico di Milano (5)</td>
<td>5</td>
<td>Italy</td>
</tr>
<tr>
<td>Kanazawa University (4)</td>
<td>4</td>
<td>Japan</td>
</tr>
<tr>
<td>Politechnika Warszawska (4)</td>
<td>4</td>
<td>Poland</td>
</tr>
<tr>
<td>Sumitomo Kinzoku Kogyo Kabushiki-gaisha (3) (Sumitomo Metal Industries Ltd.) + Amagasaki (1)</td>
<td>4</td>
<td>Japan</td>
</tr>
<tr>
<td>Centre for Surface Transportation Technology NRCC (3)</td>
<td>3</td>
<td>Canada</td>
</tr>
<tr>
<td>Corus Rail Technologies (3)</td>
<td>3</td>
<td>U.K.</td>
</tr>
<tr>
<td>Deutsche Bahn (3)</td>
<td>3</td>
<td>Germany</td>
</tr>
<tr>
<td>KRRI (2) + Korea Railroad Research Institute (1)</td>
<td>3</td>
<td>South Korea</td>
</tr>
<tr>
<td>RVD Consulting, Inc. (2) + RVD Consulting (1)</td>
<td>3</td>
<td>U.S.</td>
</tr>
<tr>
<td>Spoornet (2) + Spoornet, Materials Engineering (Rollstock) (1)</td>
<td>3</td>
<td>South Africa</td>
</tr>
<tr>
<td>Swinburne University of Technology (3)</td>
<td>3</td>
<td>Australia</td>
</tr>
<tr>
<td>TSC Inspection Systems (3)</td>
<td>3</td>
<td>U.K.</td>
</tr>
<tr>
<td>Università degli Studi di Brescia (3)</td>
<td>3</td>
<td>Italy</td>
</tr>
<tr>
<td>Vanderbilt University (3)</td>
<td>3</td>
<td>U.S.</td>
</tr>
<tr>
<td>Manchester Metropolitan University (3)</td>
<td>3</td>
<td>U.K.</td>
</tr>
<tr>
<td>All-Russ. Railway Research Institute (2)</td>
<td>2</td>
<td>Russia</td>
</tr>
<tr>
<td>Banverket (1) + Banverket (1)</td>
<td>2</td>
<td>Sweden</td>
</tr>
<tr>
<td>Booz Allen Hamilton, Inc. (2)</td>
<td>2</td>
<td>U.S.</td>
</tr>
<tr>
<td>Daneshgahe Elm va Sanat e Iran (2) (School of Mechanical Engineering, Iran University of Science and Technology, Narmak, Tehran 16884, Iran)</td>
<td>2</td>
<td>Iran</td>
</tr>
<tr>
<td>Delft University of Technology (2)</td>
<td>2</td>
<td>Netherlands</td>
</tr>
<tr>
<td>Hiroshima University (2)</td>
<td>2</td>
<td>Japan</td>
</tr>
<tr>
<td>Hunter Holiday Consulting (1) + Hunter Holiday Consulting (1)</td>
<td>2</td>
<td>U.S.</td>
</tr>
<tr>
<td>Interfleet Technology Ltd. (1) + Interfleet Technology AB (1)</td>
<td>2</td>
<td>Sweden/U.K.</td>
</tr>
<tr>
<td>Luleå tekniska Universitet (2)</td>
<td>2</td>
<td>Sweden</td>
</tr>
<tr>
<td>SP Technical Research Institute of Sweden (2)</td>
<td>2</td>
<td>Sweden</td>
</tr>
<tr>
<td>The Royal Institute of Technology KTH (2)</td>
<td>2</td>
<td>Sweden</td>
</tr>
<tr>
<td>The University of Warwick (2)</td>
<td>2</td>
<td>U.K.</td>
</tr>
<tr>
<td>University of Manchester (2)</td>
<td>2</td>
<td>U.K.</td>
</tr>
</tbody>
</table>
1.2 Safety Implications of RCF

In some operations RCF is the main contributor to broken rails and derailment. Examples include

- the previously mentioned Hatfield derailment in the United Kingdom in which 4 lives were lost and 107 people injured,
- a 2006 derailment in New Brighton, PA, where 23 tank cars derailed with several falling into a river and 20 of the cars releasing ethanol that burned for 48 hours (h) [190], and
- two RCF-caused broken rail derailments on South Africa’s Spoornet railroad in April 2004. The catastrophic rail failure was the result of gauge corner cracking initiating large (e.g., 80 percent of the head area) transverse defects.

Reference 20 notes that the loss of a length of the rail head, from a few millimeters to a couple of meters, is sufficient for precipitating a derailment and that defects that either run the length of the rail or appear in clusters are the most dangerous. The former category includes vertical split head and head-web separation, whereas RCF defects such as gauge corner cracking, transverse defects, and squats are examples for the latter. Isolated defects such as a clean rail break do not frequently lead to derailments because the gap that arises is normally bridged by the passing trains.

FRA maintains a database of accident statistics, and the frequency of rail failure related to rail defects can be extracted. For 3 years, 2006–2008, there were a total of 246 derailments as a result of cause “T220 Transverse compound fissure” and 145 derailments for cause “T207 detail fracture – shelling head check” with 40–50 percent as mainline derailments. These are the fourth and ninth primary causes of train accidents listed in the database [287]. Although the T207 incidents are clearly a consequence of RCF, only about half of the first category is due to RCF, the remainder occurring in older rail with internal defects.\(^1\) Taken together, the total of 268 RCF-caused derailments represents approximately 10 percent of all derailments over that period. In 2005 specifically, FRA reported track-caused derailments numbered 1,024 with 57 nonfatal casualties—RCF accounted for roughly 105 (10 percent) of those, with 7 (13 percent) nonfatal casualties.

The necessary joining of rails through welding is an ongoing location of rail failure, with 42 percent of total rail defects (1995) being from bolt holes and defective welds.

Analysis of FRA data from the 1980s [300] showed that one broken rail derailment occurs for every 770 defects. For contact stress-related defects (such as transverse defects and vertical split heads), the rate rises to 1 in 526. In a separate analysis [209], data for an unnamed class 1 railroad in the United States yielded that the number of rail breaks per derailment was approximately 300 in 1989.

Statistics summarizing the number of rails found damaged or broken are given in Reference 64 for several countries including France, Germany, India, Japan, the Netherlands, South Africa, and the United Kingdom.

1.2.1 (Rail) Ultrasonic No-Tests

Surface fatigue is a serious hazard because it prevents the successful transmission of ultrasonic waves to the base of the rail and back. Failure to detect the return signal results in a “no-test”

---

\(^1\) Personal communication with Michael Roney – Canadian Pacific Railway.
and the inspector is then usually required to stop and perform a hand test of the rail using a more sensitive manually operated system. This stopping and starting severely curtails the productivity of the inspection and reduces track availability for revenue traffic. Because of the risk of a rail break and derailment occurring from undetected subsurface defects, failure to achieve a test on many railroads requires removal of the affected length of rail.

1.3 Types of RCF Defects
The different types of RCF defects have been aptly reviewed by several authors over several years including References 2, 20, 57, 83, 90, 88, 174, and 175. They are summarized briefly in the following sections.

1.3.1 Rail Gauge Corner Cracking
These are thin cracks appearing at the gauge corner of the rail (Figure 1), appearing most often on the outer rails of curves and sometimes on tangent rails but infrequently on low rails. They develop as a result of high wheel-rail contact stresses coupled with sufficiently large shear stresses because of slip between the wheel and rail. Gauge corner cracking (GCC) is often regularly spaced and may occur for long lengths of track (e.g., the entire curve) or may be found in clusters. In the latter case, it is usually associated with track geometry perturbations.

![Figure 1. Well-Defined Gauge Corner Cracking from High-Speed Passenger Operation](image)

On freight railroads, the cracking when left to develop looks typically like the (dye penetrant enhanced) samples shown in Figure 2, with the gauge corner of the high rail and the field side of the low rail most heavily afflicted.
As GCC progresses, the cracks grow longer and deeper into the component and may eventually lead to pits and shells on the rail surface. Figure 3 shows examples of the different stages of crack growth on rails.
A) Moderate checking at the gauge corner of the rail.

B) Crack orientation changes due to different creep directions of leading and trailing axles.

C) High rail shelling.

D) Alternating wet/dry conditions→shallow shelling.

E) Moderate checking across the low rail surface.

F) Field side cracking eventually deteriorates into plastic flow and shelling.

G) Moderate spalling at crown of rail is difficult to remove through grinding.

H) Heavy, but shallow RCF on flattened low rail.

Figure 3. Examples of Typical High and Low Rail Checking and Shelling.
1.3.2 Wheel Tread Cracking

Cracking of the wheel surface parallels that of the rail: excessive cyclic shear stresses at the surface cause incremental plastic flow, and when the total strain exceeds the ductility limit, a crack forms. These cracks can initiate below the surface or at the surface, depending on the friction coefficient. The orientation of the cracks is perpendicular to the direction of traction, and because of the typical orientation of creepage, under wet conditions, the cracks at the field side of the wheel tend to propagate quickest [171]. Although extensive crack patterns can occur at the flange root of the wheel, it is unusual for them to propagate into shelling.

![Figure 4](image)

**Figure 4.** A) Cracks at the Field Side of the Wheel Are Typically Most Severe; B) Extensive Cracking at the Flange Root, Like That Seen Here, Is Much Less Common

1.3.3 Rail and Wheel Shelling

Shelling is the loss of large chunks of metal as a result of either:

- Surface and/or subsurface cracks linking up with other cracks below the surface such that the material above is no longer connected to the parent rail (Figure 5) or wheel (Figure 6) and drops out; or
- Collapse of the rail gauge corner because of excessive loading and shear failure along a “slip line.” Although shelling from gauge corner collapse occurs preferentially at stringers of alumina inclusions [35], even in the cleanest steel, too many high stress loads at the gauge corner will cause it to yield along a slip line [170]. This defect is often referred to as the “deep seated shell.”
Figure 5. Shelling of a Rail; Shown Are Longitudinal Sections Machined Vertically (parallel to the rail web) at Distances Ranging from 4 to 28 mm from the Gauge Side of the Rail; The Steel Is a 1984 CC CFI Rail Steel
A) Microcracks on the field/rim side of a heavy haul wheel.

B) Microcracks on the field/rim side of the high-speed wheel.

C) Well-defined cracks on the field/rim side of the wheel.

D) Deeper cracks on the field/rim side of the wheel tread with material starting to shell out.

E) Incipient shells initiating at the field/rim side of the wheel.

F) A fully shelled wheel.

Figure 6. Progression of Wheel-Tread Surface Fatigue from Microcracks to Fully Shelled [171]
A) Cross section of rail with the microstructure enhanced shows gauge corner collapse.  
B) Examples of shelling from gauge corner collapse.

Figure 7. Gauge Corner Collapse Can Lead to Deep-Seated Shells with Very Large Chunks of Metal Spalling from the Rail

Poorly restrained high rails that rotate outwards under load (Figure 8A), well-lubricated rail with a particularly high-stress gauge corner shape (Figure 8B), and even some dry rails (Figure 8C) see large chunks of metal slough from the gauge corner. The most effective treatment for this is ensuring the implementation and maintenance of proper rail shapes that eliminate the overloading, reducing the stress for each contact and distributing contacts over a wider portion of the railhead. Modeling the system as a quarter space (see Section 2.1.2.1) found that in practice it is necessary to minimize the number of wheels that touch the railhead within three-eighths of 1 inch (9 mm) of the gage face to avoid gauge corner collapse.
A) Loss of rail cant due to rail seat abrasion resulted in overloading of the gauge corner  
B) Well-lubricated rail with a sharp gauge corner  
C) Gauge corner collapse in a dry environment

**Figure 8. Examples of Gauge Corner Collapse**

In any case, shells typically develop under the gauge corner of the high rail and most commonly on heavy haul railroads. However, they have also been found on tangent track with local crosslevel errors (see Section 4.5 for examples) and on light axle load transit systems with particularly poor wheel-rail profile matching (Figure 9).

**Figure 9. Gauge Corner Fatigue on a Light Rail System Because of Flattening of Rail (with a reciprocating stone grinder) and Overloading of the Rail Gauge Corner**

### 1.3.4 Rail Transverse Defects

Transverse defects such as those shown in Figure 10 are the cause of many broken rails and subsequent derailments. These defects are initiated by shells in the presence of metallurgical inclusions, most commonly oxide stringers. Although the defect may be initiated by RCF, the development of the defect itself is not a rolling contact phenomenon. The transverse crack propagates as a result of the combination of bending stress from wheel loads, residual stress from manufacture, and thermal stress from rail contraction in winter. This process has been examined in detail by several authors [e.g., 29, 118].
Standard railroad classifications consider the rail defects listed in Table 2. Of the 22 listed, only 3 defects are directly the result of RCF.

Table 2. Rail Defect Classifications

<table>
<thead>
<tr>
<th>DTYPE</th>
<th>Definition</th>
<th>DTYPE</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>BBJ</td>
<td>Broken Base – Joint</td>
<td>TDC</td>
<td>Compound Fissure</td>
</tr>
<tr>
<td>BBO</td>
<td>Broken Base – Outside</td>
<td>TDD</td>
<td>Detail Fracture</td>
</tr>
<tr>
<td>BHJ</td>
<td>Bolt Hole Crack – Joint</td>
<td>TDT</td>
<td>Transverse Fissure</td>
</tr>
<tr>
<td>BHO</td>
<td>Bolt Hole Crack – Outside</td>
<td>TDW</td>
<td>Welded Engine Burn</td>
</tr>
<tr>
<td>BRJ</td>
<td>Broken Rail – Joint</td>
<td>HWJ</td>
<td>Head &amp; Web - Joint</td>
</tr>
<tr>
<td>BRO</td>
<td>Broken Rail - Outside Joint</td>
<td>HWO</td>
<td>Head &amp; Web - Outside Joint</td>
</tr>
<tr>
<td>DWF</td>
<td>Defective Field Weld</td>
<td>PRJ</td>
<td>Piped Rail - Joint</td>
</tr>
<tr>
<td>DWP</td>
<td>Defective Plant Weld</td>
<td>PRO</td>
<td>Piped Rail - Outside Joint</td>
</tr>
<tr>
<td>EBF</td>
<td>Engine Burn Fracture</td>
<td>SWJ</td>
<td>Split Web - Joint</td>
</tr>
<tr>
<td>HSH</td>
<td>Horizontal Split Head</td>
<td>SWO</td>
<td>Split Web - Outside Joint</td>
</tr>
<tr>
<td>HSJ</td>
<td>Horizontal Split Head - Joint</td>
<td>VSH</td>
<td>Vertical Split Head</td>
</tr>
</tbody>
</table>

Figure 10. Transverse Defects Initiated by Deep-Seated Shells, Probably from Gauge Corner Collapse
1.3.5  *(Rail) Squats*

The squat is a surface defect most commonly associated with high-speed rail and areas of high tractive effort. It is characterized as a shallow depression more or less in the center of the rail head on tangent and mildly curved track. This depression is the result of subsurface cracking, reduced strength of the material, and deposits of debris in that depression, giving it the appearance of a dark spot.

A cross section of the squat shows two cracks; a short leading crack (in the direction of travel) and a much longer trailing crack each propagating at approximately 30° from horizontal (Figure 12B). Because the stress intensity at the tip of the trailing crack is higher than on the leading crack for driving wheels [95, 143], the trailing crack grows longer and more rapidly in areas of high tractive effort. From the trailing crack, many small branch cracks tend to initiate, one of which may turn downward into the head of the rail and initiate a transverse fracture. These cracks are especially problematic because, shielded by the long trailing crack, they are often not detected by conventional ultrasonic measurement systems (see Section 1.2.1).

The causes of squats continue to be a subject of some debate, with two different mechanisms currently proposed:
1. On the Japanese railways, a thin martensitic layer exists along extensive stretches of the rail. This is postulated to be associated with the high levels of tractive effort used on those systems and microslip at the wheel-rail contact. The brittle martensite readily cracks and initiates a surface crack that then propagates, in response to the rolling contact forces, leading to trailing cracks typical of squats [112].

2. As a general principle, high tractive effort is a leading cause of excessive surface shear that initiates classical RCF cracks [43]. These too may initiate surface cracks that then develop into the leading and trailing cracks seen in Figure 12B.

The treatment in all cases has been the frequent grinding of a thin layer of metal from the rail surface to remove the most damaged surface layer where the squats are initiated. Frequent preventive grinding has proven to be an effective treatment on the Shinkansen system [148]. From a detection and maintenance perspective, Li et al. [156] have found that the dimensions and location of squats could be correlated with certain wheel-rail dynamic interactions (force and wavelength), leading the way perhaps eventually to potential early detection and preventive maintenance.

1.3.6 (Rail) Crushed Heads

Sometimes RCF cracks progress toward the center of the rail head, and with subsequent fretting at the crack faces and shearing of material, this area of lower structural integrity “crushes” under the heavy loading. The local vertical perturbation that results is then subjected to further deformation and plastic flow from increasingly heavy impact loads from passing wheels (Figure 13).

Crushed heads are typically about one-quarter to one-half meter in length. They are characterized by increased frequency and growth of large surface cracks and plastic flow to one or both sides of the running surface. The concentration of cracks is usually caused by a local stress raiser associated with improper rail grinding, lack of rail grinding, track irregularities, dirty steel, or some combination of all of these factors. Grinding can prolong the life of crushed-head–affected rail by either removing damaged material or moving the contact band away from the stress raiser [172]. Crushing is especially problematic on older rail, often having significant vertical wear, where the head may collapse entirely. The crushed head defect and its treatment are discussed in Reference 47.
1.3.7 (Wheel) Spalling

Sliding of the wheel leads to severe temperature increases at the wheel-rail contact patch, and damage to the wheel is especially acute as the thermal energy generated at the slide is distributed over a relatively long length of rail but only over a very short portion of the wheel (the opposite case occurs for the spinning wheel). The temperature on the slower moving contact (the wheel in this case) often rises beyond the transition temperature of pearlitic steel (723°C). When the wheel turns further, the contact patch cools very rapidly, and a white-phase martensite forms. This martensite is very brittle and cracks when the wheel rolls over it. These surface cracks then propagate into the wheel surface. In addition, the spall has changed the geometry of the running surface and is associated with impact. So at the wheel slide location on the wheel tread, there is a combination of brittle martensite and impact. The martensite “spalls” out, and with heavy impacting there is often little evidence of it still in the contact region. Although the mature defect might look like a regular wheel shell, its initiation mechanism, and hence the methods for its prevention, is very different. Reference 174 discusses the considerable differences between shells and spalls and how they can be distinguished.
A light martensitic contact patch at a wheel slide location with cracking starting.

Chain flats (etched in this figure) due probably to a dragging brake on a lightly loaded car, can initiate spalling around the circumference of the wheel.

Figure 14. Examples of Wheel Surface Martensite

On the rail, a completely analogous effect arises because of the spinning of the wheel on the rail. High tractive effort can leave a more or less continuous thin layer of martensite on the surface of the rail, which has been postulated as the causal mechanism of rail squats (see Section 1.3.5). On a macroscopic scale, a wheel burn is a deep layer of martensite that, upon subsequent wheel passes, develops a whole network of surface cracks, anyone of which might be the cause of a broken rail.

Figure 15. A) Thin Streaks of Martensite on the Rail Surface Because of High Creep Rates from Wheels under High Traction; B) a Spinning Wheel Can Burn a Deep Patch of Martensite into the Rail

1.3.8 (Wheel) Shattered Rim

The shattered rim is associated with a subsurface defect and adverse internal stresses. It is not a common defect—in 1999, approximately 300 were reported [262] in the North American interchange service, accounting for approximately 0.06 percent of wheel removals. However, because of the catastrophic potential of a failure, they are considered to be a problem. The
Union Pacific Railroad reported that 65 percent of its railroad wheel failures were due to shattered rims [159].

The shattered rim crack usually initiates approximately 12–20 mm below the running surface and then propagates roughly parallel with the wheel tread, eventually exiting out the rim side of the wheel. If a rim crack is discovered, then the wheel can be removed proactively; otherwise, a section of the wheel rim can separate from the wheel [262]. Many shattered rim failures develop early in wheel life—more than half fail within the first 10 percent of available rim thickness and more than 90 percent of the failures occur within the first 50 percent of the available wheel thickness.

Metallurgical examination of shattered rim failures from a high-speed train operation reported in Reference 290 found that the bulk are due to clusters of hard (usually aluminosilicate) defects and that the best treatment is ensuring steel microcleanliness and NDT testing. Reference 262 reports that in the freight experience, the key to avoiding shattered rims is to minimize voids and porosity in cast wheels and aluminum oxide inclusions in forged wheels. A combination of ensuring internal defects smaller than 1 mm, coupled with more rigorous ultrasonic inspection, has eliminated the problem of shattered rims in South Africa [177]. The same is true in North America, and the improved casting process has considerably reduced the number of shattered rims [259], but research is ongoing to optimize the inspection and maintenance plan for the hundreds of thousands of legacy wheels of older metallurgy still running in the system [159].

1.3.9 (Wheel) Vertical Spit Rim

The vertical split rim is a defect that originates from a wheel shell (Section 1.3.3) or spall (Section 1.3.6) with a significant circumferential length of the wheel tread, usually from the rim/field side, separating from the wheel. Inspection of several such failures found that the wheel typically has a hollow tread and false flange, so that rim/field side loading (such as when the wheel contacts the low rail in a curve) could cause sufficiently high tensile bending stress precipitating the failures.
Interestingly, the failures do not seem to relate strongly to the remaining thickness of the rim, nor was the type of metallurgy, wheel diameter or production process (cast versus forged) a common factor.

Less commonly, a length on the flange side can break from the wheel.

1.4 Economic Implications of RCF

The replacement of rail represents a large capital expense for any railroad, ranging from about USD375,000 per track-mile on a conventional freight railroad to USD2 million on underground mass-transit lines. Reducing the unnecessary loss of metal for fatigue related reasons can provide benefits exceeding many millions of dollars annually to a railroad. Data from North America suggest that 15–22 percent of all rail replacement is due to surface and subsurface-initiated defects [20]. For the U.S. and Canadian class 1 railroads in 2007, 20 percent of roughly 700,000 tons of rail purchased represents USD110 million, whereas 20 percent of the total rail replacement costs (includes installation costs) recorded by those same railroads amounts to at least USD220 million. The cost of RCF defects to the European rail system was estimated in 2000 to be roughly €300 million (USD417 million) annually, but following the Hatfield derailment and increased recognition of RCF in Europe, that number has certainly risen [88]. In the United Kingdom alone, extra penalty payments to train operators after the Hatfield derailment and replacement costs of switches or crossings in which gauge corner cracks were found amounted to £561 million (USD917 million) in 2000–2001 [208]. In 2004, the annual cost of RCF to network rail alone was estimated as being at least £200 million (USD327 million) and included RCF cracking of wheels [51]. In 2005, the most recent year for which cost information is available, RCF would account for approximately 105 derailments and USD30 million in FRA reportable costs, along with 7 nonfatal casualties. On the mechanical side, 22 accidents as a result of broken wheel flanges and broken rims amounted to nearly USD11 million in FRA costs in 2005. Since FRA reported costs

---

2 At USD800 per ton, as per Mike Roney via email 26JAN10.
that do not include the cost of clearing up the train wreck or the major cost of damage to the rolling stock, the true costs are likely to be considerably higher—roughly twice as much according to Reference 301.

A journal review in *Railway Age* [161] quotes an FRA source:

“Rail flaws are a top industry liability. Estimated to cost $455 million annually—including $109 million for related derailments; $70 million for train delays; $149 million for repairs and maintenance; and $127 million for inspections.”

Internationally, RCF in the late 1990s accounted for approximately 60, 25, and 15 percent of all defects found by East Japan railways, French Railways (SNCF), and Railtrack (U.K.), respectively [20]. Rail grinding internationally is a USD250–350³ million business, and it is one that is growing as knowledge and demand rises in countries such as India, China, and Brazil. In North America, the costs for rail grinding on a large freight railroad are typically on the order of USD800 per mile of track, with the North American market for rail grinding being roughly USD100–120³ million per year. Although grinding was in the 1970s initially tasked with treating corrugation, modern steels and regular rail grinding have caused RCF to supersede corrugation as the main cause for rail grinding. If RCF could be eliminated by other means, such as improved steels, profiles, or friction management, then the amount of rail grinding required annually could easily be reduced by 30 percent or more.³

The annual cost of rail inspections for defects is not known, but although not all defects are due to RCF (other causes include broken welds, base plate cracks, etc.), there is no doubt that a considerable fraction of the cost can be attributed to the problem. In the European Union, the cost of rail defects, a large percentage of which are initiated by RCF, has been estimated at €2 billion (USD2.8 billion) per year and, on this basis, became the UIC’s first World Joint Research Project [20]. The cost of research into RCF is significant in itself—for example, the European ICON project was a roughly €1.5 million (USD2.1 million) effort to address RCF [19].

Next, add to this the value of lost revenue because of track outages for inspection and maintenance, along with the safety impacts already discussed in Section 1.2, and clearly, there are opportunities for improved understanding, inspection, and treatments to significantly reduce the economic costs of RCF.

³ Numbers provided by Alan Zarembski, author of *The Art and Science of Rail Grinding* (Simmons Boardman Press, 2005).
2. Formation of RCF

There have been many efforts to model both the initiation [e.g., 28, 85, 126, 248] and propagation aspects [e.g., 12, 94, 199, 295] of surface fatigue, with the latter gaining the lion’s share of the attention. Yet, although these models were useful as “postmortems” or to understand some interesting property of a rail material or wheel-rail contact phenomenon, their lack of generality limited their contributions to the development of practical models for optimizing preventive measures, such as rail grinding. But by contributing to the general understanding of how cracks grow, they are effective in supporting the existing best practice that is based on the following general model of crack growth:

A. High stress contact between the wheel and rail leads to ratcheting of the surface material and eventual development of surface cracks by failure in ductile shear [123]. This process was estimated to take place over a period of 3–6 MGT [135]. Newly initiated surface cracks, based on metallurgical analysis, were found to propagate at an angle of 5–15° with respect to the surface and were fractions of a millimeter in length. In some cases, this ductile shear can result in thin metal particles separating from the surface, serving effectively as a mechanism of wear [78, 140].

B. These surface microcracks, which are short, propagate slowly in the initial stage, with the growth rate increasing as the crack length gets longer [95]. These cracks will continue to propagate at a relatively slow rate to the depth of maximum shear stress. This depth was known to be somewhere approximately 1 mm (using Hertzian theory) in a high traction environment and as deep as 8 mm (more typically 2–4 mm) in a heavy haul environment.

C. In a dry environment, these cracks will remain relatively dormant at that length. Yet, in moist environments, or in alternating wet and dry environments, the hydraulic propagation mechanism demonstrated by Way [292] in 1935 and quantified analytically and extended by Bower [12] contributes to continued and rapid crack growth to a depth of several millimeters. Once the crack grows large enough that it cannot be fully closed by the contact patch, hydraulic pressurization is no longer effective. The surface length of the cracks will be on the order of 15 mm, with a depth of approximately 4–8 mm (Figure 18).
Further propagation is dependent on the presence of internal stress raisers (e.g., metallic inclusions) and the combined pattern of bending, thermal, and residual stresses. Some cracks will turn downward and then may develop into a transverse defect and eventually a broken rail. Most of these cracks will propagate toward the surface and lead to shelling, with chunks of material spalling from the rail surface. Even though the rate of crack growth is not expected to be particularly rapid in this stage, the cracked and spalled surface will introduce some noise into the ultrasonic rail flaw detection signal and compromise the testing program; it constitutes an unsafe operating condition.

This shelling is in many cases considered benign because it is not generally a direct cause of broken rails. However, the dynamic forces that arise may contribute to corrugation, localized ballast damage and pumping ties, plastic flow, and crushed heads, which then precipitate further track damage and incur considerable maintenance costs.

The various factors influencing crack initiation and propagation are the subject of this section.

2.1 Crack Initiation

2.1.1 Surface Crack Initiation in Rails

RCF is, by definition, damage of components that arises because of repeated loading associated with rolling contacts. Rolling contacts are frequently used to transmit heavy loads with minimal energy loss (i.e., minimal rolling resistance). Common applications include bearings, gears, and railways. The efficiency of the rolling contact arises because of the small area of contact over which the load is transmitted. A consequence is that the stress on rolling contacts is relatively large.

A typical piece of rail in a freight or commuter system carries on the order of 3–5 million wheel load cycles every 100 MGT, whereas the wheels undergo approximately 50 million rotations every 100,000 miles. A percentage of these cycles, under conditions of high friction and high contact stress, deforms the wheel and rail metal in the direction of the applied stress [85]. The accumulating increments of deformation “ratchet” [121, 141] the surface layer. Although the effective strength of the material is increasing progressively because of strain hardening, it
cannot strain indefinitely. Eventually, its ductility is exhausted, and a surface crack forms (Figure 19). Several models for this process have been developed (see Section 2.1.8). The most popular one refers to the shakedown process and is described next.

![Image](image.png)

**Figure 19. Ratcheting in Rail Steels Associated with Contact Fatigue**

### 2.1.1.1 Shakedown

When a wheel rolls over a rail, the stress varies cyclically with a severity that is governed by the specific geometry of the contact, material properties, load pressing the components together, friction coefficient, and creepage. For a frictionless contact, the stress varies according to Figure 20A, whereas Figure 20B shows how those same stresses vary for a friction coefficient of 0.3.

The values of the principal stresses in the longitudinal plane, $\tau_{xz}/P_o$ versus $(\sigma_{xx} - \sigma_{zz})/2P_o$, are plotted in Figure 21. Johnson refers to these as the stress trajectories. This plot represents the deviatoric (the dynamic or nonhydrostatic) stress cycle that an element of material on ($z/a = 0.0$) or below (e.g., $z/a = 0.5$) the surface encounters because of the passing wheel load. If that stress cycle can be contained within the yield space (a circle that represents the strength of the material), then the cycle will be wholly elastic. If the stress cycle falls outside the yield space, then an increment of plastic strain occurs. The difference between an elastic-plastic material and one that hardens is illustrated in Figure 22. In this way, Johnson’s shakedown diagrams (Figure 23) were developed.
Figure 20. Cyclic Stresses Associated with a Rolling Contact, Nondimensionalized with Respect to the Width of the Line Contact (2a) and the Peak Normal Stress ($P_o$)

Figure 21. Plot of the Stress Trajectory ($\sigma_{xx} - \sigma_{yy}$) / $2P_o$ versus $\tau_{xy} / P_o$ for One Complete Rolling Cycle at Different Depths $z/a$ below the Surface
2.1.2 Subsurface Crack Initiation

The stress plots of Figure 20 show that the subsurface shear stresses can be very high. Under well lubricated conditions in which the friction coefficient is very low, the maximum shear stress is below the surface. Contained, subsurface plastic flow is possible and, in conjunction with metallurgical imperfections, can initiate subsurface defects. These are common in bearings where pitting develops from subsurface-initiated cracks [68, 71]. Even under higher friction
coefficients when the maximum shear stress is at the surface, there still may be sufficient subsurface stress to initiate those cracks.

On the rail, subsurface fatigue occurs most commonly as a result of poor pummeling at the gauge corner of overloaded high rails, resulting in gauge corner collapse (see Section 1.3.3). Models of subsurface cracking in wheels generally presume that a subsurface defect exists and that the combined rolling, bending, and residual stresses act to initiate the crack around that defect [59].

2.1.2.1 Gage Corner Collapse

High values of contact stress at the rail gauge corner, coupled with an excessive number of loading cycles there, contribute to a longitudinal shelling failure known as gage corner collapse. When Steele [255] looked at cross-sectional hardness profiles of rails taken from the FAST track that exhibited “unexpected shelling,” there was clearly a concentration at the extreme gauge corner. These were taken from rails that were intentionally ground to a heavy single-point contact, called “conformal” at that time but now known as “single-point non-conformal” [173] (Figure 24).

In a 1993 publication, Bower et al. modeled the problem as a Hertzian pressure distribution travelling on an infinite elastic-plastic quarter space (Figure 25). A ratcheting limit identified for incremental lateral collapse was developed by applying the shakedown limit to an optimal arc of collapse. Harder rails, gauge corner lubrication, and grinding of the gauge corner to two-point contact were identified as practical approaches to minimizing shelling from gauge corner collapse [132].
Although the problem of gage corner collapse and its solution are well known, a recent example on the Canadian Pacific Railway (CPR) shows that, even with modern maintenance practices, problems can still occur. On CPR, significant improvements to rail lubrication practices between 2001 and 2005 led to dramatic reductions in gage face and gage corner wear. Because there was little change in shape from one cycle to the next, rail grinding was dramatically reduced. Yet more importantly, the facet created between the last (45°) grinding stone and the gage face was no longer being rapidly worn from the rail. This facet was serving as a stress raiser that led to numerous small gage corner shells. The solution was to extend grinding down to 60° and to ensure that high rails in the 100-percent effective lubrication territories were being ground preventively—even if there were no visible signs of distress [249].

2.1.2.2 Dang Van Criterion

The Dang Van Criterion [42] allows for a complex rotating stress state that involves all six components of the stress tensor and where the principal shear plane may not be known. It is used chiefly for subsurface fatigue.

Dang Van suggests that crack initiation is due to excessive shear stresses, and when the effective stress \( \tau_{eq1} \) exceeds the shear strength \( k \) of the material, fatigue will initiate:

\[
\tau_{eq1} = \tau(t) + A \sigma(t),
\]

where \( A \) is a dimensionless material parameter.

The effective stresses are compared directly with the material strength in shear \( k \). As an example, the fatigue cycle of Figure 21a (at \( z/a = 0.5 \)) is shown in Figure 26 with \( A \) set to 0.32 [58] and \( P_0 = 1,600 \text{ MPa} \). The portion of the cycle that falls outside the yield space will be associated with incremental strain.
Figure 26. With the Dang Van Criterion, Plastic Strain Will Occur in That Portion of the Stress Cycle a<t<b

Example applications of the Dang Van Criterion in numerical simulation can be found in References 41 and 62. A refinement to that criterion is proposed in Reference 45.

2.1.3 Contact Stress

The calculation of contact stress continues to be an ongoing topic of study. Most commonly, the railway contact is presumed to be quasi-elastic, and a range of analytical means is applied. Such solutions presume that the size of the contacting bodies is large compared with the area of contact. Hertzian contacts presume the two bodies have a constant radius in the region of contact (i.e., they are continuous up to the second derivative) and that they are nonconforming. Hertz furthermore only considers the frictionless contact. To include the effects of traction and creepage, the methods of Timoshenko [273] or Hamilton [93] are typically used. The most popular implementation of the elastic contact calculation is Kalker’s UseTab [289] as implemented in the CONTACT software. Once the geometry of the contact, the longitudinal, lateral and spin creepages, the friction coefficient, and the elastic properties of the bodies (elastic modulus and Poisson’s ratio) are known, a lookup table is queried to give nondimensionalized values of longitudinal, lateral, and spin forces and stress. As illustrated by Toma and Sawley [275], the details of the curve-fitting, while not much affecting the calculated forces, can have a significant impact on the calculated contact area and stress. Although they did not specifically look at the wear or fatigue index, one can presume that they also were affected by the details of the curve fit. For this reason, it is now common for vehicle dynamics packages to offer the option of calculating the local radii online (e.g., once the point of contact is known, calculate the
best fit longitudinal and transverse radii for each body) instead of the older approach of precalculating contact radii over some nominal length. The former approach requires much more computing effort.

A more detailed mapping of the contact and stress conditions can be achieved using non-Hertzian elastic approaches, including approximate methods like Kik’s [146] or more complete three-dimensional solutions such as Kalker’s Contact [289] and Paul and Hashemi [204]. An example of a non-Hertzian, nonelliptical contact condition is shown in Figure 27 in which the stress distribution is elliptical in the longitudinal direction but across the rail can be highly nonelliptical. Finite element [242, 271] and boundary elements [149] are also commonly used, especially when calculating stresses in the presence of cracks, internal voids, and near the gauge corner of the rail [69].

![Figure 27. Examples of Nonelliptical Contacts Stress Distributions Calculated Using Non-Hertzian Models](image)

### 2.1.4 Traction

Rail/wheel shear forces develop as a result of a small relative slip between the rails and wheels. The amount of slip (known as creep) depends on the steering and traction demands. These creep forces, or tractions, cannot exceed the product of normal force and available adhesion (μ).

The effect of traction on damage in the wheel-rail contact can be seen in the stress cycles of Figures 21 and 22 and in the shakedown diagram of Figure 23. Tractive forces significantly increase the surface shear stresses and the probability of any wheel-rail contact (e.g., P_o/K, F_T/F_N) exceeding the shakedown limit and causing an increment of plastic flow. These can be evaluated using the models described previously.
From a quite practical consideration, the increased rates of RCF on the wheels of modern vehicles with improved wheel spin and creepage controls is believed to be attributable to the very high levels of creepage possible with these systems [284].

2.1.5 Material Strength
The effectiveness of advanced steels in resisting RCF has been discussed in numerous papers, with References 87, 129, 179, and 234 providing good overviews.

Between 1965 and 1995, the average freight car axle load more than doubled and annual tonnage on mainline track quadrupled. In spite of this, the typical rail wear-life doubled in tangent track and tripled in curves (Table 3). Improvement in metallurgical and mechanical properties of rail steel [235] can be credited with a large portion of that performance improvement. Upgraded truck design and maintenance, lubrication, and profile grinding have also contributed to the increase in rail life, although the gains in the 1995–2005 timeframe were relatively modest.

### Table 3. Demands on Rail (from Reference 124)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Average freight car axle load</td>
<td>15</td>
<td>25</td>
<td>30</td>
<td>35</td>
</tr>
<tr>
<td>Annual tonnages (MGT)</td>
<td>15–30</td>
<td>25–100</td>
<td>30–160</td>
<td>60–220</td>
</tr>
<tr>
<td>Typical rail life: (MGT)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Curve track</td>
<td>300</td>
<td>500</td>
<td>700</td>
<td>850</td>
</tr>
<tr>
<td>Tangent track</td>
<td>700</td>
<td>1,000</td>
<td>1,500</td>
<td>1,600</td>
</tr>
<tr>
<td>Rail hardness</td>
<td>250 HB</td>
<td>300–320</td>
<td>340–370</td>
<td>380–405</td>
</tr>
</tbody>
</table>

2.1.5.1 Hardness
One of the most important metallurgical properties of the rail steel is its hardness [248]. It governs wear resistance, RCF resistance, and mushrooming (plastic flow) of the rail head. For any given type of steel structure, RCF development decreases with hardness thanks to improved resistance to the ratcheting (unidirectional strain of the rail surface layers, see Figure 42B and Section 2.1) that initiates RCF. On the basis of its as-manufactured hardness, 400-HB hypereutectoid premium steel can withstand contact stresses approximately 45 percent greater than standard rail steel (275 HB) without yielding [168]. The benefit of increased hardness is recognized internationally as the survey of commonly used high-strength rails in Table 4 shows.
Table 4. Mechanical Properties of High-Strength Rails (Reference 108, Table 3.3)

<table>
<thead>
<tr>
<th>Property</th>
<th>USA, Canada, Brazil</th>
<th>South Africa</th>
<th>China</th>
<th>Russia GOSTR 51685</th>
<th>Sweden</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yield Strength, MPa (min)</td>
<td>758</td>
<td>640</td>
<td>805</td>
<td>794</td>
<td>640</td>
</tr>
<tr>
<td>Tensile Strength MPa (min)</td>
<td>1172</td>
<td>1080</td>
<td>1175</td>
<td>1176</td>
<td>1080</td>
</tr>
<tr>
<td>Elongation % (min)</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>Brinell Hardness at the surface</td>
<td>340-390</td>
<td>340</td>
<td>340-390</td>
<td>331-388</td>
<td>320-360</td>
</tr>
</tbody>
</table>

2.1.6 Thermal Softening

Relative slip at the wheel-rail interface (an inevitable circumstance as discussed in Section 2.1.4) that gives rise to temperature increases that can compromise the strength of the steel, and even lead to metallurgical transformation.

At temperatures above the ferrite-austenite transformation temperature (approximately 1,330°F, 725°C), rail and wheel steels can undergo metallurgical transformation. At lower temperatures, in the range of 930°F (500°C), thermal softening reduces the material strength [282]. The accumulation of thermal loads, for example on a wheel due to tread braking, leads to a gradual release of the surface compressive residual stresses and in severe cases (e.g., unreleased handbrake or stuck brakes) can lead to tensile stresses at the surface. Tensile stresses may only cause benign thermal cracking of the tread, while in the worst case, usually in combination with a surface defect, can lead to a crack that propagates into the wheel plate and a broken wheel derailment.

The Wheel Defect Prevention Research Consortium organized a workshop in 2008 to examine the issue. In a presentation by Stone and Cummings [261], it was noted that although there was a loosely defined “safe operating range” for wheel temperatures of between 500 and 1,000°F, a better definition of a maximum operating temperature was needed. They note that at high temperature, the resistance to shear decreases but at the same time, a reduction in the modulus of elasticity decreases contact stress. At high temperatures, the beneficial compressive residual stresses that typically develop during rolling contact can be relieved (the amount depending on time and temperature, see Figure 28), contributing to a significant reduction in hardness.

Examinations of data logs found downgrade braking events typically occur over time intervals up to 0.5 hours, which allows for significant stress relief to occur at temperatures above about 750°F. In the end, those authors suggest that temperatures up to 600°F are not contributing to wheel fatigue and further that such a temperature is unlikely to arise with properly functioning brakes. However, high wheel temperatures are likely with malfunctioning brake systems that
unevenly distribute the braking effort on a wheel-by-wheel and car-by-car basis or on heavily driven locomotive wheels [92].

![Graph showing the rate at which residual stress is relieved as a function of temperature and time.](image)

**Figure 28. The Rate at Which Residual Stress Is Relieved Is a Function of Temperature and Time [198]**

Appreciating then that excessive wheel temperatures can be an important contributing factor to RCF suggests that railroads should place greater emphasis on the monitoring and rectifying conditions that lead to high wheel temperatures. Current wayside equipment is suitable for this purpose [152] and needs only to be coupled with an appropriate data management system and corresponding maintenance process.

Reference 147 concludes that “As long as a creep value of less than 2 percent and a coefficient of friction of less than 0.6 are assumed, the calculated temperatures do not exceed 450–500°C, even if high dynamic normal forces and irregularities are taken into account,” further suggesting that creep levels of 3–4 percent would be required to achieve the 600°C mark. This clearly is not the normal case for most wheels, but powered wheels under heavy tractive effort and improperly braked wheels can see such levels.

From a materials perspective, it has been found that high temperature stability of the pearlitic structure can be improved through additions of silicon and manganese to better resist carbon diffusion and spherodization of the cementite lathes [39].

### 2.1.7 Material Embrittlement

The effect of high temperature excursions causing brittle martensitic layers and subsequent spalling has been discussed in Section 1.3.6.

The effect of low temperature on rail steels has been investigated in China [297] where it was concluded that as temperatures dropped toward -60°C, the strength of the steel increases about 10 percent, while the ductility decreases at a significantly higher rate. The higher carbon steel (U75) dropped to approximately one-third of its value at 0°C, whereas for a standard carbon rail steel (U71), the ductility dropped only about 25 percent over the same temperature range. The standard metallurgy was therefore recommended for application to the Qinghai-Tibet railway. A different set of measurements [63] conducted at +20 and -125°C showed similar results—the
0.2-percent offset yield strength increased by over 40 percent, whereas the fracture strain decreased about 30 percent at the lower temperature. Crack growth rates under three-point bending at 20 hertz (Hz) at +20 and -125°C showed that the net result being an increase in the threshold fatigue stress intensity. A third set of fracture toughness tests at temperatures from -50 to 175°C found typically a plateau between 0 and 50°C, but that at -50°C, the impact energy for wheel steels decreases about 20 percent at the tread [151]. Bainitic steels also have good toughness at low temperatures with an experimental wheel exhibiting a two- and four-fold improvement over Association of American Railroads (AAR) standard Class C wheels at +20 and -40°C, respectively [86].

The interpretation of these results from the perspective of RCF is that the higher strength at low temperatures suggests improved resistance to plastic flow and ratcheting, whereas the increased brittleness means greater rates of crack growth, especially of transverse rail cracks that lead to broken rails. This is no different than the classical tradeoff encountered by virtually all metallurgists seeking to simultaneously improve strength and toughness. In the railroad context, higher crack growth rates in cold weather suggest that measurement and inspection intervals should be decreased in the winter.

It should be noted that under the rolling-sliding wheel contact, surface temperatures can be significantly higher than ambient. The bulk temperature of rail after the passage of a 110 car heavy haul train increases by roughly 10–15°C, whereas the wheels are generally hot to the touch under the combined action of curving and tread braking, even under very cold ambient conditions.

### 2.1.8 Models of Crack Initiation

Several approaches to modeling crack initiation have been proposed and include equivalent strain, critical plane models, energy density, empirical models, and combinations thereof. These are detailed in Reference 211 and are described briefly below:

- **Equivalent strain:** this suggests that fatigue initiation should be calculated on the plane that sees the maximum shear strain range. This approach has been successful for some loading conditions but correlates poorly with the results of tension/torsion tests and other nonproportional load cycles. It fails to account for the out-of-plane mean stresses associated with nonproportional loading.

- **Critical plane:** the critical failure plane is identified on the basis of the expected mode of failure, the type of loading, and the material properties. It is possible to account for factors such as mean stresses that separate the surfaces during the primarily shear portion of the loading cycle. Unfortunately, available critical plane models are unable to satisfactorily account for material hardening.

- **Energy density:** the fatigue process is related to the cyclic plastic shear that works through the movement of microscopic dislocations. This approach works well for the multiaxial stress state in which the plastic work and strain density per cycle can be evaluated. In energy-based models, the calculated parameters are scalar quantities, and the critical plane of crack initiation cannot be predicted. However, energy density models, which use the products of stress and strain, presume that cracks will grow in the location and direction of the largest energy density values. This approach is widely used in mode 1 tensile tests in which there is a positive mean stress. A modification allows it to be applied to a system that includes a negative (compressive) mean stress such as exists in rolling contacts.
• Combined critical plane and energy density: these have been shown to fit experimental fatigue life data and crack plane orientation very well. They tend to include at least two separate components, one representing shear and the other the tension/compression portion of the cycle. The material response to those parameters is captured by applying a modifying material parameter to the terms whose characteristics are captured using basic material characterization tests, such as the tension torsion test. See, for example, Jiang and Sehitoglu [120].

• Empirical models: Kapoor [142] developed a model of the time to crack initiation because of ratcheting only that is based on the equivalent ratcheting strain per cycle (which itself has normal and shear components) encountered per cycle and a material parameter that is evaluated through twin disc testing. The results are very sensitive to the value of the empirical parameter, which itself is difficult to evaluate. Furthermore, the model can only be applied in cases where failure is known to occur by ratcheting—for low cycle fatigue another model must be applied. Fatigue life is then governed by the model which predicts the shortest time to failure.

2.2 Crack Propagation
The direction and rate of crack propagation depends on the magnitude of the modes I, II, and III stresses, along with the material’s resistance to propagation.

A) Mode I opening    B) Mode II Shear    C) Mode III Tearing

Figure 29. Crack Propagation Is by One or a Combination of Mode I, Mode II, and Mode III Stresses

The difficulties of modeling contact fatigue crack propagation in rails and wheels are summarized in Reference 56 as:

• The rolling contact loading causes a multi-axial state of stress with out-of-phase stress components and rotating principal stress directions.
• As they grow longer, cracks subjected to a multi-axial loading normally deviate into a Mode I dominated growth (or follow a weak path in the structure). This is not the case in rolling contact fatigue, due to the large confining pressures under the contact which
normally suppresses any Mode I deformation of the crack in the absence of trapped fluids, etc. Instead cracks propagate mainly in a mixed Mode II–Mode III.

- In a predominantly compressive loading, the validity of traditional fatigue models may be questioned. As an illustration, Paris’ law predicts zero crack growth under compressive loading in its original form since it employs the range of the Mode I stress intensity factor of the above point.
- Due to the compressive loading, crack face friction will control the crack propagation. The magnitude of operational crack face friction is, however, hard to quantify.
- A similar effect of the compressive loading is that crack face deflection increases with crack length and may cause complete locking between the crack faces for part(s) of the crack.
- Occasional overloads may slightly accelerate crack growth, in contrast to the behavior in tensile loading. This may lead to non-conservative fatigue life predictions.

The use of fracture mechanics to model crack growth usually involves first determining the stress field around the crack tip (the so called stress-intensity K_{IC}) and then experimental data of growth rates is introduced. For rolling contacts, this process is complicated by the fact that the stress state is multiaxial, and the propagation mode is mixed.

In modeling crack propagation, one must choose between a low cycle or high cycle fatigue approach. The differences are summarized in Table 5.

Table 5. Comparing High Cycle Fatigue and Low Cycle Fatigue Approaches to Modeling of Crack Growth

<table>
<thead>
<tr>
<th></th>
<th>Low Cycle Fatigue</th>
<th>High Cycle Fatigue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cycles to failure</td>
<td>$10^4$</td>
<td>$10^6$</td>
</tr>
<tr>
<td>Role of plasticity</td>
<td>Governed by the amount of plastic flow</td>
<td>No (macroscopic) plastic deformations</td>
</tr>
<tr>
<td>Parameter modeled</td>
<td>Equivalent Stress</td>
<td>Equivalent Strain</td>
</tr>
<tr>
<td></td>
<td>● Sines criterion [237],</td>
<td>● Morrow criterion [188]</td>
</tr>
<tr>
<td></td>
<td>● Crossland [38]</td>
<td>● Kandil, Brown and Miller [137]</td>
</tr>
<tr>
<td></td>
<td>● Dang Van [40]</td>
<td>● Jiang-Sehitoglu [120]</td>
</tr>
<tr>
<td>Static shear stress</td>
<td>Affects LCF performance</td>
<td>Has no impact</td>
</tr>
</tbody>
</table>

The subsurface stress patterns because of the normal load of the wheel on the rail can be calculated in several ways with detailed analytical methods provided in References 93 and 273 as well as using numerical approaches such as boundary element [1, 149] and finite elements [6, 16, 242, 271]. Typically, linear elastic fracture mechanics are used [5, 144, 295, 302] to assess crack stress intensity factors that include mixed-mode loading under rolling contact, crack face friction, and the presence of entrapped fluids.

Rolling contact stresses diminish with depth into the component, such that in the dry rail/wheel contact, crack growth only proceeds to a couple of millimeters in depth. In the lubricated or wet case, that propagation can extend much further (see Section 2.2.4). However, beyond the influence of rolling contact stresses, where incidentally the protective compressive residual
stresses also diminish, crack propagation is driven by bending stresses as well as residual and thermal stresses in the rail.

Resistance to crack propagation is governed by a number of material parameters. Reference 207 cites hardness, metallurgical structure, and fracture toughness to be the most important when considering contact mechanics problems.

The effects of various loading and environmental conditions on crack propagation are discussed in the following subsections.

### 2.2.1 Shear Strength

For the application of Shakedown theorems to wheel-rail contact mechanic problems, it is important to have good quantitative understanding of the values of $k$ for different materials. There are equations relating $k$ to hardness, but these are approximate and take no account either of the large amount of work-hardening at the surfaces of wheels and rails or of strain-rate effects. So although the shear yield strength of as-manufactured rail is readily correlated with hardness in the laboratory, it is the properties developed in the work-hardened rail surface layers that are believed to be critical.

As noted in Reference 182, different researchers assume widely varying values of $k$ for similar steels. $k$ can be set to $S_y/2$ or $0.577S_y$ according to the Tresca and von Mises criteria, respectively. A third approach sets $k = \text{Vicker's hardness (Hv)}/6$ (see Table 6).
2.2.2 Strain Rate Dependence

It is common for researchers to use a quasi-static test to evaluate the shear strength of steels. The tendency is to approximate quasi-static conditions using strain rates on the order of 0.002 s⁻¹. Yet the strain cycle encountered by a wheel-rail rolling contact is far from quasi-static. Consider a 1.5-inch (12 mm) diameter (D) wheel-rail contact and a train traveling at 60 miles per hour (mph) (100 km/h) or 88 feet/s (28 m/s). For surface material, the load cycle takes place over approximately 0.4 ms, with the transition from no strain to peak strain being half of that (0.2 ms). For subsurface material, the load cycle is felt both ahead of and behind the contact patch by some amount, but certainly less than plus/minus one contact patch length D. So the cycle is taking place over 1.2 ms, and the strain peaking within 0.6 ms. The actual value of strain per cycle is not well known, but laboratory tests indicate that values might average around 6 × 10⁻⁴ as accumulated shear strain reaches values of 11 within about 17,500 cycles (at 1,500 MPa and 1-percent creepage) [286]. This gives average values of strain rate of approximately 1.0 s⁻¹.

In any case, the question is then whether the rate of strain application is of any relevance to wheel-rail contact. Reference 151 looks at rate effects as they pertain to fracture toughness and found that the toughness of railway wheel steel is significantly larger at higher impact rates. Yet the case is not so clear for shear strength.

Sawley considered the problem and proposed the following approach⁴:

a) assume that \( k \) is a function of the cyclic shear yield strength \( (k_{\text{cyc}}) \), a factor because of near-surface work-hardening \( (F_{\text{WH}}) \), and a factor because of the strain-rate in rolling contact \( (F_{\text{SR}}) \):  
\[
    k = k_{\text{cyc}} \times F_{\text{WH}} \times F_{\text{SR}}
\]

---

⁴ Personal communication with the author via email, April 2010.

---

### Table 6. Increasing Hardness Corresponds to Increasing Resistance to the Initiation of RCF; Standard North American Classifications; Typical Values of Shear Strength \((k)\) Calculated at One-Sixth of the Vicker’s Hardness [170]

<table>
<thead>
<tr>
<th>Steel</th>
<th>Hardness (Brinnell)</th>
<th>90% HB value</th>
<th>Shear Strength</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>(ksi)</td>
<td>(MPa)</td>
</tr>
<tr>
<td><strong>Rail</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>“Standard”</td>
<td>260-280</td>
<td>278</td>
<td>69.5</td>
<td>480</td>
</tr>
<tr>
<td>“Intermediate”</td>
<td>320-340</td>
<td>338</td>
<td>84.5</td>
<td>583</td>
</tr>
<tr>
<td>“Premium”</td>
<td>340-380</td>
<td>376</td>
<td>94.0</td>
<td>649</td>
</tr>
<tr>
<td>“HE Premium”</td>
<td>380-400</td>
<td>398</td>
<td>99.5</td>
<td>687</td>
</tr>
<tr>
<td><strong>Wheel</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class L</td>
<td>197-277</td>
<td>269</td>
<td>67.2</td>
<td>463</td>
</tr>
<tr>
<td>Class A</td>
<td>255-321</td>
<td>314</td>
<td>78.7</td>
<td>542</td>
</tr>
<tr>
<td>Class B</td>
<td>277-341</td>
<td>335</td>
<td>83.8</td>
<td>577</td>
</tr>
<tr>
<td>Class C</td>
<td>321-363</td>
<td>359</td>
<td>89.9</td>
<td>620</td>
</tr>
</tbody>
</table>
b) The cyclic shear yield strength is equal to the cyclic yield strength divided by \( \sqrt{3} \) (from Von Mises criterion). The cyclic yield strength can be found from cyclic stress-strain fatigue data.

c) \( F_{WH} \) is equal to the surface hardness divided by the bulk hardness.

d) \( F_{SR} \) is estimated from knowledge of the strain rates in rolling contact. Some estimates placed the value as high as 2.5, although a value nearer to 1.3 gave agreement with an “empirical” curve for 220 grade rail.

e) With \( k \) known, the Shakedown diagram can then be plotted as \( P_O \) versus \( T/N \).

f) This approach has been applied to the 220 Grade of rail steel used in the United Kingdom. The \( P_O-T/N \) curve found from this method agreed closely with the empirical curve found from contact dynamic calculations undertaken for track sections with and without RCF cracks. Both these curves were significantly above the curve produced by using a traditional method of estimating \( k \).

### 2.2.3 Contact Stresses

Crack propagation models invariably predict growth rates that depend on the shear stress at the tip of the crack. For the rolling contact, these shear stresses are dependent on the specifics of the wheel-rail contact, including the contact geometry, applied load, creepage, interface friction, elastic modulus, and Poisson’s ratio. These have been covered in Section 2.1.3, where it was noted that a number of analytical, numerical, FRA, and boundary element approaches are available for calculating subsurface stresses.

The practical effect of contact stress on crack propagation is illustrated in Figure 30 where the fatigue life of tension and torsion specimens is shown to be inversely proportional to the applied stress. In rolling contact, there is a superimposed compressive stress, and the fatigue life becomes effectively infinite below a threshold stress (see Figure 31). This is consistent with the Shakedown limit discussed in Section 2.1.1.1.
Figure 30. Effect of Stress on Material Performance [18]

Figure 31. Effect of Contact Pressure on Nine Different Rail Steels [33] for a Creep Ratio of 10 Percent
However, rather than the normal stresses at the wheel-rail interface, a much more causative mechanism appears to be the shear tractions. One reason may simply be that tractions vary from a value of zero to, say, 1 percent, with 0.1 percent being typical at the top of rail and higher values appearing under curving and driving and braking. The direct role of shear tractions is recognized in the $T_y$ damage mechanism and the fatigue index of Ekberg described in Section 2.1.8. The physical tests of Clayton [see Reference 33 and Figure 40] are conceptually consistent with the wear/fatigue transition include in the $T_y$ damage relationships.

![Graphs showing the effect of slip ratio on number of cycles to failure based on laboratory testing.](image)

Figure 32. Effect of Slip Ratio on Number of Cycles to Failure Based on Laboratory Testing [33]; the Longest Life Is Achieved with Zero Slip with another Maximum Seen at Roughly 1-Percent Creep
2.2.4 Moisture and Grease

In a dry environment, surface breaking cracks will remain relatively dormant once they have grown beyond the reach of the heavily concentrated near surface stresses. Yet in moist environments or in alternating wet and dry environments, water is drawn into existing surface cracks and then squeezed toward the crack tip by the passing load (Figure 33).

![Figure 33. Lubrication of the Crack Faces Promotes Mode II Propagation, whereas Water in Properly Oriented Cracks Can Pressurize the Crack Tip and Cause Cracks to Propagate Rapidly in Mode I](image)

Numerical modeling of the fluid pressurization mechanism in Reference 8 found that:

- The Mode I stress intensity increases roughly five-fold when water is present compared with the dry case, while the Mode II value increases by 60–85 percent
- The Mode I stress intensity is higher for shallower angles (e.g., 15° versus 25°)
- The Mode II stress intensity is higher for steeper crack angles
- The addition of a braking force on the wheel in the presence of a liquid reduces the Mode I stress intensity and increases the Mode II values by modest amounts (in the order of 25 percent).

The effect of water and grease is different with respect to shelling. Water, having high surface tension and low viscosity, is readily drawn into surface breaking cracks and is available to propagate the crack tip in Mode I. This Mode I propagation has been a topic of great academic and practical interest [e.g., 5, 12, 72]. In Reference 6 it is shown that the shear Mode II stress intensity factor rises by approximately 70 percent while for the pull-apart Mode I, the factor increases nearly 10 times, such that Mode I crack propagation dominates. The authors of Reference 66 caution, though, that predicted crack growth rates under water pressurization are about 35,000 times those for the shear mode growth and are not consistent with measured rates in laboratory testing. This calls into question the assumptions used in the fluid entrapment models.

Grease has a much higher viscosity than water but can be worked into the crack over time by the passing contact load. Once the crack faces are lubricated, grease promotes crack propagation in the shear Modes II and III [138]. An example is shown in Figure 34 in which cracks grow longer for lower values of crack-face friction, and that at higher friction values (0.6 in the figure), crack growth is halted. Since grease is not drawn in volume into the crack as water is, it is unlikely that it plays as large a role as water in Mode I crack propagation.
2.2.5 Fracture Toughness

Rail steels with good fracture toughness can inhibit or prevent the occurrence of shelling and internal defects in track. Steels with good fracture toughness resist propagation of cracks from both RCF and other types of fatigue [197]. The ability of a material to resist crack growth depends on the propagation Mode (type I, II, or III) [5, 10, 97, 119].

With respect to internal defects, good fracture toughness allows for a reasonable defect detection interval, thereby contributing to the safety of operations. Controlling the morphology of inclusions [157], ensuring favorable residual stresses [197], laboratory and field testing at high [9] and low frequencies and loading rates [25] of developmental rail steels have been utilized to improve the fracture toughness of rail steel.

Microalloying elements have been shown to affect the fracture toughness [239, 240]. Molybdenum increases fracture toughness, whereas increasing phosphorus has a significant negative impact [197]. Manganese and niobium additions do not appear to influence fracture toughness [202]. Increasing the nitrogen content in steel to 0.0015 percent enhances the metal’s impact strength at subzero temperatures [191].

Fracture toughness, also called $K_{IC}$, is a commonly quoted material property (e.g., Table 7). Reference 207 notes that it is important to understand that $K_{IC}$ applies to relatively long (several millimeters) cracks, and its application to modeling of short cracks that are in the threshold regime of the $\Delta K$-da/dn function is not yet understood. Nonetheless, the Charpy impact tests remain of interest and are applied to qualify the performance of rail steels. An example is shown in Table 7.
Table 7. Impact Resistance of Rail Steels [Reference 108, Table 3-3]

<table>
<thead>
<tr>
<th>Rail Material Type</th>
<th>Impact Energy (Joules)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>V notched at 200°C</td>
</tr>
<tr>
<td>Standard Carbon</td>
<td>29</td>
</tr>
<tr>
<td>Medium Carbon Alloy</td>
<td>35–53</td>
</tr>
<tr>
<td>Heat Treated</td>
<td>56</td>
</tr>
</tbody>
</table>

In a different set of measurements using compact tension specimens (according to American Society for Testing and Materials standard 399) the average $K_{IC}$ for a test bainitic steel was 52 MPa $\sqrt{m}$, whereas for the premium pearlitic steel, the value was 41 MPa $\sqrt{m}$ [288].

2.3 Rail (and Wheel) Fatigue Failure

A rail or wheel can be defined as having failed when the accumulated damage exceeds an acceptable threshold such that it is no longer fit for service.

**Rail:** Although it is possible to grind, plane, or mill out surface defects, there is a limit whereby the cost of the treatment exceeds the expected life extension (which is a function of the remaining wear life). In addition, defects such as crushed heads, squats, and shells must be removed because of their potential for initiating a broken rail. The broken rail, while obviously a derailment risk, is also an expensive occurrence because of the unplanned maintenance and traffic disruption. North American practice is to remove a length of rail once a rate of 1–2 defects/km/year is reached [109]. All railways have instructions for remedial actions in the case of rail defects based usually on the size of the defect. The example shown in Table 8 was extracted from the rail defect standards (TES 02) of the Australian Rail Track Corporation Ltd.

**Wheel:** Failure of the wheel from RCF, such as a broken wheel or split rim, is fortunately rare. Yet, shelling of the surface leads to high impact forces that damage rails, bearings, and car components. The North American AAR Rule 41 allows for a wheel to be removed if a 1-inch diameter circle can be completely contained within the shell. The installation of wheel impact load detector and establishment of accepted limits for their removal has provided an additional approach for managing the effects of shelling, though not for its treatment. In AAR Rule 41, a wheel may be removed if it has a measured impact force of 90,000 lb (400 kN).

The march to failure of rails and wheels is complicated by the simultaneous action of wear; a process that itself limits the life of rails and wheels, but removing damaged surface material can significantly extend the life also. This competitive action is discussed in Reference 50 and is shown in the $T_{γ}$ relationship of Section 5.1 where at low $T_{γ}$ (traction force $\times$ creepage = wear energy) values the damage tends to fatigue but at higher values wear dominates. This has been described as the Magic Wear rate [131] when applied to surface fatigue of rails and wheels, with the concept being particularly potent when applied to rail grinding. The Magic Wear Rate is discussed further in Section 5.6.
## Table 8. Example of a Defect Classification and Required Action

### Classification of Rail Defects

<table>
<thead>
<tr>
<th>Type of Defect</th>
<th>Code</th>
<th>Size Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transverse defects in rail head</td>
<td>SC</td>
<td>0 to 5% (less than 10 mm) treat as TD small</td>
</tr>
<tr>
<td>(transverse fissures, transverse shelling, shatter cracks)</td>
<td>TD</td>
<td>5 to 10% (10 to 20 mm)</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>11 to 30% (21 to 30 mm)</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>over 30% (over 30 mm)</td>
</tr>
</tbody>
</table>

### Remedial Action to Protect or Correct Rail Defects

<table>
<thead>
<tr>
<th>Defect Type</th>
<th>Size</th>
<th>Immediate Speed Restriction Km/hr until defect removed</th>
<th>Plate Within</th>
<th>Remove Within</th>
<th>Other Action (Refer also to Section 4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Transverse Defect (TD)</td>
<td>S</td>
<td>7 days</td>
<td>5 days</td>
<td>5 months</td>
<td>If TD is reported in shatter cracked rail, the full length must be replaced weld to weld</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>24 hours</td>
<td>5 weeks</td>
<td>48 hours</td>
<td></td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>2 hours</td>
<td>2 hours</td>
<td>48 hours</td>
<td></td>
</tr>
<tr>
<td>TD in Shatter Cracked Rail (SC)</td>
<td>Same as TD.</td>
<td>If multiple – same as TDX</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| 2 Transverse Defect at Engine Burn (TD/EBF) | S    | 7 days                                                 | 5 months     | Where several engine burns exist in the same rail length, consideration should be given to replace full rail length, weld to weld. |
|                                           | M    | 24 hours                                               | 5 weeks      |              |                                        |
|                                           | L    | 2 hours                                                | 48 hours     |              |                                        |

| 3 Multiple Transverse Defects (TDX) | S    | 40 km/hr                                               | 7 days       | Replace full rail length – weld to weld |
|                                     | M    | 20 km/hr                                               | 24 hours     |                                        |
|                                     | L    | 10 km/hr                                               | Monitor defect and stop trains if necessary until defect removed |
3. Detection and Monitoring of RCF

Nondestructive testing of rails and wheels is of vital importance to managing risk and minimizing catastrophic failures. Rails especially must be systematically evaluated for defects that might appear benign to the naked eye but can be measured to be progressing internally to a dangerous state.

The surface condition of rail is currently assessed through a visual rail inspection by trained railway personnel, often from a hi-rail vehicle. This information, along with (internal) rail defect measurements by the ultrasonic cars dictates several of the rail maintenance processes, including some aspects of the rail relay program, defect removal through rail plugging, and rail grinding.

On some railroads, the grinding priority is driven by the severity of RCF cracks on the rail surface and the capability of ultrasonic detection cars to make successful tests on the rail. By the time they are identified in this way, the RCF cracks are usually very deep. Management of RCF requires a technology for identifying cracks in their earlier stages of development, although they are still relatively short and can be easily removed through rail grinding.

An excellent review of the state of art of nondestructive techniques for inspecting rail can be found in Reference 200. This review covers ultrasonic, magnetic induction, eddy current, radiography, longitudinal guided waves, electromagnetic acoustic transmission, acoustic emission, and others. Reference 107 provides a similarly comprehensive overview but also includes discussion about devices commercially available in Europe.

Although high-speed laser and video systems exist for scanning and recording various track features, none can yet provide a quantitative assessment of anything but the simplest of surface defects (e.g., ballast spalls). RCF cracks are an order-of-magnitude more difficult to distinguish than spalls.

Other approaches include devices for specifically measuring the depth of cracks. At present no commercially available RCF crack detection systems are in service in North America. There are low-speed hand-operated RCF detection systems available in Europe, but these can only test at speeds of 1 mph and are not feasible for the lengths of North American track. However, development is continuing on higher speed, machine mounted systems, and it seems that it is only a matter of time before they become commercially available. Then, the railroads and the rail grinding companies will have a very powerful tool for assessing and managing surface fatigue. According to one company, that tool is already available (see Section 3.1.1.2).

One more approach is to use high-frequency accelerometers or force-based measurements to detect surface damage. This approach is discussed in Section 3.1.4.

3.1 Existing Systems

A summary of the various technologies applied to rail inspection (both surface and subsurface) as found in Reference 200 is reproduced below as Table 9.
### Table 9. Technologies Applied to Rail Testing [Table 1 of Reference 200]

<table>
<thead>
<tr>
<th>NDT Technique</th>
<th>Systems Available</th>
<th>Defects Detected</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ultrasonecs</td>
<td>Manual and high-speed systems (up to 70km/h)</td>
<td>Surface defects, rail head internal defects, rail web and foot defects</td>
<td>Reliable manual inspection but can miss rail foot defects. At high speed can miss surface defects smaller than &lt;4 mm as well as internal defects particularly at the rail foot</td>
</tr>
<tr>
<td>Magnetic flux leakage</td>
<td>High-speed systems (up to 35km/h)</td>
<td>Surface defects and near surface internal rail head defects</td>
<td>Reliable in detecting surface defects and shallow internal rail head defects although cannot detect cracks smaller than &lt;4 mm. MFL performance deteriorates at higher speeds</td>
</tr>
<tr>
<td>PEC (including FGIT)</td>
<td>Manual and high-speed systems (up to 70km/h)</td>
<td>Surface and near-surface internal defects</td>
<td>Adversely affected by grinding marks and lift-off variations.</td>
</tr>
<tr>
<td>Automated visual inspection</td>
<td>Manual and high-speed systems (up to 320 km/h)</td>
<td>Surface breaking defects, rail head profile, corrosion, missing parts, defective ballast</td>
<td>Reliable in detecting corrosion, rail head profile missing parts and defective ballast at high speeds. Cannot reliably detect surface breaking defects at speeds &gt;4 km/h. Cannot assess the rail for internal defects.</td>
</tr>
<tr>
<td>Radiography</td>
<td>Manual systems for static tests</td>
<td>Welds and known defects</td>
<td>Reliable in detecting internal defects in welds difficult to inspect by other means. Can miss certain transverse defects</td>
</tr>
<tr>
<td>EMAT</td>
<td>Low speed hi-rail vehicle (&lt;10 km/h)</td>
<td>Surface defects, rail head, web and foot internal defects</td>
<td>Reliable for surface and internal defects. Can miss rail foot defects. Adversely affected by lift-off variations.</td>
</tr>
<tr>
<td>Long range ultrasonics</td>
<td>Manual systems and low-speed hi-rail vehicle systems (&lt;30 km/h)</td>
<td>Surface defects, rail head internal defects, rail web and foot defects</td>
<td>Reliable in detecting large transverse defects (&gt;5 per cent of the overall cross-section)</td>
</tr>
<tr>
<td>Laser ultrasonics</td>
<td>Manual and low-speed hi-rail vehicle systems (&lt;15 km/h)</td>
<td>Rail head, web and foot defects</td>
<td>Reliable in detecting internal defects. Can be affected by lift-off variations of the sensors, difficult to deploy at high speeds</td>
</tr>
<tr>
<td>ACFM</td>
<td>Manual systems (hi-speed system under development)</td>
<td>Surface breaking defects</td>
<td>Reliable in detecting and quantifying surface breaking defects. Cannot detect sub-surface defects. Very good tolerance to lift-off variations</td>
</tr>
<tr>
<td>AE</td>
<td>Experimental manual and high-speed systems</td>
<td>Rail breaks, wheel burns, squats, wet spots, worn rail profiles</td>
<td>Limited experiments. Cannot detect any internal defects</td>
</tr>
<tr>
<td>AEP</td>
<td>Experimental static tests</td>
<td>Surface defects, rail head internal defects, rail web and foot defects</td>
<td>Limited experiments. Can only be applied at predefined areas. Can miss non-transverse defects or small transverse defects</td>
</tr>
<tr>
<td>MAD</td>
<td>Experimental static tests</td>
<td>Broken rails, rail gaps</td>
<td>Limited experiments. Possibly capable of detecting large internal or surface-breaking defects (i.e., &gt;50 per cent of the cross sectional area)</td>
</tr>
</tbody>
</table>
Table 10. Comparison of the Three Common Detection Technologies Applied to Rail Inspection
[extension of Reference 200, Table 1]

<table>
<thead>
<tr>
<th>Eddy Current</th>
<th>Ultrasonic</th>
<th>Vision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Good at detecting surface defects</td>
<td>Poor at detecting surface defects</td>
<td>Can detect only surface defects</td>
</tr>
<tr>
<td>Near subsurface defects reasonable to detect</td>
<td>Near subsurface defects difficult to detect</td>
<td>Near subsurface defects cannot be detected</td>
</tr>
<tr>
<td>Deep subsurface defect detection is impossible</td>
<td>Good subsurface defect detection</td>
<td>Deep subsurface defect detection is impossible</td>
</tr>
<tr>
<td>Can detect through surface layers</td>
<td>Grease and films are problematic</td>
<td>Surface layers hide most flaws</td>
</tr>
<tr>
<td>Probes are less sensitive to flaw orientation</td>
<td>Signal is strongly influenced by flaw orientation</td>
<td>Detection sensitive to lighting sources</td>
</tr>
<tr>
<td>No couplant needed</td>
<td>Couplant is typically required</td>
<td>No couplant needed</td>
</tr>
<tr>
<td>Probe can be made wide and profiled to cover wear face</td>
<td>Defect must be on probe centre line</td>
<td>Wide field of view allows full coverage of rail wear face</td>
</tr>
<tr>
<td>Faster inspection speeds</td>
<td>Slow inspection speeds</td>
<td>Fast to very fast inspection speeds possible</td>
</tr>
</tbody>
</table>

3.1.1 Eddy Current

Eddy current inspection is successfully applied in a number of industries (e.g., oil pipeline, aviation, and rail), because it is robust and relatively straightforward to use.

As described in Reference 200, the eddy current approach works as follows:

Typical eddy current sensors comprise one exciting and one sensing coil. An alternating current (AC) is fed to the exciting coil in order to generate a magnetic field near the surface of the rail head. Changes in the magnetic field cause eddy currents to be induced just below the surface of the rail head. Changes in the secondary magnetic field generated by the eddy currents are detected by the search coil in the form of an induced voltage. If the inspected area is free of defects then the impedance of the eddy current sensor remains constant. However, when a near-surface or surface defect is present in the rail head, the eddy currents are disturbed causing fluctuations in the secondary magnetic field giving rise to changes in the impedance. Thus, during manual inspection for near-surface or surface damage of the rail head with eddy current systems the operator looks for any changes in the impedance signal recorded in order to detect the presence of defects.

Because the eddy currents penetrate only into the skin of the target sample, only surface and near-surface defects are examined. From a practical consideration, the eddy current system is very sensitive to changes in the distance between the coils and the target, so a consistent standoff
is required. Typically, that distance is 1–2 mm, which is difficult to achieve at high speed on rough rail and with continually varying rail head shapes. Cracks associated with tongue lipping because of plastic flow at the gauge face of the rail can be reliably detected with eddy currents (Figure 35).

![Figure 35. Eddy Current Systems Can Detect Lipping at the Gauge Corner of Rails [200]](image)

### 3.1.1.1 Walking Stick

Three examples of commercially available eddy current-based walking stick technologies are shown in Figure 36.

#### 3.1.1.1.1 TSC ACFM Walking Stick System

ACFM ... is an electromagnetic technique based around eddy current induction, but with some significant differences from conventional eddy current techniques. The probe contains an integral field generator that induces a minute current to flow in the surface of the rail head. Through use of a constant current generator in the instrument and special design of an inducer in the probe, the currents under the probe are uniform in direction and strength. They do not penetrate deep into the rail but instead run in a very thin skin (<0.1 mm deep) in the surface of the rail. Defects in the surface of the rail distort these currents and this is reflected in changes to the magnetic field just above the surface. An array of sensors is located across the rail head and monitors any disturbances in the magnetic field.

---

5 See www.tscinspectionsystems.co.uk.
The ACFM Walking Stick is an integrated trolley, containing an alternating current field measurement (ACFM) array probe, instrument and laptop PC, which is pushed along the rail by an inspector. Data is displayed in real time on the PC screen and stored as a permanent record of the inspection. State of the art software uses easy to read displays and powerful processing algorithms to automatically interpret the data. The ACFM Walking Stick does not suffer from some of the problems of conventional Ultrasonic inspection techniques where large numbers of defects can render a rail un-testable. With ACFM, the deeper the crack, the larger the response. An updated technical document [160, 281] suggests that the development of empirical correlation factors specifically for rail types of cracks have significantly improved the model.
reliability but admits that “this is a quantitative inspection method and that sufficient data must be collected to not only detect, but also determine the severity of the cracking”. Recent experience with the ACFM technique is given in Reference 84.

3.1.1.1.2 NEWT International Lizard

Lizard® system is based upon a revolutionary field gradient imaging technology (FGI) arranged in the form of an electromagnetic scanning head enabling the visualization of surface breaking defects. Its application towards detection, sizing and management of high density Gauge Corner Cracking for the rail industry where ultrasonic inspection was unable to perform, was the result of a government and industry funded project.

The system has full rail approvals for inspection of critical track components with many thousands of successful inspections performed to date. An additional FGI sensor array is available for the inspection of switches and crossings, for complete railway rail inspection. Arrays are available in both manual and electronically scanned versions and are able to "see" through coatings in excess of 8mm.

NEST Software: FGI track inspection systems can rapidly acquire data from kilometers of track with typically up to fifty thousand defects per kilometer. NEST provides an essential set of Windows™ based software to perform analytic tasks automatically. This facility for automated data processing brings other benefits, enabling better fidelity of measurement of defect clusters, and the ability to report in a format directly of use to rail engineers. NEST scans the compensated data analyzing the location and magnitude of defects, producing a defect database, from which various graphical or tabular report forms can be selected. Among other facilities, these range from a visualization of an individual targeted defect to a summary of maximum defect depth per meter of track. Reports from NEST give a basis for the planning of remedial work and a tool for assessing either the effectiveness of practices such as lubrication and remedial grinding, change of wheel/rail profile or the impact of traffic.

3.1.1.1.3 Rohmann Drasine Trolley

At the time of its testing, the Rohmann system was not specifically designed to cover the entire width of the head. Its carriage is made to hold two single transducers, and for this application, one high-depth penetrating probe and one counter probe were used. At the time of this reporting (well after the testing at CSTT) modifications to the system were being undertaken to allow the same carrier to hold a larger number of probes (probably eight in total), distributed across the head of the rail.

The test system is typically designed for the detection and evaluation of defects on or just below the surface of rails. In addition to defects along the guiding surface of rails (head checks) other inhomogenities such as corrugation, chatter marks, skid marks and weld seams can be found. Thus it is possible to track the development of damages along the surface, optimally plan the repair and monitor its success. The measuring system consists of a multi-channel eddy current test instrument and a computer evaluation and analysis network for the inspection of defects on or just below the surface of rails. It is to be integrated in a rail-inspection train. Four test channels are available per rail i.e. there is a total of eight channels for both rails.

---

6 See www.lizard.co.uk.
7 Extracted from a Rohmann Drasine technical bulletin provided by Rohmann GmbH, 10/04.
The measuring data of all channels are guided through an A/D converter and are forwarded to a measuring computer for each rail. A LAN in the inspection train forwards the measuring data to a control computer where they are merged with GPS and UT-data. Here special evaluation software provides a preliminary selection of incidents where the rail surface is damaged. All measuring data are subject to an immediate evaluation. The evaluation software can be modified to meet the requirements of the end user.

3.1.1.2 “High-Speed” Eddy Current Crack Measurements

A technology emerging from in Magdeburg, Germany, is high-speed measurements of rail surface cracks using eddy current in combination with ultrasonics [96]. The resulting test unit provides the ability to scan for both surface and subsurface defects at speeds up to 60 km/h (35 mph). The combined system, utilizing 10 ultrasonic probes at angles of +70 to -70 and four eddy current sensors focused on the gauge side 25mm of the rail head, is shown in Figure 37. Reference 96, while frequently mentioning rail surface quality, does not unfortunately look at the issue in that paper. Marketing literature for the product shows that it can be specifically applied to rail grinding and is able to measure surface fatigue at speeds up to 30 kph (18 mph).

Currently, onboard a Speno rail grinder is a system called HC Grinding Scanner that provides information on the position and “depth” of head-check type cracks [206]. The practical goals of the system are to optimize the grinding process by real-time monitoring of the depth of damage and to document rail conditions before and after grinding. Four eddy current probes per rail
provide a total 25-mm measurement width per rail. The sensors are mounted on a trolley such that the standoff distance is 1 mm, based on a UIC60 rail head shape. The standoff will vary if the rail head is worn, with the capability of the sensors limited to -1 to +2 mm. The measurement focus is typically on the gauge corner where crack development is the primary concern in Europe. Adaptation to the North American condition, with top of rail and field side cracking also being very important, should not be an overly difficult challenge. Technical information on the system is currently sparse but the practical understanding of its use is that since the signal return varies with the length of surface cracks that by shortening cracks through rail grinding it can be used to notify whether the grinding process is progressing towards the elimination of surface fatigue. Although it is difficult with eddy current to know the actual depth of damage, it is easy to visualize how this information, especially with repeated runs, could prove extremely valuable in understanding rates and cause of crack growth and the effectiveness of rail grinding practices.

![Figure 38. Example Plot of Processed Data from the Speno Rail HC Grinding Scanner System](image)

As of the end of July 2010, there is an initiative being pursued in North America to undertake a trial on several U.S. railroads to validate the technology.

### 3.1.2 Ultrasonic Measurements

It is generally supposed that ultrasonic techniques cannot see surface and near-surface defects. This is apparently not because the potential does not exist but rather that the amount of noise and scatter from the near surface signal is so large and complex that the analysis systems generally ignore the signal associated with the first few millimeters of the rail surface to focus on internal inspection. Reference 200 notes that the conventional ultrasonic probes could be used to detect
larger (>4 mm) surface-originated defects such as deep head checks and gauge corner cracking, although there are reports of ultrasonics being used to monitor surface crack growth [e.g., 102]. There is no doubt that information exists in the ultrasonic signal related to the condition of the rail surface, but to date, it does not appear to be a subject of active research.

Problems with ultrasonic approaches include that martensitic layers at the surface of a rail (e.g., because of a wheel burn) prevent effective ultrasonic testing of rail and that the presence of larger and more critical internal defects can be shadowed by smaller surface cracks during inspection. For these reasons, current best practice is to combine nondestructive evaluation with preventative rail grinding to optimize the tradeoff between maintenance cost and structural reliability [200].

### 3.1.3 Vision Technologies

A 2004 study by the National Research Council of Canada (NRCC) concluded that “although high-speed laser and video systems exist for scanning and recording various track features, none appear to provide a quantitative assessment of anything but the simplest of surface defects (e.g., ballast spalls). RCF cracks are an order-of-magnitude more difficult to distinguish than spalls” [165].

Now, MerMec (Italy) is advertising a comprehensive suite of rail inspection equipment, including the “track head-check inspection system (THIS).” The brochure supplied by the company does not say much about the system but it does include an example of the results (see Figure 39). The red highlights in the lower photo identify the cracks that were detected by the postprocessing algorithms from this image. The MerMec brochure says that:

> THIS, through the use of a no-contact optical scanning system, is capable of finding gage corner cracking in its earliest stage, when rail grinding is still a feasible option, before it leads to bigger problems.

Although there is little technical description, they do say that:

> A unique laser illumination system, developed by MERMEC Group, is used to illuminate the rail enabling THIS to determine crack dimensions from 0.15mm to 0.50mm, their angle, length and frequency.

The example in Figure 39 could be a best case scenario on rail that has apparently a flat head and well defined cracks. This would probably be sufficient for monitoring rail for safety, but it can be expected that detection of smaller cracks as required for a preventive maintenance program in North America would be a challenge for THIS.
Deutsch et al. [46] suggest in a 2004 publication that visual checks combined with eddy current are an effective approach for inspection of rail. The system they describe uses line-scan cameras to assemble a continuous image of the passing rail, which is then analyzed with relatively simple processing algorithms. It appears to be effective when used in a rail mill, but the authors suggest that it “could easily be embedded into a testing vehicle such as an intelligent train. For the presented system, it simply doesn’t matter whether the rail passes through the device or the device moves over the rail” [46]. This statement ignores the many practical implications involved in developing a field system capable of dealing with adverse environmental and vibration conditions, operating around obstructions such as fasteners, joint bars and lubricators, and distinguishing defects through visual contaminants such as grass, spiders, grease, and water.

Although machine vision systems are being applied in railway applications to look at components like joint bars, rail ties condition, and missing fasteners, its practical application to characterizing surface condition currently appears unclear.

### 3.1.4 Acceleration and Force-Based Measuring Systems

Measurement of vertical track perturbations such as squats and dipped welds have been the subject of recent study in Australia [269] and the Netherlands [49, 187]. Short length vertical geometry defects on the order of 0.1–1 mm can be detected with an “AK Car” that has axle box accelerometers sampled at 3,500 Hz. The location and severity of the defects compared favorably with the manual measurements taken with survey levels, dip gauges, and the Corrugation Analysis Trolley (CAT)°. In North America, the VTI system [256], based on the same principles, is being used by many railways to identify poor track locations. Although they

---

can measure high impact forces, their focus is generally on bridge approaches, pumping ballast, and short wavelength track irregularities.

3.2 Current Gaps

3.2.1 Surface Defects
The potential of the ultrasonic signal to provide information on the rail surface condition appears to not have yet been explored.

Eddy current systems appear to offer tremendous potential as the surface crack measurement tool of the future. Hand-held, walking stick, and train-mounted systems exist but their practical application to North American rail maintenance practices is far from assured. Demonstrations and practical experience with these are required.

Vision systems, such as MerMec’s THIS, supposedly can detect surface fatigue, but it is not at all certain whether they can do so under the wide range of environmental, rail profile, and surface conditions more typical of North America (than say Europe) and whether they can detect sufficiently small cracks at all locations across the rail head for use in a practical preventive maintenance program. Demonstrations and practical experience with these are required.

3.3 Internal Defects
A review by Clark and Singh [30] highlights some of the challenges of rail defect inspection and suggests that some emerging technologies will eventually help to overcome current limitations. They mention in particular:

- Improved ultrasonic systems with novel sensor placements and phased array applications. The latter allows an image of the defect to be created to improve in identification and sizing.
- Laser ultrasound – using laser pulses to impart ultrasonic energy into the rail, this approach allows for noncontact excitation and remote sensing, which it is anticipated will allow for progressively faster rail inspection systems. Reference 145 details an application of the same technology for detecting surface breaking cracks in railway wheels.
- Guided ultrasonics – whereby the whole cross section of the rail is flooded with ultrasonic energy, and send-and-receive sensors allow transverse defects to be readily identified.
- Magnetic sensors (induction) – for the rail web and foot show a strong ability to detect flaws, but practical application is hampered by the mechanical obstructions (joint bars, lubricators, etc.) present at these locations.
- Neural network processing of rail detection data – continues to suffer from unacceptably high rates of false positives.

From a hardware perspective, a successful RCF management system is likely to include a range of detection technologies. Eddy current can be used to detect small surface breaking cracks, whereas vision systems may find a role in detecting, measuring, and categorizing rail condition for the purposes of scheduling maintenance on well understood track. Eddy current crack measurements are likely to be applied in the calibration of vision and other systems. Ultrasonic
systems will continue to be applied to detect internal defects, ideally at a small size before they pose an unacceptable safety hazard.
4. The Influence of Operating Parameters

The key factors controlling the development of RCF, as discussed throughout Section 2, are illustrated in Figure 40. The effect of each will be discussed and approaches for mitigating their impact, if appropriate, will be given.

4.1 Metallurgy

There is no doubt that improved steel making practices have over the past three decades had a tremendous impact on the integrity of railway components. Yet, although Reference 195 optimistically suggests that a new rail type might solve RCF problems for its case study railway, for most railways, improved steel is but one (albeit sometimes powerful) contributor to minimizing RCF.

Section 2.1.5 describes the impact of material strength on RCF, with hardness playing a key role. The benefit of improved steels has not generally been associated with a reduction in required maintenance. Reference 291 notes that, although local factors and economic consideration affect the selection of rail type for an installation, the “timely execution of maintenance measures (grinding and above all lubrication of the rails) has a decisive impact on the service life of rails in curves.” Although the rates of plastic flow, wear and crack propagation are significantly lower
in softer steels, the steady increase in axle loads has not generally allowed those benefits to translate into significant reductions in maintenance.

Although one might expect the reduced amount of rail grinding and regauging to translate into economic return, Reference 30 suggests that the real economic benefit arises because of the extended life and reduced frequency of rail replacement [31]. For this reason, there is still a reluctance on many railroads to pay a modest premium for clearly superior performing rail, because the net savings may not be seen until many years (or possibly decades, depending on traffic volumes) after that initial investment is made.

Reference 291 notes that although local factors and economic consideration affect the selection of rail type for an installation, the “timely execution of maintenance measures (grinding and above all lubrication of the rails) has a decisive impact on the service life of rails in curves.” Excellent reviews of the evolution of rail and wheel steels can be found in references [180, 264]. Bainitic rail steels specifically are discussed in References 26, 33, 226, 233, 238, and 288. The properties of modern North American and European rail steels are summarized in Table 11 and Table 12, respectively.

### Table 11. General Properties of Wheel and Rail Steels

<table>
<thead>
<tr>
<th>Production Process</th>
<th>Wheel</th>
<th>Rail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cast</td>
<td></td>
<td>Rolled from ingot or continuous cast billet</td>
</tr>
<tr>
<td>Wrought</td>
<td></td>
<td>Roller straightened</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Microstructure</th>
<th>Wheel</th>
<th>Rail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eutectoid – pearlitic</td>
<td></td>
<td>Eutectoid – pearlitic</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hardness</th>
<th>Wheel</th>
<th>Rail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class B</td>
<td>277-341</td>
<td>Standard 260-280</td>
</tr>
<tr>
<td>Class C</td>
<td>321-363</td>
<td>Intermediate 320-340</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Composition</th>
<th>Wheel</th>
<th>Rail</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>0.57-0.67</td>
<td>0.72/0.82</td>
</tr>
<tr>
<td>Mn</td>
<td>0.60-0.85</td>
<td>0.80/1.10</td>
</tr>
<tr>
<td>P</td>
<td>0.05</td>
<td>0.10/0.60</td>
</tr>
<tr>
<td>Si</td>
<td>0.05 max</td>
<td>0.035 max</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Heat Treatment</th>
<th>Wheel</th>
<th>Rail</th>
</tr>
</thead>
<tbody>
<tr>
<td>controlled cooled (cc)</td>
<td></td>
<td>Standard: Plain carbon – hot bed cooled</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Intermediate: Alloxed – hot bed cooled</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Premium: Plain carbon – fully HT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Super: Head hardened &amp; micro-alloyed</td>
</tr>
<tr>
<td></td>
<td></td>
<td>On line - no reheating of head</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Off-line – head heated by flame or induction</td>
</tr>
</tbody>
</table>

62
Table 12. Prescribed Properties of Rail Steels According to UIC 860V 8 [291]

<table>
<thead>
<tr>
<th>Grade of steel</th>
<th>Chemical composition, elements in % of mass</th>
<th>Tensile strength, $R_m$ N/mm$^2$</th>
<th>Elongation $A_r$, min, %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C</td>
<td>Mn</td>
<td>Si</td>
</tr>
<tr>
<td>R0700</td>
<td>0.40-0.60</td>
<td>0.80-1.25</td>
<td>0.05-0.35</td>
</tr>
<tr>
<td>R0900A</td>
<td>0.40-0.80</td>
<td>0.80-1.30</td>
<td>0.10-0.50</td>
</tr>
<tr>
<td>R0900B</td>
<td>0.55-0.75</td>
<td>1.30-1.70</td>
<td>0.10-0.50</td>
</tr>
<tr>
<td>R1100*</td>
<td>0.60-0.82</td>
<td>0.80-1.30</td>
<td>0.30-0.90</td>
</tr>
</tbody>
</table>

* Other alloy elements such as V or Mo, Nb can be applied according to agreement between manufacturer and the buyer.

4.1.1 Hardness

Rail steels used in common service have increased in hardness from 280 HB to over 400 HB (Table 13, the 260 and 350 HT steels are now used in by far the greatest quantities), whereas for wheels, the increase has been from about 260 to 360 HB. The impact on performance can be readily seen using the shakedown approach of Section 2.1.1.1. Higher contact stresses can be endured before steel will continue to ratchet – that is for the same $P_o/k$, increases in $k$ allow for increases in $P_o$. The improved resistance to RCF associated with harder steels is ably shown by the many good works summarized in References 4, 136, and 189.

Table 13. Standard European Rail Steels and Their Hardnesses [291]

<table>
<thead>
<tr>
<th>steel grade</th>
<th>hardness range HBW</th>
<th>fracture toughness $K_{IC}$ Mpa m$^{1/2}$</th>
<th>Carbon Wt%</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>200-240</td>
<td>30 single 35 mean</td>
<td>0.38-0.62</td>
<td>C-Mn</td>
</tr>
<tr>
<td>220</td>
<td>220-260</td>
<td>30 single 35 mean</td>
<td>0.50-0.60</td>
<td>C-Mn</td>
</tr>
<tr>
<td>260</td>
<td>260-300</td>
<td>26 single 29 mean</td>
<td>0.60-0.82</td>
<td>C-Mn</td>
</tr>
<tr>
<td>260 Mn</td>
<td>260-300</td>
<td>26 single 29 mean</td>
<td>0.53-0.77</td>
<td>C-Mn</td>
</tr>
<tr>
<td>320 Cr</td>
<td>320-360</td>
<td>24 single 26 mean</td>
<td>0.58-0.82</td>
<td>1% Cr</td>
</tr>
<tr>
<td>350 HT</td>
<td>350-390</td>
<td>30 single 32 mean</td>
<td>0.70-0.82</td>
<td>C-Mn HT</td>
</tr>
<tr>
<td>350 LHT</td>
<td>350-390</td>
<td>26 single 29 mean</td>
<td>0.70-0.82</td>
<td>low alloy HT</td>
</tr>
</tbody>
</table>

European experience with rail steels is that harder rail steels can reduce the rates of crack initiation, but they require more rigorous maintenance and are less forgiving than softer steels. In one field test [196], head checks formed on new 1100 grade rails after 1 month of traffic, whereas similar cracks developed on UIC 900A grade rails only after approximately 2 years of

---

9 These steel grades can be nominally related to the UIC grades in Table 12 because of the relationship between hardness (Table 12) and tensile strength (Table 12). The Grade 700 rail includes hardnesses about BHN180-235, Grade 900 includes BHN 250-300, and Grade 1100 includes BHN 320-360 rail.
traffic. Even so, they concluded that “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”.

There is no doubt that, in terms of life extension, the greatest impact of improved steels has been its resistance to gauge face wear. Laboratory tests [e.g., 135] have shown that in unlubricated wear, head-hardened rail steels can reduce gauge face wear rates by a factor of 6 when compared with Standard rail. Under well-lubricated conditions, the head-hardened rail specimens wore at half the rate of the Standard carbon rail specimens. Recent full-scale rig tests show the wear average wear resistance of R350MHT rail is approximately 2.2 times that of R260HB rail [258], whereas the resistance of the 400HB UC (ultrahigh carbon) rail wear is double again that of the 350 MHT. Results from the field tests are less well defined, because the lubrication and/or contamination of the rail cannot be controlled to the same degree as in the laboratory [183].

Conclusion:
For a given rail microstructure (e.g., pearlitic or bainitic), increased hardness is associated with reduced wear and increased resistance to plastic flow. In situations of poor wheel-rail profile matching, the more forgiving softer steels are less likely to catastrophically fail. The notorious Hatfield derailment involved a 350-MHT rail that shattered into many broken pieces, a failure that was presumed to have been less likely with a softer rail.

Along with proper rail profiles (applied either during initial manufacture or through rail grinding), harder rail steels have played a strong role in the three-fold increase in rail life since the 1960s.

4.1.2 Cleanliness
Rail cleanliness is judged according to the amount and distribution of soft inclusions, hard inclusions, and entrapped hydrogen. In the most general terms, soft inclusions increase wear, hard inclusions are associated with subsurface RCF defects, and entrapped hydrogen is responsible for some internal defects.

Wear resistance of rail is adversely affected by soft manganese sulfide inclusions [103], with clean steel outperforming standard steel by a factor of approximately 2. Sulfide inclusions may also contribute to the initiation of RCF surface cracks [90], particularly in the softer grades of rail steel. In either case, plastic flow from asperity contact elongates the sulfide inclusion and effectively forms microscopic cracks within the surface layer, weakening the steels.

Hard inclusions or stringers of inclusions, typically of alumina-silicate composition, have been implicated in the development of subsurface initiated RCF defects [22, 267]. Inclusion stringers are strongly implicated in the formation of vertical split heads [223], deep-seated shells [267], and in particular, transverse defects from the deep-seated shells. Testing and correlation in Russia [234] found that the fatigue limit of rail steels increased with decreased inclusion size and number, especially alumina inclusions. A similar conclusion is reached in Reference 61, which shows that an increase in the size of inclusions, has a dramatic effect on the resistance to fatigue initiation (relative fatigue limit) (see Figure 41). Section 1.3.8 notes that hard alumina inclusions are implicated in the preponderance of shattered wheel rims. Field testing has shown that the formation of subsurface defects is critically dependent on the “Sugino Index” [268] a measure of the length of oxide stringers. Subsequent work by Clayton and others [132, 154, 266] suggests that the defect rate closely correlates to the so-called “Clayton Number,” which is dependent on the oxide volume fraction, hardness, and Sugino Index.
The much greater care to minimize the presence of hydrogen in the bloom has mitigated the development of hydrogen shatter cracks (also called fish-eye cracks) that lead to the formation of kidney shaped transverse failures.

![Graph showing the effect of increasing inclusion size/diameter (d/d₀) on the relative fatigue limit (Q/Q₀).]

Figure 41. The Effect of Increasing Inclusion Size/Diameter (d/d₀) Is to Dramatically Reduce the Relative Fatigue Limit (Q/Q₀) [61]

What does the future holds for further improving steel cleanliness? All rail is vacuum degassed which is effective in eliminating hydrogen and coupled with a requirement to test every batch of rail both ultrasonically and (samples) destructively has eliminated many of the vertical slit heads, split web, and other problems that have arisen in the past. Most of the wheel producers also use vacuum degassing, although the North American wheel manufacturers do not. Current thinking is that many wheel failures are related to large inclusions (e.g., refractory materials) that are not properly detected by the statistical sampling methods currently used.

While in general cleaner and cleaner steels exhibit improved properties, the focus seems to be instead on improving yield strength through alloying, heat treatments, and alternative microstructures.

### 4.1.3 Steel Type

Rails and wheels are typically composed from manganese, pearlitic, bainitic, and martensitic steels.

- **Manganese steels** are used primarily for special trackwork items that are subject to repeated heavy impact loading, such as frogs and diamonds. Manganese steels have elongation up to 40 percent and under work hardening achieve hardness of approximately 600 HB. Their greatest strength lies in their resistance to chipping and breakage, having a fracture toughness up to approximately 85 MPa/√m (compared with around 40 MPa/√m for pearlitic steel, Section 2.2.5).

- **Pearlitic Steels** have a duplex structure, consisting of iron carbide (Fe₃C) platelets with a carbon content of 6.67 percent in a soft ferritic matrix with carbon content below 0.02 percent. In rail and wheel steels the duplex structure has a net carbon content ranging between 0.4 and 1.1 percent, with eutectoid steels (0.77 percent carbon) being the most common. The hard carbides provide good wear resistance, whereas the layered structure provides inherently good resistance to deformation and crack propagation.
Bainite is also a duplex material composed of ferrite and iron carbide in a needle-like structure. Compared with pearlite, it exhibits higher toughness and hardness.

Martensitic steels have high hardness but poor inherent toughness. Martensite is produced by cooling at much higher rates than other steels and is frequently applied in tool steels. It is not generally recommended for rail or wheel steels, but limited applications have been tested, including experimental wheel and rails.

It has been found that within the class of pearlitic steel, wear resistance increases with hardness. Steel hardness rises as the carbon content is raised and as the lamellar structure is refined, either by alloying or—more usually—by accelerated cooling given by head-hardening processes. Because wear is perceived to be the main cause of rail replacement, all rails in North America (and many worldwide) are produced with about 0.75 weight percent carbon and 0.9 weight percent manganese. Rail intended for tangent track and mild curves is routinely produced in the naturally cooled condition, giving a hardness of approximately 300 HB. Rail intended to have especially good wear resistance in curved track is typically produced in the head-hardened condition, with hardness in the range 340–390 HB.

When engineering rail steels to better resist RCF, properties such as hardness, fracture toughness, metallurgical cleanliness, and the state of residual stresses are of primary consideration. The next generation of harder and tougher steels is being pursued through improved steel making facilities [37, 198], head-hardening practices [14], and better inspection techniques [46, 215]. Improvements to resistance in wear and fatigue have occurred by three primary means:

- Improving cleanliness: With the focus primarily on reducing interstitial hydrogen and sulfur and alumino-silicate inclusions, cleaner steels have significantly reduced the number of catastrophic failures such as shattered rims, tache ovals, vertical split heads, and others.
- Alloying: An example of how the addition of small quantities of alloying elements into the steel microstructure can be used to improve material performance is provided in literature available at one rail manufacturer’s Web site:

Additions of chromium, molybdenum, nickel, silicon, and many other elements can be added to the steel content to improve material properties. Griffin's patent-pending micro-alloy Class C steel contains chromium, silicon, and molybdenum to improve high temperature material properties. The micro-alloy composition is still within the AAR specification for Class C steel. Chromium is a solid solution strengthener in the ferritic zone of the fine pearlitic microstructure, which increases the yield and tensile strength. In addition, chromium is a strong carbide former and provides hardenability in quenched and tempered.
steels, both, which improve wear resistance. However, the amount of chromium used is limited due to adverse effects on ductility. Molybdenum is also a solid solution strengthener of the ferritic zone, strong carbide former, and increases hardenability with minimal effects on ductility. Furthermore, molybdenum suppresses temper embrittlement. Silicon is added to increase strength with little sacrifice in ductility and promotes "cleaner" steel by allowing for more uniform ferrite grains.

- Microstructural refinements: refining pearlite spacing increases the hardness of pearlitic steel which in turn improves wear resistance. Fine micro-structural refinements to avoid detrimental products at the grain boundaries are being pursued in the USA [213]. Manufacturers continue to refine the pearlitic steels for improved performance through increased carbon content and advanced head hardening processes.

4.1.3.1 Pearlitic Steels

Further improvements in the RCF performance of pearlitic rail steels is coming from ongoing refinements to the microstructure. Sawley\textsuperscript{10} suggested that one should avoid proeutectoid ferrite (grain boundary ferrite) because this was implicated in squat defects in older British rail steel, which had a carbon content of approximately 0.55 weight percent. He noted a suggestion from Paul Clayton who, on the basis of microstructural observations, believed that the grain boundary ferrite acted as planes of weakness when deformed at and near the rail surface. That is, cracks appeared to follow the ferrite planes. An older draft European rail specification [23] included a requirement that “No continuous closed ferrite network …shall be observed below 0.5 mm depth measured anywhere on the rail head surface…. ” Modeling by Franklin et al. [75] supports this recommendation. Recent work in North America reports that “pro-eutectoid cementite (a hard product) is one of the most detrimental constituents for crack initiation” [213] and that eliminating this constituent and nonmetallic inclusions, coupled with further refinements of the pearlitic lamellae, additional wear and fatigue improvements can be achieved.

4.1.3.2 Bainitic Steels

Laboratory studies consistently show that, for a given level of hardness, pearlitic steels are more resistant to RCF than are other structures such as bainite and martensite [34, 136]. However, although pearlitic steels have a maximum initial hardness of about 400 Brinell, bainitic steels can be manufactured with hardness approaching 460 HB.\textsuperscript{11} Although it is generally agreed that the wear performance of bainite is poorer than pearlite, the view is by no means unanimous [e.g., 27, 32, 226]. Pointner [207] explains that this is probably due to confusion in defining the bainitic structures themselves, with a variety of structures being called that same name. Upper and lower bainite are two classical forms, where upper bainite has generally poor properties and lower bainite having excellent strength, toughness, and RCF resistance, even when compared with fine pearlite. The RCF performance of bainitic structures has been demonstrated in both laboratory [299] and field [27, 86, 226] to exceed that of pearlite. Work continues [e.g., 17] to explain the performance of different rail steels.

\textsuperscript{10} Personal communication with the author in 2005.

\textsuperscript{11} Personal communication via email from D. Szablewski, AAR 07DEC10.
4.1.4 Layered Steels

The addition of coatings onto wheels and rails has been a topic of research for at least the past couple decades. Clayton refers to plasma spray coatings and self-lubricating metal coatings and bonded lubricants [32] for the purposes of reducing wear. Hard deposits [192] and surface hardening [186, 205] have been applied to rebuilding thin wheel flanges or for hardening the wheel flange face, the area subject to highest wear. Resistance to RCF has become a more recent objective.

In Europe, an extensive project was undertaken to evaluate the performance of a “two-material” rail [77, 100]. A commercial laser-cladding technique applied hard surface coatings of approximately 3 mm (one-eighth inch) thickness to 10-meter lengths of rail. The cladding material was applied to the gauge side of the rail over a width varying from 0.68 inch (17 mm) to approximately 1 inch (26 mm). The initial HV of the rail and two materials trialed were 295, 390, and 530, respectively. After approximately 7 MGT of heavy haul traffic, the surfaces hardened to 360, 540, and 670 HV, respectively. The width of the applied layer was found to be critical. The thinner band was found to “slide” toward the gauge corner, and a longitudinal crack developed at the boundary with the uncoated surface steel. The wider running band exhibited no such problems. It appears that the testing and monitoring did not continue beyond the 2003 end date of the study. A summary of the project [73] found that there was insufficient evidence that the added cost of the layering process could be justified by reduced maintenance costs (although it appears that interest remains for applications in noise reduction). A subsequent proposal to the European Union to fund investigation into the repair and maintenance of coated rails was unsuccessful.

4.2 Traction Coefficient

Creep or creepage refers to a small speed difference between the contacting wheel and rail. For example, positive or negative creepage and microslip exist between wheel and rail due to driving torque or braking of a wheel. Similarly, creepage, also in the longitudinal direction, develops because of rolling radius difference, usually from lateral displacement of the wheelset in curves. Yaw misalignment or angle of attack, also present in curves, results in lateral creepage and microslip that gives rise to traction force perpendicular to the direction of travel. The relationship between creepage and traction force, regardless of direction, is shown in Figure 43A. This figure also shows the approximate extent of microslip within the contact patch (a circular patch is depicted for illustrative purposes). For free rolling of a wheelset with a cylindrical tread, there is no creepage and the microslip within the contact patch is limited to the exit region of the contact patch. When the relative velocity between wheel and rail is approximately 1 percent or larger, the microslip overwhelms the entire contact patch. The amplitude of traction force compared with normal force is then equal to the limiting friction coefficient, and the contact patch is said to be in the state of full or saturated creepage. The level of creepage at the wheel-rail interface depends on the curving and traction demands. Especially in sharp curves, where the restrained wheelset takes a large yaw angle to the rail, lateral forces rise asymptotically toward the limiting friction value (Figure 43B). Note that the L/V exceeds μP in Figure 43B because of the wheel tread conicity; even in the absence of creepage forces, the wheel tries to roll the rail outward as a result of gravitational loading.

Since the tractive force has a limiting value of μP, an angle of attack and lateral creepage in curves reduces the maximum tractive effort that can be achieved in the longitudinal direction. This is the premise behind steerable trucks on locomotives—the reduction in yaw angle allows
for greater longitudinal traction. However, from the perspective of RCF, what matters is the total tractive effort (the vectorial sum of longitudinal and lateral creepages). Figure 44 shows that, even though the longitudinal traction coefficient has typical values of less than 0.3, when the lateral creepage is included the net traction nearly doubles [164].

As noted in Section 2.1.1, surface tractions have a profound effect on the distribution of compressive, tensile, and shear stresses within the contact patch. The most notable effect concerns the position of surface and subsurface shear stresses. Once the traction/normal force ratio exceeds the value of approximately 0.26, the peak shear stress is at the surface and acts over a very thin layer. This stress rises rapidly with increasing traction and accounts for the rapid rate of RCF crack development under high tractions. Although a doubling of wheel load increases the peak normal stress by approximately 27 percent, a doubling of traction coefficient from 25 to 50 percent increases stress by 50 percent. Thus, the higher the traction force, the faster the initiation of microcracks.

Tractive forces at the contact patch are governed by a wide range of factors including friction conditions, wheel and rail profiles, cant deficiency, track curvature, wheel loads, wheelbase, and suspension parameters. The resulting longitudinal and lateral creepages impact the rate and orientation of surface cracks (see Figure 45). Surface cracks initiate perpendicular to the direction of the average resultant creep vector. Longitudinal cracks are thus generated by lateral creepage in sharp curves. Inclined cracks are generated by mixed longitudinal and lateral creepage in intermediate or mild curves (SCLR in Figure 45). Observations of the fatigued wheel-tread, coupled with an understanding of curving requirements and wheelset kinematics allows the following general statements to be made:

- The direction of plastic flow is opposite at either side of tangent track running band.
- The lateral component of plastic flow as a result of shear stress from lateral tractions always points towards tangent-track wear-band (center of the wheel tread).
- The direction of the longitudinal component of plastic flow as a result of shear stress from longitudinal-traction forces (when in contact with rail) always points in the direction of travel when at flange root and opposite to direction of travel when at the rim side of the wheel.

Microcracks develop most rapidly when the wheel-rail interface is subject to a high friction environment. Although this is typical of dry climates, the simultaneously high wear rates exterminate many of those cracks early in life. However, in wet climates, microcracks may be generated almost instantly during the dry period that follows flushing of the wheel-rail interface by water (rain or snow) when friction (\(\mu\)) reaches levels of 0.6–0.7. Subsequent propagation can be very rapid in the next moisture cycle and depends on the directionality of longitudinal plastic flow and associated orientation of cracks. Any water in track becomes trapped on the rim side of the wheel and squeezed out of the flange root area of wheel tread. It is from this mechanism that cracks on the rim side of the wheel usually propagate faster than those in the flange root.
Figure 43.  A) Schematic of Full Traction Curve for Free Rolling, Rolling-Sliding, and Sliding Contacts;  
B) L/V versus Degree Curvature for the Leading Wheelset with a Mildly Worn Freight Wheel 
Profile against the RE136, 141AB, and NRC-H2 Rail Grinding Template (top of rail friction is 0.45, 
and the gauge face value is 0.3)
Figure 44. Traction Forces Measured with Instrumented Wheelsets Running on the Amtrak Northeast Corridor
Figure 45. Surface Traction and Subsurface Plastic Flow – Longitudinal and Lateral Traction Directions for RCF Crack Initiation at Different Locations on the Wheel Tread; One or Several Distinct Bands of Cracks May Appear as a Result of Contact with Tangent Track (TT) or the High Rail (HR) or Low Rail (LR) of Sharp Curves (SC) and Mild Curves (MC)

This mechanism applies equally well to crack development in rails. Since the creepages on the wheel surface shown in Figure 45 are equal and opposite on the rail, the crack orientation is similar on the rails, but with material flow in the opposite direction. So whereas the water fluid mechanism exacerbates crack propagation at the field side of the wheel, it is worst at the gauge corner of the outside rail and top of low rail.

4.3 Wheel Loads
Normal contact stress is a function of four main factors; wheel diameter, wheel load (including possible dynamic loading), the transverse rail profile and the transverse profile of the wheel. According to Hertz’s elastic contact equations, the doubling of wheel load will increase contact stress by approximately 27 percent, a tripling increases contact stress by 44 percent. If, with respect to the strength of the material and traction coefficient, this increase is sufficient to place the material beyond the fatigue boundaries, then the consequences are indeed dire. For example, a 10 percent increase in contact stress (from 3,320 to 3,640 MPa) decreased the fatigue life by a factor of about 2 for nontreated and heat-treated rail steels [235] under pure rolling conditions in the laboratory.

Although the nominal static wheel loads are generally set by commercial concerns and limits on track infrastructure (often bridges), management of dynamic loads has the potential to reduce dramatically the incidents of broken rails [158], broken wheels and wheel shelling in particular. Dynamic wheel loading associated with flat spots or out-of-round wheels, poorly straightened and aligned rail, insufficient tie support, rail and wheel corrugation, dipped welds, poorly maintained turnouts, etc., can be detected and managed. Uneven loading, poor cant selection [106] and poor truck/bogie wheel load equalization capabilities are further contributors to wheel and rail damage. As noted in Reference 20, the reduction of dynamic loads and more even
spreading of stress is an approach to either limiting rail damage or allowing for greater static wheel loads.

4.4 Truck/Bogie Characteristics

**Truck Suspension:** a truck with stiff plan view suspension resists displacement of the wheelset with respect to the truck frame. The more flexible the suspension, the greater the potential for favorable steering moments to reduce the yaw angle in curves and thereby reduce RCF. However, a more flexible truck has a greater ability to respond to unfavorable steering moments and increase the yaw angle [278], especially in the case of trucks that have been poorly maintained and are running with a number of worn-out components. A number of trucks have been developed to improve upon the limited curving performance of the standard three-piece truck but with improved stability in tangent track [229, 230, 245].

A study from the United Kingdom [217] to examine the potential of modifications to the primary yaw stiffness for reducing RCF:

> ... demonstrated potential cost benefits for the GB rail industry from modifying vehicles to reduce PYS. The cost-benefit for the Cl357 with a 33% reduction in bush stiffness was particularly robust, as the change to the vehicle required no extra maintenance. The cost benefit calculation suggests possible overall saving to industry of around £60 thousand ($97 thousand) from year 3 onwards for Cl357. The possible savings indicated for Mk4 suggest a possible overall saving to industry between £0.5 million ($0.8 million) to £2.5 million ($4 million) from year 3 onwards.

In a freight environment, work has been performed in the Brazil, North America, China, South Africa, and elsewhere to quantify the benefits of improved trucks with respect to RCF.

The work in Brazil [283] found that a frame-braced truck should theoretically halve tractions in intermediate (873 m radius) curves and reduce Tγ values to below the damage threshold level required to initiate RCF. The field service results showed that tread and flange wear were 30–50 percent lower for frame braced trucks compared with standard trucks, while nearly 60 percent of test wheels on standard trucks showed RCF damage compared with only 3.5 percent for the frame braced truck over the same 85,000-mile (136,000 km) interval. The net effect on wheel life is projected to be a four-fold increase.

In North America, the CPR found that self-steering trucks on its (captive) coal fleet considerably improved wheel performance, providing a 36 percent increase in average wheel life in a system where the primary cause of wheel removals was shelling [185]. Besides the wheel improvements, CPR also found fuel savings measured during controlled tests. Wider adoption of advanced trucks in North America is stymied by the fact that the bulk of the benefits are seen to fall to the infrastructure owners, but the costs of their implementation falls to the rolling stock owner [31].

In China, approximately 75 percent of freight car bogies are of a frame-braced design [104]. Continually developing understanding of the effects of bogie characteristics on the wheel-rail interaction is the foundation for subsequent design, manufacture, and maintenance of bogies [298].

In South Africa, high rates of wheel wear prompted the introduction of self-steering mechanisms in 1975 that included rubber pads at the journal boxes coupled with cross-bracing linkages [230]. Development of the self-steering “Scheffel bogie” has been ongoing [e.g., 81] and is a key feature of the South African approach to minimizing RCF [79].
Swedish engineers, after evaluating the potential of improved trucks to minimize infrastructure damage and maintenance costs on their ore lines, experimented with radial self-steering trucks in the late 1990s [193]. Many complications arose related to snow and ice [81] and their effects on the steering linkages and braking systems that contributed to unsatisfactory performance of those trucks in operation. That has not curbed Swedish enthusiasm for trucks that are more “track-friendly,” as illustrated by ongoing research into double-link suspension on two axle wagons [241].

4.5 Track Geometry

As noted by Smith [244], “local features of track geometry can enhance stresses by dynamic effects and can therefore be instrumental in initiating fatigue cracks at particular locations along the track.” On the Network Rail system, modeling found that the probability of RCF clusters arising because of track geometry variations was strong for shallow (1,000–2,250 m radius) curves at balance speed [44]. Field investigations found that discrete alignment errors (e.g., bad weld) or short sections of poor track geometry were the cause of local patches, or clusters, of RCF [265]. In contrast, the addition of track perturbations into dynamic models reduces the calculated fatigue damage on wheels because the distress is more evenly spread across the wheel tread [48].

The additional sideways load on the truck from cant deficiency changes the orientation of axles. Increasing cant deficiency in freight railroads has been demonstrated through modeling to contribute to a reduction in RCF generation, primarily because of the reduced overall creepage that arises as the truck realigns itself in curves to address the unbalanced lateral force [44]. In the case of high-speed trains running on mild curves with large cant deficiency, both the leading and trailing wheelsets offset heavily to the outside rail, sometimes to the extent of flanging (depending on the profiles). This lateral shift is much greater than when running at balance speed in the same mild curve. The result is increasing (longitudinal) tractions with cant deficiency and high potential for RCF of the mid-gauge position on the rail.

On freight railroads, track geometry error such as tight gauge and crosslevel errors introduce localized RCF defects (see Figure 46). Such errors are associated with increased contact stress (because of increases in dynamic wheel load) and increased concentration of wheel passes, both of which increase fatigue rates in the vicinity of the geometry error.
4.6 Wheel-Rail Profiles

Wheel-rail contact varies in position, amplitude, and consequence with the profile shapes and degree of track curvature. The rail profile at any location in the track has to conform to the combination of new and worn wheel shapes that pound the rail at that location. In tangent track, the focus is to promote stability and sometimes to protect against crushed heads or gauge-corner defects. The low-rail design must minimize contact stress and ideally improve steering. The high-rail design must provide sufficient relief to avoid gauge corner defects (especially in high axle load systems) and yet promote steering to control wear and RCF.

Besides reducing RCF of the rail, a family of rail profiles that contact the wheel at different running bands will provide a significant benefit to the wheel. Rail profiles designed to spread wear on the wheel slows the development of a false flange or geometrical stress raiser [251]. A properly designed system of wheel-rail profiles that controls stress and wear provides for durable, stable, and optimized wheel-rail performance.

It should be noted that, since wheel and rail profiles cannot be changed instantaneously, the transition from current to new shapes must be properly considered during profile design and subsequently managed during implementation.

4.7 Wear – Material Attrition

As noted in Section 2, newly initiated surface cracks are a fraction of a millimeter deep and propagate at an angle of 5°–15° to the surface. These newly initiated cracks generally propagate at a relatively slow rate, then much faster at the intermediate length of approximately 5–10 mm and depth of approximately 1–3 mm. At greater length and depth, where the contact stresses are less intense, crack propagation proceeds at a slower rate [95].

As these cracks are propagating into the surface, at the same time they are being truncated by either natural wear or machining processes such as wheel re-truing, abrasive brake shoes, or rail grinding. Regular re-truing of wheels to an optimized shape has increased wheel life by more than 50 percent in Australia [109, pages 4.3-4.5], South Africa, and Canada [130]. Abrasive
brake shoes, such as those with cast iron inserts, can be effective in scrubbing metal from the wheel tread to remove incipient fatigue cracks as well as wheel-slide damage. One manufacturer reports\textsuperscript{12} that after field testing involving three unit trains operating over the same route for 3 years, 20 wheelsets were replaced for shelling on the one trainset completely equipped with brakeshoes having cast-iron inserts, whereas on two otherwise identical trainsets with standard composite brakeshoes, the average number of wheelsets removed for shelling was 119.

Some railways have dealt temporarily with RCF issues by simply turning off the rail lubricators. The resulting very high rates of wheel-rail wear can scrub visible RCF from the surfaces but at the expense of increased plastic flow, uncontrolled wear, unfavourable profile changes and increased fuel consumption. Furthermore, upon restarting of the lubricators, the extensive system of surface cracks that developed under dry friction propagates rapidly with grease contamination.

A more effective approach is to use lubrication to control wear and through rail grinding frequently remove a small amount of the rail surface material to truncate existing cracks and remove very shallow damage. The optimal strategy is to remove just the right amount of metal to control both surface and subsurface crack initiation and to remove short cracks while the rate of propagation is still slow. This metal removal rate has been called “The Magic Wear Rate” \cite{131} (Figure 47).

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure47.png}
\caption{Controlling Wear at “The Magic Wear Rate” Maximizes Rail Life}
\end{figure}

\textsuperscript{12} http://www.rfpc.com/upload/Brochure/TG_Brochure.pdf.
5. Management of RCF

Good reporting and control systems are essential for the safe and effective management of RCF. Unfortunately, inspection reports are often filled out by persons who are not experts in rail failure and who are often under severe time pressure to repair the track. Yet nonetheless, a good record will contain at least the following management information for each defect: geographical location, damage type, route, track (e.g., 1 or 2), rail (e.g., east/west, north/south, high/low, or left/right), rail stamp (metallurgy, year of manufacture), and current state of rail wear. These reports can then be used to assemble a statistical picture of rail failure causes and trends and hence provide clues to appropriate remedies. Inspection and maintenance schedules can be adjusted to be location specific and maintenance actions prioritized to ensure the most efficient use of resources [20].

In the Netherlands, statistical methods are used to develop a “damage number”

\[ D = D_r \times D_s \times D_p \times D_t \]

where \( D_r \) is related to the curve radius, \( D_s \) to sleeper type, \( D_p \) to rail profile, and \( D_t \) to traffic type [101]. The parameters \( D_r, D_s, D_p, \) and \( D_t \) are determined through statistical correlation with a large database that warehouses RCF visual inspection records. On the basis of the damage number, grinding priorities are assigned. The inclusion of a few additional parameters (such as grade, metallurgy, weld quality, prior maintenance practices, etc.) with regular updating might provide a useful and ongoing approach to managing RCF in North America.

Several other visions are emerging for technologies that may considerably improve the reliability and accuracy of RCF detection and treatment, reduce the cost, and improve safety with respect to RCF. They are discussed in the following sections.

5.1 RCF Damage Model

Computer-based predictive models of several forms are being derived. In Europe, extremely comprehensive systems are linking multibody vehicle dynamic, wheel-rail contact, track, crack initiation and propagation and deformation models to predict the evolution of RCF on the rail [e.g., 116, 293]. These have since migrated to North America [e.g., 216].

The current state of the art with respect to modeling the development of RCF relates damage to the sliding energy dissipated at the interface. The so-called wear number \( (T\gamma) \) is simply the shear force or traction \( (T) \) at the contact patch multiplied by the normalized sliding distance (creepage \( \gamma \)). For low values of \( T\gamma \), there is insufficient energy at the contact region to cause damage, and the fatigue life is essentially infinite. As \( T\gamma \) increases, fatigue cracks can initiate, and the fatigue life decreases. At very high values of \( T\gamma \), wear dominates and existing RCF can even be removed. There is a value of \( T\gamma \) where the wear rate just balances the rate of fatigue initiation, and the fatigue life is again infinite. This series of circumstances is illustrated by the RCF damage function in Figure 48. The RCF damage function is simply the reciprocal of the number of cycles to fatigue, such that a value of 0 implies infinite fatigue life and a negative value represents conditions under which existing fatigue is being worn away. The specific curve shown in Figure 48 was derived for steels commonly used in the United Kingdom. FRA is currently sponsoring work to calibrate the curve for class B (passenger car) wheels running on premium rail steels. Since \( T\gamma \) can be readily calculated in various multibody dynamics models, the damage model can be easily incorporated in a post processor.
As noted in Reference 48, the shortcomings of this model are as follows:

- It must be recalibrated (against field data) for different material pairs and perhaps operating conditions.
- Contact pressure, which can be expected to have a strong influence on both RCF and wear, is not explicitly considered.

In addition, this model in no way includes the effect of water and oil on the propagation rates of cracks. Furthermore, rates of wear and possibly fatigue on rails and wheels with contaminant layers at a given friction coefficient (e.g., 0.3) will be very different from the same rates of damage under a noncontaminated contact of the same friction coefficient. Jendel [117] developed environment compensation factors to account for this difference, which could be included if a more advanced model is considered.

### 5.2 Managing Wheel-Rail Profiles

#### 5.2.1 Wheel Profiles

It is unfortunately the case that the current wheel-rail profiles specified in North America are generally poor in that they provide a nonconformal two point contact against most of the worn high rails and so have poor steering and high slip rates, which exacerbates both wear and fatigue.

The design of a wheel profile for a railway needs to match the characteristics of the rolling stock (suspension, axle load, speed) and the track (percent of curves, maximum curvature, track gauge, metallurgy). An optimal wheel profile designed for the system will provide the following benefits:

- Improved steering in curves to reduce wear, L/V forces, and RCF. The Cartier Railway Company reduced wheel shelling by 60 percent by adopting a custom wheel profile [130].

---

**Figure 48.** The RCF Damage Function (1/N_f, the number of cycles to failure) as a Function of the Wear Number T_f [15]
CPR increased the life of coal fleet wheels by 18 percent using an improved wheel shape [185].

Control of creep forces was achieved by matching wheel profile curving ability (conicity) to the curving requirements. A wheel profile designed for the U.K. railways [171], called the WRISA2 (and now the “P12”), showed a considerable improvement with respect to shakedown (Figure 49), not so much by reducing contact stress but rather by improving curving to reduce the wheel-rail creepage forces.

![Figure 49. A Dynamic Shakedown Plot Summarizes Contact Conditions for Three Different Wheel Profiles against the Same Rail Shape As-Is Negotiates a 1,500-Meter, High Cant Deficiency Curve](image)

Besides reductions to rail fatigue, other benefits were designed into the wheel shape including:
- Improved stability and reduced hunting by providing an appropriate tread slope over the tangent running band to control effective conicity in tangent track [278].
- Reduced damage to the field side of the wheel and the low rail as a result of lower tread wear rates and by lightly rolling off the field side of the profile to minimize hollowing.

Implementation of a new profile can be accomplished with modest incremental effort by re-truing during the regular turning cycle.

A new wheel profile eventually wears to a shape that is governed by the profiles of the rails over which it runs. Poor rail profiles or low-strength steels limit the benefits that a wheel can achieve by itself. Yet when implemented with matching rail profiles, the benefits are multiplied. The grinding of rail profiles in tangents and curves that reduce hollowing and help to maintain the wheel’s designed shape throughout its life will maximize the performance of an improved wheel profile.

The pummeling approach discussed in Section 5.2.4.1 can also be applied to the analysis and design of wheel profiles. Reference 169 gives an example of how it is used on the Amtrak high-speed rail system to model wear on various new and worn wheel shapes. Many others have applied multibody dynamics codes to wear simulations [e.g., 155, 225] and profile designs, with increasingly detailed models of the track being used.
5.2.1.1 Wheel Hollowing
A survey of North American worn wheels in 1998 found that nearly half of the measured profiles showed some hollowness, with more than 6 percent of the wheels exhibiting more than 3 mm of hollowness [201]. Subsequent analyses with multibody dynamics models showed that the economic cost (wear, rolling resistance) of hollow wheels to the rail infrastructure merited the removal of such wheels once they reached 3 mm of hollowness [224]. Unfortunately, that work did not examine the contribution of hollow wheels to rail or wheel fatigue, and although several publications [e.g., 173, 279] hint that consequences could be considerable, it does not appear that any quantitative analysis has yet been undertaken.

The South African experience with wheel profile monitoring systems [79] notes that errors in the measurements are common. Sunlight and dirt on the wheel can result in profile measurements with errors that range from obvious to difficult to detect. Statistical methods are required to avoid an excessive number of false-positive values.

5.2.2 Rail Profiles
That rail and wheel profiles have a direct impact on contact stress and steering, and thereby to RCF, has already been discussed in Section 4.6.

In the remainder of this section, we focus on the approaches taken in practice to develop wheel and rail profiles, starting with shapes designed initially to treat rail surface problems and progressing to the latest designs, based on detailed computer analysis.

5.2.3 Rail Profile Shapes Developed from Practical Field Experience
The shape to which rails are ground has also been an evolutionary process. In the earliest days of rail grinding, removal of corrugation from the top of rail was the focus, and the result tended to be a squaring of the rail head, with a progressively flattened top and a sharp gauge corner. This corresponded well with the shape that develops naturally as a result of plastic flow, contact fatigue, and wear (Figure 50).

The fixed motor angles on rail grinders meant effectively that the same shape would be applied to all rail. Work in Australia in 1978 [280] found that intentionally applying asymmetric rail shapes to the ground rail in curves to shift the high-rail contact band to the gauge and the low-rail...
contact band to the field could exploit the rolling radius difference available with coned wheels. This improved the wheelset curving ability and reduced gauge-face contact, especially in mild curves. They also found that shifting the contact to the field in tangent rail reduced a tendency for periodic side cutting of the rails caused by body hunting of the ore cars. Soon thereafter, Speno offered a rail grinder to North America that allowed the grinding stone angles to be adjusted from within the cab. Profile grinding was now a realistic possibility. A technique quickly adopted was to relieve the field side of the low rail, as it was proved in tests to reduce the rate of corrugation development by 40 percent, simply by minimizing contact with the “false flange” of hollow wheels [219]. A defensive strategy evolved in North America, where it was seen that tighter contact bands would protect rails from excessively worn wheels, and would pass the load through the axis of the rail web, reducing eccentric loading of the rail.

Another important development in rail grinding spawned from the empirical observation that the occurrences of deep shelling and transverse defects, frequently associated with gauge corner spalling, were dramatically reduced through deep relief of the high rail gauge corner. In standard carbon rail, the steel would eventually flow back to fill in the gap between the wheel flange-root and the rail gauge-corner while in premium steels, a strong two-point contact would prevail for some time (resulting in high gauge-face wear rates). This practice increased lateral loading but was generally effective at mitigating the risk of gauge corner defects.

A strategy developed on the Canadian National Railway is to simply grind more metal off the rail where the cracks appeared and leave the rail crown to a specified radius. Over several cycles, the cautious, frequent treatment of lightly cracked rail will eventually iterate the shape to one that exhibits no concentration of cracks. As a general philosophy, the procedure has merit. At the time when the rail is ground, the ideal rail-head shape will exhibit a uniform distribution of fine microcracks across the running band, with no localized band of more severe cracks. This shape would have all the rail material bearing its share of the stress applied by the passing wheels. This practice requires experienced Rail Grinding Supervisors to pre- and postinspect the track at all locations on the railway system to ensure it is successful, because the rate of surface damage can vary wildly over small distances with curvature, grade, friction conditions, metallurgy, axle load, cant deficiency, and vehicle characteristics. Railways were looking for a more general procedure for quality-assuring the derived rail shape.

The first step towards a generalized grinding procedure that controlled rail surface fatigue was developed in 1991. After observing how worn rail profiles naturally developed in service and overlaying worn wheel profiles onto these rails, Kalousek [133] presented in 1991 a set of eight rail templates for use on North American heavy haul systems. Starting with the tangent template, the four high-rail templates mandate progressively larger amount of metal removal (relief) from the gauge corner of the rail in steps of 0.5 mm, whereas the three low rail templates provide progressively greater field side relief (Figure 51). The various degrees of relief allow this family of templates to be used for a wide range of track features, including wide track gauge, dynamic rail rotation on spiked track, and different rail hardness. The eight NRC templates became the de facto standard for grinding in North America. Both British Columbia Rail (BCR) and CPR took the practice one step further, ordering their new rail with the NRC-tangent shape already installed. The newest North American standard rail shape, the AREMA 141AB, follows the same path by placing a slightly modified NRC tangent shape on the heavier 141-pound rail section.

Railways in North America found that the frequent (e.g., 8–12 MGT intervals on sharp curves) preventive grinding to the NRC 8-inch rail templates reduced the need for the severe
undercutting of the low-rail field side and high-rail gauge corner that was previously required to compensate for the profile changes of Figure 50 that led to high-stress contact conditions. As a result, rails that are frequently ground can use rail profiles that more closely match typical worn wheel profiles, which in turn are often specific to the suspension type, axle load, percentage of curves in the track, cant deficiency, etc. The wheels that run, for example, on a dedicated heavy haul coal line are almost certainly going to have a different shape than the wheels on a high-speed, double-stacked intermodal car. Greater refinement of the rail profile shapes is required when trying to match with specific wheel profile distributions and track and vehicle characteristics.

5.2.4 Computer Modeling for Rail Profile Optimization

Any section of rail is exposed to a wide range of wheel shapes from vehicles with varying suspension characteristics, speeds, and axle loads. The best shape for the rail will match optimally that population of wheels and vehicles (Figure 52), as compared with simply matching the nominal unworn profile on a new vehicle. All other things being equal, stresses are minimized when wheels and rails conform in shape as much as possible, with the possible exception of the high curvature surfaces at the wheel-throat/rail-gauge-corner contact and the convex curvature of the false flange on a hollow wheel. With a broad contact, the concentration of wheel load over a narrow path is avoided. Precise shaping of the rail head to minimize stress and promote a spread of contact is one technique for minimizing rail grinding effort and maximizing rail life. The move of railways worldwide from a defensive reshaping of rails to a more synergistic wheel and rail profile matching was highlighted at the 1999 Moscow conference of the International Heavy Haul Association (IHHA), whose theme was clearly conformal wheel and rail contact [110].

Determining the best rail shape for a given track segment is far from trivial. On the one hand, a broad conformal contact promotes low stresses that spread any fatigue damage across a wider portion of the rail head. On the other hand, some regions of wheel-rail contact have sharp contacting radii and promote unfavorable rolling radius differences that should be minimized.
Contact between an individual wheel and rail should not be too broad, because the far edges of
the contact band that extend down the gauge corner will produce excessive spin creep (which
contributes to excessive wear, instability within the contact patch as load jumps from one edge of
the contact to the other, rail corrugation, and wheel-rail noise).

Several approaches have been taken to develop optimal rail shapes [e.g., 133, 222]. These tend
to focus on designing new shapes that satisfy the criteria of sufficient rolling radius difference
for curving [e.g., 232], control of contact stress [173, 243], stability [167], and wear
minimization [246]. Although initial designs used static analytical analyses, multibody vehicle-
dynamics codes are increasingly being used in the process [e.g., 243]. Most of the referenced
papers presume a single profile (a designed wheel for example) must be directly matched by
optimal shapes on the opposing body (rail profiles for example). Yet, in practice, any train
wheel encounters a range of rail profiles with shapes that vary along the track, from tangent
through spiral to curve. Similarly, any piece of rail encounters a distribution of wheel profiles—
from unworn to very worn, new to hollow, wide flange, and thin flange. Pummeling analysis
was designed to deal with specifically this issue.

5.2.4.1 Pummeling

Understanding and controlling the distribution of contact across the running surface of wheel or
rail has been described as pummeling [279]. References 47 and 173 describe a wheel-rail
interaction model developed specifically for designing optimal rail profiles that consider the
wide range of worn wheel and rail shapes that the optimized rails and wheel shapes, respectively,
must be designed to accommodate. The NRC Profile Optimization (Pummeling) Model can use
either a quasi-static curving model or a multibody dynamics model to apply many (sometimes
thousands) measured worn wheel profiles to a truck characteristic of that fleet and derive
distributions of contact stress, fatigue damage, stability, and curving performance. By iterating
the shape of the rail, the model is used to reduce the concentrated loading of wheels on any point
of the rail surface without introducing hunting in tangent track or increasing wear in curves.

For example at CPR, a major concern is “crushed heads” on significant lengths of tangent track
that are at advanced service lives exceeding 1.2 billion gross tons. Although these defects were
shown to initiate at the gauge corner of the rail [128], the “obvious” approach of grinding down
the gauge side of the rail to inhibit further initiation would not address the existing defects,
which extend over the ball to the field side of the rail. A rail profile was needed to lightly
relieve the gauge side, spread contact broadly over the field, and minimize stress levels overall. The
desired shape was engineered using a pummeling analysis [173], which involves determining the
distribution and amplitude of stress and wear by modeling the contact between a rail and over
1,000 measured wheel profiles. Figure 53 shows how the new rail profile provides a distribution
of stress spread broadly across the field side of the rail head and avoids (although not
completely) the stress concentrations found with the NRC-TT profile.
Pummelling was further applied to the design of profiles to address wear and fatigue in curves as sharp as 12° under axle loads of up to 36 tons. The final result for CP was a set of only four rail templates [47], which contrasts with the original set of eight NRC templates. There is only one template for all high rails, one for all low rails, with the remaining two templates applied to tangent rail.

A family of generic rail profiles has been developed for Intercity Rail Systems [167] that considers a range of optimization criteria including control of RCF. Optimal rail profile designs for high-speed [222] and transit [184, 246] systems have also been published.

An approach very similar to pummeling is used in the TTCI WRTOL system for identifying track sections with unfavorable rail profiles [296].
Figure 53. The Improved CP-TF2 Rail Profile Spreads Wear across the Rail Head, Minimizing Contact at the Gauge Side where cracks Initiate, Placing Most of the Load through the Center of the Rail, but also Allowing the Field Side Metal to Bear a Portion of Stress; This Contact Distribution Is Believed to Be the Best Compromise between Controlling Initiation while Managing Propagation on Existing Defects

5.2.5 Managing Wheel Profiles

The impact of wheel and rail profiles to RCF has already been discussed in Section 4.6. Despite the availability of tools for on-line measurement of wheel and rail profiles, little is being done to exploit the full capabilities of the system. Such systems are being used in both preventive and corrective modes to catch wheels that are approaching or reaching threshold values of thin flange and high flange or to look for other problems such as excessive back-to-back spacing. Yet, although the full profile information is captured, rarely is that information used to examine any number of potential issues relevant to wheel/rail performance and safety such as flange angle, gage-face angle, wheel hollowness, grooving, asymmetric and diagonal wheel wear, and oscillating rail gage face wear. That said, tolerances on wheel and rail profiles and their geometrical interaction remain to be developed. Even less frequently are the rail and wheel data analyzed together to examine wear, RCF, stability, and derailment, although progress is starting to be made [176, 296]. Wheel-rail interaction indices suitable for evaluating risk and maintenance requirements are available (such as effective conicity [65, 296], RCF damage [15], wheel-climb [166], conformality, and pummeling [173]) but more can be done to developed improved criteria and support their adoption by the industry.

5.3 Managing Forces

Force at the wheel-rail contact affects not only the normal load in the contact patch but also the relative position of the wheel on the rail (and hence the contact geometry). As noted in Section 4.5, track geometry perturbations are often locations of RCF clusters on rail, while at the
same time, poorly performing wheels and trucks are associated with poor wheel life and can reasonably be assumed to be causing concomitant damage to the track. The identification and monitoring of track sections, cars, and wheels with high forces facilitates proactive maintenance and (when required) removal of problem components.

5.3.1 Wheel Impact Load Detector
The Wheel Impact Load Detector typically consists of rails strain gauged to measure vertical force [e.g., 285]. The length of rails instrumented is usually at least one full wheel circumference. Impact loads on the wheel are responsible for progressive and deep shelling of the wheel tread and shattered rims [262] and are a leading contributor to broken rails [158]. The wheel impact detector is now commonplace in the North American rail system, with 86 units deployed (as of 2008, Reference 277). Within the North American interchange railway system, regulations have been developed to ensure that the most damaging impacting wheels can be removed from service (see Section 2.3 for details). Examples of work done to reduce impacting wheels to minimize damage include References 127 and 276.

5.3.2 Skewed Truck Detector
One characteristic of a poorly performing truck is the poor axle alignment in both curves and tangent track. In curves, warping or parallelogramming of trucks [294] contributes to high wheel-rail creepage and concomitant lateral and longitudinal forces. The high tractions, as noted in Section 2.1.4, are a leading contributor to RCF. Measurements of axle angle-of-attack and lateral offset using laser-based systems [113] can be analyzed to assess the probable creepage and damage associated with a particular truck. Furthermore, a diagnostic assessment of the data has proven effective in identifying truck maintenance concerns including excessive center-plate/center-bowl, high wedge wear and excessive center-plate/center-bowl friction. Undoubtedly, removal of the poorest performing trucks will reduce the stress-state and excessive damage.

5.3.3 Truck Performance Detector
The truck performance detector uses strain gauged rails to measure vertical and lateral forces at each wheel as the train passes over at revenue speed [294]. Subsequent processing of the force data is used to identify trucks/bogies that may have maintenance concerns or are at high risk of derailing because of rail rollover or gauge widening. Typical problem conditions include excessive side bearing clearances and high center-plate/center-bowl friction. Although the TPD does not provide a measure of RCF-related damage, there is no doubt that removing poorly performing trucks will reduce wheel tread damage and also rail damage.

5.3.4 Instrumented Wheelset
A railway Instrumented Wheelset (IWS) is a standard railway wheelset that has strain gauges affixed at specific locations on the axle and wheels [36, 181]. When installed under a railcar, the IWS measures in real-time the forces acting at the wheel-rail interface. It is the only tool for directly measuring forces arising at the wheel-rail contact, whereas other systems such as axlebox accelerometers and strain-gauged sideframes can only infer force. The measurements of steady or quasi-static force components (including lateral forces) that have a strong contribution
to RCF can only be collected with the IWS. The IWS is therefore a powerful tool for comparing the performance of different car types to wheel and rail RCF.

5.3.5 **Acceleration-Based Systems**

On the freight railroads, rail grinding preinspection of track is often undertaken by a trained person whose tool bag includes a “seat-of-the-pants meter.” This refers to the vibration that transfers from the track to the body of the hi-rail vehicle and eventually through the seat of the driver. It would stand to reason that accelerometers, mounted on that hi-rail or on some other vehicle, could be trained also to pick up the same vibrations.

Accelerometers, typically mounted on the axle box are being used to detect short (e.g., 0.1 m) and longer (e.g., 1 m) wavelength defects such as squats, dipped welds, misaligned joints, and crushed heads [156, 256]. However, there does not appear to have been a successful effort to apply this to detection of the more general case of RCF.

5.3.6 **Vehicle-Dynamics Simulations**

As models of the wheel-rail interaction steadily improve in quality, there is increasing usefulness in applying them to predictions of wear and fatigue. The effects of track geometry, wheel and rail profiles, friction, and various truck characteristics can be readily qualified. With better material performance characterization and techniques for accumulating stress at the rail and wheel surfaces, the credible predictions of damage rates that follow will improve efforts to optimize maintenance processes and rail and wheel life.

5.3.6.1 **Wheel-Rail Contact Inspection System**

A system developed in North America for large-scale analysis of the wheel-rail interaction is discussed in Reference 296. The Wheel Rail Contact Inspection (WRCI) System performs real-time contact analysis between an onboard collection of approximately 200 “representative” wheel profiles and rail profiles that are collected at approximately 3-meter intervals. Location is tracked using GPS. Track gauge and curvature are also considered in the analysis. Of particular interest are the conicity and “conformity” measurements that can be used to assess stability and stress, respectively.

5.4 **Nondestructive Rail Inspection**

Ultrasonic measurement of rail for internal flaws has been an effective and longstanding tool for managing risk associated with rail fatigue. A 40-MGT freight line, for example, will typically be inspected two to three times annually. Much heavier trafficked lines, such as the 140 MGT coal lines in North America, may be inspected every month. In Australia’s heaviest axle load systems, inspection is as frequent as every week.

Ultrasonic and induction measurement cars typically operate at 25–45 mph, although some are now operating at 60 mph. Because the vehicle often stops for the required manual verifications, the average speed is often much lower.

Other nondestructive methods to inspect for surface defects include magnetic particle, dye penetrant [254], and eddy current [206] inspection.

New technologies being explored to improve the inspection for surface and subsurface defects include [20]:

87
- Low frequency eddy current,
- Neural network analysis of signals,
- Longitudinal guided waves [3], and
- Laser excitation of ultrasonic systems that allow for non-contacting probes [145].

### 5.5 Friction Management

Section 4.2 notes the considerable effect that friction has on the rates of crack initiation and propagation. In addition, friction affects lateral wheelset loads and steering forces, thereby affecting the attitude of the wheelset to the track. Furthermore, there are safety implications associated with friction, with insufficient friction leading to sliding wheels [150], wheel burns on rails, long braking distances and stalled trains [253] and excessive friction contributing to hunting [163], wheel climb [166], and rail roll-over derailments [212].

The friction between clean steel on clean steel is roughly 0.3 [105, p. 29]. However, the wheel-rail contact patch is anything but clean, having various films and deposits on both surfaces. If friction is not actively managed but instead is left to the whims of “Mother Nature,” the coefficient of friction can range tremendously, from a low of near zero to a high of approximately 0.7 (Figure 54).

Although tractions can be managed by controlling creepage through better profiles and flexible or steering trucks, modifying the properties of the interfacial layer has proven to be a very effective approach to controlling wheel-rail forces and the damage that arises there.

Friction management is the process of controlling the friction levels to values most appropriate for the operating demands (Figure 55).

- At the contact between the rail gauge face and the wheel flange, there are typically very high levels of slip, which in the absence of a protective film can lead to catastrophic rates of wear. Low-cost lubricating films are generally applied.
- At the top of the rail it is necessary to ensure sufficient friction for pulling and braking of trains. This is achieved using a friction modifier.
Figure 54. The Coefficient of Friction at the Unmanaged Wheel-Rail Interface Can Vary Widely, from 0.05 to 0.7

Figure 55. Friction Management Involves Two Different Interfaces (top-of-rail/tread-of-wheel and gage face/wheel flange) with Different Requirements for Each
5.5.1 Lubrication
Lubrication of the rail gauge face/wheel flange can reduce gauge face/flange wear by 95–100 percent [253]. Lubrication can also contribute to reductions in RCF initiation by reducing tractions. However, lubrication of the gauge face decreases the steering moment that can be developed by a wheelset and increases the leading-axle angle-of-attack in curves [114]. The consequential increase in lateral creepage and lateral force increases top of rail wear on the inside rail, RCF initiation on both high- and low-rail top surfaces, and the formation of deep-seated shells on the high-rail gauge corner. Generally speaking, grease or oil contamination of existing surface cracks further increases RCF propagation [74], although on clean rail the reduced tractions delay crack formation [196]. Poor equipment or improper settings on lubricators may also cause grease to migrate to the TOR, thereby compromising traction and braking. Gauge-face lubrication, by preventing wear of the gauge corner, combined with significantly longer grinding cycles has resulted in deep gauge-corner shelling on premium rails [252]. Thus, gauge-face lubrication that matches the grease application rate to type of grease, dispensing equipment and track conditions must be used in conjunction with a preventive rail-grinding program that regularly removes metal from the rail gauge corner.

5.5.2 Top-of-Rail/Wheel-Tread Friction Control
Although the goal of lubrication is to minimize friction, friction control refers to a process of manipulating the characteristics of the interfacial layer to provide an intermediate friction coefficient. At a value of 0.3–0.35, friction modifiers ensure sufficient friction for braking and pulling trains but not so high as to promote RCF and increase the risk of safety problems. The benefits of top-of-rail friction control have been demonstrated for noise [53], corrugation suppression [55], lateral force, and rail wear reduction [54, 125, 257]. With respect to RCF, recent work by Stock et al. [257] using a full-scale test rig have shown that friction management dramatically reduces the rates of RCF formation for both standard 260 HB and head-hardened 350 HB European steels (see Figure 56).
There are a number of approaches for managing friction that include solid stick (Figure 57 left) and liquid (Figure 57 right and Figure 58) products.
Friction modifier sticks apply products directly to the running surface of the wheel.

Wayside systems dispense a product to the running surface of the rail which is picked up by the wheel and dispersed along the rail length.

Figure 57. Solid Stick and Fluid Friction Modifier Products Developed by Kelsan Technologies for Control of Wheel-Rail Friction

A jetting device (essentially a high tech sander) has been developed in Japan for ejecting a granular friction modifier product into the wheel-rail contact patch [111]. At the time of that publication (2008) the system was still in the experimental stage. Subjects for further exploration include quantifying the effect of TOR-FM on rail grinding and wheel RCF.

5.5.3 Improved Performance Trucks/Bogies

Section 4.4 reviewed the experience of several freight railroads that adopted premium trucks. In all published cases a properly aligned flexible truck has been found to reduce the rates of rolling contact fatigue. Although replacement of trucks is not usually a viable approach to treating an
RCF problem, component upgrading (e.g., center-plate/center-bowl liners, constant contact sidebearers, shear pads between wheel bearing and pedestal roof and frame bracing) are options and for new purchases, advanced trucks are being increasingly considered.

In North America, the M-976 criterion is designed to reduce rail “stress” from 286K freight cars by limiting rolling resistance in curves. At the same time, the AAR stability criterion is twice as tight as the regular Chapter 11 limit, halving the allowable lateral acceleration values. To achieve these values, two main modifications were required. The first is the use of an elastomer pad between side frame and axle adapter to reduce bending stiffness to decrease curving resistance. The second change is to considerably increase warping stiffness, usually through special wedge designs. Along with the mandatory application of constant contact sidebearings, the wedge design is key to achieving the required stability.

The impact of improved trucks on rolling contact fatigue of rails has not been documented. Since the reduction in wheel fatigue has been well established, it seems logical that there should be a parallel reduction in stresses that promote rail fatigue. However, this is a difficult thing to establish in practice. While improved trucks might be progressively adopted, over the duration of any feasible test there are often simultaneous changes in friction management, rail grinding, metallurgy, and operational practices that make any RCF reduction from improved trucks virtually impossible to assess.

5.6 Rail Grinding

Rail grinding is a well established process for reducing the rate of RCF development and for removing damaged (e.g., cracked) surface metal. The three primary functions of rail grinding are to:

a) Remove already cracked and damaged surface metal. There are three reasons for this:
   I. To remove cracks before they propagate deep into the head and cause more damage (e.g., shelling and broken rails)
   II. To improve the capability of ultrasonic testing (signals cannot pass through cracks)
   III. Cracked material has poor strength and permits continued deterioration of the profile
   In the preventive mode, the objective is to remove a thin layer of heavily deformed surface material that may not yet have cracked but is likely to do so in the very near future.

b) Reprofile the rail back to a desired transverse profile. This (ideally optimal) shape is required to control and manage the performance of the wheel-rail contact (e.g., contact stress, steering, and stability).

c) Correct (minor) surface defects such as minor squats and wheel burns, rail corrugation, and dipped welds.

References 20, 231, and 252 give a good overview of current practices, state of the art, and ongoing developments.

The best practices related to grinding intervals and amount of metal to remove at each cycle are continually evolving. Although there is a consensus that grinding does extend rail life [24, 227], the tradeoff between the cost of frequent grinding and the increase in rail life is often debated.

The current best practice in the North American heavy haul railway has evolved over a period of about three decades. Rail grinding was introduced in the 1960s when heavier axle loads on soft steels lead to rapid formation of low rail corrugation that reduced rail life to as little as 50 MGT on North American railroads. Annual rail grinding with heavy multiple passes was introduced with the hope of adding an extra year of life to the rail. Yet even with this heavy grinding,
corrugations were not being controlled. So some railroads started grinding twice a year [247]. But even so, it became apparent that the effects of corrugation were never fully removed and that deep cracks had developed in the troughs of the corrugation. Although the impact of grinding was positive, it was determined that this corrective approach (i.e., grinding after the defect became severe) could be improved upon by further doubling the grinding frequency. This evolved into the concept that the cycle should be based on tonnage rather than time. Once the philosophy of regular, tonnage-based grinding of asymmetric profiles developed, practitioners started thinking of the process as being “preventive.” Regular grinding applied transverse profiles onto the rail head at a defined cycle, and corrugations were gradually brought under control.

On commuter railroads and metros, the developments in rail grinding have been somewhat haphazard. In some railroads, there has never been any grinding done at all, either through ignorance or because it was simply impossible to fit a rail grinder into the confined spaces. Other impediments have been the unavailability of suitable grinders, concerns about dust and fire, and low annual tonnages (meaning on some lines that preventive grinding could be a once a decade requirement). The historical practice has been that an out-of-face rail grinding program is undertaken only once rail conditions have deteriorated so badly, and the need vocally championed such that the only option has been grinding or replacement. Within the past couple decades, the benefits of rail grinding have been more widely spread and appreciated, with some agencies now either buying their own grinder or else contracting out the service for regular rail grinding programs.

In the preventive mode, rail grinding is a process of controlled artificial wear and through fine-tuning can be applied to restore the desired profiles and achieve the required depth of metal removal with minimal grinding effort and steel wastage. “Fine-tuning” means both determining and applying the “Magic Wear Rate” – that is, the combined amount of natural and artificial wear required to just remove the existing and incipient cracks that are contained within a thin skin of metal at the surface (see Section 4.7). CPR’s Magic Wear Rate in sharp curves has over time evolved to be about 0.001 inches (0.025 mm) per MGT of traffic, providing them approximately 750 MGT wear life based on 0.75 inches (19 mm) of allowable wear. At a 25-MGT grinding interval, this translates into approximately 0.6 mm of vertical wear each cycle, of which about 0.2–0.3 mm is typically removed during grinding—the rest being natural wear. After each preventive grinding cycle, a surface free of cracks with healthy, work-hardened material remains. This contrasts sharply with the corrective grinding practice in which heavy, multiple-pass grinding of the deformed and cracked surface eradicates the work-hardened layer—leaving behind soft steel with some residual cracks (the deepest cracks would not be removed) that is substantially more vulnerable to plastic deformation and RCF. Regular grinding also pushes the depth of maximum shear stress progressively deeper into the rail head, preventing that stress from dwelling at any particular spot for an extended period. This is especially important for steels that have more numerous inclusions, although it is relevant for even the cleanest steels.

5.6.1 Beyond Preventive Grinding

A hybrid grinding procedure was pioneered on the Burlington Northern Santa Fe Railroad (BNSF) [254]. BNSF, faced with a corrective grinding practice, poor surface condition, and increasing defect counts, was seeking a method to get back to the preventive grinding practice from which they would regressed several years earlier. They already leased three 88-stone rail
grinders to service their 34,000-mile system and were not prepared to increase their grinding budget to get to Preventive Grinding. Installing improved rail profiles while simultaneously maximizing metal removal on each grinding pass enabled an 88-stone production rail grinder to “catch-up” on surface fatigued rail using 15-MGT intervals. Two more class 1s have initiated similar preventive-gradual programs in the past 3 years as the North American rail industry continues to move toward preventive grinding. CPR moved one step further by paying close attention to the installation of low-stress profiles and tuning patterns and speeds to maximize grinder efficiency, and CPR successfully extended their previous 18-MGT intervals to 25 MGT on their timber-sleepered track while retaining the single pass grinding strategy. This step removed a full grinding cycle from the program, saving USD440,000 annually in rail grinding costs without compromising rail life. The rail surface condition at 25-MGT intervals and improved profiles was judged by the CPR Supervisor of Rail Grinding and Testing to be in even better condition than under the prior regime. It is worth noting that extending the cycle further to 37 MGT (a grinding cycle was missed on a high-curvature subdivision because of production gangs working in the area) resulted in gauge-corner spalling of the premium rail in most sharp curves. This suggests that 25 MGT may be an upper limit for the rail grinding interval with premium steel (e.g., head-hardened, microalloyed rail with hardness above 360 Brinell) and optimized rail profiles. This contrasts with the opinion of other experts who, based on their different experiences, suggest that the rail-grinding interval should be no less than 25 MGT [e.g., 228]. The wide range of conditions and the particular economics of a given railroad must be considered when developing a rail grinding interval designed to maximize rail life and assure effective use of the grinding budget.

Although field trials do continue in Europe and North America to determine the optimal grinding procedure and metal removal rates, incredible advances in computing power realized over the past decade have made more substantial modeling efforts possible. Models of the vehicle-track interaction that use full multibody dynamics simulations are being combined with finite element simulations of crack initiation and propagation. Besides being useful for quantifying the impact of vehicle and track dynamics on rail deterioration, a key goal has been to analytically determine the optimal wear rate by simultaneously calculating the rates of wear and crack growth to find a level where they balance.

A successful integrated model must be based on accurate representations of the relationships between a daunting number of parameters. Current vehicle dynamics models are quite mature, but practical crack initiation and propagation models are relatively primitive. Evans and Iwnicki [67] have reported on the extensive use of various vehicle dynamics models to analyze the effects of many vehicle, track, and wheel-rail parameters on contact stresses and the exceedence of plastic shakedown limits for predominant rail steels. They conclude that rails are characteristically overstressed by various amounts, leading to an unspecified role of wear (or artificial wear through rail grinding) in mitigating fatigue. However, a notable development in crack growth prediction models is presented by Kapoor and his colleagues [139] from Sheffield University (see Figure 59). They suggest that crack initiation occurs early in life through ratcheting (phase I), and then, those cracks propagate at an increasing rate until they extend through the depth of maximum shear stress (first half of phase II). The crack extends at a relatively slow rate thereafter as it grows outside the region influenced by rolling contact stresses (unless lubrication or moisture is present) – second half of phase II. Further growth is dependent on the pattern of combined bending and residual stresses; these same stresses control whether the cracks drives up to the surface or down into the rail head (phase III).
The Sheffield model schematically captures the stages of crack development discussed at the beginning of this section, but with greater clarity and detail than had been available in previous publications. Its practical implementation will depend on realistic quantification of the initiation and crack growth rates for a multitude of influencing factors. Friction conditions, wheel loads, suspension characteristics, cant deficiency, metallurgy (including cleanliness, hardness, and residual stresses), bending and thermal stresses, presence of surface films and wheel/rail profiles are only some of the factors at play. The impacts of numerous operational and geometrical parameters are being addressed by coupling the crack model with a vehicle-track model of the wheel-rail interaction. The Whole Life Rail Model (WLRM) [293] is one program (another is found in Reference 116) to build a comprehensive wheel-rail interaction model for the prediction of damage and risk associated with surface fatigue. Other properties have been included through the use of empirical relationships, at the moment restricted to a single metallurgy. The crack model has been “calibrated” against an extensive rail condition database assembled in the United Kingdom.

Among the many outputs of WLRM is a prediction of the amount of metal that must be removed through rail grinding to maintain the rail surface condition at its current state. In several examples shown in the report, the wear rates to maintain cracked rail in curves are approximately 0.01–0.015 mm/MGT. CPR’s experience with premium steels on well maintained rail with minimal or no cracking has arrived at a Magic Wear Rate of approximately 0.008, 0.016, to 0.025 mm per MGT on tangent track, mild curves, and sharp curves, respectively. It is encouraging then that the WLRM numbers appear to be in a similar range, although perhaps a little low compared with North American experience.

5.6.2 The Benefits of Rail Grinding

In combination with improved metallurgies, preventive grinding has contributed to a two-fold increase in system rail life and a four-fold increase in system rail fatigue life over the past 25 years [252].
Strong examples of the benefits of rail grinding in North America come from BNSF [254] and CPR [219]. The documented benefits of preventive grinding were found to include much better surface conditions, reduced total rail wear rates, improved grinder productivity, reduced grinding costs, improved rail inspection, and reduced rail defects. At the Canadian Pacific, the better surface condition reduces risk of rail failure and enabled them to extend their rail wear limits by nearly 5 mm to contribute to an 80 percent increase in wear life and quadrupling of fatigue life.

Preventive grinding of rail to remove 0.08 mm every 40 MGT effectively eliminates rail squats as a problem for Japanese railways [112].

5.6.3 Quality Assurance

Rail grinding offers the opportunity to apply custom rail profiles and metal removal rates that match the requirements of a specific section of track. Thus, the shape on the high rail can be different than that of the low rail, and the depth of material cut can be different. Once the forward speed of the train is controlled, the pressures applied to the grinding motors, the angles of the motors, the type of stone and the number of passes, and the rail grinder offer great flexibility for managing rail [134]. However, it is very easy to improperly apply grinding and either exacerbate the problems that one was trying to avoid or simply fail to achieve anywhere near the benefits possible through rail grinding. Excessive grinding intervals, improper template selection, poor grinding pattern design and implementation, insufficient machine maintenance/setup, bad recordkeeping, and failure to remove sufficient metal all limit the effectiveness of rail grinding.

The old axiom that “you can’t manage what you cannot measure” is as valid for RCF as it is for wear or any other damage issue. From the perspective of rail grinding, the key parameters of concern are:

- Accuracy of the installed rail profiles. Although contact mechanics suggests that these should be within ±0.004” (±0.1mm), in practice tolerances of ±0.010 inch (±0.25 mm) are typically applied.
- Effective elimination of surface defects (surface fatigue and corrugation). Visual inspection for surface fatigue (often with dye penetrant or magnetic particle enhancement) and surface roughness measurement for corrugation (e.g., using the corrugation analysis trolley) are current approaches.
- Appropriate facet width and angle between stones to eliminate stress raisers [173].
- Productivity and economics [e.g., 98, 252, 254].

In Europe, the acceptance criteria for rail grinding are governed by European Standard EN 13231-3.

In North America there is a Best Practice for Rail Grinding developed by AREMA Committee 4 – Rail Maintenance. AREMA Committee 2 - Rail measurement is taking on the task of determining how to assess profile quality using rail profile measurements. Unfortunately for the rail industry there is still missing a methodology for automatically assessing the surface condition of the rail. Vision systems (with the appropriate interpretive algorithms) or crack measurement technologies (e.g., eddy current) are needed to fulfill this role.

5.7 Friction Transformation Processing

A study completed in 2006 for the U.K. Rail Safety and Standards Board (RSSB) investigated whether it is possible to heal existing surface fatigue cracks through the application of
deformation and heat [220]. In this process, a rapidly rotating steel disc spins against the rail surface, heating it beyond approximately 1,200°C as it moves slowly forward. In laboratory tests with laboratory manufactured cracks, the technique proved capable of sealing the crack tips, but because of rapid cooling, the steel transformed into brittle martensite. The need to heat treat the rail for approximately 30 s after processing limits the speeds to the order of 0.3 km/h.

Figure 60. The Friction Transformation Process “Heals” Cracks in a Laboratory Test
6. **Recommendations**

**Modeling**

1. Current predictive models rely on quasi-static measurements of fundamental properties (such as fracture toughness and yield strength) to arrive at a shear strength value for use in models. However, simple calculations show that the short duration of the wheel-rail contact involves an impact scenario. Further characterization and performance assessment of rail steels is required. This includes the development of repeatable testing methodologies that mimic the true state of stress and the short loading duration of the wheel-rail system.

2. Numerical models have vastly improved in the past decade, but many simplifications are still made in the input variables.
   - Further field measurement and characterization of friction characteristics and their implementation in VTI models is required.
   - Distributions of vehicle characteristics, wheel-rail profiles, and material properties to better mimic the actual conditions should become common in predictions of safety and damage.

3. Although the effects of grease and water on crack prediction are generally understood, their profound effect on crack initiation and propagation remains to be adequately quantified.

**Monitoring and Inspection**

4. Explore the potential of near-surface noise signatures from current ultrasonic inspection systems for assessing the severity of rail surface deterioration.

5. Vision systems offer the potential to identify surface cracks and support a fatigue management system
   - As high-speed imaging systems have steadily improved, such as for joint bar inspection, they should next be extended to looking at the condition of the running surface.
   - The MerMec THIS is possibly the leading commercial system available utilizing machine vision to visually identify surface cracking. It may be worth asking the supplier to bring this tool to North America for evaluation.
   - Appropriate interpretive algorithms should be developed to exploit the surface characterization information that seems inevitable. Correlating surface distress with subsurface damage and providing practical maintenance actions are among the requirements.

6. Eddy current crack measurement systems should continue to be developed for railway applications. Validations for North American steels and correlations of crack length to depth are required. Their reliability and practicality as monitoring tools remain to be understood.

7. Accelerometers on railway cars, especially near the axle box, are able to detect cases of deep RCF surface defects but must be coupled with an appropriate reactive program. The instrumented wheelset, by measuring forces and (with limited reliability) the position of the contact point, can provide the detailed information needed for a preventive/predictive maintenance program.
8. Improved NDT tools, statistical sampling and analysis approaches are required to better detect small, nonmetallic inclusions that are believed to be the cause of some wheel failures such as shattered and broken rims.

Management

9. As high-strength materials continue to be developed, modeling and performance testing must also develop for quantifying the potential economic benefits. In combination with proactive monitoring and maintenance practices, significantly extended rail life is possible.

10. Work should continue toward developing rail steels that maximize cleanliness and minimize proeutectoid ferrite. Bainitic steels, especially lower bainite, appear to offer improved resistance to RCF. Cleaner and cleaner steels may continue to improve resistance to subsurface failure in both rails and wheels.

11. Although the solid contribution of wheel-rail profiles to minimizing RCF and techniques for their design are well understood, practical approaches for managing profiles are not fully developed. Management tools exist, but their analytical capabilities remain simplistic, and their penetration into the rail industry poor. Tolerances on wheel and rail profiles remain to be developed. Although various wheel-rail interaction indices suitable for evaluating risk and maintenance requirements have been developed, more can still be done to develop improved criteria and support their adoption by the industry.

12. Friction management – the ability of friction management to reduce RCF has been demonstrated recently and the theory is sound, but the practical applications and economic and safety cases need still to be convincingly demonstrated for a range of systems.

13. Rail grinding is a well-established method of removing surface defects and profiling rail. The optimal rail grinding interval and metal removal as a function of specific operating conditions (e.g., axle loads, metallurgy, track curvature, and rolling stock characteristics) is not known except by generalized, “best practice” principles. The integration of VTI, rail defect, and rail wear data into rail grinding decisions remains an area for development. The mechatronic rail grinder—one that measures profile and cracks at the lead end of the train, computes the required speed and patterns, and measures the results on the tail end—remains a theoretical and practical challenge.

14. The contribution of advanced (e.g., M976 compatible) trucks to reducing RCF of wheel and rail remains to be quantified. For higher speed trainsets, the trend toward higher yaw stiffness (which tends to increase rates of RCF) must be offset by careful adjustment of other suspension parameters.
7. References


23. CEN/TC256/WG4 Draft Standard Railway Applications — Track Heavy Rail, Part 1: Flat Bottom Symmetrical Railway Rails 46 kg/m and Above, 1997 (European Committee for Standardization).


