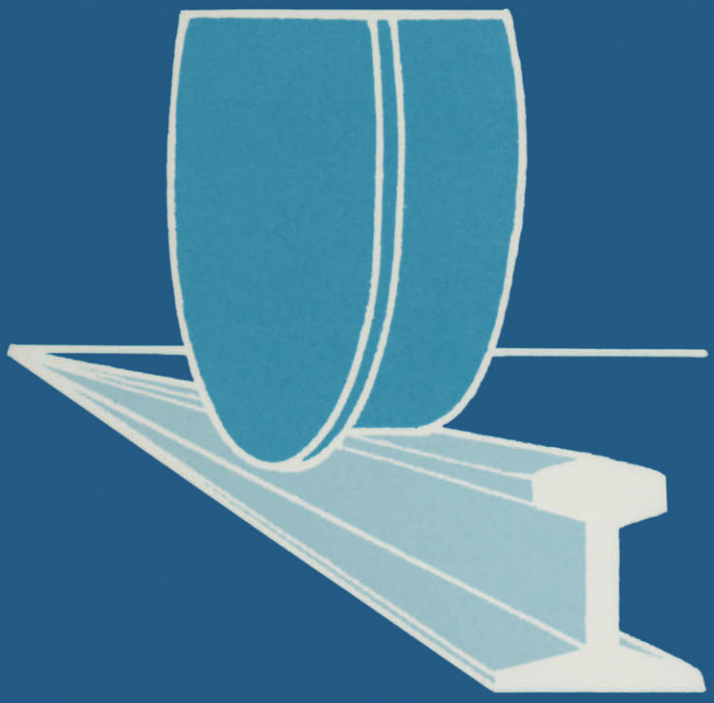
J.J. Kalker, D. F. Cannon and O. Orringer
editors

Rail Quality and Maintenance for Modern Railway Operation



SPRINGER-SCIENCE+BUSINESS MEDIA, B.V.

RAIL QUALITY AND MAINTENANCE FOR MODERN RAILWAY OPERATION

Rail Quality and Maintenance for Modern Railway Operation

International Conference on Rail Quality and Maintenance for Modern Railway Operation Delft June 1992

edited by

J.J. KALKER
T.U. Delft,
Delft, The Netherlands

D.F. CANNON
BR Research,
Derby, United Kingdom

and

O. ORRINGER
US Department of Transportation,
Cambridge, MA, U.S.A.

Springer-Science+Business Media, B.V.

Library of Congress Cataloging-in-Publication Data

Rail quality and maintenance for modern railway operation / edited by
J.J. Kalker and D.F. Cannon and O. Orringer.
p. cm.
Papers from the International Conference on Rail Quality and
Maintenance for Modern Railway Operations, held in June 1992 at the
Technical University, Delft.
1. Railroads--Rails--Congresses. I. Kalker, J. J. II. Cannon,
D. F. (David F.) III. Orringer, Oscar.
TF258.R325 1993
625.1'5--dc20 93-893

ISBN 978-90-481-4267-5 ISBN 978-94-015-8151-6 (eBook)
DOI 10.1007/978-94-015-8151-6

Printed on acid-free paper

TABLE OF CONTENTS

CONTACT PHENOMENA

RESIDUAL STRESSES

CRACK PROPAGATION

INTRODUCTION

In April 1990 a conference was held at the Cracow Institute of Technology, Cracow, Poland. The title of that conference was "Residual Stresses in Rails – Effects on Rail Integrity and Railroad Economics" and its themes were the measurement and prediction of residual stresses in rails, but, as the sub-title suggests, the intention was also to provide a link between research and its application to the practical railway world. At the Cracow conference there were 40 participants with 5 railways and 5 rail makers being represented and 25 papers were given. The Cracow conference was a success, and by March 1991 its off-spring, "The International Conference on Rail Quality and Maintenance for Modern Railway Operations", was conceived and birth was ultimately given in June 1992 at the Technical University, Delft. It turned out to be some baby, with 112 delegates from 24 countries taking part! As with its predecessor, the conference was to provide a forum for the exchange of ideas between research investigators, rail makers and railway engineers. A cursory examination of the list of participants suggests that about 57% were from the railway industry, 34% from universities and other research institutions and 9% from the steel industry. Bearing in mind that some of the railway industry participants were from their respective research and development organisations the balance of interests was about right.

The aim of the Delft conference was rail integrity, with the emphasis being on rolling contact fatigue, which is a failure mode increasingly recognised as a threat to the reliability of modern heavy freight and high speed railway systems. A measure of this recognition can be seen from the large number of conference participants from the railway industry and that substantial research programmes are being undertaken in this area in Europe and North America.

The conference began with two keynote papers representing European and North American views and then each session was introduced by an overview. The session subjects were Rail Maintenance, Rail Metallurgy and Processing, Contact Phenomena, Residual Stress, and Crack Propagation. 28 papers were presented within the sessions and the conference concluded with a general discussion which was led by a panel of railway engineers. The conference proceedings provide an indication of the development of our understanding of the problem and what rail steel technology and maintenance procedures have to offer to combat it.

Thanks are due to many for making this conference a success and they include; the United States Federal Railroad Administration, the European Railway Research Institute (formally ORE), Professor J J Kalker and Mrs van der Windt–Kröse and the authorities of the Technical University of Delft, Mrs Komen–Zimmerman and her staff at ASD, members of the Scientific and Local Conference Committees and our many sponsors who generously provided financial support. The most thanks must however go to the authors of the papers and the delegates who attended and participated in the conference.

David F Cannon
Member of the Conference Scientific Committee
Chairman, ERRI Committee D173 "Rail Rolling Contact Fatigue"

Derby, England
February 1993

The conference organisers wish to thank the following sponsors for their generous financial support:

British Rail Research

British Steel Track Products

Fatigue Technology Inc.

Loram Rail Ltd.

Métallurgique et Minière de Rodange-Athus

N.V. Nederlandse Spoorwegen

Speno International SA

Thyssen Stahl AG

Unimétal

US Department of Transportation

Van Welzenes Spoorbouw B.V.

RAIL DEVELOPMENT AND FUTURE REQUIREMENTS

INTERNATIONAL CONFERENCE ON RAIL QUALITY AND MAINTENANCE FOR MODERN RAILWAY OPERATION - DELFT JUNE 1992

FUTURE RAIL REQUIREMENTS

C O Frederick, British Rail Research, Derby, England

Introduction

There have been dramatic improvements in the quality of rail steels as a result of techniques such as the Basic Oxygen process, vacuum degassing and continuous casting leading to much better control over inclusions and hydrogen. Also there is now a much better understanding of how to achieve very fine pearlitic microstructures which will give high strength and wear resistance eg mill heat treatment. These developments and improved inspection at the manufacturer's works have largely defeated problems caused by internal defects in rails and now the emphasis must move towards solving problems which arise at the rail surface. This is shown by a levelling off in Tache Ovale failures and a rise in squat failures (see Fig 1). If material is not lost by wear at the surface, the material accumulates fatigue damage and surface crack initiation and propagation becomes more likely. This dilemma could now limit further improvements in wear resistance.

As rail quality improves and rail lives extend, the rate of removal of rails tends to diminish and the problems faced by railways are largely those of maintaining safety whilst using rails made by earlier processes. In this area, new more precise techniques using ultrasonics are being introduced which have raised testing speeds to 70 km/hr with on-board defect recognition in real time. This paper describes how this system can be used to measure rail depth as well as detecting defects and emphasises the difference between ensuring safety and reliability and managing the stock of rails to give maximum economic benefit. The paper also illustrates the high dynamic forces which can arise at high speeds if the rail surface contains small amplitude waves and describes briefly the current theory concerning the formation of waves by periodic wear. Finally, the potential benefits of using bainitic steels for crossings are described.

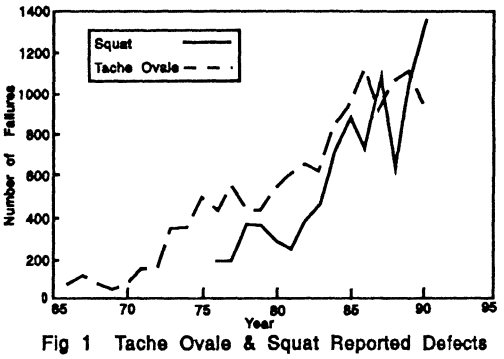


Fig 1 Tache Ovale & Squat Reported Defects

J.J. Kalker et al. (eds.), Rail Quality and Maintenance for Modern Railway Operation, 3–14.

Rail Metallurgy

Buying rails for a passenger railway is a long term investment. There can be few purchases where quality is so important. A large proportion of rail defects stem from some inherent feature of the rail; it may be hydrogen flakes, oxide inclusions or high residual stresses. Over the years, these problems have been recognised and rail manufacture improved. However, service conditions have become more severe so rail defects are still a problem; particularly since rails bought in the 1960s are still in use in the main lines. This can be illustrated by looking at the Tache Ovale problem on BR as shown by the rail failure statistics for rails made in different years (see Fig 2). In 1959 the bulk of BR was jointed track and Taches Ovales were not a serious problem. In the 1960's there was a campaign to convert the main lines to cwr and the rails installed at that time have subsequently seen much more intense traffic at higher speeds and axle loads. The result was a much more severe Tache Ovale problem. In 1974, BR started to introduce concast rails made from oxygen steel and later vacuum degassing was added to the process. As a result of these measures the Tache Ovale problem was expected to diminish. However, because of the long term nature of rail problems, only now is there an overall reduction in Tache Ovale failures. The rails that are bought today may be a source of concern in 30 years time so a detailed examination of their performance is important if we are to recognise the potential problems and prevent them arising. What are the future problems? A small number of Tache Ovales are being reported in post 1974 rail and a proportion of these have been sent to BR Research for examination. Out of a sample of eleven, three were found to be in rail made by the ingot route, three had initiated at the

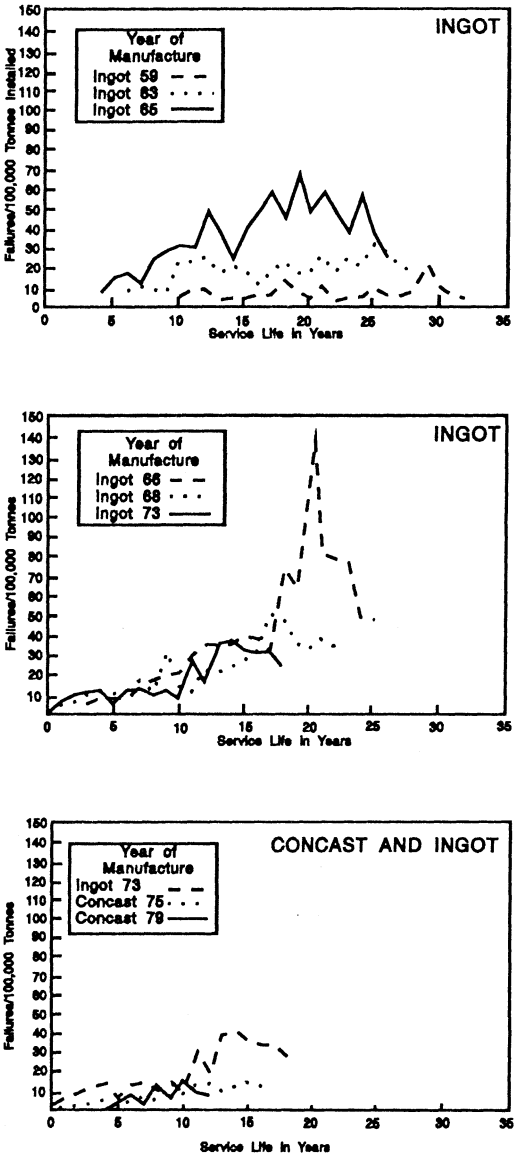


Fig 2 Reported Tache Ovale Defects for 100,000 Tons of Rail Installed

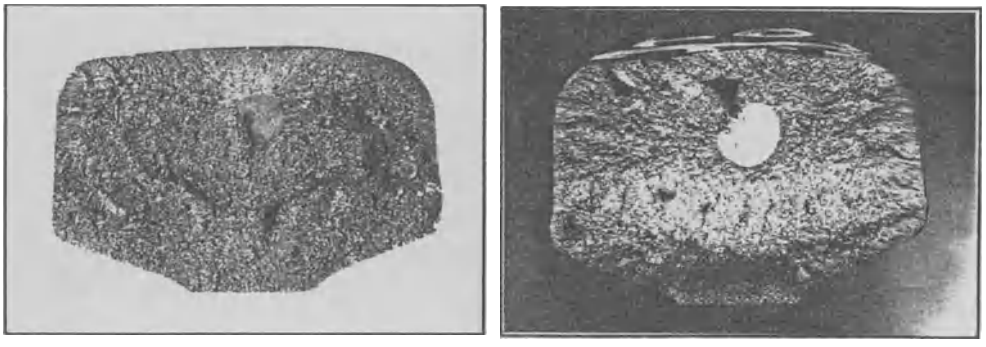


Fig 3 Tache Ovale Defects Due to Longitudinal Cavities

surface of the rail, one had initiated at a weld due to lack of fusion and four had initiated at longitudinal holes in the head of the rail (see Fig 3). Thus of the eleven, only four could be taken as representing a possible future Tache Ovale problem due to rail manufacture. It is possible that these are characteristic of the continuous casting process since the reduction of bloom to rail is typically only a factor of 10 whereas the reduction of an ingot is typically a factor of 30. As a result, the grain size of continuously cast rails is noticeably larger than ingot rails and fine longitudinal cavities are more likely. Nevertheless, it is clear that the delayed cooling and reheating of continuously cast blooms and the steel making process are producing very low hydrogen levels which should give a dramatic reduction in Tache Ovales.

The continuous casting process initially produced silicate stringers of up to 0.5 mm diameter but the process has been progressively improved and no service problems have as yet been caused by these oxides. This may be because BR specifies that oxygen reduction in the steel must be done using a silicon based process as opposed to an aluminium based process. Historically, BR's rails have been silicon killed and the rails have not suffered from shelling defects. It has been reported elsewhere (Ref 1) that alumino-silicates and oxides are particularly prone to cause shelling. This avoidance of the use of aluminium in steel making is now being perpetuated in the new Euro code.

Rail Surface Problems

Rail makers have introduced new processes which can generate very fine pearlitic structures. The result is a very high UTS and high wear resistance. Mill hardened rail steel produced by British Steel Corporation is a good example (Ref 2). Because the high strength is attained by heat treatment as opposed to alloy additions, the steel is weldable and because the process does not require reheating of the rails, it is not much more expensive than the process for normal grades of rail. Thus there is likely to be a progressive move towards the more widespread use of such rails. However, there are problems if they are used indiscriminately. These problems arise mainly at the rail surface.

When a relatively soft rail is used, the transverse profile of the rail rapidly becomes conformal with the wheel profiles and any bumps in the longitudinal surface profile tend to get flattened by the wheels. The result is a general lowering of contact stresses. If a hard rail is used there will be some lowering of contact stresses but it will be slight and the process will take much longer than for the soft rail. Furthermore, the rate of metal removal from the running table of a hard rail by wear and oxidation will be less than for a soft rail. The surface flow of the steel and the contact stresses are constantly producing small surface fissures lying at a small angle to the surface (see Fig 4). Sometimes these small fissures grow into cracks and form head checks (on the gauge corner) or squat defects (away from the gauge corner) (see Fig 5). It is clear that the principal factor which decides whether a small fissure becomes a crack is the balance between the rate of propagation and the rate of metal removal by wear and oxidation. A hard rail loses on both counts; contact stresses stay high and metal removal is low. Thus hard rails are more prone to surface cracking (see Fig 6).

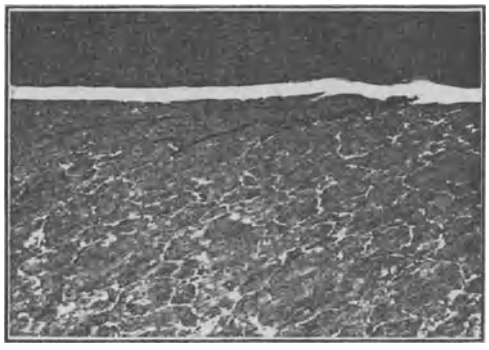


Fig 4 Small Surface Fissures

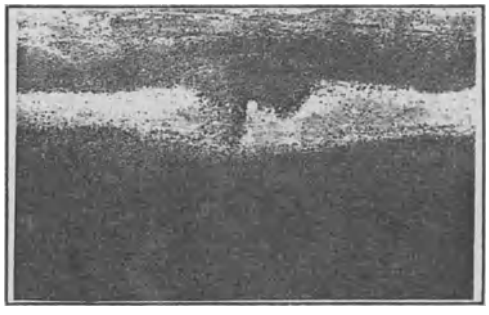


Fig 5 Developing Squat Defect

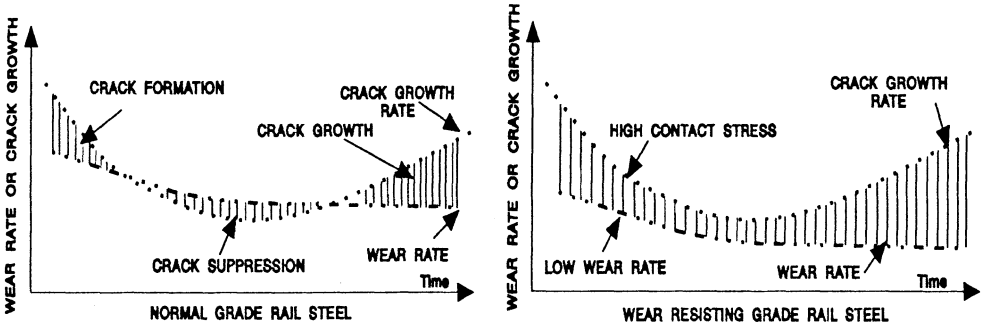


Fig 6 Schematic Illustrating Tendency to Grow in Wear Resisting Rail

It is possible to grind hard rails to special profiles so as to reduce contact stresses and this may solve the problem but there is still the low metal removal rate to consider. As speeds rise, contact stresses will rise due to dynamic effects. This can be illustrated by reference to some calculations which were done for the APT; even the residual roughness of a used rail

is sufficient to raise dynamic wheel loads by 30% (see Fig 7). These calculations were done for BS113A section rails; a heavier section such as UIC 60 will give a greater increase in dynamic loads and contact stresses. In general, dynamic wheel loads due to shortwave rail roughness will rise linearly with speed and superimposed on this are the effects of longer wave roughness. If the rail is ground, it may be argued that there is no shortwave rail roughness. However, the wheels themselves may not have perfectly smooth circumferential profiles, so an increase in dynamic contact stresses must be expected. Contact fatigue data indicates damage as being proportional to the cube of contact stress (Ref 3) so these increases in contact stress are potentially very serious. There are two factors which could make the problems certainties as opposed to possibilities; one is corrugation of the rails or wheels and the other is the presence, over long periods, of water or lubricating fluid on the running table of the rail. The former increases contact stresses and the latter increases the rate of crack propagation. In trials on the West Coast Main Line where rails were ground to remove corrugations, within two years the rails developed serious squat defects (Ref 4). This was because small cracks had developed due to the corrugations and when the corrugations were removed the cracks remained.

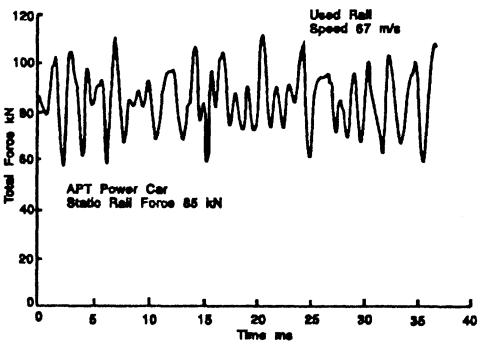


Fig 7 Wheel/Rail Forces Resulting from "Rail Roughness"

Since corrugations can play an important part in leading to rail fatigue failures, the choice of rail steel for high speed main lines should aim to reduce the corrugation problem. As will be described below, the dynamics of the track are important but the material properties of the rails also play a part. When a rail first enters service, the wheels tend to smooth the longitudinal rail profile by plastic deformation, then the corrugation process begins to work on the residual roughness and to cause wave deepening and propagation as a result of periodic wear. A soft rail steel will suffer extensive plastic deformation and profile smoothing, therefore delaying the onset of periodic wear but once the wear process is established, it can be rapid, Ref 4. A hard rail steel suffers little plastic deformation and smoothing so that periodic wear soon gets underway, however, the wear resistance slows the deepening process. The preferred situation is most likely to be the use of harder rail steels, but these must be ground soon after entering service to give a smooth longitudinal profile and a transverse profile designed for low contact stress. In fact, it has been shown that a low contact stress profile can also reduce side wear whilst still keeping conicity to a value which gives good vehicle ride stability (Ref 5).

The use of mill hardened rails poses a problem when it is necessary to weld the rails. Although it is possible to make sound welds, it is inevitable that some parts of the weld or heat affected zone will cool at a faster or slower rate than the original heat treatment. The result is a

variation in hardness across the weld (see Fig 8). In the case of thermit welds, the coarser grain size provides another source of metallurgical variation. These variations lead to variations in wear rate and, over a period of time, to rail surface roughness which can act as a starter for rail corrugations and also raise contact stresses. This problem applies to flash welds, thermit welds and repair welds and will probably prevent mill hardened rails being used for high speed lines until the necessary technology is developed to give uniform wear resistance. Until then it would seem that wear resisting grade A at 90 kg/mm^2 UTS is likely to be at the limit of hardness for high speed main lines.

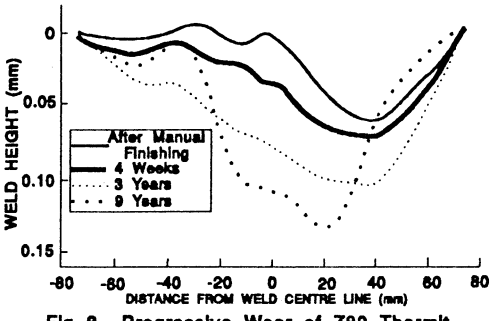


Fig 8 Progressive Wear of Z80 Thermit Weld in Grade A Rail

Rail Straightness

It has been known for a long time that irregular welded joints are a source of impact loads on high speed lines and can lead to a moulding of the ballast so that the geometry of track is shaped to suit the geometry of the weld (Ref 6). In the days of bolted joints, the rail specification said that the rail ends should not turn down, now with welded track it is perhaps more important that they don't turn up but tight limits are being applied to both in the most recent codes; the limit on upsweep being 0.4 mm over 2 metres and 0.3 mm over 1 metre. Furthermore very tight tolerances are being specified for general running surface flatness ie 0.3 mm over 3 metres and 0.2 mm over 1 metre (peak to trough). In fact, very small waves in the rail can produce large dynamic augments at high speed. This has been reported as occurring in high speed running when 1.7m waves in the rail profile gave a wave passing frequency which approached the resonant frequency of the unsprung mass of a wheelset vibrating on the track resilience ie a frequency of about 50Hz. These waves in the rail have been attributed to incorrect setting of the roller straightener (Ref 7) and could typically have a peak to peak amplitude up to 0.6mm. Most recent practice has reduced this to a value of 0.3mm with 0.1mm being achievable. However, the rails which are in service commonly exhibit these waves and they are difficult to remove by grinding. Similar problems can occur at lower speeds when running over older track. In 225 km/hr trials of the class 91 locomotive, wheel load fluctuations were observed at sleeper passing frequency in some places. These were eventually traced to small waves in the rail surface with a depth of only 0.05mm. It is likely that these waves were the result of a periodic wear process since they were most evident on old concrete sleepered track (see Fig 9).

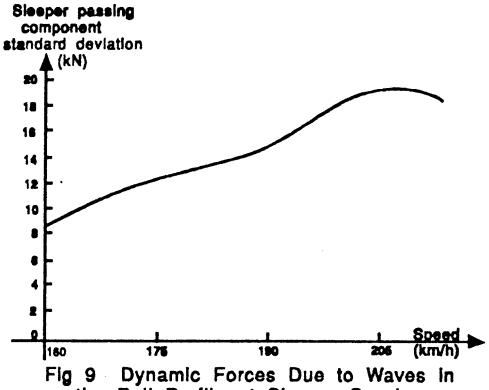


Fig 9 Dynamic Forces Due to Waves in the Rail Profile at Sleeper Spacing

Rail Management, Safety and Reliability

It is now possible to calculate the cost of a disruption to service and the costs of removing a defective rail so that it is also possible to calculate the number of defective rails per year which can be tolerated on a given route. The calculation balances the cost of renewing the rails against the cost of disruption caused by failures. This policy should not adversely affect safety since ultrasonic inspection of the track should consistently find rail defects before they cause a sudden rail failure. Furthermore since rail failures follow a Weibull distribution (Ref 8), it is possible to predict in advance the rate of rail failures on any given track if the present failure rate is known. Thus it is possible to review the rail failure rates on any part of the railway and plan ahead for the renewal of the rails. It is also possible to calculate the optimum time to cascade rails from main lines into secondary lines. This calculation will depend on the transfer value of the rails. A high transfer value will encourage the manager of the main line to renew rails earlier so that more serviceable rail becomes available to the manager of the secondary lines. It is possible, therefore, to calculate the optimum transfer value for the railway as a whole. BR Research have written software to do all these things, but for the process to be fully effective, it must be possible for rails to move from one part of the railway to another and the decisions as to which rails are to be renewed or cascaded must be taken by the managers of the profit centres. It is necessary therefore to set up a Rail Market so that managers can offer rails for cascading and see what deals are available.

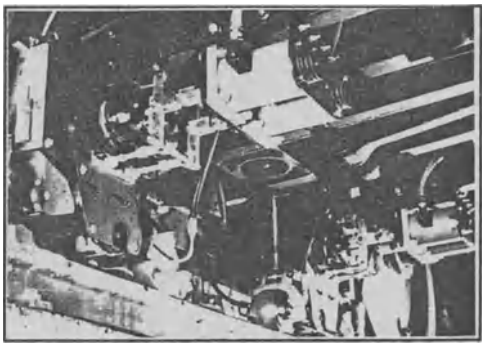


Fig 10 Wheel Probes on Ultrasonic Testing Unit

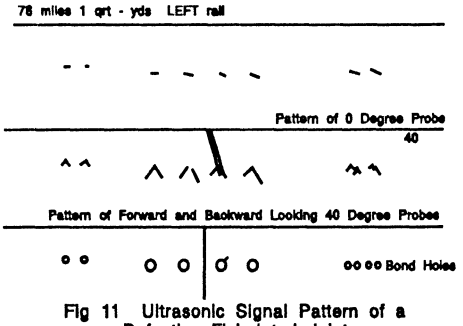


Fig 11 Ultrasonic Signal Pattern of a Defective Fishplated Joint

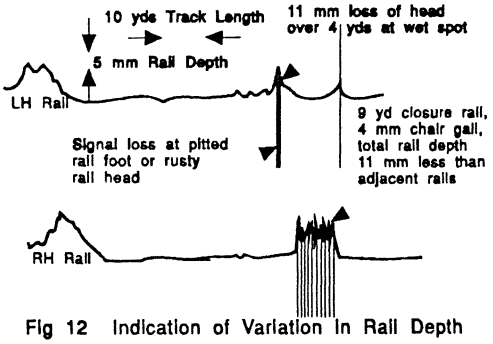


Fig 12 Indication of Variation in Rail Depth

To preserve safety and reliability it is essential for a railway to maintain an effective non destructive testing programme for its rails. BR uses both manual testing and vehicle based testing. The Ultrasonic Testing Unit has recently been brought into service on

BR (Ref 9). This Unit has been designed to operate at 70Km/hr on welded track. It may operate at slower speed on track with bolted joints so as to maintain probe contact if the track is very rough. However it uses resilient wheel probes which in general give very good contact with the rail (see Fig 10). Defects are identified and sized automatically as the vehicle runs and the position of the defect is marked on the track. There were some initial difficulties since the rail was found to contain many non standard features eg multiple drilled holes to fix bond wires and these could not be properly recognised by the automatic system. This problem has been overcome by outputting a picture of each defect indication so that the nature of the indication is clear and only the true defects need be reported (see Fig 11). This system can examine 600 miles of track in a week and is proving very reliable. It is finding bolt hole cracks down to 2mm long.

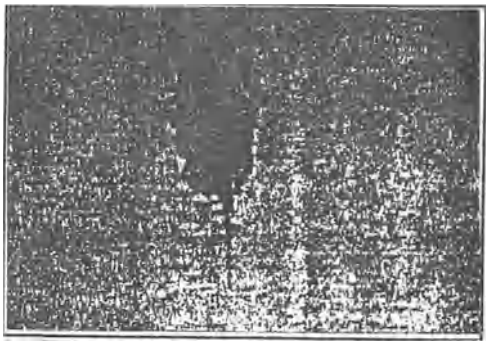


Fig 13a Fatigue Crack Starting from a Corrosion Pit on the Underside of the Rail Foot

In addition to finding defects, the UTU can measure rail depth and detect rail foot corrosion (see Fig 12). This is proving very useful for the examination of rail in tunnels. Sometimes a tunnel which is generally dry will have one or two locations where water drips from the roof. At these places, the rail will suffer a localised loss of head height and probably exhibit corrosion pitting of the rail foot. This combination is very dangerous if the line carries high speed traffic since the loss of head height causes a large dynamic force and bending moment at a place where the rail foot may be weakened by corrosion pitting. Corrosion pits can combine with tensile residual stresses introduced in manufacture (by the roller straightener) to give a marked reduction in the bending fatigue strength.

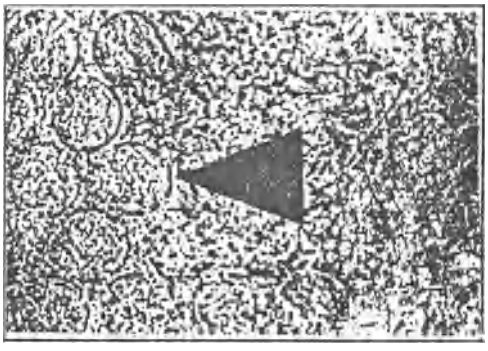


Fig 13b Corrosion on the Underside of the Rail Foot due to a Pimpled Rail Pad. (Transverse Corrosion Pit Arrowed)

The use of rail pads with transverse bars or pimples is particularly unfortunate since these can trap moisture and encourage corrosion to form a trough transverse to the axis of the rail (see Fig 13). Only a small fatigue crack is needed to cause sudden rail fracture and these conditions have been known to cause loss of a length of rail.

Neither the UTU nor routine manual ultrasonic testing provide a fully satisfactory means of testing thermit welds. This remains a problem to be solved, the availability of the UTU may however allow more emphasis to be given to the manual testing of thermit welds.

Crossing Steels

Cast 14% manganese steel is a very tough steel and cracks can become very long without causing a sudden fracture. Unfortunately, this steel is very prone to fatigue cracking. This is partly because it is exceedingly difficult to produce sound castings and partly because the steel has an inherently low fatigue strength. The situation is made worse by the opaqueness of cast manganese steel to ultrasound which means that radiography must be used to look for defects and this is too expensive for every casting to be examined. A further problem is that if it is necessary to weld the cast steel to pearlitic rails, special precautions must be used. To deal with these problems, BR Research working with Edgar Allen Engineering Co Ltd and the British Steel Corporation first developed a low carbon version of cast manganese steel which had the toughness of the original Hadfield composition but which could be thermit or flash welded to itself. This had the advantage that the centre casting of a crossing could be smaller and less likely to contain defects. Then a completely new cast steel was developed - a bainitic steel (Ref 10). This steel had high fatigue strength and adequate toughness. Furthermore, this steel can be cast easily, can be examined ultrasonically and can be thermit or flash welded to pearlitic rails without difficulty. Although a hardness of 330 HB could be achieved, it was decided that this was not enough and an increased hardness was obtained by adding copper to the composition giving a hardness of 400 HB. These steels benefit from very low sulphur levels and need very low hydrogen levels. To achieve these, the Argon-Oxygen Decarburisation steel making route has proved important. The strength of the steel increases with carbon level and there is a corresponding reduction in ductility; 0.12%C is the normal target level. Further increases in strength are obtained by adding up to 2% copper and precipitation hardening. These crossings have very attractive features and should compete against cast manganese on price. Their life should be long since they should be resistant to fatigue and can be weld repaired with only nominal preheat. Approximately 70 bainitic crossings are in service on BR; some have been in track for 10 years.

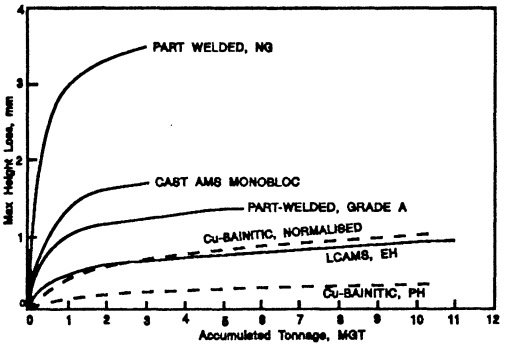


Fig 14 Deformation Performance of Crossing Types at Westhouses Test Site

Fig 14 shows loss in height of crossing noses with tonnage under consistent traffic conditions for different types of crossing steels and methods of construction ie part welded Vees in normal grade and grade A steel, cast AMS with and without explosive hardening and Cast Bainitic Steel as normalised and precipitation hardened.

Rail Corrugations

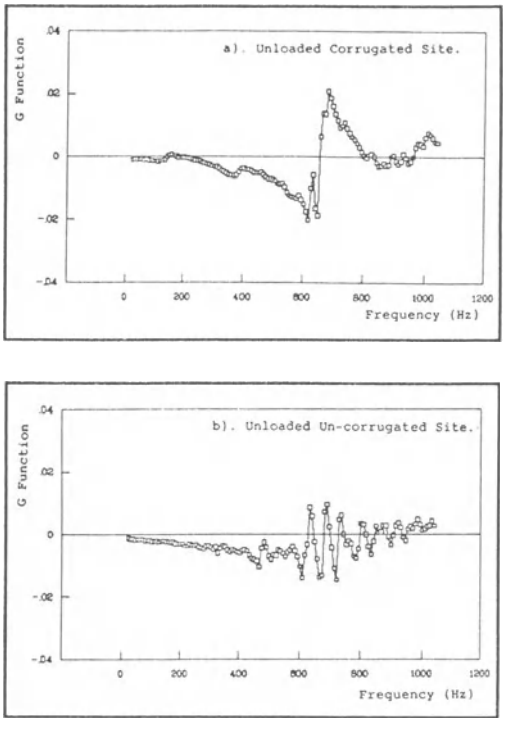


Fig 15 Some Examples of G(ω) Based on Measured Dynamic Response of the Track

It is the author's opinion that future railway tracks should be designed to minimise the problem of rail corrugations. Current theories to explain the formation of corrugations fall into two classes, large creepage and small creepage (Ref 12, 13). The former is most appropriate for heavily curved track and the latter for straight track. Thus the latter is likely to be most important for future high speed railways. In essence, the formation of corrugations is believed to be a periodic wear process which is initiated by rail roughness and wheel/rail creepage. Lateral creepage ie creepage caused by the wheelset having an angle of yaw, seems to be the most damaging form of creepage. The dynamics of the track are important since they control the relationships between rail roughness and fluctuations in vertical force, lateral force, creepage and wear. If the fluctuations in wear are such as to deepen an existing wave in the rail profile, then corrugations at that wavelength will deepen and propagate along the rail. Early in the formation process, the rate of deepening is proportional to the existing depth and thus the process is exponential

$$\text{ie Depth} = \text{Initial Depth } e^{CN,\gamma G(\omega)}$$

where C is a constant depending on Wear Resistance and contact patch parameters and where N = Number of Axles and γ = Creepage. The coefficient of this exponential relationship $G(\omega)$ is a function of the vertical and lateral track dynamics. $G(\omega)$ depends on frequency ie wave passing frequency (running speed/wavelength). A plot of $G(\omega)$ against ω thus shows the frequencies at which corrugations may be expected to deepen ($G(\omega)$ positive) and the frequencies at which they may be expected to diminish. ($G(\omega)$ negative).

The function $G(\omega)$ can be calculated from measured dynamic responses of the track. Dynamic measurements at corrugated and uncorrugated sites have led to values of $G(\omega)$ which fit the observed corrugations (see Fig 15) since a broad positive peak denotes growth which is not sensitive to modest speed variations and narrow peaks do not indicate growth since the wave passing frequency is sensitive to speed. The correlation has not been so good,

however, when the dynamics of the track are measured with the track loaded. Work is still progressing to try and understand the reasons for discrepancies and why the measured track dynamics vary from place to place. Nevertheless, considerable progress has been made.

Conclusions

Much has been achieved in improving rail manufacture but higher speeds and more intense traffic mean that rail problems will become more concerned with fatigue of the running surface. The trend towards using heavier section and harder steel for rails in plain line will need more use of rail grinding procedures to control surface damage and also improved welding procedures which do not give large hardness variations along the running surface.

Even very small waves in the rail surface can give large increases in wheel/rail force at very high speeds. Rail corrugations will have to be controlled by grinding on existing lines. On new lines, an effort should be made to minimise rail corrugation when designing the track, but a better understanding of the corrugation process is still required.

The renewal or cascading of rail should be planned well in advance using computer programs which set the cost of service disruptions against the cost of renewal. To gain long life from the rails, particularly in situations where corrosion is a problem, rails are needed with low tensile residual stresses in the foot - ideally a compressive stress is required!

Bainitic steels have much to offer for cast crossings and should provide an attractive alternative to the Hadfields high manganese composition.

Acknowledgement

The author would like to thank the Director, British Rail Research for permission to present and publish this paper.

References

Ref 1 "Laboratory Investigation of Transverse Defects in Rails"
Marich S, Cottam JW and Curcio P
Procedings of Heavy Haul Railways Conference, Perth, Western Australia, September 1978

Ref 2 "Rails for the '90s"
Hodgson WH and Preston RR
Presented at the International Symposium on Accelerated Cooling of Rolled Steel, Winnipeg, Manitoba, August 1987

Ref 3 "Development of a Failure Parameter for Rail Squats" Sinclair JC and Allery MBPA
British Rail Research, Technical Memorandum TM MF 211, March 1991

Ref 4 "Corrugation Research on British Rail"
Frederick CO and Bugden WG
Presented at the Symposium in Rail Corrugation Problems, Berlin, June 1983

Ref 5 "Service Experience with the LCS Rail Head Profile"
BRR Technical Report RR MF 21, March 1992

Ref 6 "The Effect of Rail Straightness on Track Maintenance"
Frederick CO
Presented to the Conference on Advanced Techniques in Permanent Way Design, Construction and Maintenance, Madrid 1981

Ref 7 "Summary of Conditions Concerning the Straightness of Rails"
ORE D148 Report, September 1983

Ref 8 "The Effect of Increasing Axle Loads on Rail Fatigue Life"
Zarembsski AM et al
Procedings of Rail Technology Seminar, Nottingham University, September 1981

Ref 9 "Infrastructure Monitoring Systems for British Rail"
Rail International, Schienen der Welt, June 1992

Ref 10 "Developments in Railway Crossings"
Jones EG
Procedings of Third International Heavy Haul Railway Conference, Vancouver, Canada 1986

Ref 11 "A Rail Corrugation Theory Which Allows fir Contact Patch Size"
Frederick CO and Sinclair JC
Presented to Conference on Rail Corrugations, Berlin, April 1991

Ref 12 "Short Wave Corrugations - An Explanation Based on Slip-Stick Vibrations"
Clark RA, Scott GA and Poole W
ASME Winter Meeting, Chicago, December 1988

RAIL DEVELOPMENTS AND REQUIREMENTS FOR HEAVY HAUL RAILWAYS

Daniel H. Stone
Executive Director, Chicago Technical Center
Association of American Railroads
Chicago, Illinois, USA

INTRODUCTION

The one constant of North American railway heavy haul service is the quest for ever increasing axle loads. During the 1960's, the maximum axle loads permitted in interchange increased from 23.9 tonnes (26.25 tons) to 29.9 tonnes (32.9 tons). This change resulted in an increase in rail wear and defect formation which, in turn, stimulated an increase in research and development activity to extend rail life in the heavier axle environment. Further the change to 29.9 tonne (32.9 ton) axle loads was made without a thorough examination of the increase in rail costs. To establish the costs associated with the operation of trains with heavy axle loads the Facility for Accelerated Service Testing, known as the FAST experiment, was begun on a 7.6 km (4.8 mi) track loop as a cooperative program between the Association of American Railroads and the Federal Railroad Administration at the Transportation Test Center near Pueblo, Colorado [1]. The FAST experiment and it successor, the Heavy Axle Load Program, have continued to quantify the system, including rail, costs encountered in operating trains with axle loads of 29.9 tonne (32.9 ton), and more recently, 35.8 tonne (39.375 ton) axle loads. Concurrently, major findings have been developed based on the operation of Australian iron mining railways which operate at axle loads up to 32.5 tonnes (35.7 tons)[2].

Axle loads of the above magnitudes require a rail fabricated from a high-strength steel with excellent resistance to wear, deformation, and fatigue. Additionally, the rail must be weldable by either flash-butt or alminothermit processes. Finally, practically all heavy-haul railways employ profile grinding and lubrication to minimize wear.

J.J. Kalker et al. (eds.), Rail Quality and Maintenance for Modern Railway Operation, 15–26.

This paper will review the effects of 29.9 tonne (32.9 ton) and greater axle loads on fatigue, deformation, and wear and the methods developed to provide acceptable rail serviceability under heavy axle loads.

FATIGUE AND DEFECT FORMATION

As axle loads increase, the rate of fatigue defect formation naturally increases. Figure 1 shows the Wiebull distribution of defects produced in rails subjected to various axle loads [3,4]. Increasing stress causes an exponential decrease in fatigue life. For example, at 200 million gross tons (mgt), the 25-percent increase in axle load from 23.9 tonnes to 29.9 tonnes increases the rate of defect occurrence by a factor of 2.5. Notice also that defect rates may be reduced by the use of a heavier rail section which effectively reduces bending stresses. Three methods have been employed to cope with the potential of increasing defect formation rates, 1) the use of higher strength rails, 2) the use of rails with superior microcleanliness, and 3) rail grinding to decrease the level of contact stress. In fact, it is common practice in North America to employ all of the above methods simultaneously.

During the last 20 years the American Railway Engineering Association's specification for standard rail steel has increased the requirement for minimum hardness from 243 Bhn to 285 Bhn. Additionally, a new intermediate strength grade has found wide usage with its requirement of a minimum hardness of 300 Bhn. Finally, the strongest and most durable rails are those that are either heat-treated or alloyed in order to produce rails with surface hardnesses of 350 to 370 Bhn.

Among the many alloy rail steels which have been developed are 1-percent Cr [5,6], Cr-V [5], Cr-Mo [7], and Cr-Cb-V [8]. While alloy rail steels have been developed which have strength and hardness levels comparable with heat-treated rails, they have not gained wide acceptance, presumably due to difficulties associated with weldability. Further, certain Cr - V alloy rail steels have shown a tendency to shatter under impact as shown in Figure 2 [9]. The reasons for this behavior are not well understood, but vanadium may have the effect of rendering the steel more strain rate sensitive as observed in Figure 3 [10].

Heat-treated rails are produced by one of three methods, head-hardening, full heat-treatment, or more recently, in-line hardening after rolling. While processes for head hardening and oil quenching were developed in the early 1960's, in line hardening was not practicable until the advent of digital control systems. Such control systems sense rail temperature,

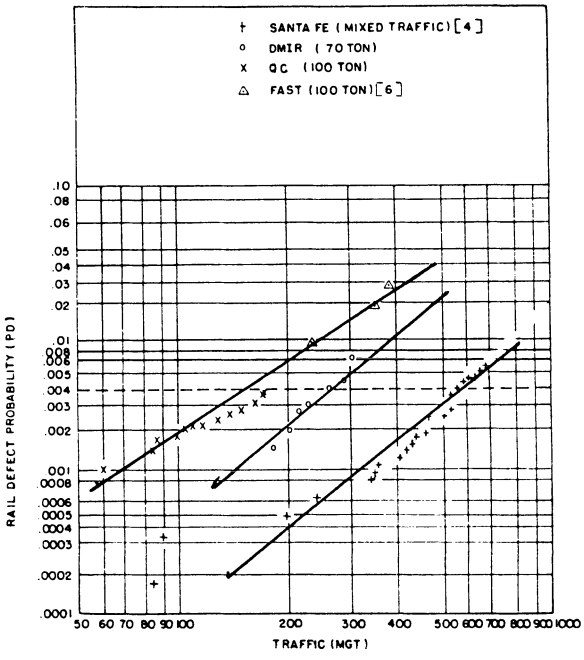


Figure 1. Wiebull distribution of defects in rail under various axle loads [3].

and either adjust the speed of the rail or the quenching spray intensity to ensure the development of an extremely uniform longitudinal hardness. While the strength and hardness of heat-treated and alloy steels are comparable, the heat-treated steels, which are reheated, have a finer grain size with an attendant increase in toughness [11].

Producing rails with improved cleanliness is a more recent development. The production of clean rail steels may have resulted more from the manufacturers converting from open-hearth steel poured into big-end-down open-top ingots to vacuum degassing as a requirement for the more economical continuous casting. Whatever the degree of motivation, today's rails are rolled from cleaner steel than they were 20 years ago. It was at that time that both Marich [12] and Sonon *et.al.* [13] demonstrated that complex oxides could act as fatigue crack initiation sites for transverse defects. Further, sulfide inclusions have been believed to be benign as crack initiation sites, and Heller [14] has shown that extremely low sulfur content can render rail steels susceptible to hydrogen flakes, Figure 4. Perhaps the most important effect of heavy axle loads is that, as axle load increases, the size of inclusion that can initiate a fatigue

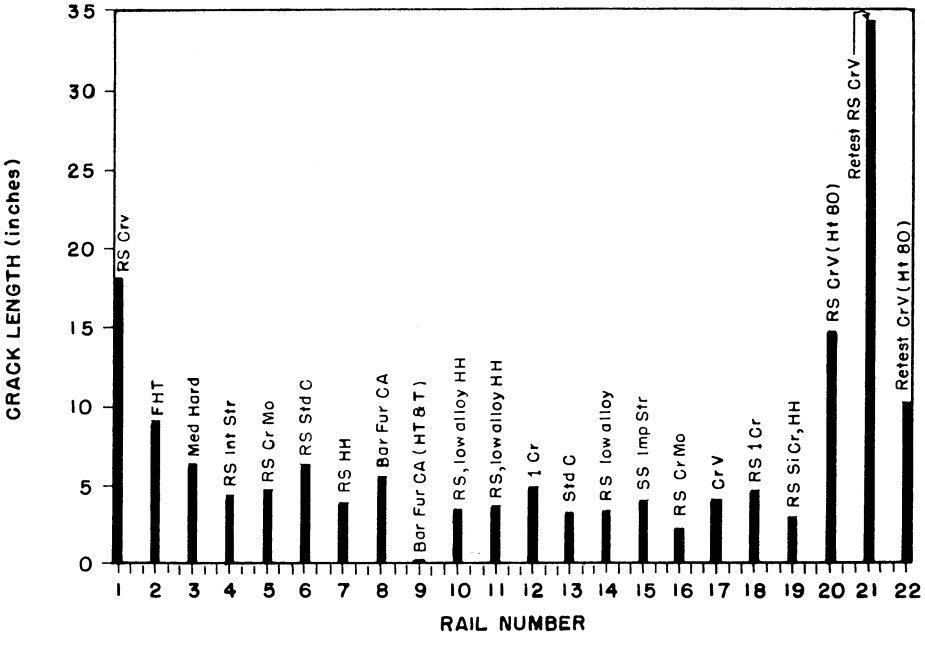


Figure 2. Web crack lengths in various rails subjected to drop tests. [9]

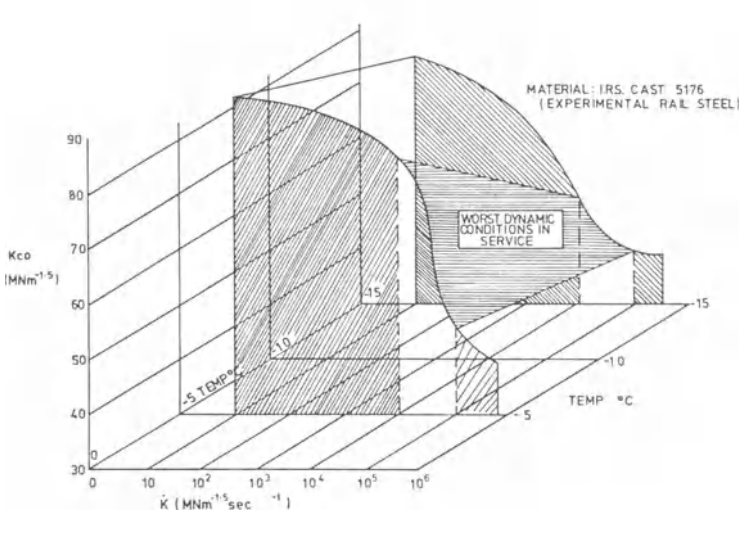


Figure 3. The effect of temperature and strain rate on the dynamic fracture toughness of a rail steel. [10]

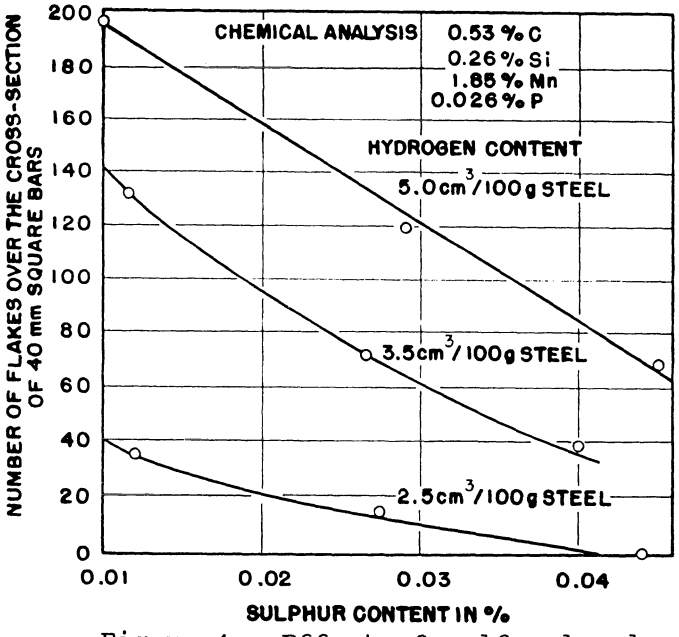


Figure 4. Effect of sulfur level on hydrogen content and flake occurrence. [14]

crack decreases, Figure 5 [2]. An analysis of rail defects from the FAST experiment have led Clayton and Brave [15] to conclude that the defect occurrence rate appears to be directly proportional to the logarithm of the volume fraction of oxide inclusions as well as a measure of the clustering of these inclusions divided by some function of the hardness. Sugino *et.al.* [16], in the results of a failure analysis of fourteen North American rails, have shown long alumina clusters to be important contributors to the formation of rail defects, while manganese sulfides play no role in the defect formation process.

PLASTIC DEFORMATION

The extremely high surface pressures in the contact area of the wheel cause the rail to deform and work-harden under the running surface when axle loads are above 200 kN (45,000 lbs.) as shown in Figure 6 [17]. While work-hardening allows the steel to cope with heavy axle loads and minimize wear, it develops a residual stress in the rail head, which can increase the rate of fatigue crack initiation and growth. It feeds material into the gage face where high wear rate mechanisms may be acting.

As would be expected, increasing the rail yield strength

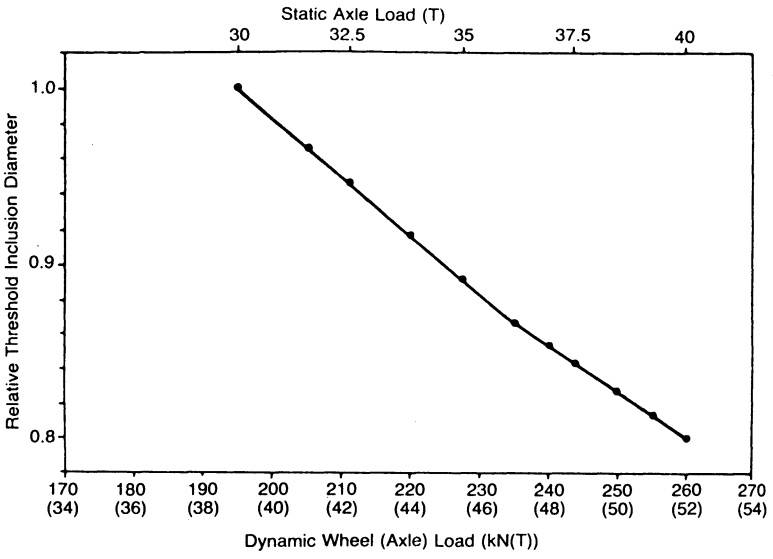


Figure 5. Influence of axle load on relative inclusion threshold diameter. [2]

minimizes the amount of plastic flow. In a series of cyclic compressive tests, the resistance to plastic deformation of standard and high-strength rail steels have been evaluated [18]. Figures 7 and 8 show that deformation resistance increases with increasing hardness and yield strength. The wavy pearlite microstructure developed in the work-hardened test specimens, Figure 9, is the same as seen in work-hardened rails.

The residual stresses, developed by plastic flow, are of primary consideration in the development of transverse defects as outlined in an extensive interpretive review by Steele [19, 20]. The effect of residual stresses can be easily seen in Figure 10, where residual stresses tend to operate in crack opening modes both for the longitudinal shell and the sometimes resulting transverse defect [21].

Long wave length corrugations are another manifestation of plastic flow in heavy haul service. While the mechanism of corrugation formation continues to elude researchers, there are a number of characteristics that have been observed. During the FAST experiments corrugations developed from rail joints of battered welds in the high rail of curves [22]. Zarembski *et.al.*[23], in a report covering a series of field measurements on North American railroads, produced the following four conclusions:

1. The overall distribution of corrugation

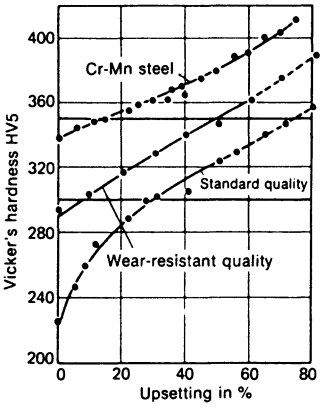


Figure 6. The work hardening behavior of rail steels [17]

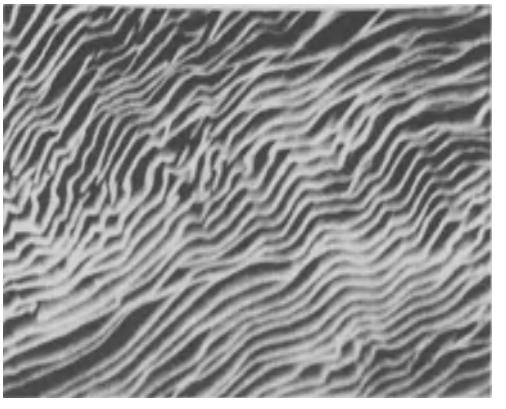


Figure 9. Microstructure of deformed rail steel. (etched in nital, 4900X)

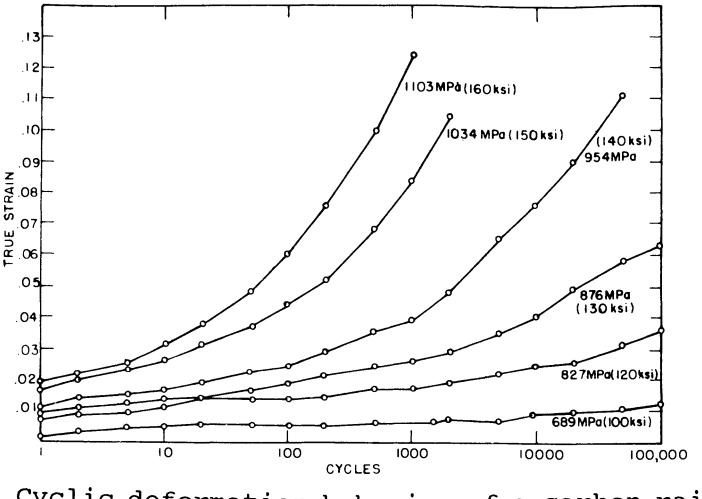


Figure 7. Cyclic deformation behavior of a carbon rail steel.

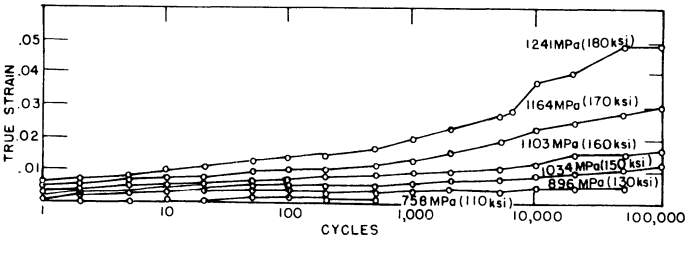


Figure 8. Cyclic deformation behavior of a heat-treated rail steel.

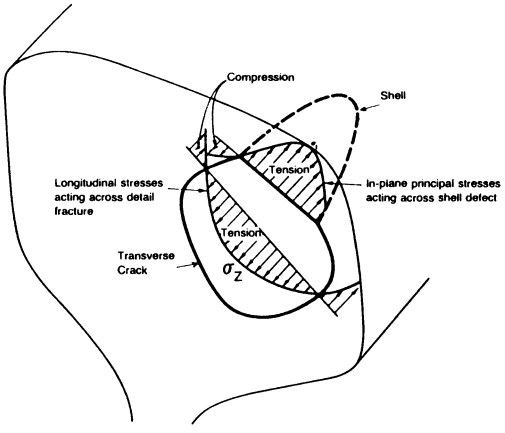


Figure 10. Diagrammatic distribution of residual stress in a rail head. [21]

wavelengths were similar in the five railroads studied. The greatest percentage of corrugations were in the 300- to 600-mm (12- to 24-inches) range.

2. Noticeable differences exist in the distribution of corrugations between wood and concrete tie track. The strongest corrugations in concrete occurred between 150- to 300-mm (6- to 12-inches).

3. In all cases, the width of the range of corrugation wave lengths ranged from 150- to 1200-mm (6- to 48-inches).

4. There appears to be a strong correlation between increasing wheel load and deeper corrugation depth.

As plastic flow is a necessary condition for corrugation formation, increasing the rail's yield strength can limit corrugation formation [24].

WEAR

Wear remains the major cause for rail removal. In the recent past, attempts to limit wear, especially in curves of 2° or greater, were made by employing high-strength rail steels. The use of these heat-treated or alloy steels at least tripled rail life. However, the strength of pearlite has reached its limit with heat-treated rail steels that have tensile strengths in the range of 1300 MPa (195 ksi). Additionally, Kalousek *et.al.* [25] have shown that sulfide inclusions, benign with respect to defect initiation, tend to increase the rate of adhesive gage face wear.

Initial FAST experiments dramatically demonstrated the effectiveness of lubrication in curtailing wear. Steele and

Reiff [26] showed that the wear rates of standard rail were reduced by a factor of 8, while the wear rates of high-strength rails were reduced by half. Such reductions have been seen in service with the added benefits of reduced energy consumption and reduced wheel wear. However, problems remain with insuring that the proper amount of lubricant is applied to the proper location on the rail.

Finally, the Australian heavy-haul railways developed grinding procedures that assisted proper curving, and substantially reduced gage face wear [27]. Many grinding profiles have subsequently been developed and applied in North America which have reduced rail wear. Unfortunately, none of the profiles has been designed with a total systems approach where the total wheel and rail wear costs are minimized without adversely affecting vehicle dynamic behavior.

CONCLUSIONS

1. Increasing axle loads have placed severe demands on the existing rail steels' ability to resist fatigue, wear, and plastic deformation.

2. The use of stronger steels with better microcleanliness, combined with profile grinding and lubrication, have helped North American Railroads to economically cope with increasing axle loads.

3. Further improvements in rail steel strength will have to be accomplished with steels of a different microstructure.

4. Further improvements in microstructure must not be made with sulfur levels too low as to promote the reappearance of hydrogen flakes.

5. Rail profile grinding procedures should not be developed without regard to wheel wear and vehicle dynamics.

REFERENCES

1. R. P. Reiff, "Introduction to the FAST/HAL Program", *Proceedings of the Workshop on Heavy Axle Loads*", (Pueblo, 1990) paper 5.

2. S. Marich and U. Maass, "Higher Axle Loads are Feasible - Economics and Technology Agree", *Third International Heavy Haul Railway Conference*, (Vancouver, 1986) pp. 18 - 31.

3. A. M. Zarembski, D. H. Stone, T. R. Wells and R. A. Armstrong, "Effect of Increasing Axle Loads on Rail

Fatigue Life", *Rail Technology*, C. O. Frederick and D. J. Round, editors, (Nottingham, 1981) pp. 151 - 161.

4. R. A. Armstrong, T. R. Wells, D. H. Stone, and A. M. Zarembski, "Impact of Car Loads on Rail Defect Occurrences", *Second International Heavy Haul Railway Conference*, (Colorado Springs, 1982) pp. 129 - 135.

5. W. Heller, E. Koerfer, and H. Schmedders, "Production of Special-Grade Naturally Hard Rails in Germany and Experience Gained in Operation", *Ibid.*, pp. 170 - 177.

6. W. H. Hodgson, J. K. Yates, and R. R. Preston, "The Development of a Second Generation of Alloy Steel Rails for Heavy Haul Applications", *Ibid.*, pp. 207 - 215.

7. Y. J. Park and F. B. Fletcher, "Fatigue Behavior and Fracture Toughness of Standard Carbon and High Strength Rail Steels", *Ibid.*, pp. 199 - 206.

8. S. Marich and P. Curcio, "Development of High-Strength Alloyed Rail Steels Suitable for Heavy Duty Applications", *Rail Steels - Developments Processing and Use, ASTM STP 644*, D. H. Stone and G. G. Knupp, editors, American Society for Testing and Materials (Philadelphia, 1978) pp. 167 - 211.

9. R. K. Steele and D. H. Stone, "Developments in Railroad Rail", *A.R.E.A Proc.*, 707, *87*, pp. 311 - 358.

10. K. Morton, D. F. Cannon, P. Clayton, and E. G. Jones, "The Assessment of Rail Steels", *Rail Steels - Developments Processing and Use, ASTM STP 644*, D. H. Stone and G. G. Knupp, editors, American Society for Testing and Materials (Philadelphia, 1978) pp. 80 - 98.

11. Y. J. Park and I. M. Bernstein, "Mechanism of Cleavage Fracture in Fully Pearlitic 1080 Rail Steel", *Ibid.*, pp 287 - 302.

12. S. Marich, P. Cottam, and P. Curcio, "Laboratory Investigation of Transverse Defects In Rails", *First International Heavy Haul Railway Conference*, (Perth, 1978) paper 303.

13. D. E. Sonon, J. V. Pellegrino, and J. M. Wandrisco, "A Metallurgical Examination of Control-Cooled Carbon-

Steel Rails with Service Developed Defects", *Rail Steels - Developments Processing and Use, ASTM STP 644*, D. H. Stone and G. G. Knupp, editors, American Society for Testing and Materials (Philadelphia, 1978) pp 99 - 117.

14. W. Heller, L. Weber, P. Hammerschmid, and L. Weber, "The Action of Hydrogen in Rail Steel and Possibilities of Low Hydrogen Melting", *Stahl und Eisen*, 1972, **92**, pp. 934 - 945.

15. P. Clayton and G. Brave, "FAST High Tonnage Loop Trials 1985 - 1987", To be Published.

16. K. Sugino, H. Kageyama, C. Urashima, and A. Kikuchi, "Metallurgical Improvement Of Rail for the Reduction of Rail-Wheel Contact Fatigue Failures", *Wear* 114 (1991) pp. 319 - 328.

17. W. Heller, "Fabrication, Properties, and Operating Behavior of Rail Steels", *Railroad Track, Theory and Practice*, F. Fastenrath, editor, Frederick Ungar Publishing (New York, 1977) pp. 112 - 138.

18. D. H. Stone, S. Marich, and C. M. Rimnac, "The Deformation Behavior of Rail Steels", *Transportation Research Record 744, Railroad Track and Facilities* (Washington, 1980) pp. 16 - 21.

19. R. K. Steele, "Recent North American Experience with Shelling in Railroad Rails - Part I Observations", *A.R.E.A Proc.*, 722, *90*, pp. 311 - 345.

20. R. K. Steele, "Recent North American Experience with Shelling in Railroad Rails - Part II Analytical Treatment", *A.R.E.A Proc.*, 723, 90, pp. 395 - 408.

21. R. C. Rice, R. Rungta, and D. Brock, "Post Service Rail Defect Analysis", Battelle Columbus Laboratory, June 26, 1983.

22. J. Hannafious, "FAST/HAL Rail Performance Test", *Workshop on Heavy Axle Loads*, AAR (Pueblo, 1990) paper 10.

23. A. M. Zarembski, G. Izbinski, S. N. Handal, and W. M. Worthington, "Corrugation Behavior in the Freight Railroad Environment", *A.R.E.A Proc.*, 712, *88*, pp. 303 - 332.

24. L. E. Daniiels and N. Blume, "Rail Corrugation growth Performance", *Second International Heavy Haul Railway Conference*, (Colorado Springs, 1982) pp. 294 - 318.

25. D. M. Fegrado, J. Kalousek, M. T. Shehata, and A. Palmer, "The Effect of Sulfide and Oxide Inclusions on the Wear Rates of Standard and Alloy Rail Steels", *Third International Heavy Haul Railway Conference*, (Vancouver, 1986) pp. 143 - 151.

26. R. K. Steele and R. P. Reiff, "Rail: Its Behavior and Relationship to Total System Wear", *Second International Heavy Haul Railway Conference*, (Colorado Springs, 1982) pp. 227 - 276.

27. S. T. Lamson and B. H. Longson, "Development of Rail Profile Grinding at Hamersley Iron", *ibid.*, pp. 372 - 379.

RAIL METALLURGY AND PROCESSING

RAIL METALLURGY AND PROCESSING

W H HODGSON
BRITISH STEEL TRACK PRODUCTS

ABSTRACT

Improved rail production techniques provide new options for the rail user. Steelmaking improvements offer tighter composition control, lower levels and more controlled oxide and sulphide inclusions, greater control of segregates and freedom from surface defects. New roller straightening developments offer straighter rails with significantly flatter running surfaces. Ultrasonics, laser and eddy current equipment give printed information on surface flatness, straightness and interior condition. These have come along at a time when dramatic changes have occurred on the rail heat treatment scene. Rails can now be hardened during the actual rolling process. Rails of simple composition can be produced in the hardness range 240/400 BHN as required. The rail user can now choose hardness and geometry requirements over a wide range to cover all aspects from high speed to heavy haul with equal ease and availability and all under the latest quality assurance systems.

Introduction

The production of rails has seen very slow development over a large number of years but momentum began to develop in the early 1970's. The late 80's saw a rush of developments where long sought after ideas finally reached fruition. These developments were wide ranging covering many aspects of railmaking and brought it to the forefront of technology. Improvements in quality, properties and line geometry have provided dramatic increases in rail performance under almost all conditions of usage.

Rail profile grinding, lubrication and other track improvements have greatly added to expectations of even longer life. All this, coupled to the mill heat treatment developments, tilt both technical and economic advantages away from tradition to the point where rail specifications need to be re-written so that railroads can take full advantage of these improved products.

J.J. Kalker et al. (eds.), Rail Quality and Maintenance for Modern Railway Operation, 29–39.

Development Aims

Those at the forefront of rail manufacturing technology have responded to the needs they foresaw among the railways of the world. Around twenty years ago such needs could be seen as means to avoid premature failure, to increase resistance to surface damage and the need for improvement in straightness and line geometry. This led to developments on tighter control for chemistry, segregation, inclusion types, pipe, blister, hydrogen and surface condition. Rails containing these improvements would be straighter, flatter, longer and harder. By 1992 logical developments have in general met every one of these objectives to a very high degree whilst still retaining common steel prices. For the first time in many years there is no longer any obvious large volume of development work waiting to be carried out. However, railroad demands are ever increasing and much work has still to be done but with at this stage less obvious direction. The one area of failure has been in the field of residual stresses. Foot stresses particularly still do not allow rails to perform to their full potential.

The following is largely a survey of the state of the art as several leading manufacturers reach a similar stage of development by differing routes.

THE BLOOMMAKING PROCESS

Steelmaking/Casting

Today's rail steels are manufactured in either basic oxygen or electric arc processes. The most prefered route is basic oxygen steelmaking where liquid blast furnace iron is given a partial refining to rail steel. This is followed by full refining, temperature and chemistry control in a ladle arc secondary steel making process followed by vacuum degassing. The combination of these three processes gives extremely fine control of all aspects of steel condition. The liquid steel is then preferably continuously cast under gas shrouded conditions to give a sound low segregation bloom free from surface defects. Due to pick up of hydrogen from the tundish at the beginning of pouring all manufacturers require some blooms to be slowly cooled for further hydrogen removal.

Where scrap melting is used rail manufacture can be successful but for economic advantage and for best technical effect the process could best be followed by the same system as for basic oxygen i.e. use the arc furnace for melting and partial refining but do full refining in the ladle arc system followed by vacuum degassing. Again for heat treated rails the system can be fully effective but great care is required in scrap selection and additions to give the desired consistency of tramp elements.

Oxides

During steelmaking the decision has to be taken as to the level and type of oxide to be found in the final rail. During the 1970's finished steel oxygen developments of 75ppm were common but today's developed processes give around 12ppm. Oxides used to be heavily effected by aluminium giving brittle broken aluminium silicates. These gave rise to shelling and fatigue damage. Today's clean processes use preferably specially reduced aluminium alloys with no actual aluminium addition. This gives freedom from the alumina type inclusions which were so troublesome.

Sulphides/Segregation

In today's premium rail production sulphur averages 0.015% as compared to 0.035% just a few years ago. A bottom limit of 0.008% is usually applied to avoid the re-introduction of hydrogen sensitivity. Segregation has been brought under strict control and it is proposed to reflect this in the new European standard.

Hydrogen

This element is unusual in that it can not be measured by the customer in the finished cold rail. Ultrasonics can not detect its initiation. Its effects, in the form of hydrogen flakes and consequent fatigue, can show up twenty years later so hydrogen has always been a subject of nervous concern. Today's steelmakers have no longer reason to pass such concerns to the customer. Low hydrogen steelmaking techniques are well understood and if these are followed by proper vacuum degassing the general hydrogen level should be adequate for rail purposes. However, pick up of hydrogen from the tundish at the beginning of teeming is unavoidable and some bloom cooling is required to overcome this. Hydrogen diffusion conditions are well investigated and there is no reason for todays rails ever to shatter crack. Most manufacturers work to 2.5ppm maximum but techniques are available to achieve half that level if necessary.

Chemical Control

All rail specifications use very wide limits for most elements. They are all based on yesterday's steelmaking techniques. Railmakers still argue for wide limits claiming the necessity to have several options to achieve the same properties. It is often even claimed that tight controls are unnecessary as current rails perform well enough. However, it is only by offering an improved range of chemistries that the benefits can develop. If for example carbon/manganese/sulphur tolerances were reduced to only half their current limits flash butt, thermit and surface repair welding techniques could all be considerably modified.

RAIL PRODUCTION

Rolling

Rail steel blooms typically 4-6 tonnes in weight and around 260 mm x 360 mm are reheated to an even temperature taking precautions to contain decarburisation. Blooms are rolled to rails by either the caliber or universal processes. Each system has its advantages and disadvantages but these differences only effect the manufacturer and have little influence on the consumer. Hot rolled rails are transferred onto cooling banks to transport the rail to arrive cool at the roller straightener entrance in a condition correct for straightening. Slow cooling of rails for hydrogen removal is no longer necessary and should not be required for stress relieving.

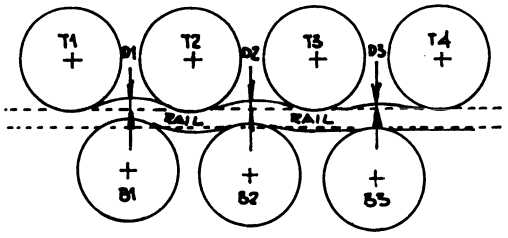
Hot rolling mills have made gradual improvements over the years but significantly the manufacturers are not offering closer rolling tolerances although today's product must be more consistent.

Finishing

The finishing of rails has seen very considerable development and currently offers a significantly improved product. Finishing consists of roller straightening, gag-pressing, cutting to length, drilling, inspecting and packaging for delivery.

Roller Straightening

Roller straightening equipment and techniques have been modified in ways which are highly significant for many customers. The straightening machines have usually 7 or 9 roller configurations, figure 1. Primary plane straightening is usually followed by a similar but smaller machine carrying out secondary plane straightening. Today's straightening machines have many tasks including straightening, surface flatness improvement, running surface contour and residual stress control. To achieve today's best performance figures the straightening rollers have to be set at the required distance apart to give the balance between rail crush, straightness and residual stresses.



ROLL ARRANGEMENT - LAMBERTON STRAIGHTENER SHOWING DEFLECTIONS D1, D2 & D3.

Figure 1.

Rotating components have had to be re-designed to give some form of cone fitment, figure 2, to remove all play. The same cone fitting system should be used on roller machining equipment to give the best effects. Stiffer machines with these new improved tolerances are given full instrumentation to allow accurate control and recording of roller adjustments. With the rolls themselves at around 550 BHN the combination has for the first time allowed accurate adjustment of straightness, flatness, residual stress while at the same time ensuring that the machine can be reset and maintained at the original settings. After 50 years the "art" is finally being taken out of roller straightening.

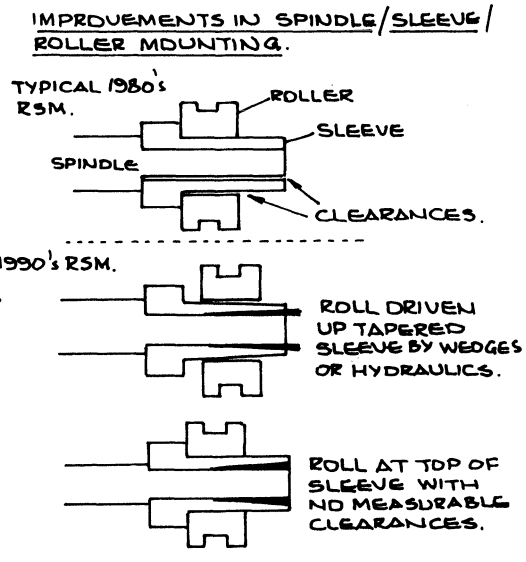


Figure 2.

The straightened rail has the non-straightened ends corrected to new standards by presses which are more and more computer aided. Better machines can measure the rectifying requirements at the rail end and suggest a setting. However for extreme accuracy at this time the operator still makes the final judgement.

INSPECTION

Ultrasonics

Multi-probe ultrasonics systems covering 60% or so of the rail area to a 1.5mm flat bottom hole standard are the norm, figure 3. These machines test rail along its complete length, spray defective areas and record results.

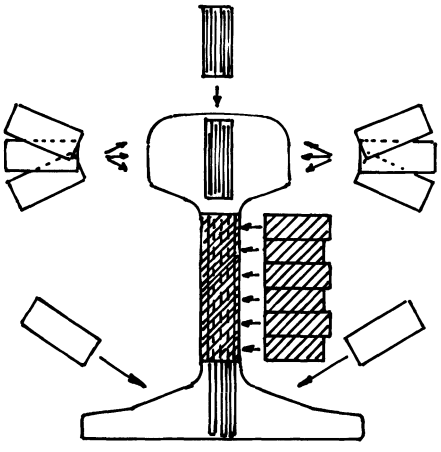


Figure 3.

In general the practice of clean steelmaking has kept ahead of the improvements in ultrasonic testing. While ever more advanced testing equipment is being developed and the developers are rightly proud of their ability to detect 2 mm flat bottomed holes it must be remembered that this only equates to inclusions 4-6 mm and then only in the more effective part of the beam.

Modern steelmaking almost never produces inclusions more than a fraction of this size. The role of all this ultrasonics today is to pick up entrapment from powders and covers. Unfortunately there is no quick technique for the detection or measurement of small steelmaking inclusions.

Surface Flatness

Automatic equipment using direct contact or lasers has been developed to very sophisticated levels. When coupled to adequate computers this type of equipment is capable of measuring waves in the running surface or side of head of a rail. Such equipment first showed that as rolled rails contained many waves. Roller straightening greatly reduced these waves but installed new ones.

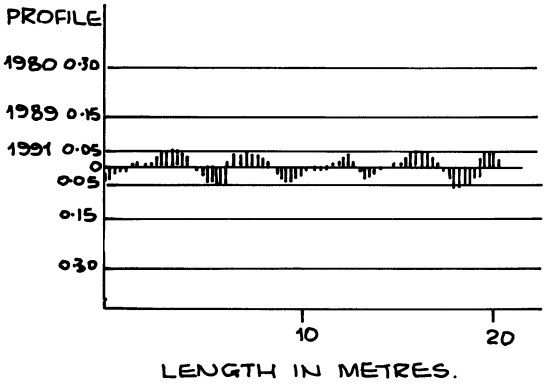


Figure 4.

Roller straighteners were first given improved tolerances, some time later cone fitments and all of the time more and more instrumentation. The fully modified straightening equipment gives wave tolerances to dimensions unthought of just a few years ago, figure 4. Again production is beginning to get ahead of measurement but at least today's rails are capable of being produced with the majority having surface flatness $\pm$ 0.05 mm and 95% $\pm$ 0.08 mm. The rail is not so stiff in the lateral direction but wave depths as low as $\pm$ 0.2 mm are possible.

In the rolling mill inspite of the use of harder rolls, roll pass lubrication and high pressure water sprays of 200 bar, surface roughness of the running surface has still not been improved to required levels. These improved levels are required for noise and corrugation control.

Surface Quality Assessment

Today's rails contain very few defects of steelmaking origin. By 1970 rails contained 12% surface defects when assessed against a very slack standard. By 1980 the defects were down to 4% and by 1990 to below 0.4%. With rails of such relatively good surface visual inspection reveals only around 20% of the seam type defects known to be present. Unfortunately the eye cannot differentiate between 0.2 mm and 2 mm depth even on this 20%. Automatic equipment has been available for many years, figure 5, but this generally tests to a depth of 0.5 mm. There is no evidence that seams undetected by such equipment ever give trouble in service.

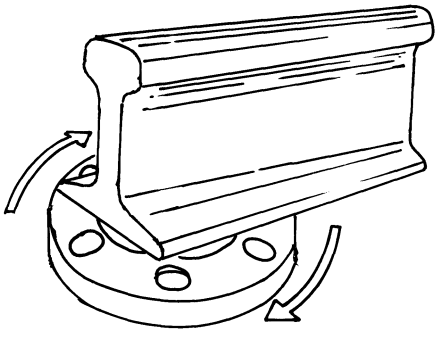
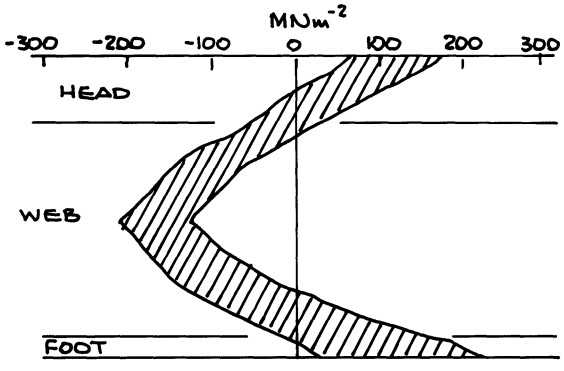


Figure 5.

Residual Stress

Longitudinal residual stresses are of ever increasing interest to the rail user but unfortunately there is no automatic or quick way of testing. The Debro ultrasonic strain measuring equipment is a useful comparator but no more. Much more accurate results are obtained by the slow strain gauge or web saw opening tests.

The majority of residual stresses are put into the rails by the roller straightening machines. As indicated earlier these machines are now better understood and more controllable and can give reduced stress for a given rail strength than before. The main area of interest on residual stresses depends to some extent on the rail section and strength.



SCATTER BAND OF LONGITUDINAL RESIDUAL STRESSES MEASURED BY ORE D156 IN UNUSED ROLLER STRAIGHTENED GRADE 90A RAILS.

Figure 6.

Tall hard rails tend to centre interest on the compressive web stresses as web splitting can occur while on the less tall predominantly European designed rail the tensile residual stresses tend to increase foot fatigue. The stress pattern is shown in figure 6. The web saw opening test has been found to relate closely to web splitting potential but not to foot fatigue. Separate tests are therefore required for AREA and European Rails.

RAIL HARDNESS

Although work is progressing in several areas including low and high carbon bainitics and even martensitic structures the vast majority of rails for the foreseeable future will have a pearlitic microstructure. Other things being equal the finer the pearlite the more wear resistant the rail. The finer the pearlite the harder the rail and provided the pearlite remains true then the hardness figure becomes the main requirement in this area. Fully pearlitic rails fall into the BHN hardness range 250-410. The rails within this range are manufactured by a variety of processes. Naturally hard rails containing carbon and manganese have a maximum hardness of around 300 BHN. Above that level some form of assistance from small quantities of alloying elements is required.

For harder naturally hard rails 1% chromium/vanadium was used but due to the inherent low fracture resistance this rail is much reduced in general usage. For significant hardening above 300 BHN rails are manufactured around the world by a variety of heat treatment processes. Some processes re-heat but all use some form of accelerated cooling and there are a variety of techniques aimed at following the cooling curve, figure 7.

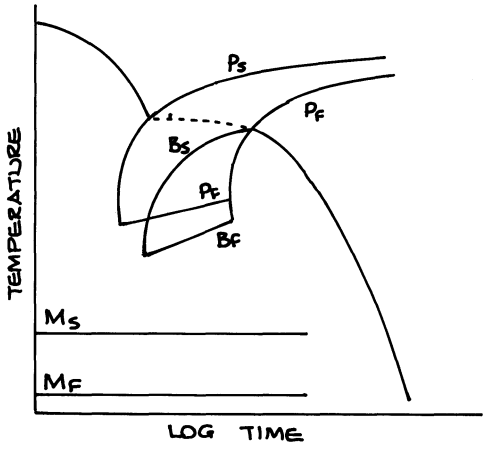


Figure 7.

The area between PS and PF is the one which taxes the mind most. Rail shape, dimensions and the heat of reaction of pearlite all combine to make following this line a near impossibility. Bramfitt (1) has recently reviewed heat treatment techniques. Typical hardness pattern for off-line head hardened and continuous on-line treated rails are shown in figure 8.

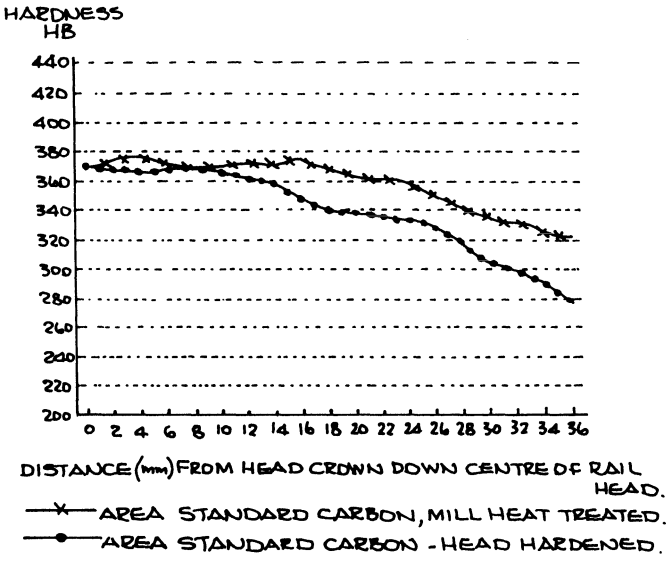


Figure 8.

Fully Heat Treated Rails

Fully heat treated rails have been successfully produced for many years by heating a carbon rail to 850°C prior to oil quenching. This gave a hardness of 340/370 BHN with the head largely fine pearlite. Tempering of the rail at 450°C removed any martensite and most residual stresses. This rail is still produced in quantity in Poland, CIS and USA.

The process gives rails which perform well but the hardness range is restricted and due to having to quench the whole rail equally at a fixed rate of extraction.

Head Hardening

This name was given to the other major off-line re-heating process where originally the head of the rail was heated and air quenched to form a region of fine pearlite in the top portion of the head. Hardness was usually in the range of 360-400 BHN. The process has been further developed by pre-heating the rail and using many electric inductors to give a better heat and therefore stress balanced to the quenched rail. Some processes now even fully re-austenatise the rail prior to air quenching the head to give fine pearlite. Both pre-heating and re-austenatising remove the initial roller straightening stresses. However, in the main rails are re-straightened after treatment. These processes can now deep harden rails but by today's standards this route is slow and inherently costly.

On-line Heat Treatment

By 1985 several producers were experimenting with new forms of on-line heat treatment processes which do not require reheating or double handling of the rail. Currently there are several systems in use which have been given general descriptions in the literature.

a) **On-line air cooled**

Rails are rolled and grouped together while still hot. These are clamped for restriction and on reaching a suitable temperature below 800°C are given an air blast to increase their cooling rate. This process gives a hardness of around 340 BHN to carbon/ manganese rails and 370 BHN if a small amount of chromium is added. The process appears to give good hardness control and should be much less costly than the re-heating system. Rails can be treated at full mill rolling rates.

b) **Head Quenching On-line**

This process takes the hot rails after the exiting the finishing pass, inverts the rail and dips the head into a tank containing circulated aquaias polymer. The head of the rail is quenched to form fine pearlite. Operation is at full mill production speed and the process gives a hardness of 340/370 BHN.

c) **Continuous On-line Heat Treatment**

Hot rails are taken directly from the rolling finishing pass without delay into a water cooled treatment system. Rails are passed through the spray system at a rate depending upon rail temperature. The process varies speed and waterflow along the length to give a constant fine pearlite and hardness along the head of the rail. The web and foot also receive spray cooling and can be hardened as necessary but usually not to the hardness of the head. While the rail enters the machine at any austenatite temperature 700-1000°C it exists at a mean temperature of 600°C. This plus the fact the rail is guided but not significantly restrained leaves the rail with an inherently low residual stress when cool. The system works at full mill rolling rate with the rails continuously moving along their normal route. This must be inherently the most cost effective system providing only that the space requirement is consistent with plant capability. Treating at 130 tonnes per hour the system requires 100 metres of space for carbon rails and even more for alloy grades.

The Future

The integrity of todays rail is much more assured. In the main premature failures should be very rare. The railroad can therefore better plan for the total life of the rail.

Wear resistance can be varied by hardness and lubrication to cover a very wide range. This should allow the railroad to select two or three types of rail best suited for their needs to give a planned long life in each situation. Particularly if head contour grinding becomes the norm wear rates can be very low for most applications without risk of running service fatigue. It should be possible to know how many MGT a rail can carry in each situation with a high degree of certainty. This same grinding when used purely as defense against fatigue or to provide the correct head contour would be the current best form of defense against noise and corrugation. In fact the position may well be reversed if noise legislation makes grinding necessary it would be economically wrong not to introduce harder rail.

Ever flatter running surfaces have been produced for high speed rails but the spin off must be reduced stresses in all except light rail systems.

Manufacturers are still finding it difficult to dramatically reduce rail foot tensile stress. A breakthrough in that area would allow rails to reach their full potential.

The new on-line heat treatment techniques should eventually offer a new range of pearlitic rails. Current developments have all been with standard rail chemistries. Some of these new heat treatment techniques are not limited to treating such chemistry. However, it will be some time before cooperation between manufacturer and user develop such rail. When the next stage of development is complete it would be anticipated that all heavy rail would be heat treated to give the best all round combination of properties.

Reference

(1) Bramfitt - Mechanical Working and Steel Processing Proceedings Conference Cincinnati 1990.

A REVIEW OF BAINITIC STEELS FOR WHEEL/RAIL CONTACT

P.Clayton, R.Devanathan, N.Jin and R.K.Steele*
Oregon Graduate Institute of Science and Technology
*Association of American Railroads

ABSTRACT

The potential of bainitic steels for applications involving wheel/rail contact is reviewed. Particular attention is paid to wear resistance because this is where bainitic steels have been perceived to be inferior to pearlitic steels. Five new experimental, low carbon bainitic steels have been produced to answer some of the unresolved questions.

HISTORICAL PERSPECTIVE OF RAIL AND WHEEL STEEL DEVELOPMENT

Fatigue, fracture, wear and deformation, in their various external forms, have provided the incentive to experiment with alternative materials. For such critical components as wheel and rail, however, the acceptance of new alloys has involved much time, doubt and suspicion. The introduction of the steel rail was no exception.

The first Bessemer rail was used by the Midland Railway at Derby Station in 1857 where it proved itself beyond any doubt (1). Variable quality and the normal human reaction to personal experience as the only reality, fuelled a controversy of iron versus steel that lasted until the end of the century.

The early steel was low in C and Si but very high in S, P and inclusions. Hackney (2) gives details of tests on rails containing 0.16% C and 0.3% P. With experience, steelmakers learned how to maintain a P level below 0.1% enabling C contents as high as 0.5% to be utilized for wear resistance (3). The subsequent mainstream development of rail steel can be viewed as the reduction of free ferrite, interlamellar spacing and inclusion content, to produce cleaner steel of higher strength.

The last decade has seen a burst of activity with the development of on-line head hardening and significant innovations in clean steel practice. The latter issue will no doubt continue to receive considerable attention. However, with interlamellar spacings as low as 100nm it is conceivable that this approach to strength has almost run the full course. There are other options, precipitation hardening, particularly with vanadium, has been studied extensively (4).

There have been notable diversions from the mainstream development, usually to address concerns about toughness. In the USA in the 1930's (5), 0.3%C 3%Cr rail was produced to provide tougher steel with high hardness and strength. Cost, presumably, brought about its demise since the early experiments were encouraging.

An improved fracture toughness rail steel was developed in Britain in the early 1970's (6). Carbon content was reduced and austenite grain size controlled by nitrides. Cost and poor wear resistance were the primary reasons for discontinuing the program.

J.J. Kalker et al. (eds.), Rail Quality and Maintenance for Modern Railway Operation, 41–51.

Hadfield's manganese steel offers an unusual combination of strength and toughness. In wrought and cast form it has become the universally accepted standard for switch and crossings operating under severe conditions. The material is no panacea, however, since it is difficult to cast and machine, and only achieves full resistance to deformation and wear as a result of significant plastic strain. It is also high in alloy content and difficult to weld. Periodically, a new approach is sought either through a replacement or methods of improvement. Attempts have been made to increase the work hardening rate by additions of V, W and Mo (7). Aluminum raises the solubility of C in austenite and a high work hardening rate has been attained in a 1.75%C 1.3%Al alloy (8). There is an accompanying enhanced risk of carbide formation and general improvements in properties were not achieved.

Bainite has been proposed as an alternative to pearlite (9) and austenite (10). A major obstacle to the evaluation of experimental data has been the inadequate characterization of bainitic microstructures, of which four types are now commonly recognized.

BAINITE

The term bainite refers to microstructures formed by the decomposition of austenite at a temperature, which depends on alloy composition, within the broad range 250-650°C. The two classical forms, upper and lower bainite, have lath structures. This is a common characteristic with the shear transformation product martensite. The formation of bainite also involves features associated with diffusion and an academic debate over the significance of the two mechanisms still continues (11).

Two non-classical forms of bainite have been identified and termed carbide free and granular bainite, following the work of Habraken and Economopoulos (12). These forms of bainite are usually found in continuously cooled low carbon steels. Optical microscopy does not provide the necessary resolution to identify them and transmission electron microscopy has to be used.

There have been several attempts to categorize bainitic structures. Ohmori et al (13) referred to three bainitic forms B_I, B_{II} and B_{III} with the fourth type being frequently mentioned in the literature:

B_I	Carbide-free bainite with packets of heavily dislocated ferrite laths and inter-lath martensite-austenite (MA).
B_{II}	Upper bainite with packets of heavily dislocated ferrite laths and inter-lath cementite.
B_{III}	Lower bainite exhibiting packets of heavily dislocated ferrite laths with cementite present within the laths at an angle of 55-60° to the longitudinal axis.
Granular	Heavily dislocated lath-free ferrite and MA islands.

A new classification of bainitic structures has been presented by Bramfitt and Speer (14). They conclude that the ferrite of granular bainite has a lath substructure and divide all bainitic forms into three categories:

B_1	Acicular ferrite with intralath precipitation of carbide.
B_2	Acicular ferrite associated with interlath particles or films
B_3	Acicular ferrite with a discrete island constituent.

Superscripts are employed to identify second phases and so granular bainites become $B_3^{m,a}$ and lower bainite B_1^c. While this approach has not yet received wide attention it has considerable merit although it could create some confusion with the evaluation of low carbon structures.

Normally, a homogeneous bainite transformation is difficult to achieve in low alloy steel because bainite is prevented by the formation of ferrite and pearlite. The addition of 0.5% Mo and 0.002% B is one way of retarding these reactions to produce bainite over a wide range of cooling rates.

Further control of the austenite to bainite transformation is exercised through the use of the alloying elements Ni, Cr and Mn which act to depress the bainite start transformation temperature (B_s). As this temperature is lowered, the strength properties are usually increased without any significant loss of ductility.

Several empirical equations have been established (15-18) to relate B_s and alloying additions. The relations tend to be specific to the group of alloys used in the derivation. In all of them, however, C has by far the largest effect on B_s.

Hundreds of thousands of tons of bainitic steels are manufactured each year (14). A favorable balance of strength, toughness and cost has led to their use in many applications, including gas transmission lines, steam turbines and pressure vessels. The potential of bainitic steels for tribological situations has been limited by poor wear resistance.

WEAR RESISTANCE

There are several different mechanisms of wear. A material may show excellent resistance in one situation and indifferent performance in another. Identification of the dominant mechanism in practice and its simulation in a laboratory test is usually a necessary prerequisite to the successful adoption of new alloys.

In rolling/sliding contact wear tests, Bolton and Clayton (19) identified three regimes of wear that they labelled, types I,II and III. Danks and Clayton (20) demonstrated that gage face wear of an unlubricated rail in heavy haul traffic is type III in which an initial break-in period results in roughened surfaces that suffer a form of abrasion by wear debris particles. This is the most severe of the three and has many similarities to sliding wear. For high speed systems with lighter axle loads, type II is usually considered appropriate.

The earliest bainitic rail service tests were conducted with two steels (21), H1 and H2 in Table I. The low carbon steel had a coarse acicular bainite and the second an acicular bainite with finely dispersed carbides. In curved track they wore 50% faster than conventional pearlitic steel rails.

Ichinose, et al.(22) observed that a low carbon bainitic steel, I1 in Table I, wore about ten times faster than a pearlitic steel of the same hardness. The laboratory tests were conducted at a slide/roll ratio of 10% and contact pressures of 550 MPa.

Masumoto et al (23), using a disk on disk machine at 9% slide/roll ratio and contact pressures of 550 MPa, found that ferrite-pearlite steels were more resistant to wear than low carbon bainite. Kalousek, et al (24) also concluded that bainite wears faster than pearlite from laboratory tests with a CrMo eutectoid steel.

It is difficult to draw any firm conclusions from the above work because the microstructural characterization is rather vague. Furthermore, the investigations have not systematically examined wear resistance with respect to test conditions, chemical composition and other mechanical properties.

Clayton et al (25) conducted an examination of a range of materials, B1-B9 in Table I. The tests were carried out under sliding conditions. Figure 1 shows that the softer bainitic steels compared favorably with pearlitic steels of the same hardness.

Devanathan and Clayton (26) examined the behavior of two low carbon bainitic steels. One, D1, was taken from an experimental cast frog while the other (D2) was as-rolled, Table I. These steels were tested with a slide/roll ratio of 35% and contact

pressure in excess of 500 MPa to generate type III wear. The 0.04%C steel (D1) was also tested in two heat treated conditions. The material was austenitized at 1025°C and cooled at 2 and 600°C/second to generate hardness values of 267 and 317 HB, respectively.

Although the most recent head hardened rail steels exhibit a resistance to break-in at pressures greater than 1220 MPa (27) the steady state wear rate of D1, heat treated to give more retained austenite and a hardness of 267 HB, compared favorably with current rail steels, Figure 2.

Garnham and Beynon (28) tested similar bainitic steels under type II conditions with results no better than BS 11 rail steel. Sawley et al (29) also tested a low carbon bainitic steel, F in Table I, under type II conditions and showed that it did not perform as well as a BS5982 R8T wheel steel.

Clayton and Devanathan (30) investigated a eutectoid CrMo rail steel (L), heat treated to give a hardness of 337 HB for two microstructural conditions. The bainite was upper bainite and the pearlite had an interlamellar spacing of 50nm. Under type III conditions the bainitic steel wore appreciably faster than the pearlitic steel, confirming the conclusion of Kalousek et al (24). However, when the CrMo rail was heat treated to lower bainite of 54 HRC, resistance to break-in to type III wear was better than head hardened pearlitic rail steel.

MECHANICAL PROPERTIES

The mechanical properties of bainite result from four contributions; packet and lath size, dislocation substructure, solid solution hardening and dispersion hardening. Separating these effects for any given steel is often difficult.

The tensile strength and fatigue resistance of the nine bainitic steels used in the study by Clayton et al (25) were strongly related to carbon and chromium content. Two of the steels exhibited excellent impact resistance. Since none of the steels contained any Ni, the results conflict with the work of Callender (31). He concluded that to obtain high impact resistance a bainitic steel should contain 3% Ni and no second phase particles.

Sawley has suggested (32) that the discrepancy could be explained by unusually small prior austenite grain size in steels B1-B9. Naylor et al (33) found that the prior austenite grain size controlled packet size and this could be an influential crack growth barrier. It is not easy to detect the prior austenite grain boundaries in many of the bainitic steels. Optical microscopy indicates that for those that can be measured, Table II, prior austenite grain sizes are very variable for B1-B3 ranging from 6 to 20μm. This compares with 25μm for D1 (and D2 for which only a few grains were measurable) and around 20μm for Callendar's steels, C1-C4. Modern head hardened rail steels have grain sizes as low as 10μm compared with 50μm for standard section rail of two decades ago.

However, while grain size is undoubtedly a factor (34), the data of Table II suggest that it is unlikely to be the only explanation for the high impact resistance of B7. On the other hand a bainitic wheel steel without Ni (29), F in Table I, exhibited poor impact properties. The grain size of this material was 50/60μm.

In a simulated field experiment, a low carbon bainitic steel frog performed as well as explosively hardened austenitic manganese steel (35). In laboratory rolling contact fatigue (RCF) and deformation tests (27) low carbon bainitic steels equalled pearlitic steels of the same hardness.

DISCUSSION

The goal of any steel development program, whether for rails, wheels or frogs, is to produce the correct combination of properties. For bainitic steels the objective is a deformation, fatigue and wear resistant steel that is tough, easily manufactured and welded, at an economic cost. Since the toughness of current wheel and rail steels is low by any standard, its improvement may be seen as a worthwhile bonus.

The first question is whether the existing knowledge is sufficiently consistent to identify the optimum compositions for bainitic steels or to halt their development altogether. While some of the information is possibly contradictory there is more than sufficient encouragement to continue the quest.

The wear resistance of several different bainitic steels has been disappointing (22,23,28,29) under relatively mild test conditions. In laboratory tests more representative of non-lubricated heavy haul track conditions, the low carbon bainitic steels perform well.

Their wear resistance is inversely proportional to hardness, Figure 2. This is counter to most experience but hardness is not a good indicator of wear behavior except for a narrow range of test conditions and materials. In this case a decrease in hardness is associated with an increase in retained austenite. The influence of retained austenite on wear resistance has not been studied extensively but there are data which show it can be beneficial in abrasive situations (36,37,38). Dong et al have observed a positive benefit of retained austenite in the RCF behavior of some highly alloyed steels (39).

In high Si, carbide free bainite, retained austenite films between the laths were beneficial to impact resistance (40). Blocky austenite, however, exhibited a greater degree of instability and transformed to untempered martensite. This would not be expected to be a problem in very low carbon steels but the carbon content of the MA phase can be as high as 10 times the average (41).

Callendar's data (31) support the view that Ni is a strong influence on impact toughness. Grain size may well be an influential factor in the behavior of the B series of steels (25), which did not have any Ni. The best of these steels, B7, outperformed the best of the steels tested by Callendar, Figure 3. Even if small grain size explains the behavior of the former, further evidence of the effect of Ni on impact resistance and fracture toughness is required because of the cost. In the event that Ni proves to be very significant a cheaper alternative might be to use grain refinement of the austenite.

For one set of steels, Cr was associated with good strength, fatigue and wear resistance and did not impair impact resistance (25). Since none of the empirical equations shows Cr to be particularly potent in controlling B_s there is some doubt about a direct link between the two. Callendar (31) used Cr in preference to Mn for depressing the B_s because it has less effect on martensite start temperature (M_s). The concern was that if the M_s were too low, martensite could form when the alloy was welded. However, there seems little reason to suspect that very low carbon untempered martensite would be deleterious.

It would be economically advantageous to use as little Ni and Cr as possible. Mn is only about one third the cost of Ni and two thirds the price of Cr, per weight %. Although Mn is limited in higher carbon steels to prevent segregation this does not apply to low carbon steels. Heller's rail steel (21) contained 4.5% Mn with 0.07%C and the FAMA steels (42) contain high Mn levels.

Finally, there is the question of carbon level. The evidence indicates that for low carbon bainite the absence of carbides is beneficial. However, this still leaves a choice to be made within the 0.04-0.15% range. Also, there would seem considerable incentive to explore further the properties of high carbon lower bainite.

A pilot matrix of five new low carbon, Mo-B alloys, J1-J5, has been prepared, Table I, to answer the following questions:

a) Can the performance of the 0.04%C bainitic steel D1 be reproduced.
b) Is retained austenite beneficial to overall mechanical properties.
c) Can good impact properties be obtained without Ni.
d) Does carbide free offer better properties than granular bainite.
e) What is the optimum carbon content.
f) Does Cr affect wear resistance.

Alloy J4 has greater Mn and less Cr than D1, while J5 has 4% Mn and neither Ni nor Cr. This pattern is repeated for 0.1% C in J3 and J2. The latter is similar to the original B7 alloy, Table I, with the addition of 2% Ni. Although 0.2% C would normally lead to carbide formation, the Si content of J1 suppresses this reaction to produce a high level of retained austenite (43).

Figure 4 shows that, in the as-rolled condition, a range of bainitic structures has been achieved with carbides present only in J2. These steels will be tested for wear, deformation, fatigue and impact resistance. Heat treatments can subsequently be used to vary the microstructures to quantify the effects of retained austenite and prior austenite grain size.

CONCLUDING REMARKS

Pearlitic steel has evolved over a period of 135 years and Hadfield's steel has been used for more than a century. Bainite was only identified as a separate constituent in the 1930's and has received relatively little attention.

Recent research provides some encouragement that bainitic microstructures could offer engineering alternatives to pearlite and austenite. In particular, very low carbon bainite appears to perform well under severe contact conditions. A new series of alloys has made to clarify the relationships between microstructure and wear, deformation, fatigue and toughness properties in these steels.

High carbon lower bainite steel remains virtually unexplored.

ACKNOWLEDGMENTS

This work has been supported by the Association of American Railroads and the Federal Railroad Administration. Bethlehem Steel donated the most recent five heats of bainitic steel.

REFERENCES

1 D.Brooke, The advent of the steel rail, 1857-1914, J.Transport History, 7, Part I, 1986 18-31.
2 W.Hackney, Manufacture of steel, Proc Inst Civil Engrs, 1875
3 C.P.Sandberg, On rail-joints and steel rails, Proc Inst Civil Engrs, 84 1885-6, 365-398.
4 F.B.Pickering and B.Garbarz, Strengthening of pearlite formed from thermomechanically processed austenite in vanadium steels and implications for toughness, Mater.Sci.Technol, 1989, 5, 227-237.
5 G.B.Waterhouse, Chromium steel rails, Met Prog, Apr 1934, 32-35.
6 T.G.Collinson, W.H.Hodgson and V.J.McNeely, Initial development of fracture tough rail steels, Rail Steel Symp, ISI, London 1972.
7 W.B.F.MacKay and R.W.Smith, Unpublished work, Queens University, Ontario, Canada.
8 B.K.Zuidema, D.K.Subramanyam and W.C.Leslie, The effect of aluminum on the work hardening and wear resistance of Hadfield's manganese steel, Met Trans A, 18A Sept 1987 1629-1639.

9 A.L.DeSy, J.Dilewijns and G.Goethals, Development of a high strength Cu-Ni-Cr-Nb rail steel, Rail Steel Symp, ISI, London 1972.

10 K.Morton, D.F.Cannon, P.Clayton and E.G.Jones, The assessment of rail steels, Rail Steels- Developments, Processing and Use, ASTM STP 644, D.H.Stone and G.G.Krupp, Ed. ASTM, 1978, 80-98.

11 J.W.Christian and D.V.Edmonds, The bainite transformation, Phase Transformations in Ferrous Alloys, Ed. A.R.Marder and J.I.Goldstein, Met Soc of AIME, 1983 293-325.

12 L. J. Habraken and M. Economopoulos, Bainitic microstructures in low carbon alloy steels and their mechanical properties, Transformation and Hardenability in Steels, Climax Molybdenum, 1967, 69-108.

13 Y. Ohmori, H. Ohtani, and T. Kunitake, Bainite in low-carbon low-alloy high strength steels, Trans ISI Jap., 11 (1971) 250-259.

14 B.L.Bramfitt and J.G.Speer, A perspective on the morphology of bainite, Met Trans, 21A, April 1990, 817-829.

15 W.Steven and A.G.Haynes, The temperature of formation of martensite and bainite in low alloy steels - some effects of chemical composition, JISI 183 1956 349-358.

16 R.L.Bodnar and K.A.Taylor, Structure/property relationships in medium carbon bainitic steels for thick sections, Trans ISS, Aug 1990, I&SM 47-63.

17 R.L.Bodnar, T.Ohhashi and R.I.Jaffee, Effects of Mn, Si and purity on the design of 3.5NiCrMoV, 1CrMoV and 2.25Cr-1Mo bainitic alloy steels, Met Trans, 20A, Aug 1989, 1445-1460.

18 R.Devanathan, Wear behavior of bainitic steels, PhD Thesis, Oregon Graduate Institute of Science and Technology, 1991.

19 P.J.Bolton and P.Clayton, Rolling-sliding wear damage in rail and tyre steels, Wear, 93 (1984) 145-165.

20. D.Danks and P.Clayton, Comparison of the wear process for eutectoid rail steels: field and laboratory tests, Wear, 120 (1987) 233-250

21 W. Heller and R. Schweitzer, Hardness microstructure and wear behavior of steel rails, 2nd Int. Heavy Haul Railway Conference, Colorado Springs, CO, 1982, 282-286.

22 H. Ichinose, J. Takehara, and M. Ueda, High strength rails produced by two stage flame heating and slack quenching, 2nd Int. Heavy Haul Railway Conference, Colorado Springs, CO, 1982, 178-186.

23 H.Masumoto, K.Sugino and H.Hayashida, Development of wear resistant and anti-shelling high strength rails in Japan, Heavy Haul Railway Conference, Sept 1978, Perth, Inst Engineers, Australia.

24 J. Kalousek, D. M. Fegredo, and E. E. Laufer, Wear resistance and worn metallography of pearlite, bainite and tempered microstructures of high hardness, Wear of Materials, ASME, 1985, 212-232.

25 P.Clayton, K.J.Sawley, P.J.Bolton and G.M.Pell, Wear behavior of bainitic steels, Wear, 120 (1987) 199-220

26 R. Devanathan and P. Clayton, Rolling/sliding wear behavior of three bainitic steels, Wear 151 (1991) 255-267.

27 P.Clayton, R.Devanathan, V.Dikshit, D.E.Christensen, D.Danks and R.K.Steele, Recent research on wear, deformation and fatigue of steels in rolling contact, 34th Mech Working and Steel Proc Conf, Montreal, October 1992, ISS-AIME.

28 J.E.Garnham and J.E.Beynon, Dry rolling-sliding wear of bainitic and pearlitic steels, Wear, In press.

29 K.J.Sawley, J.A.Benyon and E.G.Jones, Bainitic steels for railway wheels, 9th Int Wheelset Congress, Montreal 1988.

30 P.Clayton and R.Devanathan, Rolling/sliding wear behavior of a chromium-molybdenum rail steel in pearlitic and bainitic conditions, Wear, In press.
31 W.R.Callender, Bainitic steels for railway applications, PhD Thesis, Sheffield University, England, 1983.
32 K.J.Sawley, Private communication, 1992.
33 J.P. Naylor, The influence of the lath morphology on the yield stress and transition temperature of martensitic-bainitic steels, Met Trans 10A (1979) 861-873.
34 K.J.Irvine and F.B.Pickering, The impact properties of low carbon bainitic steels, JISI, June 1963, 518-531.
35 E.G.Jones, Developments in railway crossings, 3rd Int. Heavy Haul Railway Conference, Vancouver, BC, 1986, IB-15.
36 M.S.Bhat, V.F.Zackay and E.R.Parker, Alloy design for abrasive wear, Wear of Materials 1979, ASME New York, 286-291.
37 N.J.Kar, Investigation of the role of microstructures on the two-body wear resistance of steels, Wear of Materials 1983, ASME New York, 415-425.
38 W.J.Salesky and G.Thomas, Design of medium carbon steels for wear applications, Wear of Materials 1983, ASME New York, 298-305.
39 Z.Dong, W.Fu-Xing, C.Qi-Gong, Z.Ming-Xin and C.Yin-Qian, Effect of retained austenite on rolling element fatigue and its mechanism. Wear, 105 (1985) 223-234.
40 H.K.D.H.Bhadeshia and D.V.Edmonds, Bainite in silicon steels: new composition-property approach, Met Science, 17 Sept 1983, 411-419.
41 T.Bojarski and T.Bold, Structure and properties of carbide free bainite, Acta Met, 22, Oct 1974, 1223-1234.
42 H.Martensson, The physical metallurgy of a high strength low carbon manganese steel, Proc ICSTIS, Supp Trans ISIJ, 11 (1971) 1072-1076.
43 H.K.D.H.Bhadeshia, Private communication, 1991.

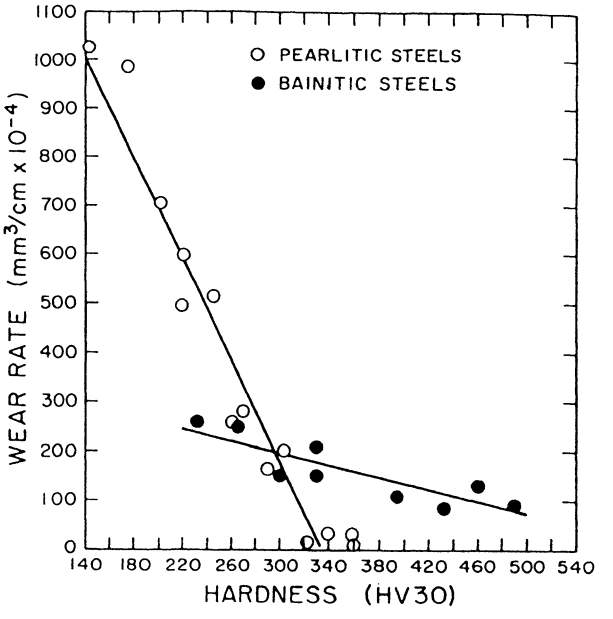


Figure 1. Sliding wear rate versus hardness for pearlitic and bainitic steels.

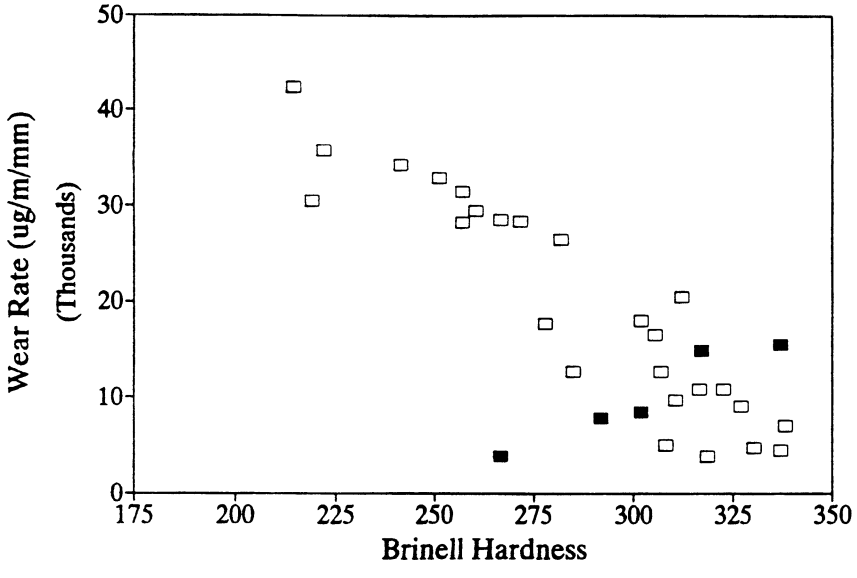


Figure 2. Rolling/sliding wear rate versus hardness for eutectoid pearlitic and low carbon bainitic steels.

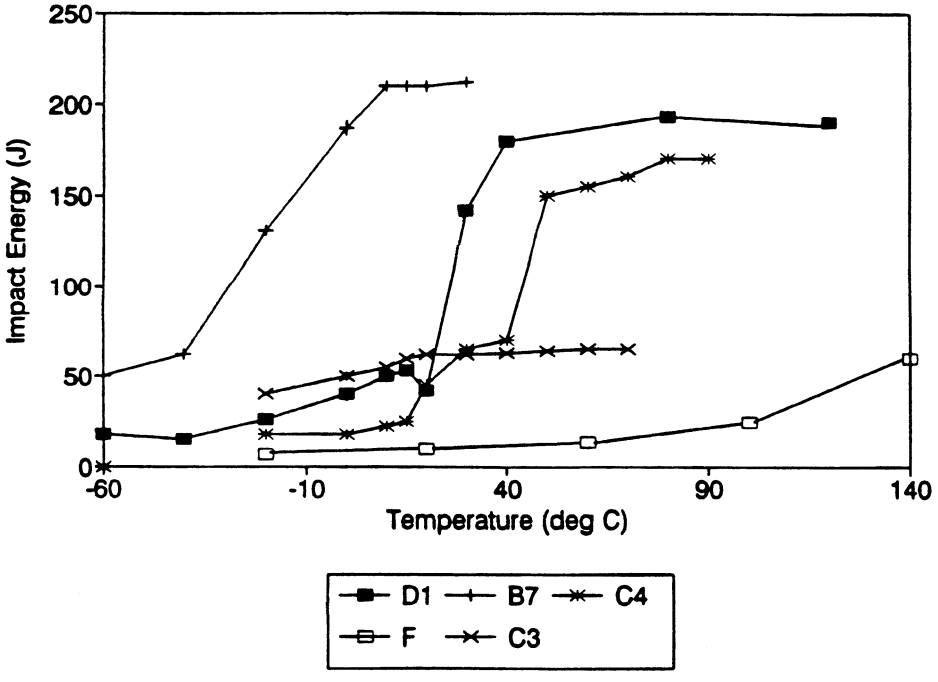


Figure 3. Charpy impact energy versus temperature for five low carbon bainitic steels in Table I.

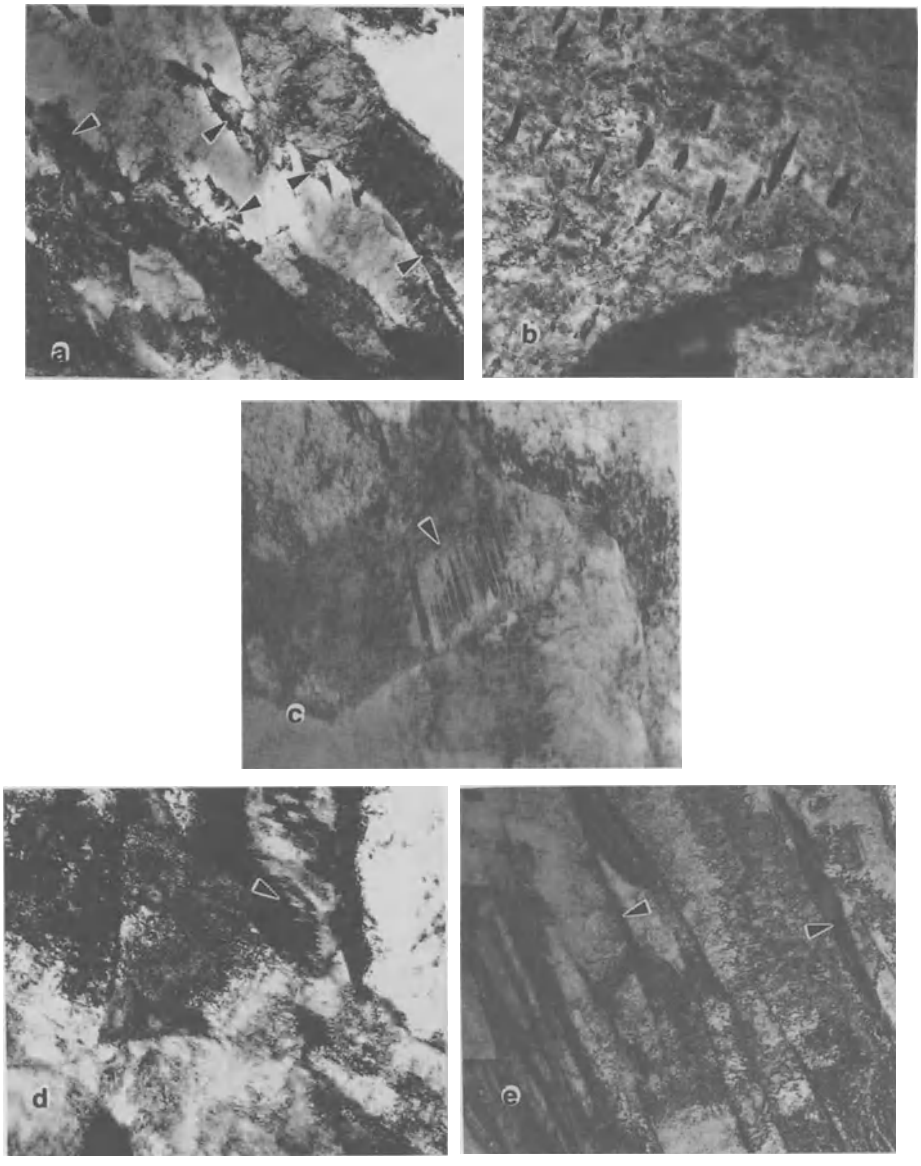


Figure 4. As-rolled microstructures of J series steels.

(a) Steel J1, Carbide free bainite/granular bainite, arrows indicate retained austenite, Mag. 40,000X
(b) Steel J2, Upper bainite, Mag. 20,000X
(c) Steel J3, Granular bainite, arrow indicates twinned martensite, Mag. 60,000X
(d) Steel J4, Granular bainite, arrow indicates twinned martensite, Mag. 30,000X
(e) Steel J5, Carbide free bainite, arrows indicate retained austenite, Mag. 12,000X

Table I. Chemical composition of reviewed steels.

		Element (Wt%)											
Steel	Reference	C	Mn	Si	Cr	Ni	Mo	S	P	B	Al	Ti	Hardness
H1	21	0.30			2.70		0.20						460 HB
H2	21	0.07	4.50				0.50						345 HB
I1	22	0.33	1.20	0.31	1.25	0	0.20	0.015	0.014	0.0029	0.28	0	345 HV10
M1	23	0.33	1.20	0.88	1.17	0	0.20	0.008	0.018	0	0	0	307 HB
K	24	0.72	0.81	0.28	0	0	0.21	0.022	0	0	0	0	
B1	25	0.09	1.01	0.21	0	0	0.50	<0.010	<0.005	0.0029	0.03	0.03	232 HV30
B2	25	0.21	1.99	0.22	0	0	0.50	<0.010	<0.005	0.0026	0.03	0.02	331 HV30
B3	25	0.30	1.49	0.21	0	0	0.50	<0.010	<0.005	0.0027	0.27	0.02	266 HV30
B4	25	0.09	1.53	0.21	0.95	0	0.49	<0.010	<0.005	0.0028	0.02	0.03	301 HV30
B5	25	0.19	1.98	0.20	0.95	0	0.49	<0.010	<0.005	0.0028	0.02	0.02	330 HV30
B6	25	0.29	1.98	0.21	1.02	0	0.50	<0.010	<0.005	0.003	0.03	0.03	460 HV30
B7	25	0.09	2.01	0.24	1.96	0	0.50	<0.010	<0.005	0.0029	0.03	0.03	395 HV30
B8	25	0.19	1.52	0.22	2.00	0	0.50	<0.010	<0.005	0.0025	0.03	0.03	433 HV30
B9	25	0.29	1.20	0.23	1.98	0	0.50	<0.010	<0.005	0.0030	0.03	0.03	490 HV30
D1	26,28	0.04	0.80	0.19	2.76	1.93	0.25	0.009	0.009	0.0023	0.03	0.03	29 HRC
D2	26,28	0.11	0.57	0.27	1.68	4.09	0.58	0.026	0.008	0.0023	0	0	35 HRC
F	29	0.08	1.60	0.26	1.40	0	0.46	0.024	0.013	0.0018	0	0	280 HV30
L	30	0.71	0.88	0.21	0.57	0.1	0.21	0.020	0.005	0	0	0	
C1	31	0.12	0.48	0.26	3.54	0.06	0.49	0.022	0.010	0.0030	0.04	0.04	334 HV30
C2	31	0.13	0.55	0.27	2.69	1.47	0.51	0.024	0.011	0.0030	0.04	0.04	305 HV30
C3	31	0.11	0.99	0.30	1.48	3.21	0.52	0.029	0.012	0.0030	0.05	0.04	364 HV30
C4	31	0.08	0.51	0.16	3.51	<0.02	0.51	0.005	0.008	0.0020	0.03	0.04	327 HV30
J1		0.21	2.00	0.97	2.01	0	0.51	0.010	0.013	0.0030	0.02	0.04	40 HRC
J2		0.13	4.04	0.28	0	0	0.51	0.008	0.013	0.0030	0.02	0.04	35 HRC
J3		0.09	2.04	0.28	1.98	2.03	0.50	0.011	0.013	0.0030	0.02	0.04	35 HRC
J4		0.03	2.10	0.29	2.10	2.10	0.53	0.009	0.014	0.0030	0.02	0.04	27 HRC
J5		0.04	4.10	0.29	0	0	0.51	0.009	0.012	0.0030	0.02	0.04	26 HRC

Table II. Charpy impact data.

Steel	Austenite Grain Size (μm)	Impact Energy (joules) (-15°C)	(15°C)	(60°C)
B1	6	12	18	212
B2	14	10	14	25
B3	20	12	14	16
C1	15	23	30	55
C2	18	32	50	60
C3	23	56	63	65
C4	18	13	25	150
D1	25	27	48	200

An Off-Line Heat Treatment Process Results in Quality DHH Rails at Sydney Steel Corporation

Theodore E. Burke, Ajax Magnethermic Corp., Warren, Ohio, USA
Blair George, Sydney Steel Corp., Sydney, N.S., Canada
Brian E. St. John, Permatrack Systems Inc.; Vancouver, B.C., Canada

Introduction

A joint development on the part of Permatrack Systems Inc., Vancouver, B.C., Canada, and Ajax Magnethermic Corporation, Warren, Ohio, USA, has resulted in a unique process for heat treating railroad rails. The process equipment is owned and operated by Sydney Steel Corporation, Sydney, N.S., Canada.

The heat treatment of rails has been done for over a decade, ranging from fully hardened to only head hardened. The alloys of rail steels have also varied through the years to accommodate flash butt welding, improved wear surface, and quench characteristics.

The head hardening processes are about as varied as the number of rail mills. Many of them accounted for the heat treatment distortion in rails by mechanical rework. The roller straighteners and presses to do this are usually the same ones used to correct rail straightness in the finishing mill for unhardened rail.

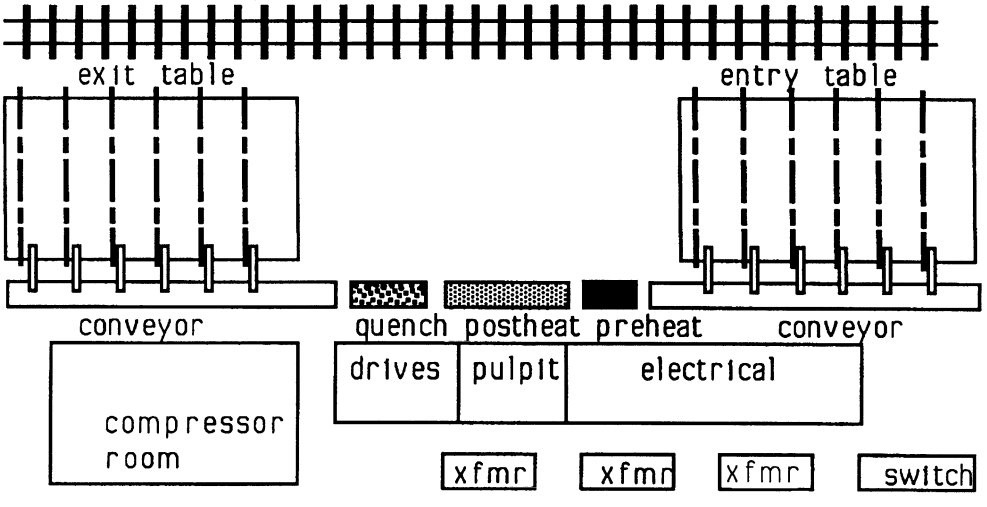
The Permatrack/Ajax, P/A, rail process equipment is intended to:

1. Head harden carbon and low alloy rail steels.
2. Head harden to standard and deep hardening patterns.
3. Head harden all AREA, ARA, and UIC railroad rail sizes.
4. Harden rail without further roller straightening.
5. Maintain small sectional changes that would not require special rolling mill adjustment.
6. Produce rail with minimal residual stresses.
7. Reheat the rail with maximum efficiency.
8. Harden rail at 5 fpm (1.5 m/minute) as supplied and 10 fpm (3 m/minute) with an upgrade of heating and cooling modules.

J.J. Kalker et al. (eds.), Rail Quality and Maintenance for Modern Railway Operation, 53–60.

Equipment

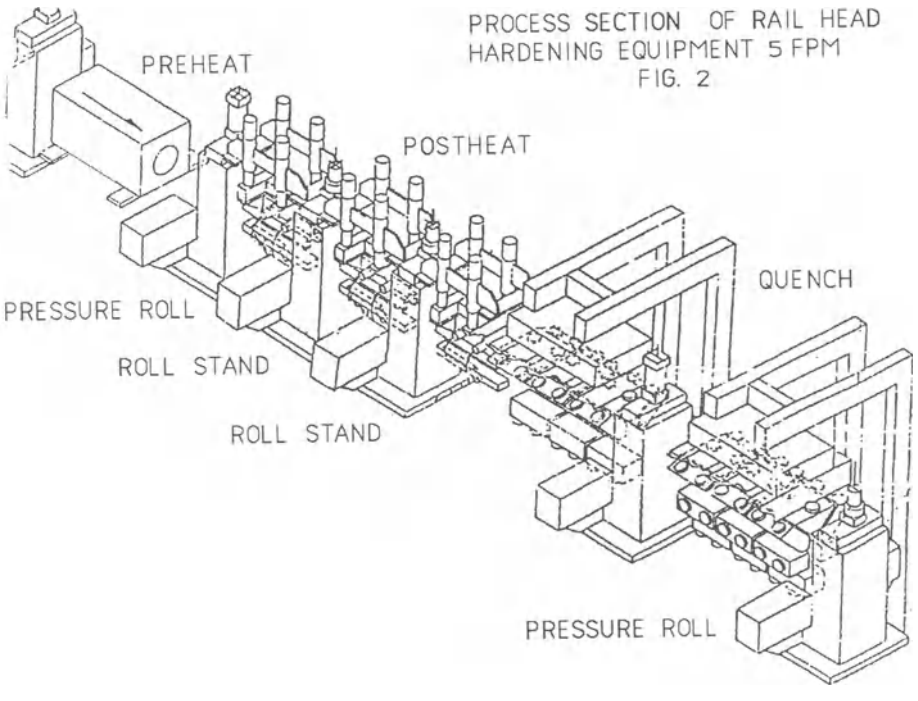
The building equipment layout is shown in Figure 1. Separate rooms are used for compressors, electrics, and control, whereas the process section, handling, and inspection are in an enclosed storage building with a trunk line inside it.



Sydney Steel Corporation Head Hardening Plant Layout FIG. 1

The equipment operation is presently manned by two operators who also perform line adjustments, routine maintenance, and rail handling functions.

Line adjustments for changing rail sizes are based on keeping the rail's vertical cross sectional centerline on the equipment centerline. Side rolls are adjusted in or out to contain the rail's flange. In this way symmetrical heating and cooling is maintained. Asymmetrical heating and cooling about the rail's neutral axis (head to flange) is done to maintain the rail's "surface" straightness. The pass line for all rails is the same. Resetting the line for a change in rail size or alloy requires about twenty minutes.



The process section for a rate of 5 fpm (1.5 m/minute) is about 30 feet (10 m) long from start of heating to end of quench and will accommodate a minimum rail length of 36 feet (12 m). Presently the storage table will handle rails 84 feet (26 m) long.

Four pinch rolls along the line drive the rails and keep them butted. Other rolls are containment rolls and powered conveyor rolls. In no case do the rolls deform the rail.

The chain conveyors keep the rails side by side for maximum storage capacity and ease of picking with the crane magnets. The conveyor never slides under the rail base.

Process

Rails are delivered from the finishing mill functions of straightening to the entry storage table of the head hardening equipment. Rails are run butted through the equipment and are transferred to an exit storage conveyor prior to final inspection.

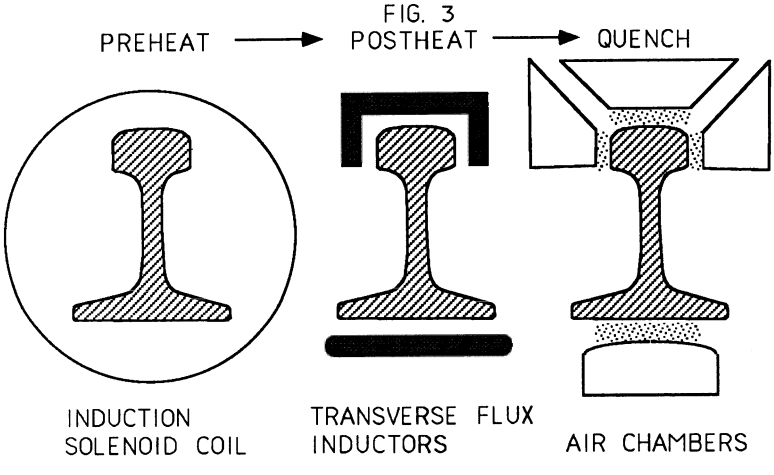
The sequence of processing goes as follows:

1. The preheat induction solenoid coil fully austenitizes the rail.

2. The postheat induction transverse flux coils for the rail's head and flange soak the rail thermally and provide straightness control.

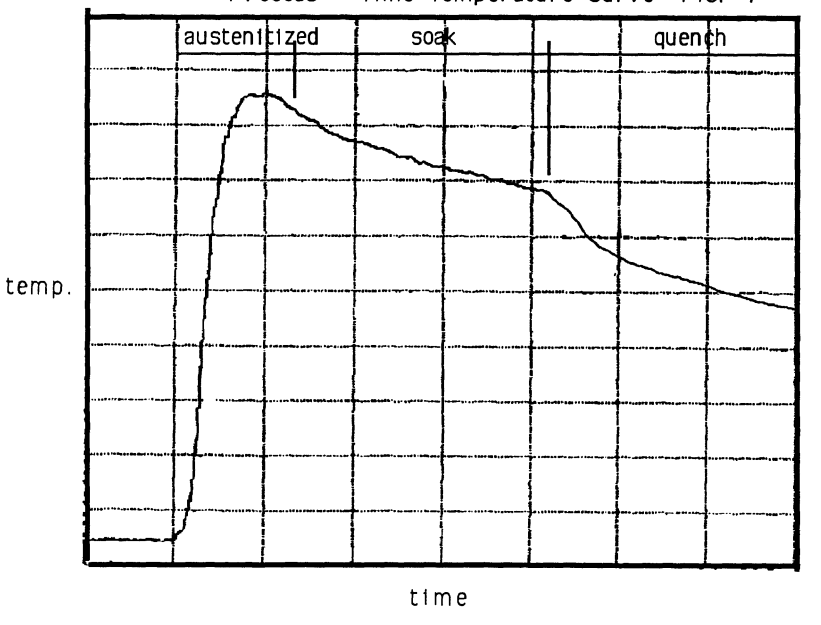
3. The air quench chambers have independent top, sides, and bottom volume control.

FIG. 3

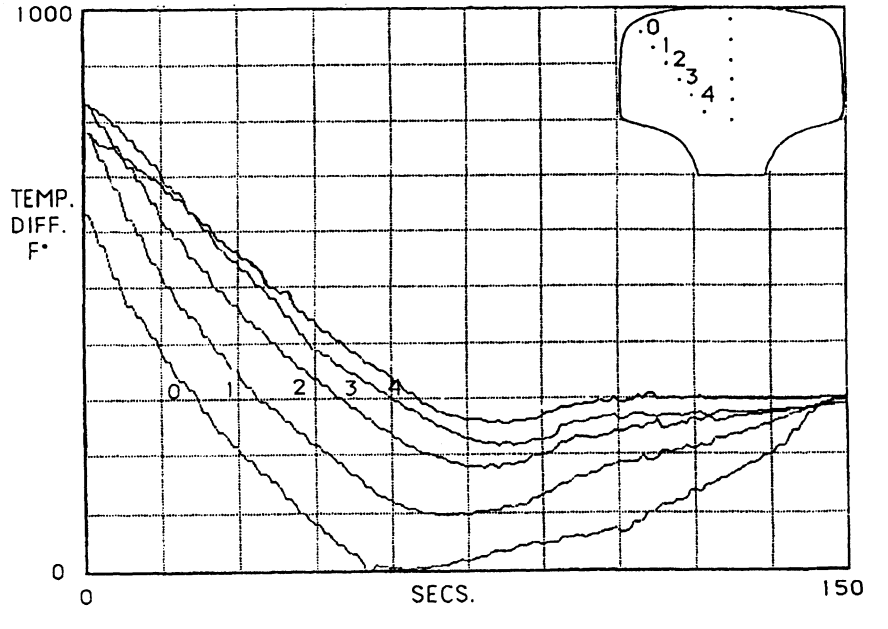


A PLC controls processing variables. Heating uses PID loop control, getting feedback from two color pyrometers and then adjusting power supplies if needed. Parameters are stored and selected from the PLC for different rail sizes and alloys. Process functions, fault monitoring, and SPC information are also accounted for by the computer.

Process - Time Temperature Curve FIG. 4



MAXIMUM AIR COOLING RATE FIG. 5



Results

The results presented are compiled from commissioning and customer orders, totaling about 5,000 tonnes. Some of this has gone to the heavy haul railways, but as yet none has completed test at FAST. During the period from commissioning to full production three different alloys and five different rail sizes were processed in small orders for Sydney Steel's customers. The line exhibited tremendous flexibility to handle all combinations with line adjustments only. At this point no special rolling is given to rails to be head hardened. However, rails rolled at the "tighter" end of the normal specification are preferred. Without roller straightening to compress the rail the growth in rail height is close to the maximum allowable.

Processed rail is intended to meet all AREA specifications. The DHH hardened rails change dimensions to be within that specification, including:

1. Cross Section Dimensions for:

    a. Head Width
    b. Height
    c. Flange Width
    d. Asymmetry
    e. Fish
    f. Base Concavity

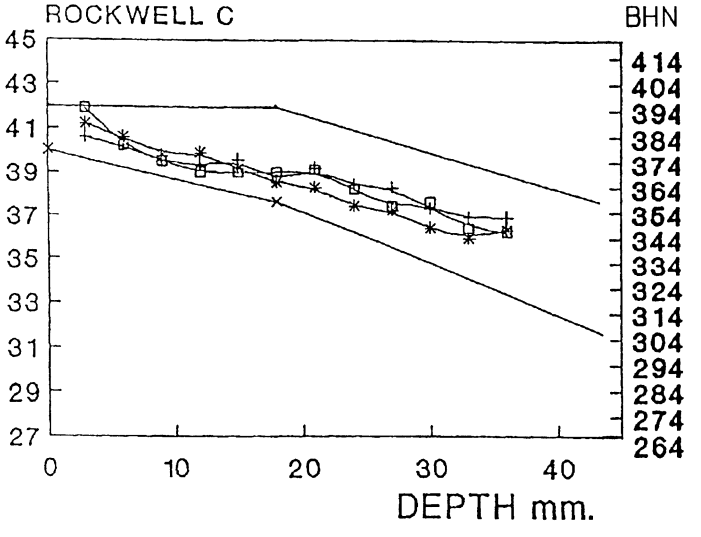
2. Overall Dimensions for:

    a. Length
    b. Sweep
    c. Surface
    d. Hooks
    e. Droops
    f. Depressions

Processed rails have the following characteristics:

3. Metallurgical

a.	Cross sectional hardness pattern	See Figures 6 & 7
b.	Microstructure is fine pearlite with mean intercept spacing averaging	266 nm
c.	Chemistry: .79 C, .87 Mn, .50 Cr, .72 Si	Typical
d.	Average hardness 6 points in the head	378 BHN
e.	Average running surface 200 samples	384 BHN



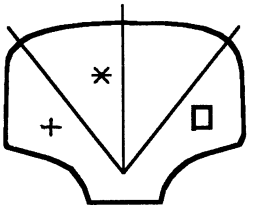
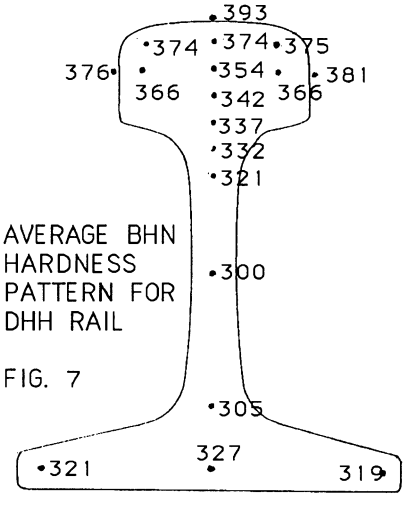


FIG. 6



AVERAGE BHN HARDNESS PATTERN FOR DHH RAIL

FIG. 7

4. Physical

a.	Average Yield Strength	119.5 ksi	824 MPa
b.	Elongation	13%	13%
c.	Ultimate	188.0 ksi	1296 MPa
d.	Residual, Crack Opening	-0.039 to -0.078 in.	-1 to -2 mm
e.	Drop Test, 3 ft. supports, 40 K-Ft.	2.65 in./6th drop	67.3 mm/6th drop
f.	Stress Intensity	63 ksi $\sqrt{\text{in.}}$	69.3 MPa $\sqrt{\text{m}}$
g.	Charpy V-Notch	7.0 ft-lb.	9.5 N-m

Conclusions

The significant advantage of this head hardening equipment is the lack of roller straightening after processing. New information cites "micro-cracking" due to roller straightening as detrimental to the rail's web. Preheating of the whole rail in the P/A process erases the stress memory in the rail and allows control over the final straightness. To date about 30% of the rails processed need slight end pressing for droops of a few thousandths of an inch; the remaining 70% are ready for shipment. We believe the droops can be improved to half that amount. Flexibility is also a key advantage of this process equipment.

This off-line process has the distinct disadvantage of higher operating costs when compared to the on-line process. Efforts have been made in the P/A process to maximize heating efficiencies by using an induction solenoid coil to bring the rail to austenitization.

The on-line hardening process requires the equipment to match the mill speed. In some cases the off-line hardening as a parallel process to other mill functions may prove to be an advantage. One such case would be the hardening of rail strings after welding. Hardening the weld joint would reduce the need for alloying. Hardening of welded rail segments has been done with this equipment and the results are very encouraging.

RESIDUAL STRESSES IN RAIL

W H HODGSON
BRITISH STEEL TRACK PRODUCTS

1) Introduction

One legacy left by the manufacturing route used for the shaping of metals is the presence of residual stresses arising from non-uniform plastic deformation. A knowledge of their magnitude and distribution within the product is important with regard to subsequent performance.

The significance of residual stresses with respect to rail life has been recognised for some time (1, 2). Railway administrations operate rolling stock at increasingly higher axle loads and speeds, placing greater demands on the rail. British Steel Track Products is in the forefront of developments of high quality wear resistant rails produced to tighter straightening tolerances. The different process routes adopted for their manufacture, such as on-line and off-line mill heat treatment have an effect on the character of the residual stresses in the product. The method of straightening adopted has a major influence and further modifications arise particularly in the rail head during service. The resultant residual stress distribution not only influences the fatigue resistance of the rail but also under particular conditions may be sufficient to drive web fracture. Therefore, the measurement of residual stresses in rails is an important factor in the evaluation of further processing improvements. The results of process improvement trials are described.

2) Measuring Techniques

The mechanical methods used in engineering stress analysis include sectioning, hole drilling, trepanning and the removal of surface layers, enabling the measurement of released strains to be made by strain gauges. The residual stresses are then calculated involving the use of elastic moduli.

J.J. Kalker et al. (eds.), Rail Quality and Maintenance for Modern Railway Operation, 61–73.

Many railway administrations have found sectioning methods, positions as Fig. 1, to provide comparable results between different test houses.

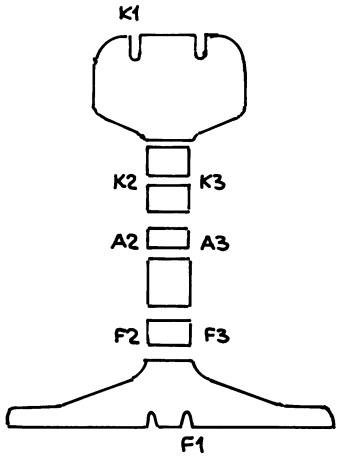


Figure 1.

A longitudinal saw cut along the rail web is sometimes employed to provide a qualititive indication of the residual stresses in the web by measuring the gap opening, Fig. 2. Such an inspection of through hardened rails is specified in the GOST 18267-82 Standard (3).

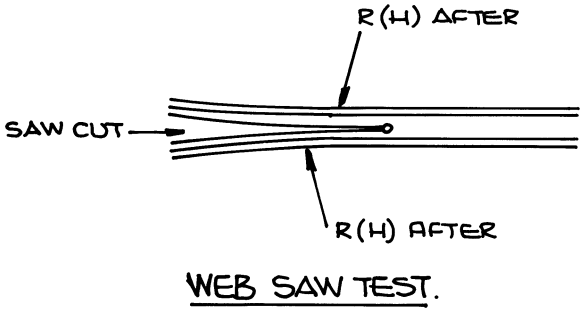


Figure 2.

In the whomper test the wedge is driven into the web of a rail to induce web cracking (4). Web stresses cause propagation and the longer the crack the higher the stress.

X-ray diffraction is a non-destructive technique but only measures strain in terms of lattice parameter changes st the surface of the product. The specialist technique of neutron diffraction has been applied to rail samples not more than 12mm thick for the measurement of residual stress but is expensive (5).

Use has been made in a number of organisations of an ultrasonic stress meter (Debro) which relies on the change in the velocity of an ultrasonic wave due to the presence of locked-in stresses (6). This device can measure longitudinal stresses near to the surface and around the periphery of the rail, although said to be useful for production quality control is very much 'surface sensitive'.

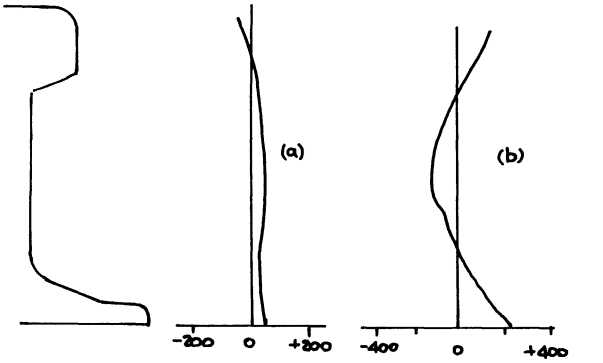
3) General Observations

Residual stresses are elastic stresses and the maximum value which can be reached is the elastic limit of the material. A complete analysis of the state of residual stress in three dimensions is complicated. However, because of symmetry a consideration of the residual stress in one direction may be sufficient to obtain an indication of the distribution pattern through the product. Comparisons between rails are frequently based on the measurement of longitudinal residual stress.

a) Rail Condition Before Straightening

When rail is cooled from an austenitising temperature both thermal and transformation stresses are produced. Unavoidable temperature gradients in the product result in non-uniform plastic deformation. Therefore, the conditions controlling the rate of cooling of the steel are important in determining the final state of stress.

Conventionally rolled rails are cooled comparatively slowly on cooling banks and their straightness is influenced by the degree of precurvature of the hot rail and uneven cooling across the section. However, the levels of residual stress are low (approx. 50 N/mm2), as shown for a Grade 90A rail in Fig. 3(a).

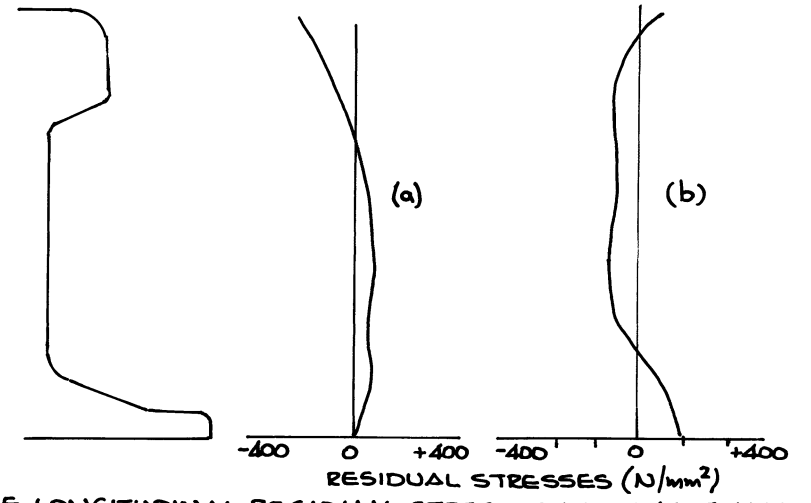


SURFACE LONGITUDINAL RESIDUAL STRESSES IN GRADE A RAIL

(a) AS ROLLED. (b) ROLLER STRAIGHTENED.

Figure 3

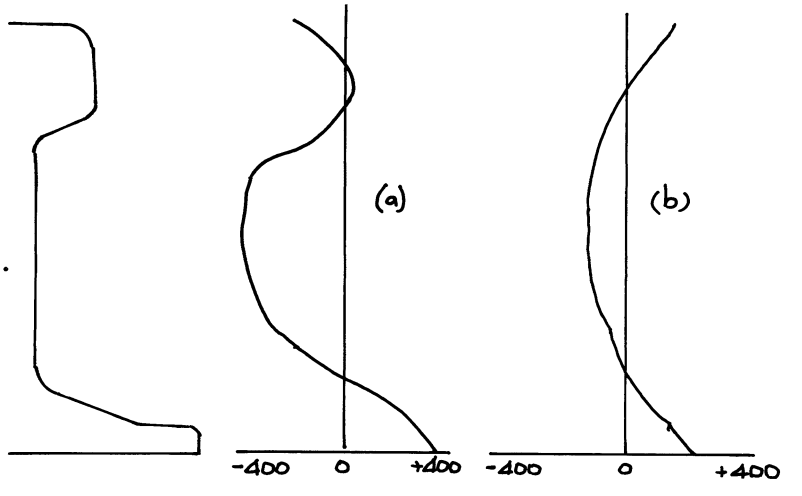
In general, an increase in the rate of cooling results in the development of higher stresses. This situation is complicated by the superposition of transformation stresses as austenite transforms to pearlite.



SURFACE LONGITUDINAL RESIDUAL STRESSES IN ON-LINE HEAT TREATED RAIL (132 RE) (a) AS PROCESSED (b) AS STRAIGHTENED.

It is therefore not surprising to record higher residual stresses in heat treated rails, as illustrated in Figs. 4(a) & 5(a) following on-line and off-line processing.

Fig. 5.



SURFACE LONGITUDINAL RESIDUAL STRESSES IN EARLY OFF-LINE HEAT TREATED RAIL (132 RE) (a) AS PROCESSED (b) AS STRAIGHTENED.

The magnitude of the residual stresses depends upon the processing conditions used. Both types of heat treatment give compressive stresses in the rail head and tensile stresses in the rail foot. Increased understanding has led to considerable modification of stresses within these processes but these are not discussed here.

b) **Rail Condition After Straightening**

The rail is passed through staggered rollers to plastically deform and straighten the product by a combination of bending, shear and roll contact stresses. This operation is of particular importance because it is the final deformation process of the production route which determines the residual stress pattern.

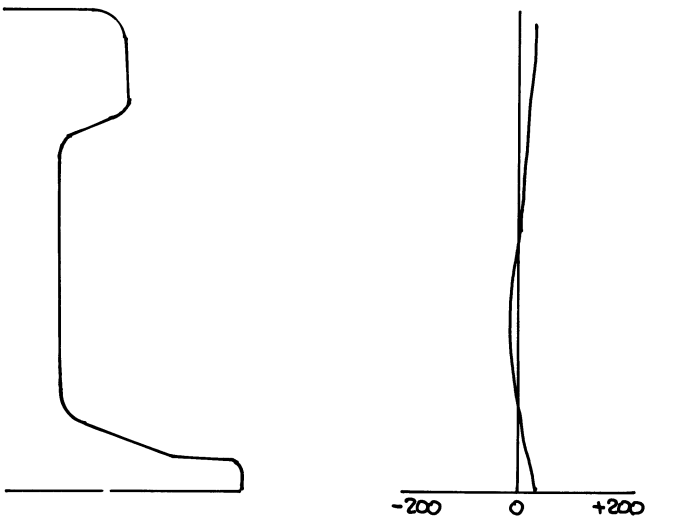
Although an over simplification, the dimensional changes produced by roller straightening result in a rail with the head and foot preferentially shortened with respect to the web. Longitudinal residual tensile stresses are produced in the head and foot and compressive stresses in the web of 100-300 N/mm2 for plain carbon and alloy rails. The corresponding effects on heat treated rails are illustrated in Figs. 4(b) & 5(b), where roller straightening has converted residual compressive stresses to smaller tensile stresses near the rail running surface and reduced the magnitude of the compressive in the web. Vertical stresses in the head and web of the rail have also been measured but have relatively low values compared to the longitudinal stresses.

The problem of unfavourable residual stresses from roller straightening has received wide attention. Apart from the rail material and geometry, the most important factors are the roll loads and displacements inside the straightener, the spacing of the rolls and the correct sequence of heavy and light passes through the machine.

The best solution appears to be based on the optimisation of the overall process of rail manufacture, rather than exclusive reliance on the roller straightening operation for the redistribution of residual stress.

c) **Other Straightening Methods**

An alternative method of straightening is stretch straightening which involves pulling the rail such that strains in excess of the yield strain occur across the cross section and along its entire length.



THE EFFECT OF STRETCHING ON RESIDUAL STRESSES IN GRADE A RAIL.

Figure 6.

Both the French rail maker Unimetal and British Steel in collaboration with British Rail have stretched rail on an experimental basis (7, 8). In the British work a permanent strain of 0.8% was the minimum required to produce a rail of reasonable straigthness with essentially no residual stress, as shown for BS11 Grade 90A rails in Fig. 6.

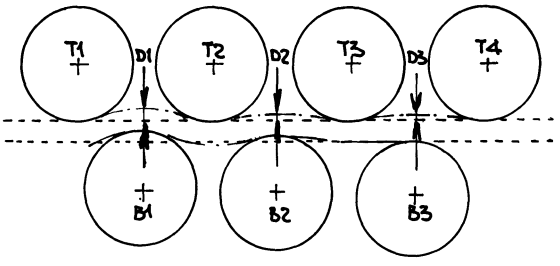
The method is slow and due to the effects of the grips there is additional yield loss, also tight control of the process is necessary to avoid tensile fracture of high strength rails. The process has only a minor effect on rail waviness and does not correct the crown profile. It is therefore at best a very expensive addition to the roller straightening process.

4) **Control of Stress During Roller Straightening**

As indicated earlier roller straightening alters the stress within the rail, it can however, both increase and decrease residual stresses and British Steel has carried out considerable tests to evaluate and control the situation.

A very considerable amount of finite element modelling has been done but this will be reported at a later date. Without such modelling the roller straightener effect on residual stresses is difficult to follow and much "black art" has been used. Because of this and the lack of sufficient proof of logic within the "black art" British Steel used an experiment design technique based on. Taguchi to unravel the situation.

The roller straightening configuration used was single plane as shown in Fig. 7. The trials were designed initially to determine the effect of roller straightening deflections on the web saw opening tests on 136lb rails.



ROLL ARRANGEMENT - LAMBERTON STRAIGHTENER SHOWING DEFLECTIONS D1, D2, & D3.

Figure 7.

The rails had been off-line heat treated and showed an average web saw opening of 2mm prior to the straightening trial. The rails were also tested by a "nick and break" bending test to determine whether subsequent fracture was likely to be horizontal along the web or vertical. The settings were as shown in Fig. 8.

DEFLECTIONS 1/16 inch		WEB SAW ORIGINAL	WEB SAW	BREAKAGE
B1	B2	1.5 - 2.0mm	mm	
12	8		3.69	Vertical
12	7		3.49	Vertical
12	6		4.17	Vertical
10	8		3.72	Vertical
10	7		3.67	Vertical
10	6		3.53	Vertical
8	8		3.37	Vertical
8	7		3.35	Vertical
8	6		2.92	Vertical

B1 = 58% contribution to stress levels.
B2 = 1%
Original web saw = 22% contribution to stress levels.

Figure 8.

Mathematical analysis of the results showed B1 setting to have a 58% contribution to stress level and ingoing stress to have a 22% effect.

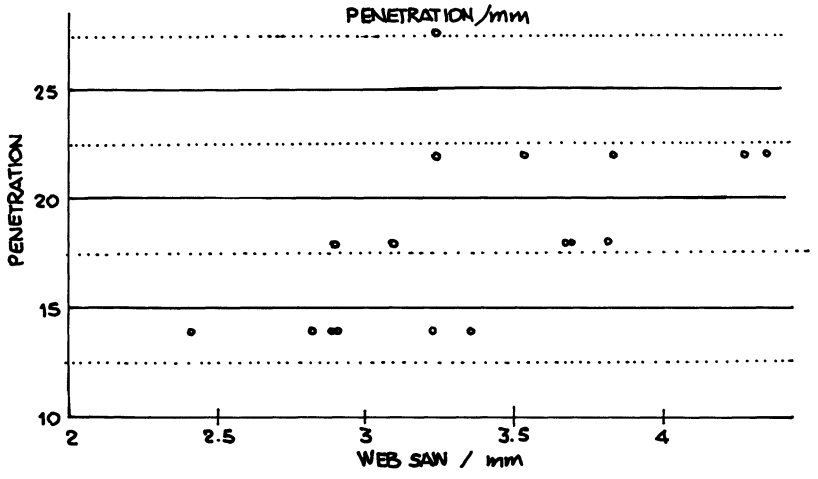


Figure 9.

The results of a second series of tests omitting initial stress (but still around 1.5 - 2.0mm) showed B1 again to have the major effect with B2 and B3 marginal. By doing successive trials it is possible to arrive at the lowest final stress state for any ingoing stress on a particular rail section Fig.9.

	B1 (1/16ths	B2 of an	B3 inch)	Simulated Breakage	Web Saw Opening (mm)
1	13	9	5	Clean break in RSM	3.4
2	13	7	4	Unknown	3.7
3	13	5	3	Split web	4.7
4	11	9	4	Clean break in RSM	3.1
5	11	7	3	Clean break	4.0
6	11	5	5	Clean break	4.7
7	9	9	3	Clean break	3.9
8	9	7	5	Split web	4.5
9	9	5	4	Split web	5.0

B1 % contribution = 17%
B2 % contribution = 79%
B3 % contribution = 4%

Figure 10.

For rail with high ingoing residual stress a different final indication was achieved. Rails produced by the same off-line process were roller straightened to give a web saw opening of 5mm. The trial and breakage tests were repeated Fig. 10. This time the controlling factor was B2 which controlled 79% final of stresses. Web saw reductions of between 0 and 2mm were obtained on the single trial. Again repeat trials were required to give optimum results but the work shows that roller straightening can considerably reduce stress. A second straightening can be advantageous.

During the trial rails of three different sections were given the same straightening procedure. Strain gauge and web saw opening testing was carried out on the straightened rails. Fig. 11 gives a selection taken from the results and this indicates that other things being equal an increase in rail height progressively increases roller straightener stresses.

Rail	Head K1 -N/mm	Web A2 +N/mm	Foot F1 -N/mm
BR113lb	102	162	167
UIC60 Kg	171	169	200
136lbs CN	166	191	264

Stress variation with section for fixed roller straightening deflection.

Figure 11.

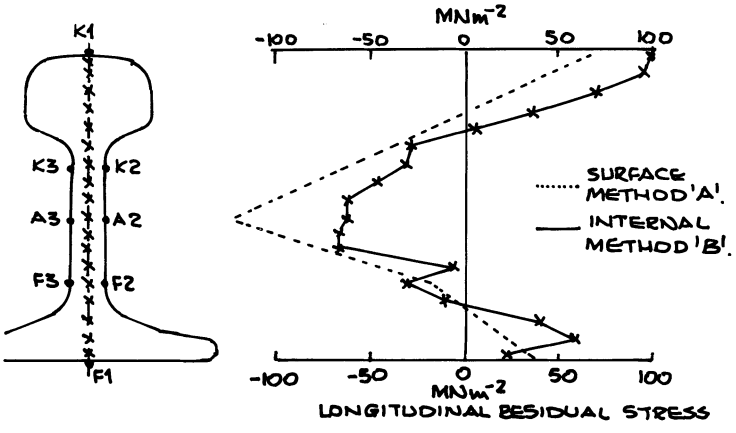
Mathematical modelling had indicated that second plane straightening would reduce web saw opening. A further trial was carried out to investigate this Fig. 12.

This work showed that second plane straightening under optimum settings could give web saw reductions of over 40%.

	Actual Deflections (mm)			Web Saw Opening (mm)		
Run	B1	B2	B3	before	after	
1	18	10	1	2.14	1.82	-0.32
2	18	7	-1	2.52	1.76	-0.76
3	18	4	-3	2.38	1.23	-1.15
4	14	10	-1	2.23	1.94	-0.29
5	14	7	-3	2.41	1.76	-0.65
6	14	4	1	2.26	2.13	-0.13
7	10	10	-3	2.20	1.71	-0.49
8	10	7	1	2.29	2.19	-0.1
9	10	4	-1	3.31	2.67	-0.64

Figure 12.

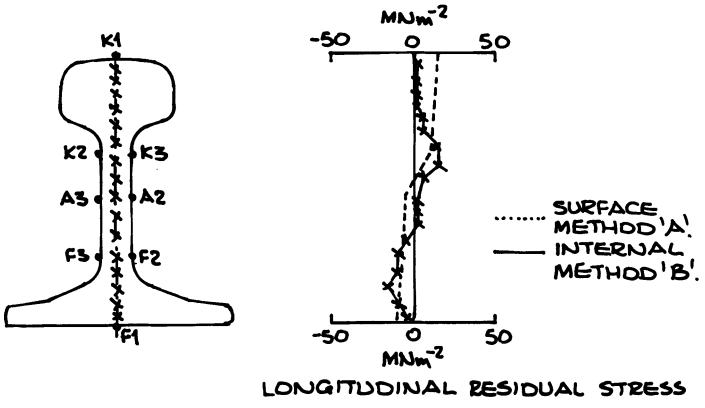
5) **Residual Stress Changes With Service**



LONGITUDINAL RESIDUAL STRESSES DETERMINED BY METHODS 'A' & 'B' IN UNUSED ROLLER STRAIGHTENED GRADE 90A RAIL STEEL.

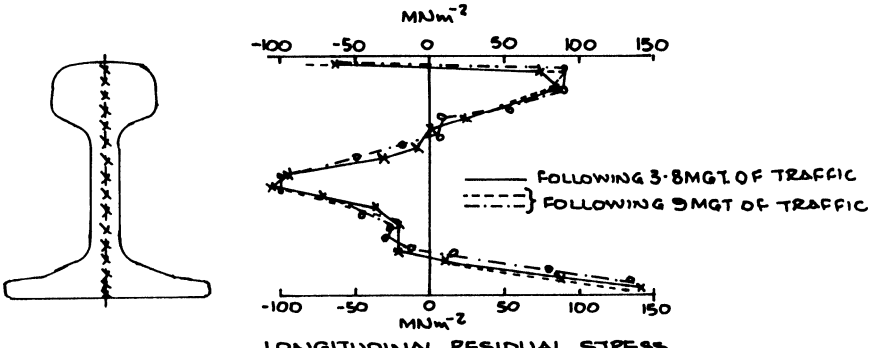
Figure 13.

Work carried out earlier for ORE showed that roller straightened rails Fig. 13 and stretched straightened almost stress free rails Fig. 14, under went considerable changes in track Fig. 15 & 16. The most obvious was the development of compressive stress at the top of the head.



LONGITUDINAL RESIDUAL STRESS DETERMINED BY METHODS 'A' & 'B' IN UNUSED STRETCH STRAIGHTENED (0·8%) GRADE 90A RAIL STEEL.

Figure 14.



LONGITUDINAL RESIDUAL STRESSES DETERMINED BY METHOD 'B' IN ROLLER STRAIGHTENED GRADE 90A RAIL AFTER 3·8MGT. & 9MGT. OF TRAFFIC.

Figure 15.

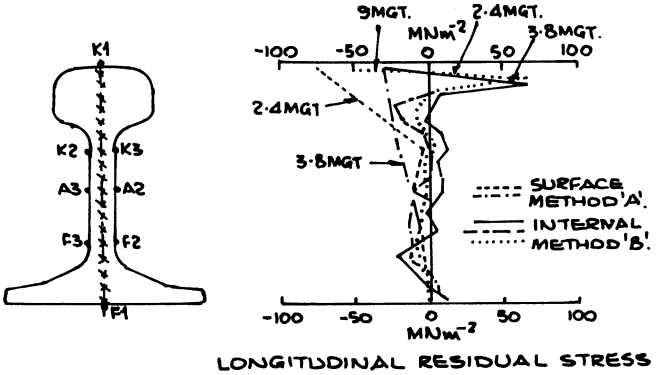
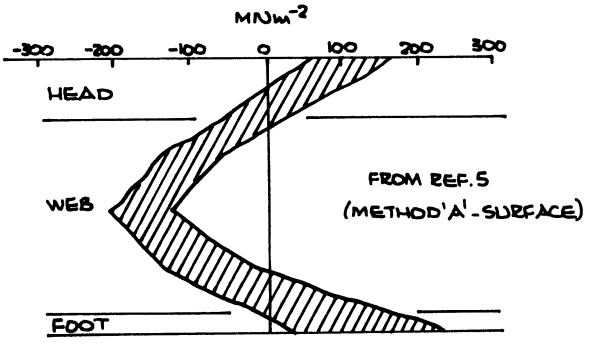


Figure 16.

LONGITUDINAL RESIDUAL STRESSES DETERMINED BY METHODS 'A' & 'B' IN STRETCH STRAIGHTENED (0·8%) GRADE 90A RAIL AFTER TRAFFIC LOADING.

Figs. 17 & 18 show the overall scatter band changes for head, web and foot between new and used rails.



SCATTER BAND OF LONGITUDINAL RESIDUAL STRESSES MEASURED BY ORE D156 IN UNUSED ROLLER STRAIGHTENED GRADE 90A RAILS.

Figure 17.

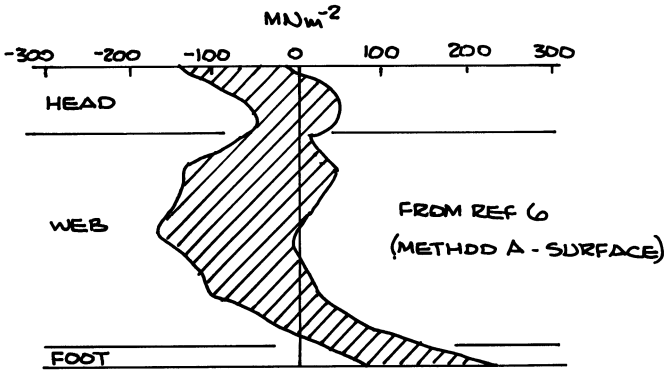
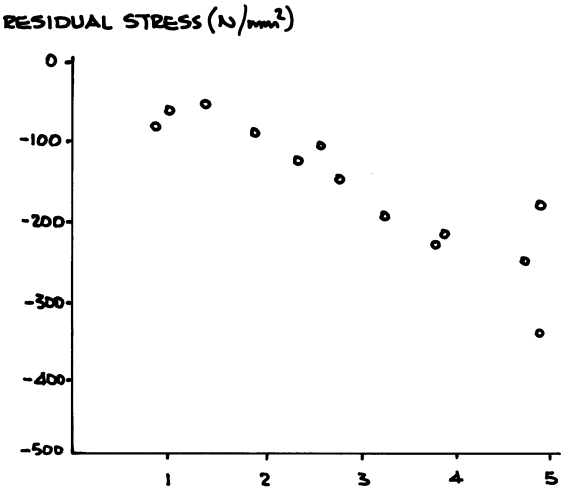


Figure 18.

SCATTER BAND OF LONGITUDINAL RESIDUAL STRESSES MEASURED BY ORE D156 IN USED ROLLER STRAIGHTENED GRADE 90A RAILS.

6) **Residual Stress Measurement For Service**

The work done by British Steel indicated that different tests were required for different service problems. The work initially compared Debro, whomper, web saw and strain gauge testing. Debro was dropped at an early stage because of inaccuracy and machine unserviceability. The indications were that the web saw and strain gauges gave consistent results. The whomper gave variable but still usable results. Comparisons between web saw opening work and strain gauge work showed good web agreement Fig. 19.



LONGITUDINAL WEB STRESSES AND GAP OPENING IN WEB SAW TEST.

Figure 19.

However, web saw and foot strain gauges gave little correlation. Therefore it was considered that when web stresses were important the web saw could be used as a general and relatively cheap test. Unfortunately for foot fatigue, only the slow and expensive strain gauge test gave adequate indication of stresses.

Fortunately as the roller straightening machines can now be fully instrumented and their role in control is better understood quality assurance procedures can be used which greatly reduce the need for repeat testing.

7) Summary

Residual stresses within the new rail are known to influence service behaviour. They are caused by thermal and mechanical deformation. Attempts must be made to control thermal strain during manufacture. To a large extent this can be achieved. However, roller straightening is the only means of accurately straightening rails for the foreseeable future and this process can have a very significant influence on final stress.
Trials carried out by British Steel have shown that straightening machines can be set to maintain low stresses to some degree or even to reduce already high ones. Particularly compressive web stresses can be controlled by accurate settings and by secondary straightening. Unfortunately, at this time foot stresses remain at levels too low for maximum fatigue life, particularly on tall rails. The work is continuing.

8) References

1. ORE Report No. C53/R86/E 'Behaviour of the Metal of Rails and Wheels in the Contact Zone', October 1970.

2. Allen, R.J. and Moreland, G.W., 'The Significance of Defects in BS11 Rails: Bolt Hole Failures and Tache Ovales', ISI Conference on rail Steel developments, London, 1972.

3. GOST 18267-82 'Through Hardening in Oil of Rails, Types P50, P65 and P75, for Wide Gauge railways' USSR, 1982.

4. Steele, R.K. et al, 'Catastrophic Web Cracking of Railroad Rail', AAR Paper C-I-RP-1142 May 29, 1990, Chicago, USA.

5. Cannon, D.F. et al, 'Neutron Diffraction Determinations of Residual stress Patterns in Railway Rails', Int. Conf. 'Residual Stresses in Rails: Effects on Rail Integrity and Railroad Economics', Cracow, 3-5 April 1990.

6. Deputat, J., Szelazek, J. and Kwaszczynska-Klimek, A., 'Experience in Ultrasonic Measurements of Residual Rail Stresses', ibid.

7. Deroche, R.Y. et al, 'Stress Releasing and Straightening of rails by Stretching', Paper No. 82-HH-17, Proc. 2nd Int. Heavy Haul Conf., Colorado, USA, 1982.

8. Cannon, D.F., 'An Evaluation of Stretch Straightened Rail by British Rail and the British Steel Corporation', Report No. 2301/25461/1, ORE Committee, D156, Sept. 1986.

RAIL MAINTENANCE

STRATEGIES FOR MAXIMIZING RAIL LIFE

R. K. Steele, Association of American Railroads
Chicago Technical Center
Chicago, IL 60616

ABSTRACT:

Strengthening of pearlitic rail steel and use of lubrication have been the traditional methods of maximizing rail life. The future seems to hold potential for microstructural modification and profile control/metal removal. Improved resistance to wear and rolling contact fatigue is possible with high hardness, near eutectoid, bainitic microstructures. Alternatively, similar results can be achieved by lowering the carbon content significantly and raising the alloy content. Reduction of oxide content will improve shelling performance while reduction of sulphide content will improve dry wear resistance and perhaps ductility and impact toughness. Rail grinding profile control can influence the contact stresses and the creep forces responsible for head checking and corrugation. Care must be taken with profile selection to avoid concentration of work hardening on the gage corner to minimize shell occurrence.

INTRODUCTION:

The continuing trend toward greater wheel loads[1] in North America necessitates a continuing effort to coax more life out of rail. This necessity applies to both new rail received from the steel mill and older rail already laid in track. The parameters which control the life of rail are wear, plastic deformation including corrugations, and fatigue, both of the rolling contact variety and the internally initiated type. Much has been written[1][2][3] about means of improving rail performance. The entire topic cannot, of necessity, be reviewed in the limited space of this paper. However, some the newer innovations will be addressed in some detail.

Historically, the two most traditional strategies for extending rail life have been strengthening of the rail steel by alloying and/or heat treatment, and lubrication of the wheel/rail interface, particularly the gage face/flange interface. Both strengthening and lubrication enhance wear and deformation resistance while strengthening offers the additional advantage of improvement in fatigue resistance. At the present time, we probably are approaching the maximum strengthening achievable commercially while retaining the *preferred* pearlitic microstructure. In a similar fashion, the benefits of lubrication have been well recognized. But, our lack of understanding of how lubricants achieve their benefits limits significant

[1] *Over the next several years unit train maximum wheel loads are expected to climb from 33 kips to nearly 36 kips. Some cars with 39 kip wheel loads currently are in service.*

J.J. Kalker et al. (eds.), Rail Quality and Maintenance for Modern Railway Operation, 77–97.

advancements in this area. The engineering problem has devolved to finding the best combination of lubricants and application systems. Little more will be said of this here except to note that lubrication interacts with other strategies and to say that perhaps there are more permanent methods possible to achieve wheel rail lubrication than those currently utilized.

Having so cavalierly dismissed the traditional approaches, what remains? The answer appears to this writer to be microstructural modification and carefully controlled metal removal and profile control.

MICROSTRUCTURAL MODIFICATION:

This strategy has two component parts: replacement of the matrix pearlitic microstructure with some other constituent such as bainite or tempered martensite and (b) modification of the non metallic inclusion content.

Most usually the attempts to introduce alternative microstructures[4][5][6][7] have been applied to near eutectoid composition carbon rail steels. The results of such efforts have been to confirm that pearlite of the same hardness as the alternative constituent microstructure has significantly better wear resistance. Kalousek etal[7] propose that the maximum useful hardness of conventional rail steel for wear resistance is near HRC38-40. More recent work by Clayton and Devanathan[8] has shown that chromium molybdenum rail steel heat treated to microstructures of upper and lower bainite, i.e. hardnesses near 49 and 59 HRC, exhibits remarkable resistance to Type III severe wear at very high contact pressures (1575, 1645 MPa) with very large creepages (35,50%). Thus, the development of alternative constituent microstructures in conventional rail steels may be beneficial if the hardnesses are high enough.

There is, however, an alternative approach to the use of conventional near eutectoid composition rail steels. And that is the use of low carbon steels containing significant amounts of chromium (~ 1.6 w/o) and nickel (~ 1.2 w/o). Work on this approach by Devanathan[9] has shown that a low carbon steel (0.04 w/o) having a hardness of only HRC 28 could have wear and rolling contact fatigue resistances better than those of premium pearlitic steels. The main difficulty with such a low carbon composition was its initially high plastic deformation rate. After a period of cycling, however, the deformation rate dropped significantly.

The second approach to microstructural modification has focused on the non metallic inclusion content. Generally the non metallic inclusions can be divided into two categories, (a) those (oxides) which are not readily deformed at hot working temperatures and (b) those (sulphides) which are deformed in the hot working process. The oxides, most especially aluminum or silicon based oxides[2] are associated with internal fatigue crack initiation[10][11][12]. The sulphides, on the other hand, are associated with wear[13][14] or ductility exhaustion[15] related processes.

It is easy to argue that inclusions have a deleterious effect on rail performance and therefore their presence should be minimized. But the essential question is how much are the railroads willing to pay for the reduction of inclusion

---

[2] *In North America the most troublesome oxides generally are of the complex calcium aluminate type suggesting their occurrence from exogenous sources in the steel making practice.*

content. To answer that question, one needs to know the relationship between inclusion content and rail performance. Work in Australia[16] has related the oxide inclusion diameter to the defect occurrence rate. Below a threshold near 100μm, relatively small changes in inclusion diameter have large effects on defect rate. Tests conducted at the Facility for Accelerated Service Testing (FAST) under 100T capacity cars have suggested that both the volume fraction of aluminum oxide and the stringer length have a direct effect on the shell defect rate[17]. Hardness has been shown to have an inverse effect. Figure 1 shows the shell defect rate as a function of the product of oxide volume fraction and stringer length with the defect rate corrected[3] for hardness variations. Somewhat embarrassingly, the data fall remarkable well along a straight line having a slope of approximately 0.34 (d log (defect rate)/d log (cleanliness parameter)=0.34). Remarkably, studies of steel cleanliness upon roller bearing life[19] have yielded a log life v.s. log inclusion length slope of approximately - 0.33. The roller bearing life studies suggest that modification of the morphology of the inclusions by use of rare earth additives may prove more effective than relatively more costly efforts to reduce the overall oxide content. The low slope value means that large changes in the amount and/or cluster length of oxides must be made to produce only modest improvements in fatigue life and reductions in defect rate.

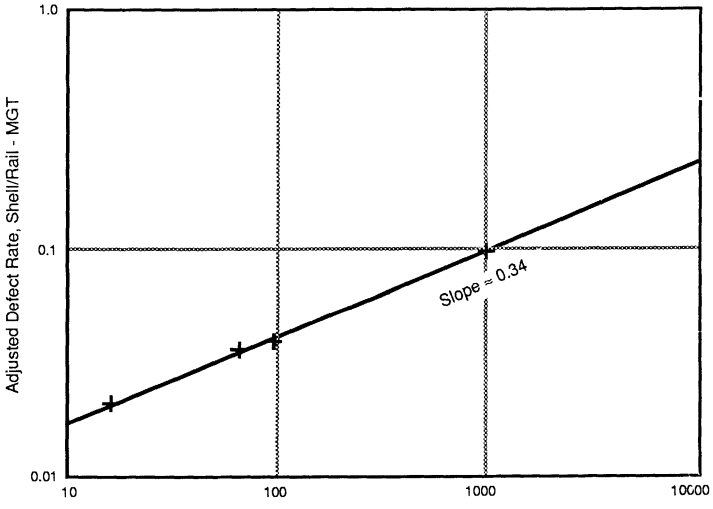


Figure 1: RELATIONSHIP OF ADJUSTED (FOR HARDNESS VARIATIONS) DEFECT RATE AND THE OXIDE CLEANLINESS PARAMETER

[3] *The adjustment to the defect rate was made using S/N data for rail steels of different hardness[18] and assuming that S/N life was inversely proportional to the defect rate.*

The similarity of behavior of rail defect rate and bearing life as functions of metallurgical cleanliness is reassuring. But Clayton[17] has observed that the FAST rail defect data also suggest that the rail steels tested which were metallurgically cleaner exhibited a higher tendency to have those shells which did form turn to transverse type defects. This behavior is illustrated in Figure 2. The unfortunate implication of this, if it really is true, is that by reducing oxide content, we perhaps will have replaced a mostly benign defect (the shell) with a dangerous defect (the detail fracture). The data of Sugino etal[20], have been examined to see if there had been any tendency for metallurgically cleaner rails to develop a greater proportion of transverse defects; such tendency has not been found.

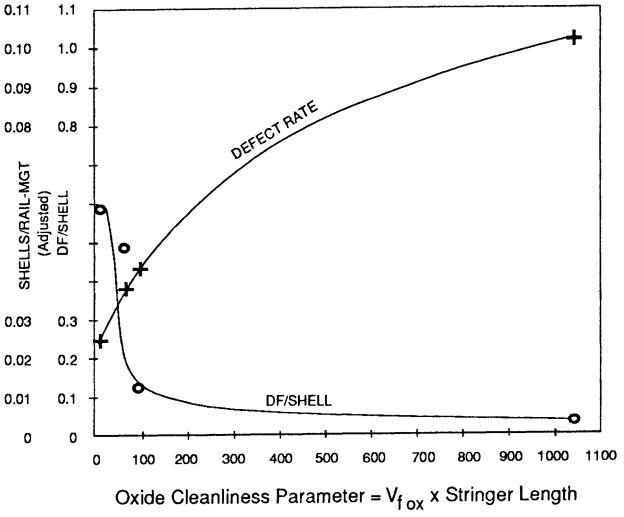


Figure 2: VARIATION OF THE ADJUSTED DEFECT RATE AND THE DF/SHELL RATIO WITH OXIDE CLEANLINESS PARAMETER

As noted previously[13][14][15] sulphides have their effect on deterioration processes which involve significant levels of plastic deformation[4]. Generally the means by which sulphides control local fracture processes is by void initiation, growth, and coalescence. In the case of sulphides, decohesion between the inclusion and the surrounding matrix is the mechanism of void initiation. The plastic strain necessary for decohesion of a sulphide inclusion is believed to be negligible[21]. The laboratory studies of Kalousek etal[14] and the theoretical analysis of Clayton and Steele[15] suggest that a factor of 10:1 decrease in sulphide inclusion content should yield a 3.5 to 5:1 decrease in dry (Type III) wear rate.

There may be other beneficial effects of sulphide control as well. Although sulphides appear to have no effect upon internal fatigue crack initiation, recent research work with HY-80 ship plate steel[21] has shown that ductility and

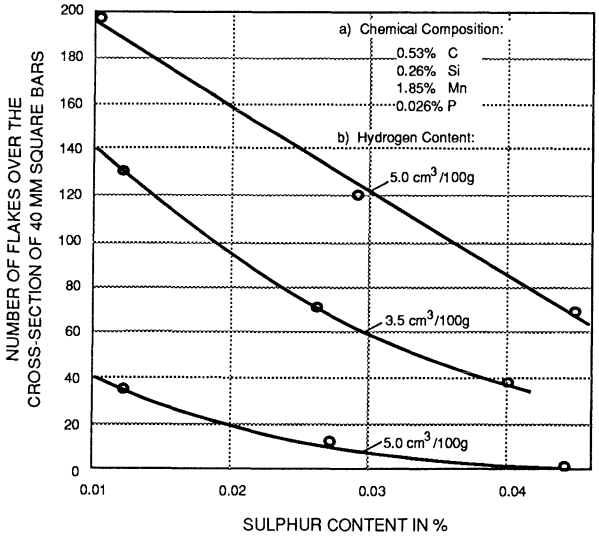
---

[4] *Debris generation from Type III (severe) wear occurs by cracking along shear deformation bands.*

charpy fracture toughness are adversely affected by the presence of sulphides. The fact that impact energy decreased as the inclusion aspect ratio increased suggests that there may be benefits to morphological adjustment using rare earth additions or alterations in the hot rolling practice. It is not clear to this writer that a systematic study of the effects sulphide content and morphology has ever been published for rail steels.

There remains in this writer's mind an intriguing question: does sulphide content play any discernible role in the rolling contact fatigue (RCF) process? Type III wear and RCF bear some resemblance metallographically to each other. Both involve cracking along the shear bands which result from plastic deformation under repeated wheel passages. In the case of Type III wear, the crack progression along a shear band is quite small (a mm or less) before a debris particle is produced. In RCF, the shear band crack growth may be much greater (many mm's) and the crack may extend through the work hardened region and turn to propagate into the base metal. At least for the shear band part of the propagation, this writer is inclined to wonder if sulphides accelerate RCF.

Whenever we think of sulphide reduction (i.e. reduction in the sulphur level) we should keep in mind the role that sulphur and manganese play in the hydrogen flaking problem. The interrelationship of hydrogen flake occurrence, and the hydrogen and sulphur levels is best illustrated in Figure 3 from the work of Heller etal[21]. The basic message is: if manganese is to be maintained at a high level (to achieve strengthening) and the sulphur level is to be reduced, then extraordinary measures must be taken to keep the hydrogen level low. Before undue haste is made to reduce sulphur levels further than those already achievable, perhaps we should have a clearer quantitative idea of the effects of such reductions on rail characteristics (specifically impact toughness, ductility, and RCF resistance).



SOURCE: REFERENCE 22

Figure 3: FLAKE FORMATION AS A FUNCTION OF SULPHUR WITH VARIOUS HYDROGEN CONTENTS FOR A GRADE C RAIL STEEL

METAL REMOVAL/PROFILE CONTROL:

In recent years considerable attention has been paid to the contact profile between the wheel and the rail. Grinding generally has been the means by which control of rail profile has been achieved. It can control three parameters:

(a) wheel/rail contact locations
(b) contact stresses
(c) fatigue damage accumulation

The three parameters are not usually affected independently of each other.

The introduction of asymmetric rail profiles[23] provided a means by which rolling radius difference (between the inside and outside wheels on an axle) could cause wheelsets to negotiate low curvatures ($\leq 3°$) without flanging. In the process, the approach prevented wheel contact on the smaller profile radius at the gage corner (high contact stress).

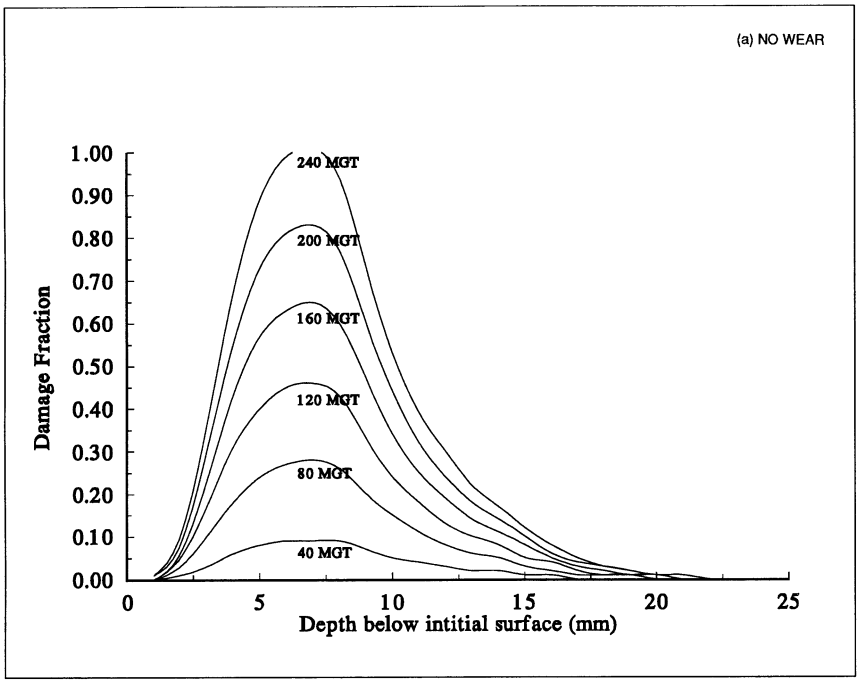
Surface removal of metal by grinding obviously has the ability to remove damaged metal resulting from rolling contact fatigue. Likewise corrugations are readily removed. What may not be so obvious is that the internal accumulation of fatigue damage leading to shell formation can, theoretically at least, also be controlled by shifting the point of maximum damage accumulation to greater depths. This is illustrated in Figures 4a and b for two grinding conditions, no metal removal (zero wear rate) and optimal metal removal (maximum fatigue life), input into the AAR three dimensional fatigue model, PHOENIX. At zero wear rate, the depth at which the maximum damage rate occurs remains fixed beneath the surface. But at the optimal wear rate, the depth at which the maximum damage rate occurs shifts to greater depths beneath the initial surface with increasing tonnage.[5] Figure 5 shows how the optimal grinding (wear) rate would be predicted to decrease as the fatigue life percentile increases. To this writer's knowledge, no systematic experiments have ever been made to show the metal removal from the surface does delay shell initiation and that the optimal grinding rate should decrease with increasing fatigue life percentile. Tests currently underway on the FAST HTL are intended to provide such verification but are not yet far enough along to do so. However, previous tests on the old FAST loop have suggested[24] that some dry running in conjunction with lubricated running could significantly lengthen the shelling fatigue life of rail in curves.

Turning now to the benefits of profile control, we should recognize that the first efforts in this regard seem to have been made on the wheel side of the W/R interface.[25][26] The intent of such work was to use rolling radius difference to help wheel sets steer themselves through curves and thereby minimize train resistance. Generally the approach recommended was to adopt a wheel throat profile close to that of a worn wheel. It is not clear that the effects of such a profile on rail fatigue behavior were ever considered although the effect on side wear of rail was expected to be beneficial.

The works of Lamson and Longson[23] and, somewhat previously, of Townend etal[27] have served to focus attention on the benefits of rail (head) profile

---

[5] *However, the depth for maximum damage rate beneath the current surface diminishes with increasing tonnage.*



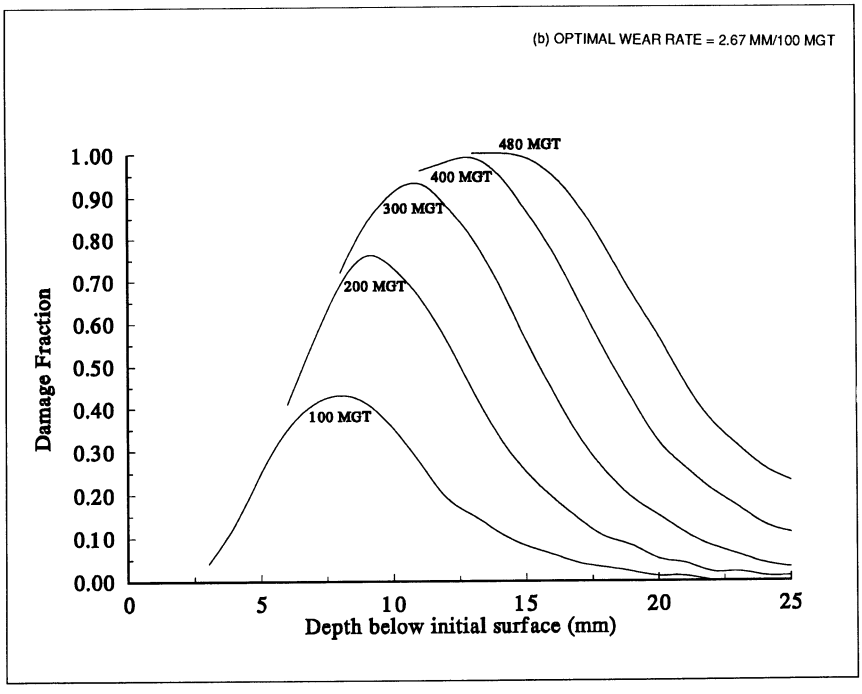


Figure 4: PHOENIX PREDICTIONS OF THE GROWTH OF THE FATIGUE DAMAGE FRACTION FOR NO WEAR AND FOR OPTIMAL WEAR CONDITIONS/FIRST PERCENTILE LIFE

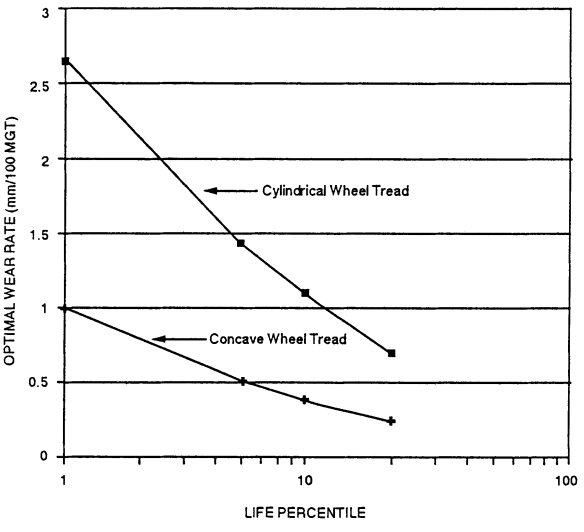


Figure 5: CHANGE IN OPTIMAL WEAR RATE WITH LIFE PERCENTILE

control. The asymmetric profile, applied in curvatures up to approximately three degrees, was able to limit side wear. This it did by positioning the high rail contact region slightly to the gage side of the center of the ball while the low rail contact was positioned slightly to the field side of the center of the ball. The approach is said[23] to be effective for control of "shellings[6] and corrugations.... for any degree of curvature."

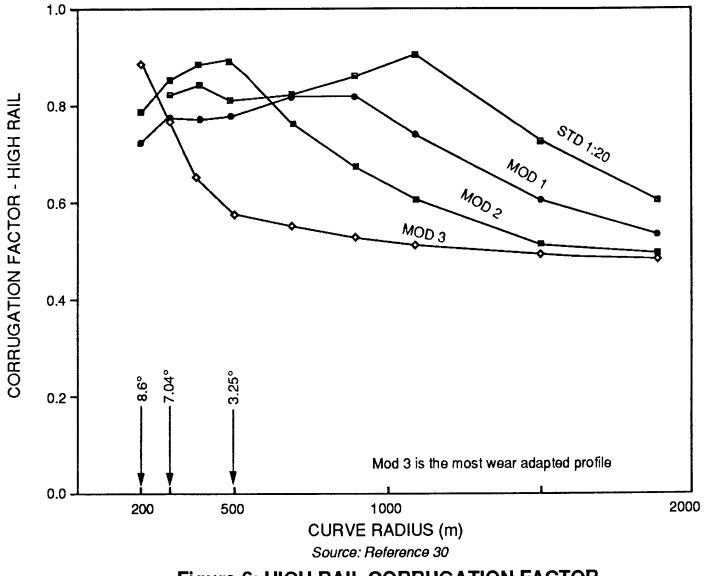
Kalousek etal[28] have described a number of different rail profile configurations developed on the basis of tests made on the Canadian Pacific Railroad and the British Columbia Railroad. The profiles required the removal of metal preferentially from the gage side of the ball on high rails (and on rail in tangent track) and from both the field and gage side of the ball on low rails. The amount of metal removed from the gage side of the ball of new rails depended on the original new profile[7] and increased with sharper track curvature. The basis for the selection of these profiles seems to have been the control of rolling contact fatigue damage that developed on the ball and at the high rail gage corner. No mention is made of side wear measurements nor internal fatigue crack occurrence (shelling).

Epp and O'Rourke[29] and Epp[30] have examined theoretically the effect of wheel profiles on both wheel and rail performance. In these studies, a series of wear modified wheel profiles as well as the old AAR 1:20 profile have been considered on partly worn 60 kg/m rail[26]. The predicted effects of wheel profile on the partly worn rail are perhaps best illustrated by Figures 6 and 7. In the first of

---

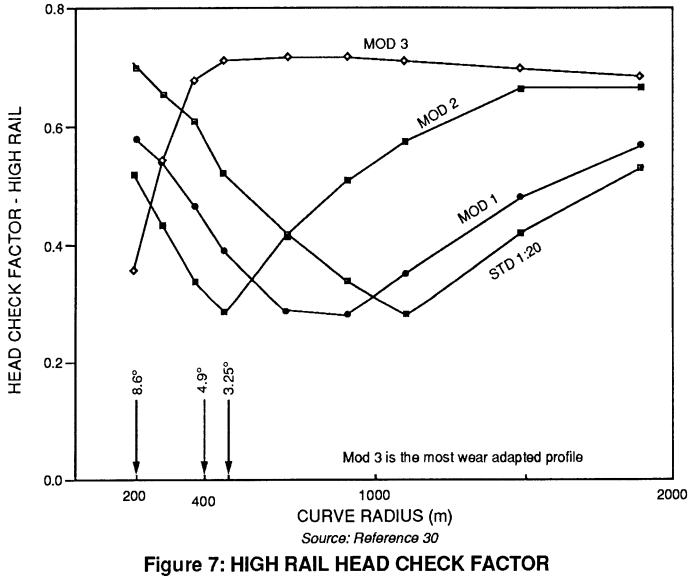
[6] *It is not clear whether the word shellings means internally initiated shells and/or surface initiated head check/spalls.*

[7] *The 14" crown radius 136RE section with 9/16" corner radii requires 50 - 100% more gage side removal than does the 10" crown radius 132RE section with 3/8" corner radii.*



Source: Reference 30

Figure 6: HIGH RAIL CORRUGATION FACTOR



Source: Reference 30

Figure 7: HIGH RAIL HEAD CHECK FACTOR

these, the lateral creep force component has been used to estimate corrugation tendency. The results suggest that a wheel profile most closely approaching a worn wheel provides the greatest resistance to corrugation occurrence, except at curvatures above about 7° where the most wear modified wheel profile exhibited somewhat higher corrugation tendency. Figure 7 portrays the effect of wheel profile upon head checking tendency calculated from longitudinal creep force. Here, the most wear modified wheel profile was predicted to have the highest head checking tendency up

to about 3-1/4° of curvature (500 m radius), above which the predicted tendency fell. Unfortunately, these studies do not make clear the degree of conformality of the most wear modified profile with the rail profile.

The FAST observations that dry running in conjunction with periods of lubricated running improved the fatigue life of rail in curved track and the predictions of the PHOENIX model have prompted a series of profile tests at FAST. In all these tests, great care was taken to assure that the gage face was very well lubricated. The focus of the FAST test was primarily on shell/DF occurrence and upon wear behavior because FAST has not ever been a prolific generator of RCF damage even though dry periods have created considerable head checking.

The results (Table I) from the first of these tests[17] showed, surprisingly, that the number of internally initiated defects i.e. shells, could be as high or higher in the section of a test curve ground to a two-contact profile[8] comparable to those used in revenue service. In addition, the gage face wear rates of the two-contact ground rails were found to be about two or three times those of rail which was left without grinding (Figure 8).

Table I: SHELLS FROM THE 33 KIP WHEEL LOAD HTL TESTS

Rail ID	Hardness BHN	Shells in 90 MGT			
		Section 25 (6°)			Section 3 (5°)
		Seg A	Seg B	Seg C	
A	284	1	4	0	11
B	285	43	143	9	19
C	290	0	0	0	0
D	296	0	11	11	0
E	302	0	0	0	0
F	309	0	0	0	0
G	321	10	2	0	0
H	324	0	0	0	0
I	343	0	0	0	0
J	371	-	0	-	0
K	390	0	-	-	-
L	294	-	-	-	4

Segment B high rail was two - contact ground
Segments A and C high rails were not ground at all.

Source: Reference 17

[8] *The profile ground onto the high rail removed metal preferentially on the gage side of the running surface to create two contacts, one on the gage side and the other near the center of the ball; the low rail was left unground.*

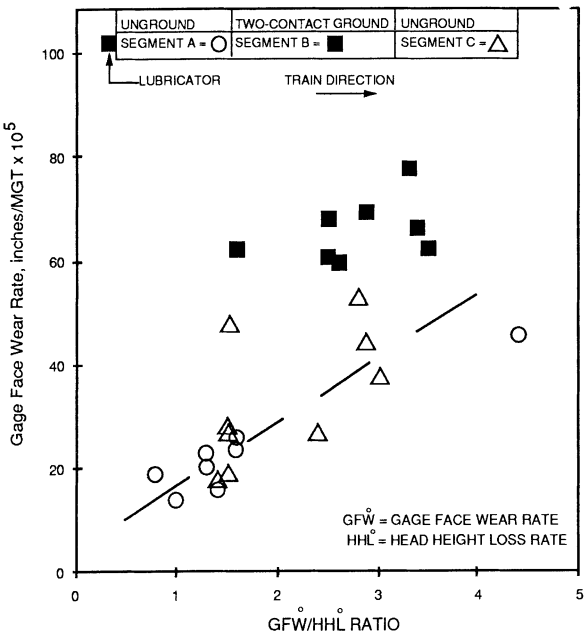


Figure 8: LUBRICATED WEAR RATES IN THE HTL (33 KIP WHEEL LOAD) TEST

The surprising results of the first experiment prompted a second in which the following four profiles were intended to be utilized:

(1) As worn, dry, subsequently ground conformally on the ball to the as-worn profile at 25MGT intervals

(2) Ground conformally on the ball to the FAST as worn profile and subsequently ground to the same profile at 25MGT intervals.

(3) Ground asymmetrically (two-contact) to relieve the gage corner typical of revenue service profiles and subsequently ground to the same profile at 25MGT intervals.

(4) As manufactured with no subsequent grinding.

The rail was 133RE, HB300. Although the first two segments were intended to be ground conformally to the FAST as-worn profile, this was not done initially. Instead, the profile of the third segment (two-contact) was ground onto the ball in the first two segments. Only after 50MGT was this situation corrected. The low rail was ground (at 25MGT intervals) to position the wheel path slightly to the field side of the center of the ball.

The results were again a surprise because the two contact asymmetric profile again exhibited the greatest number of shells as shown in Table II. One-half

of the rail in second, third, and fourth segments was B rail (from the top of the ingot) while that in the first segment was entirely C rail (more from the center of the ingot). Thus, it was thought that there might be a metallurgical cleanliness difference. But quantitative metallographic examinations of specimens taken from each rail suggested that the rail in the two-contact ground segment containing the greatest of defects was perhaps the cleanest rail metallurgically in the entire test. The shells which formed under the gage corner were biplanar in nature.

Table II: DEFECT AND RAIL INFORMATION FROM THE SECOND PROFILE EXPERIMENT (39 KIP WHEEL LOADS)

	TYPE PROFILE							
	DRY WORN CONFORMAL		GROUND CONFORMAL		TWO CONTACT		AS ROLLED	
DEFECTS/RAIL	0	0	3	6	0	13	0	2
CUT/INGOT I.D	C28	C30	C32	B32	C33	B13	C17	B14
ASTM E45 MICROCLEANLINESS RATING								
SULPHIDES	2.5	2.5	3	3	3	2	3	3
OXIDES	1	1	1	1	4	<1	1	2

In addition, rail side wear measurements were made in each segment. The first three segments initially had essentially the same gage face wear rates. But when the ball profile grinding pattern on the first two segments was corrected to yield a conformal profile the gage face wear dropped to about 1/10th of what it had been with the two contact type profile. This is illustrated in Figure 9.

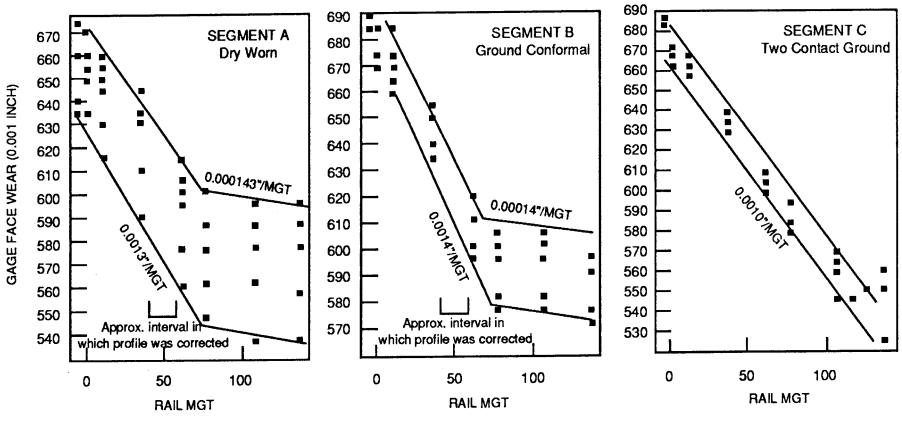
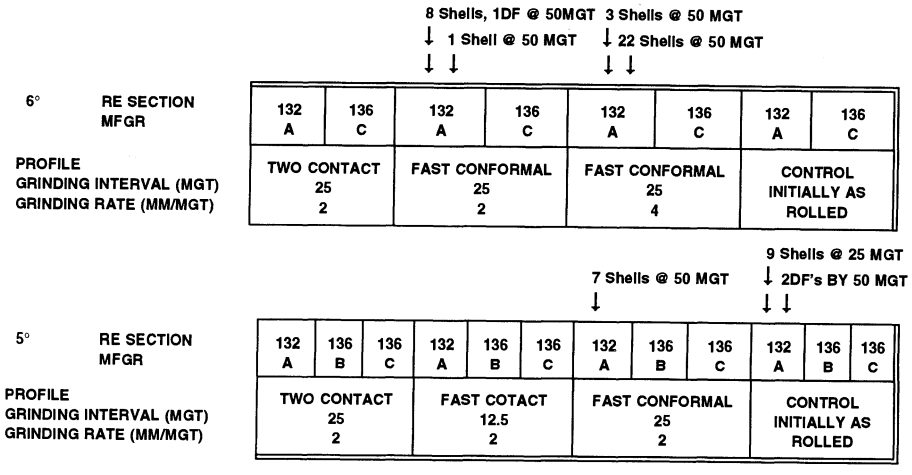


Figure 9: GAGE FACE WEAR IN THE PROFILE TEST ZONE HTL/HEAVY AXLE LOAD TEST

Recognizing the small scale of the second test, a grander test involving much larger quantities of rail in both 5° and 6° curves currently is underway. Standard carbon (~ 300 BHN) rails from three different manufacturers are incorporated into the test[9]. Three different profiles are utilized: (1) as-manufactured, (2) ground to conformality on the ball with the FAST worn wheels, and (3) ground to a revenue service profile essentially the same as that proposed by Kalousek etal[28] for sharp curves.

The first results have shown a rash of shell defects occurring at the gage corner of the conformally ground and the as-manufactured profiles (see Figure 10) for the 132RE rail. The cause for *this* surprising behavior apparently arises from a change in the average wheel profile shape just after the start of the test[10] such that many wheels in the train now were not at all conformal with the ground rail profiles which they were supposed to match. The contact conditions with the newly introduced wheels tended to carry the wheel load on the gage corner of the rail. The tendency for this to occur was greater for the newly ground 132RE in contrast to the newly ground 136RE rails (Figure 11).



Source: Jon Hannafious/Transportation Test Center

Figure 10: LAYOUT OF THE THIRD RAIL PROFILE TEST WITH LOCATIONS OF RAIL DEFECTS DEVELOPED BY 50 MGT

[9] *The rails from two of the manufacturers were 136RE while those from the third was 132RE.*

[10] *A derailment of the FAST train slightly after the start of the test caused the replacement of about1/3 of the wheels in the train with revenue service worn and AAR1B wheels. These wheels had more metal in the throat region than the customary FAST worn wheel.*

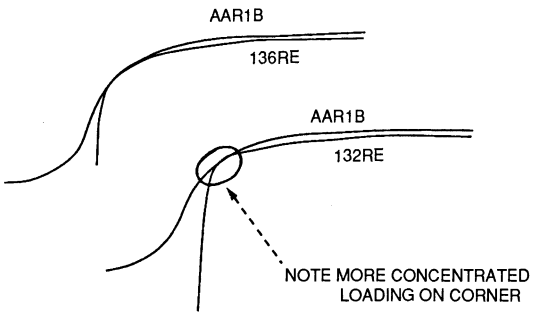


Figure 11: COMPARISON OF CORNER LOADING OF CONFORMALLY GROUND 136RE AND 132RE RAILS

The shelling behavior of ground rail (with both the two contact and the conformal profiles) as opposed to the apparent shelling resistance of the dry worn rail has prompted an examination of the plastic deformation distribution within the rail heads. Hardness maps have been made on cross sections of a number of rails from the FAST HTL (High Tonnage Loop) test as well as from rails from revenue service exhibiting shell occurrence. Figure 12 shows the concentration of work hardening that has occurred in the vicinity of the gage corner for a two contact profile (a) and a FAST conformal profile (b) upon which revenue service worn (and AAR1B) wheels had run. Figure 13 shows the more broadly distributed work hardening in a dry worn rail which exhibited no shells (c) and two 136RE rails (a and b) conformally ground in the current experiment that have exhibited far greater shell resistance than the 132RE test rail. The dry worn rail exhibited significantly deeper work hardening (by about 50%).

Several instances of unexpected shelling have occurred in revenue service rails (high rails of curves) which have been ground to two contact type profiles.[11] Those that have become available to us have been hardness mapped. Figure 14 illustrates the concentration of work hardening that has occurred in the vicinity of the gage corner in each case. The shell/DF occurrence on the fully heat treated rail (b) was particularly surprising because it had been in track (6° curve) for only 35MGT (from installation and grinding).

The means by which concentration of work hardening into the corner region of a rail facilitates shell/DF initiation is not exactly clear, at least to this writer. To better understand the problem, X-ray diffraction methods have been

---

[11] *Post mortem examination shows that grinding marks remain on the gage side of the ball thereby indicating that the initial profile was two-contact. But the exact configuration of the two contact profile cannot be inferred from the post mortem examination.*

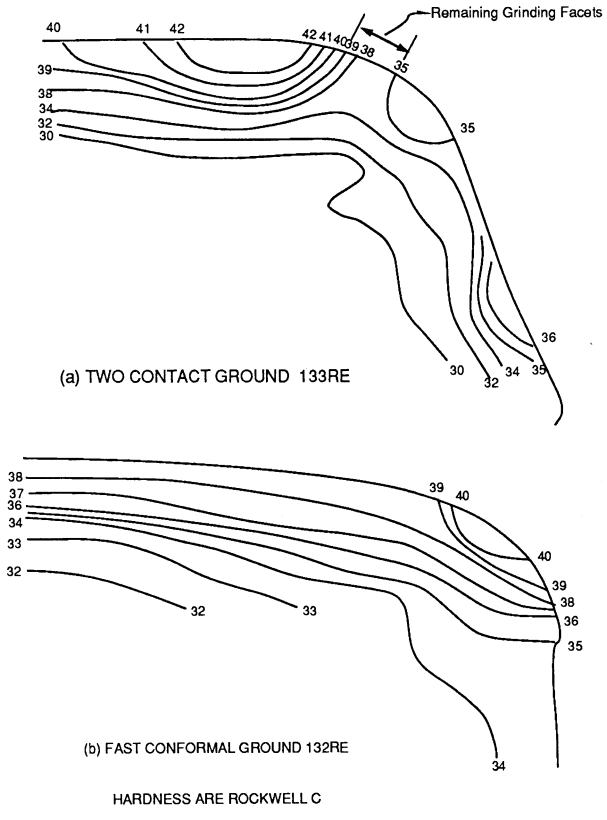


Figure 12: DEFORMATION PATTERNS OF TWO FAST HIGH RAILS WHICH EXHIBITED UNEXPECTED SHELLING

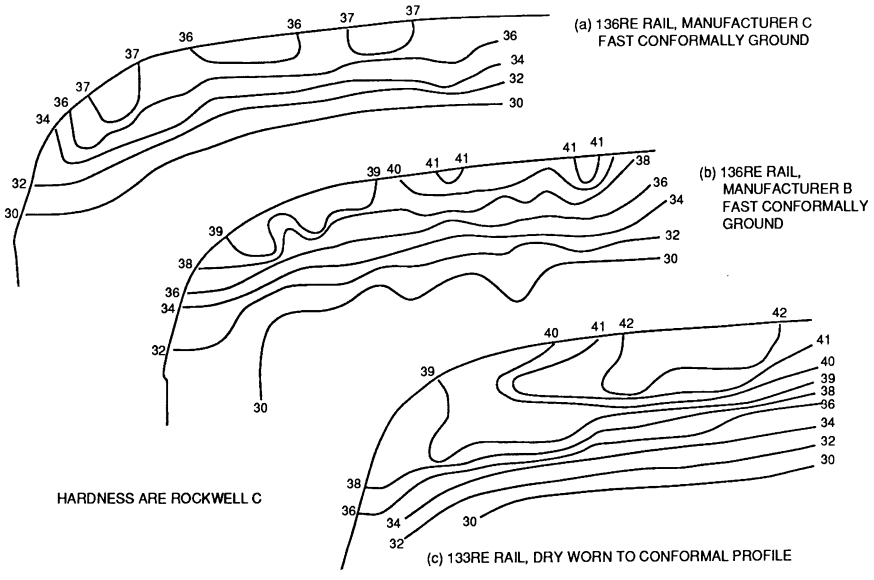


Figure 13: DEFORMATION PATTERNS OF THREE FAST HIGH RAILS EXHIBITING GREATER RESISTANCE TO SHELL/DF OCCURENCE

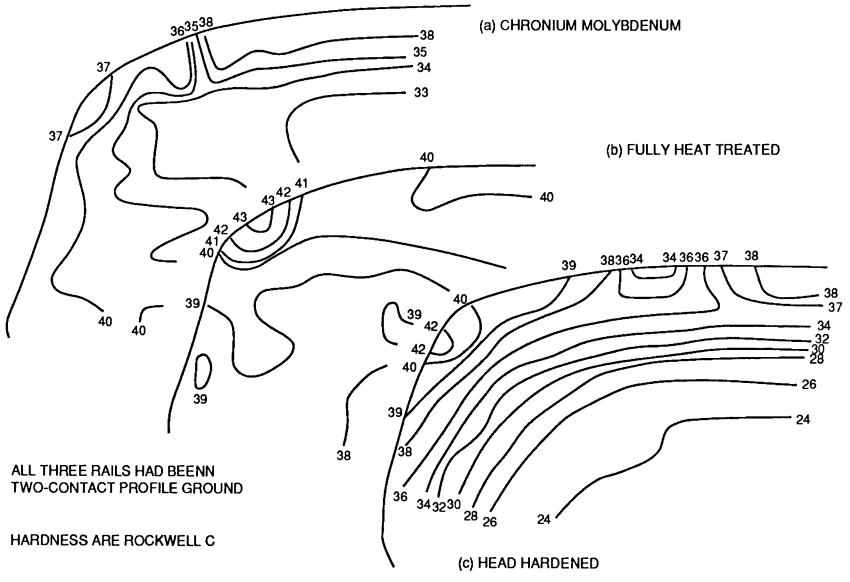


Figure 14: DEFORMATION PATTERNS OF THREE REVENUE SERVICE RAILS EXHIBITING UNEXPECTED SHELL/DF OCCURRENCE

applied[12] to determine residual stresses on transverse cross sections taken from the two FAST rails, one dry worn (which did not shell) and one two contact ground (which did shell profusely). The residual stress contours for the vertical residual stresses are shown in Figure 15. Those under the corner of the dry worn rail (a) were very much like those that have been reported by Groom[31] for unground rails taken from revenue service. Eventually those residual stresses would be expected to contribute to shell initiation[32] usually with the shell predicted to form at a location of slight tensile residual stresses. The two contact profile (b) appears to have a very different configuration of vertical residual stresses indeed. The location of shell initiation appears to be very near the zero residual stress contour at a hardness level near HRC33. At that same location under the corner of the dry worn rail, the hardness is a much more fatigue resistant HRC39. Even were a shell to form just to the tensile side of the zero stress contour under the corner, the hardness of the origin region would still be HRC 38 - 39.

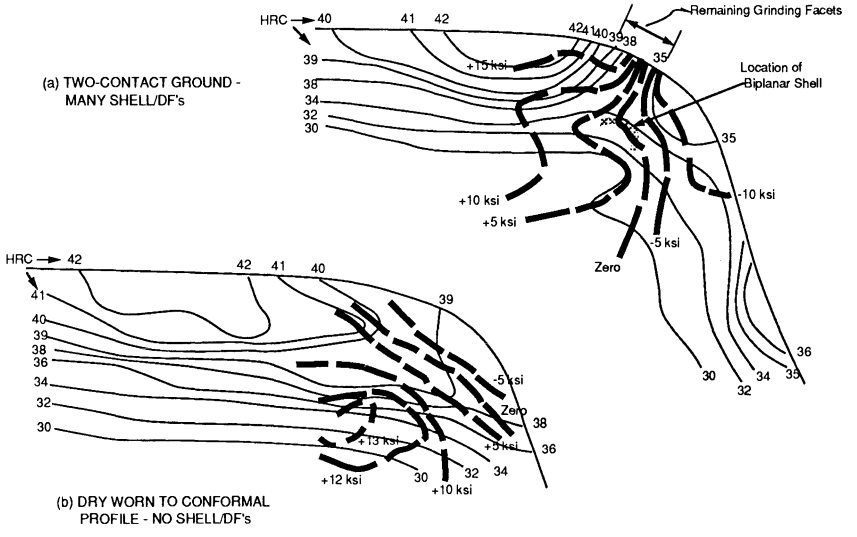


Figure 15: VERTICAL RESIDUAL STRESS CONTOURS AND HARDNESS CONTOURS FOR TWO RAILS FROM FAST HTL TESTS

CONCLUDING REMARKS:

The railroad civil engineer is close to being able to tailor the material characteristics of the rail to meet the performance demands placed upon rail under different service conditions. Our inability to specify the contact conditions with exactness and our, as yet, incomplete knowledge of how (and why) rail materials respond to those conditions as they do are remaining barriers to achievement of

[12] *Studies conducted at the Technology for Energy Corporation, Knoxville, TN.*

tailored performance. Still though, today there are many more manufacturing options than have ever existed before so the tools for tailoring the performance are pretty much available.

Improvement in the non metallic cleanliness and possible modification of the morphological character of inclusions of rail steels offer the potential to reduce internal fatigue crack initiation (aluminum oxide modification) and improve wear resistance (manganese sulphide modification) and possibly toughness and ductility as well. The change of microstructure from pearlite to bainite offers the possibility to far greater wear resistance under very heavy wheel loads, but the hardnesses must be very much greater than those customarily used (near eutectoid carbon steels) or the carbon content must be reduced significantly while the alloy content is increased.

Grinding is an important tool by which the wheel contact location and level of contact stress can be controlled. In addition, the controlled removal of surface damaged metal and the repositioning (over time) of critical damage accumulation regions within the rail can delay the initiation of both rolling contact and internal fatigue cracking. However, the choice of what wheel and rail profiles work best together is dependent upon track curvature and lubrication. A profile combination that works well to control corrugation may not be entirely successful for controlling head checking and spalling. Rail profiles that unload the gage corner of the high rails in curves may not be as successful in suppressing shell/DF formation as one might hope. Premature shell initiation has been observed with both two-contact and conformal type profiles where work hardening has been concentrated to the gage corner region of the high rail of curves. A somewhat tentative conclusion at this moment is that greater resistance to shell/DF initiation can be achieved by using practices that avoid concentration of work hardening in the gage corner region.

REFERENCES:

(1) Steele, R. K., "Performance Requirements of Railroad Rail", AAR Report #R-793, January 1991.

(2) Mutton, P. J., and Marich, S., "Rail and Wheel Materials for High Axle Load Operations", Third International Heavy Haul Railway Conference, 13-17 October 1986, Vancouver, B. C. Canada.

(3) Ghonem, H., Kalousek, J., Stone, D. H., and Laufer, E. E., "Aspects of Plastic Deformation and Fatigue Damage in Pearlitic Rail Steel", Second International Heavy Haul Railway Conference, 25-29 September, 1982, Colorado Springs, Colorado, USA.

(4) Heller, W., and Schweitzer, R., "Hardness, Microstructure and Wear Behavior of Steel Rails", ibid (3).

(5) Ichinose, H., Takehara, J., and Ueda, M., "High Strength Rails Produced by Two-Stage Flame Heating and Slack Quenching", ibid (3).

(6) Kalousek, J., Fegredo, D. M., and Laufer, E. E., "The Wear Resistance and Worn Metallography of Pearlite, Bainite, and Tempered Martensite Rail Steel Microstructure of High Hardness," Proceedings of the International Conference on Wear of Materials, pp. 212-231, K. C. Ludema (Ed), Vancouver, 14-18 April 1985, ASME, New York, 1985.

(7) Mutton, P. J., "The Influence of Microstructure on the Wear Behavior of Rail and Wheel Materials," Master of Applied Science Thesis, University of Melbourne 1985.

(8) Clayton, P., and Devanathan, R., "Rolling/Sliding Wear Behavior of a Chromium-Molybdenum Rail Steel in the Pearlitic and Bainitic Conditions." To be published in WEAR.

(9) Devanathan, R., "Wear Behavior of Bainitic Steels", Ph.D. Thesis, Oregon Graduate Institute, October 1991.

(10) Sonon, D. E., Pellegrino, J. V., and Wandrisco, J. M., "A Metallurgical Examination of Control-Cooled, Carbon Steel Rails with Service-Developed Defects", Rail Steels-Developments, Processing, and Use, Stone, D. H., and Krupp, G. G. Eds, ASTM Special Technical Publication, 644, 1978.

(11) Rice R. C., Rungta, R. and Scott, P. M., "Post-Service Rail Defect Analysis", Fourth Technical Report to USDOT/FRA, Cambridge MA, by Battelle, Columbus Division, June 1984.

(12) Marich, S., Cottam, J. W., and Curcio, P., "Laboratory Investigation of Transverse Defects in Rails", Proceedings of the Heavy Haul Railways Conference, Institution of Engineers, Australia 1978, Session 303.

(13) Clayton, P., "Wear Behavior of Pearlitic Steels with Particular Reference to Rails," Ph.D. Thesis, 1977, Brunel University, England.

(14) Fegredo, D. M., Shahata, M. T., Palmer, A., and Kalousek, J., "The Effect of Inclusion Type and Control on the Wear of an Alloy Rail Steel", CANMET Report MRP/PMPL 85-2 (OP-J), Energy, Mines and Resources Canada, Ottawa, K1A 0G1, Canada; also ibid(2).

(15) Clayton, P., and Steele, R. K., "Wear Processes at the Wheel/Rail Interface," AAR Report R-613, September 1987.

(16) Marich, S., and Beolen, R., "Revised Rail Selection Criteria," Fourth International Heavy Haul Railway Conference, 1989, Brisbane, Australia, 11-15 September 1989, pp. 25-34.

(17) Clayton, P., "Fatigue Behavior of Rail Steels in a 33 kip Wheel Load Experiment at FAST", AREA Bulletin #731, vol. 92 (1991), May, pp. 189-217.

(18) McEvily, A. J., and Minakawa, K., "Metallurgical Evaluation of FAST Rail Steels," Final Report, University of Connecticut, Institute of Materials Science, August 1981, USDOT Contract #DOT-TSC-1551.

(19) Stover, J. D., Kolarik, R. V. II, and Keener, D. M., "The Detection of Aluminum Oxide Stringers in Steel Using an Ultrasonic Measuring Method", 1989 Mechanical Working and Steel Processing Proceedings.

(20) Sugino, K., Kageyama, H., and Newell, H. W., "Detection Method for Harmful Inclusions in Rail Steels", AREA Bulletin 716, vol. 89, May 1988, pp. 230-259.

(21) Biswas, D. K., Venkatrama, M., Narendramath, C. S., and Chatterjee, U. K.," Influence of Sulphide Inclusion on Ductility and Fracture Behavior of Resulfurized HY-80 Steel", Met Trans A, vol. 23 A #5, May 1992, pp. 1479-1492.

(22) Heller, W., Weber, L., Hammerschmid, P., and Schweitzer, R. "Zur Wirkung von Wasserstoff in Schienenstahl und Möglichkeiten einer wasserstoffenarmen Erschmelzung," Stahl und Eisen, vol. 92, #19, 1972, 14 Sept., pp. 934-945.

(23) Lamson, S. T., and Longson, B. H., "Development of Rail Profile Grinding at Hamersely Iron", ibid (3).

(24) Steele, R. K., and Reiff, R. P., "Rail: It's Behavior and Relationship to Total System Wear.", ibid (3).

(25) Koffman, J. L., "Heuman Tyre Profile Tests on British Railways", Railway Gazette, 2 April 1965, pp. 279-283.

(26) King, B. L., "Tyre Profiles," Third International Wheelset Conference, Sheffield, England, July 1969.

(27) Towend, P. H., Epp, C. J., and Clark, P. J., "Bogie Curving Trials, Rail Profiling and Theoretically Modelling to Reduce Rail Deterioration and Wheel Wear on Curves," Proceedings of the Heavy Haul Railway Conference, Perth Australia, 1978.

(28) Kalousek, J., Sroba, P., and Hegelund, C., "Analysis of Rail Grinding Tests and Implications for Corrective and Preventative Grinding", pp. 193-204, ibid (16).

(29) Epp, C. J., and O'Rouke, M. D., "Wheel/Rail Profiling and Lubrication", pp. 402-408, ibid (16).

(30) Epp, C. J., "Wheel/Rail Contact: Energy, Lubrication, and Alternative Technologies," Lubrication on the Rail/Wheel Interface, Caulfield Campus Monash University, 12-13, November, 1991.

(31) Groom, J. T., "Residual Stress Determination: Reliability Analysis of Rail", Battelle Columbus Laboratory Report, Contract DOT-TSC-1426, 26 November 1979.

(32) Steele, R. K., and Joerms, M. W., "A Fatigue Analysis of the Effects of Wheel Load on Rail Life", AAR Report R-689, September 1988.

NONCONTACT ULTRASONIC INSPECTION OF TRAIN RAILS FOR STRESS*

Raymond E. Schramm, A. Van Clark, Todd J. McGuire,
B. James Filla, Dragan V. Mitraković,§ and Patrick T. Purtscher
Materials Reliability Division
National Institute of Standards and Technology
325 Broadway
Boulder, Colorado 80303 U.S.A.

ABSTRACT: This paper discusses acoustic techniques to quantify stress, the construction of transducers, and initial measurements on a short test section of railroad track. Our goal is to develop instrumentation useful for nondestructive testing during field inspection for potentially dangerous conditions, such as those generated by thermal stress. The application of unconventional noncontact transducers (EMATs) takes advantage of minimal surface preparation and the elimination of fluid couplants. This approach depends on precise (1 part in 10^4) timing of signal arrival. Elastic changes caused by stress generate very small, but detectable velocity changes. However, other factors, notably changes in metallurgical texture, may similarly alter the velocity. First tests on a rail section under applied compressive loading showed a sensitivity comparable to that seen with conventional piezoelectric transducers. If an all-EMAT system proves practical, it should be possible to design an automated device to scan at reasonable speeds.

INTRODUCTION: Increasing traffic, speed, and axle loads make rail integrity ever more important (1). While one problem is the size and nature of metallurgical defects (2), another is stress. This may be a remnant of straightening during fabrication or the result of thermal expansion; the result can be flaw growth or even track buckling. Structural calculations may establish tolerable stress for train rails, but nondestructive evaluation is necessary to determine that field conditions are within acceptable limits.

Investigators have long known many effects of stress on the propagation of sound (3), and the literature has suggested both methods and equipment for this purpose (4-8). Recently, ultrasonic probes for residual stress have come to market. They rely on very accurate timing measurements, potentially difficult with fluid couplants due to possible thickness variations, especially in a dynamic environment. Relative or differential stress measurements are simple and accurate, while absolute values have more uncertainty due to the effects of metallurgical texture.

* Contribution of the NIST; not subject to copyright in the U.S.A.

§ NIST Guest Researcher, on leave from the University of Belgrade

J.J. Kalker et al. (eds.), Rail Quality and Maintenance for Modern Railway Operation, 99–108.

Electromagnetic-acoustic transducers (EMATs) are noncontact devices to generate and detect ultrasonic signals directly inside metal. A liftoff on the order of 1 mm is possible since, unlike conventional piezoelectric transducers, the sound does not pass through a couplant. This simplifies the signal by eliminating interfaces that may cause reflections or mode conversions. Within some constraints, the EMAT designer can choose the wave type, frequency, and polarization and propagation directions.

Our experience with noncontact EMATs on railroad wheels indicates they would likely be useful in rail scanning operations also (9). Initial data from laboratory tests on a loaded rail section suggest the accuracy and repeatability possible. Measurements on surface-skimming pressure-waves yield near-surface stress while the birefringence of shear-horizontal waves reflects the through-thickness average stress.

ACOUSTIC METHODS: Our goal is to examine three methods to measure stress by ultrasound. In each case, the measured effect is a change in sound velocity due to elastic strain. The two main difficulties are the small size of the change and the effect of grain structure. We have begun preliminary tests on two methods.

Birefringence: Birefringence is the variation of the propagation velocity of a shear wave as the polarization direction rotates. For elastic waves, there are two principal causes: stress fields and metallurgical texture, or preferred grain orientation.

There are several ways to probe this anisotropy (10, 11). Figure 1a illustrates the use of shear-horizontal (SH) waves. A transducer on the rail head launches a 2 MHz tone burst; the wave reflects from the opposite face and returns to the transducer (pulse-echo technique). Velocity measurements are necessary with the polarization vector oriented along and normal to the rail length. Since the two signals travel the same distance, the difference in the two velocities reduces to a difference in arrival times. The fractional change between them is the birefringence

$$B = (t_{\perp} - t_{\parallel})/[(t_{\perp} + t_{\parallel})/2], \qquad [1]$$

t is the transit time of waves polarized in the $\perp$ and $\parallel$ directions. The functional relationship between birefringence and stress is (12)

$$\sigma_{\parallel} - \sigma_{\perp} = (B - B_0)/C_A. \qquad [2]$$

B_0 is the birefringence due to the metallurgical texture and can be measured on an unstressed specimen. B_0 is often a major portion of the total birefringence B. C_A is the stress-acoustic constant for shear waves; two values in the literature for railroad wheels are -7.6×10^{-6}/MPa for wrought wheels (12) and -9.5×10^{-6}/MPa for cast wheels (13). This small value dictates very accurate timing measurements (typically one part in 10^4 for a stress resolution of about 20 MPa). The σ's are stresses in the two principal directions; $\sigma_{\perp}$ is likely negligible. The calculated stress is the average value over the volume of metal traversed by the wave.

In practice, the largest uncertainty is the value of B_0, the metallurgical contribution. As shown by our earlier work on railroad wheels (9), it may be possible to collect a sufficient data base of measurements on unstressed specimens to find a statistically satisfactory value to use for each manufacturer's product. Another approach might be to use internal calibration along a stretch of track; measurements on lengths known to experience little or no stress due to their physical layout can be compared to those on lengths where stress buildup is likely.

Surface-Skimming P-Wave Velocity: Longitudinal or pressure (P) waves have a larger stress-dependent acoustic constant than shear waves and are more stress sensitive (14). Launched along the surface, they are surface-skimming P-waves (SSPWs). Conventional piezoelectric transducers generate them by critical refraction (14-16).

Large stress gradients are not uncommon. Because SSPWs travel near the surface they seem a likely candidate to probe these localized stress fields. Figure 1b shows the placement of transducers (pitch-catch) on the side of the rail head.

A rigid fixture maintains a set path length, so the problem becomes the measurement of the transit time between the two receivers. Comparing measurements from a test specimen and an unstressed reference specimen, we can calculate the stress from (16)

$$\Delta\sigma = [E/(C_A \cdot t_0)]\Delta t. \qquad [3]$$

$\Delta\sigma$ is the difference in stress between test and reference specimens, E is Young's modulus, C_A is the appropriate stress acoustic constant, t_0 is the travel time in the reference specimen, and Δt is the change in travel time due to stress in the test specimen. C_A for carbon steel is -12.6×10^{-6}/MPa (17).

The birefringence calculation comes from two timing measurements made at the same specimen temperature. For the SSPW method, measurements may be made at different times on different specimens. Since conditions may vary, it is necessary to consider the velocity dependence on temperature (16, 18):

$$V = V^0 - (dV/dT)\Delta T. \qquad [4]$$

V and V^0 are the velocities measured at ambient and reference temperatures, dV/dT is the speed change constant, and ΔT is the difference from the reference temperature. For forged steel, dV/dT [in m/(s·°C)] is 0.55 for P-waves and 0.38 for shear waves (18); a temperature change of only 1 °C introduces a timing change on the order of the necessary precision.

Combination of Waves: A major obstacle to calculating the stress from acoustic measurements is the large effect on velocity of the specimen microstructure, mostly the metallurgical texture noted above. Even when possible, comparisons against a

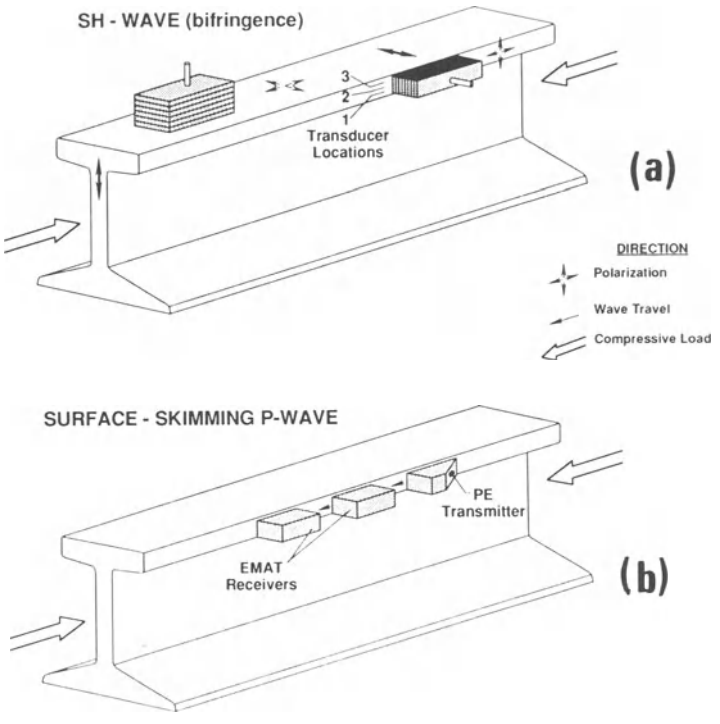


Figure 1. Transducer placements for the two types of measurements.

a. Birefringence. With the transducer on top of the rail head, the sound wave propagated through the web and reflected from the base. When the transducer was on the side of the rail, the signal remained within the head. We made measurements at three locations through the width of the head.

b. Surface-skimming P-wave hybrid array. A rigid fixture assured a constant separation between the two EMATs serving as receivers. Measurements were made only on the side of the rail head.

nominally unstressed reference leave unanswered the questions on how similar their textures are and whether the reference stress is actually zero.

Theoretical and experimental work on steels (19, 20) has taken advantage of the different elastic properties that combine to control the velocities of longitudinal and shear waves. Combining measurements of P-waves and SH-waves polarized along and normal to the stress field will minimize the effect of texture. Experimentally, this combines the shear-wave measurements (Fig. 1a) with a velocity measurement done using a transducer that generates P-waves propagating normally to the surface.

Yet another approach to this problem is to take advantage of the large difference in the magnitude of the stress acoustic constants for shear and longitudinal waves (6). The approximation is that any change in the velocity of a shear wave propagating in

the stress direction is due only to elastic changes resulting from texture. This requires measuring the velocities for the two wave types in both a stressed specimen and an unstressed reference. Comparison should yield the stress independent of any texture differences between the two specimens.

TRANSDUCERS AND ELECTRONICS: EMATs generate and detect ultrasonic energy directly in an electrically conductive specimen. A high-current radio frequency signal in a coil induces an eddy current in the test piece. This interacts with an externally applied magnetic field to produce a mechanical force. Reception is the inverse process; electrons vibrating in the magnetic field generate a current in the specimen which induces a signal in the coil. EMATs are very inefficient so the receiver electronics must amplify small signals and introduce little noise.

Virtually any wave type and radiation pattern are possible, depending on the coil-magnet configuration (21). The unit in Fig. 2a generates an SH-wave that travels normal to the specimen surface. The polarization direction is normal to the straight sections of the coil. This EMAT stacks two coils at right angles to generate and receive signals polarized along and normal to the rail length. An external switch connects only one coil at a time to the electronics. A large permanent magnet and a pole piece generate an active area of about 10 mm square. We electronically tune and operate this EMAT at 2 MHz.

To realize their full benefits in the SSPW array, EMATs would be advantageous as the transmitter and both receivers. For this initial phase, we found it convenient to use a conventional piezoelectric (PE) transmitter and wedge shown in Fig. 1b. The PE device gave us a stronger signal with less noise than the inefficient EMAT. Future research will eventually allow us to incorporate an EMAT transmitter.

Our current SSPW receivers are EMATs using a large (26 mm × 52 mm) permanent magnet atop a meanderline coil with a period of 12 mm (Fig. 2b). They detect a 0.5 MHz P-wave traveling on the surface.

To apply the combination-of-waves approach to through-thickness measurements, we will use the SH-wave EMAT in Fig. 2a; a vertical-incidence P-wave transducer to operate on ferritic steel also will be necessary and will demand some development. The coils probably will have the same pancake format, but the magnetic field, in this case, must be parallel to the surface. Design of this EMAT is now in progress.

These techniques require timing signal arrivals to a few nanoseconds in several tens of microseconds. Commercial timers that average about 100 arrivals can readily achieve this if provided accurate start and stop signals. For the SH-wave system, the start signal corresponds to the beginning of the transmitter pulse. The stop signal comes from a special digital gate designed to detect the zero crossing of the selected cycle in the receiver signal (9). For the SSPW system, we use a dual digital gate so

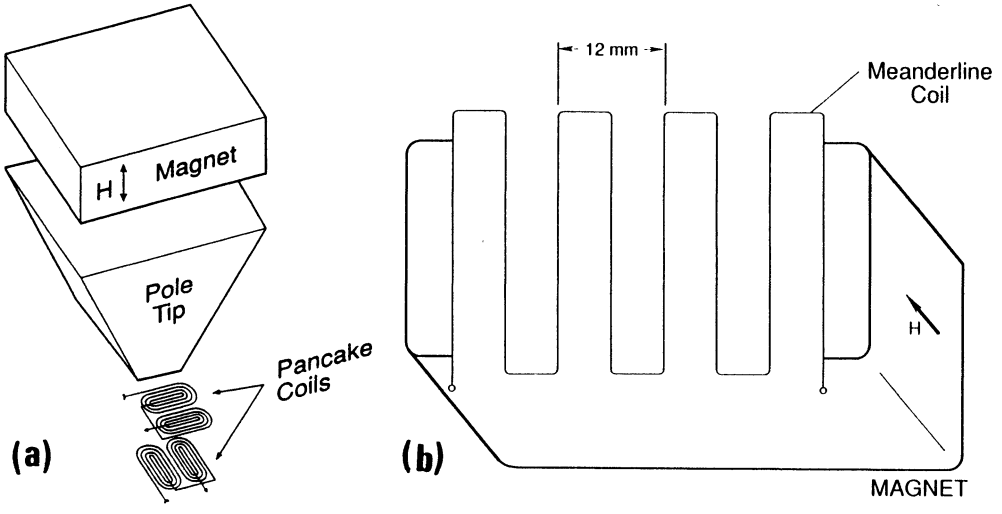


Figure 2. EMAT configurations for the two ultrasonic probes used here.
a. SH-wave propagating normal to the surface. The two orthogonal coils generate the two polarizations required by the birefringence method. This transducer operates in a pulse-echo mode at 2 MHz.
b. P-wave propagating along the surface. The 12-mm period of the meanderline coil fixes the operation at 0.5 MHz in steel. This transducer operates as a receiver in the pitch-catch mode.

the first receiver starts the timer, while the second stops it. A window is set to the selected cycle and the circuit generates a TTL pulse as the acoustic signal crosses zero. For the birefringence method, there is only one set of signals.

EXPERIMENTAL PROCEDURES: To measure the load-strain response, we used three pairs of strain gages attached to the head, web, and base of a 1-m length of 67.5 kg/m (136 lb/yd) rail taken from service. A hydraulic testing machine with a load capacity of 1000 kN applied compression along the length. Using a nominal elastic modulus of 200 GPa (29 Msi), the strain gave the load-stress response in each of the three locations.

The birefringence tests were repeated several times with the transducer in four locations (Fig. 1a). With the EMAT on the top of the rail head, the SH-wave traveled through the entire depth of the rail and reflected from the base; the transit time was nearly 118 μs. With the EMAT on the side, the signal traveled the thickness of the head and back; the transit time was ~48 μs. The lines identified as 1-3 are 1 cm apart and indicate the center of the transducer for three data sets.

The P-wave tests were conducted solely on the side of the head. The transmitter was a commercial 0.5 MHz PE transducer mounted on a polymer wedge designed for critical refraction of P-waves along the surface. Springs in the array fixture pressed this unit into the rail. The acoustic couplant was a commercial aqueous gel. The two EMAT receivers were locked together ~10 cm apart; their magnets held the system on the rail. The transit time between receivers was ~17.4 μs.

EXPERIMENTAL RESULTS: The strain gage data showed that, because of the irregular cross section, there was a stress distribution among the three locations. The web stress was 97% of the head stress, while the base stress was 86%. Since the SH-wave from the top of the rail head traversed all three regions, we used this information and the fractional distance of each section to calculate an average stress (Fig. 3). At incremental loads, we made timing measurements for the birefringence and SSPW methods.

With the SH-wave system the echo arrival for the two polarizations was a measure of the birefringence (Eq. 1). Figure 3 shows how ($B-B_0$) changes linearly with the stress. The stress-acoustic constant for this method is the slope (Eq. 2) of these data. The values for the three EMAT positions on the side of the head (table below) vary by about 20% but are in reasonable agreement with the previous reports cited above (12, 13). Fitting all the side-of-head data to a single line, $C_A = -9.6 \times 10^{-6}$/MPa. The table also shows that B_0 changes rapidly with position; this indicates a large gradient in the texture through the depth of the rail head. When the EMAT is on the top, the effect of texture on the birefringence is some average of the head, web, and base texture.

The SSPW data in Fig. 4 are the average values from five complete load-unload cycles. The stress-acoustic constant is again the slope (Eq. 3), and the value is similar to prior measurements (17).

SUMMARY: Ultrasonic velocity measurements are a sensitive indicator of stress state. The major question is whether these procedures will be useful outside the laboratory environment or to any other than specialized technicians. EMATs may point to an affirmative answer because:

1. Their signals are sufficiently regular and repeatable that it will likely be possible to automate the timing system.
2. They eliminate the need for both fluid couplant and extensive surface preparation, both of which can introduce errors if not carefully controlled.
3. Their noncontact capabilities might allow for a dynamic measurement while scanning rails.

The major needs in further studies are:

1. Develop an all-EMAT system for the SSPW. This may entail transducer design and/or electronic improvements.

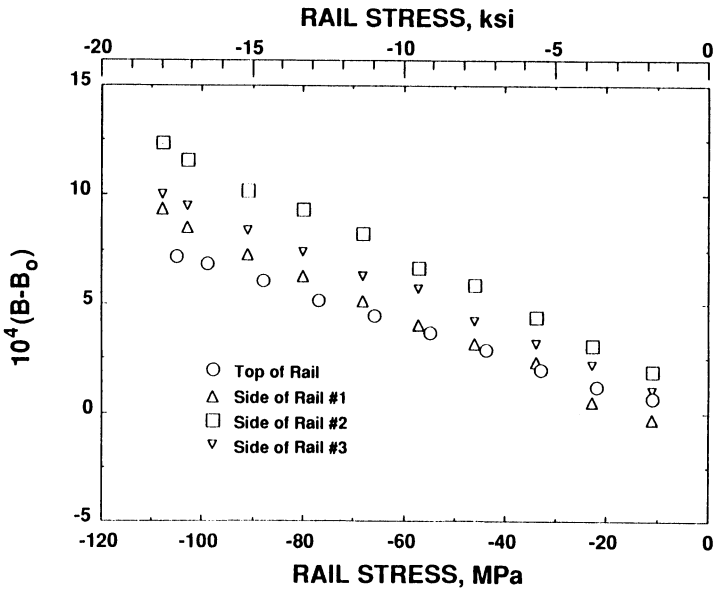


Figure 3. Change in birefringence as a function of compressive stress applied along the rail length. Measurements included EMAT placement on the top of the head as well as the side of the head at three locations through the thickness.

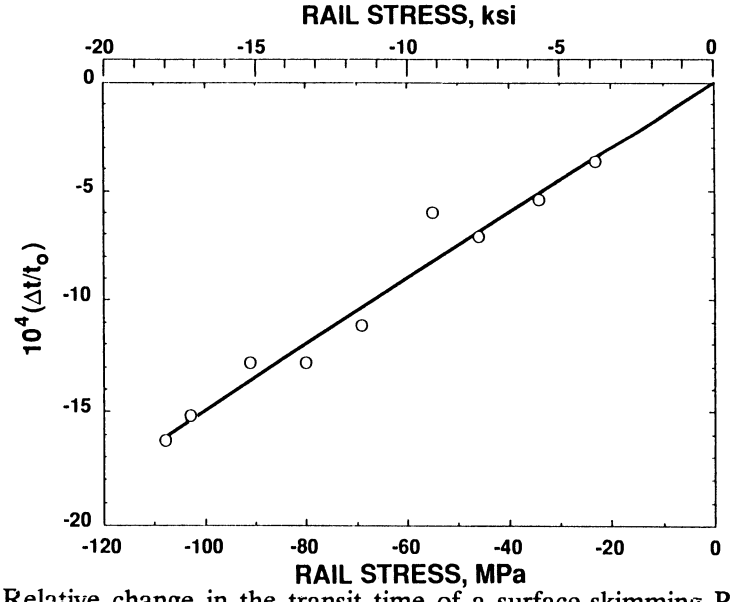


Figure 4. Relative change in the transit time of a surface-skimming P-wave as a function of the compressive stress applied along the rail length.

2. Electronic circuitry to automate the signal timing. This will eliminate a large portion of the operator variability and improve speed and accuracy.
3. Learn how best to deal with the effects of metallurgical texture.

METHOD	EMAT LOCATION	UNSTRESSED BIREFRINGENCE, B_0 ($\times 10^{-4}$)	STRESS-ACOUSTIC CONSTANT, C_A ($\times 10^{-6}$/MPa)
Birefringence	Top of Head	+18.2	-6.8
	Side of Head:		
	Position 1	-50.2	-7.9
	2	-38.5	-11.6
	3	-20.6	-9.3
SSPW	Side of Head	---	-14.7

ACKNOWLEDGEMENTS: This work was sponsored, in part, by the U.S. Department of Transportation, Federal Railroad Administration, Office of Research and Development, Washington, D.C. We appreciate the cooperation of the Transportation Test Center, Association of American Railroads, Pueblo, Colorado.

REFERENCES:

1. Orringer, O., "Rail Testing: Strategies for Safe and Economical Rail Quality Assurance," Rail Replacement and Maintenance Management, Transportation Research Record 1174, Transportation Research Board, National Research Council, Washington, D.C., 1988, pp. 28-42.
2. Nishida, S., Urashima, C., and Sugino, K., "Analysis of Surface Effects on the Rail Head," Role of Fracture Mechanics in Modern Technology, G. C. Sih, H. Nistani, and T. Ishihara (eds.), Elsevier Science Publishers B.V., Amsterdam, 1987, pp. 761-774.
3. Hughes, D. S. and Kelly, J. L., "Second-Order Elastic Deformation of Solids," Phys. Rev., V. 92, Dec. 1953, pp. 1145-1149.
4. Egle, D. M. and Bray, D. E., "Application of the Acousto-elastic Effect to Rail Stress Measurements," Mats. Eval., V. 37, 1979, pp. 41-46, 55.
5. Bobrenko, V. M., Kutsenko, A. N., and Sheremetikov, A. S., "Acoustical Tensometry. I. Physical Principles (Review)," Soviet J. NDT (translated from Defektoskopiya), V. 16, Feb. 1980, pp. 120-134.
6. Brokowski, A. and Deputat, J., "Ultrasonic Measurements of Residual Stresses in Rails," Proc. 11th World Conference on Nondestructive Testing, V. 1, Las Vegas, Nevada, 1985, pp. 592-598.
7. Leon-Salamanca, T. and Bray, D. E., "Mean Travel-Time for Zero-Force Determination in Railroad Rails Using P-Waves," Proc. 15th Symposium on Nondestructive Evaluation, D. W. Moore and G. A. Matzkanin (eds.), NTIAC, San Antonio, Texas, 1985, pp. 276-282.

8. Szelążek, J., "Ultrasonic Probeheads for Measurement of Surface Skimming SH Wave Velocities in Steel Product," Proc. 12th World Conference Non-Destructive Testing, J. Boogard and G. M. vanDijk (eds.), Elsevier Science Publishers B.V., Amsterdam, 1989, pp. 977-979.
9. Schramm, R. E., Clark, Jr., A. V., Mitraković, D. V., Schaps, S. R., McGuire, T. J., "Report No. 23 - Residual Stress Detection in Railroad Wheels: An Ultrasonic System Using EMATs," NISTIR 3968, May 1991.
10. King, R. B. and Fortunko, C. M., "Residual-Stress Measurements Using Shear-Horizontal Waves from Electromagnetic-Acoustic Transducers," NBSIR 84-3002, March 1984. (This is a collection of published papers.)
11. Pao, Y. -H., Sachse, W., and Fukuoka, H., "Acustoelasticity and Ultrasonic Measurements of Residual Stress," Physical Acoustics, V. XVII, Academic Press, New York, 1984, pp. 61-143.
12. Fukuoka, H., Toda, H., Hirakawa, K., Sakamoto, H., and Toyo, Y., "Acoustoelastic Measurements of Residual Stresses in the Rim of Railroad Wheels,"Wave Propagation in Inhomogeneous Media and Ultrasonic Nondestructive Evaluation, V. 6, G. C. Johnson (ed.), ASME, NY, 1984, pp. 185-193.
13. Iwand, H. C., "A Comparative Analysis Using Barkhausen Noise Analysis, Ultrasonic Birefringence, and Saw Cutting Techniques in Determination of Residual Stress in Railroad Wheels," Thesis, Univ. of Nebraska, 1988.
14. Egle, D. M. and Bray, D. E., "Measurement of Acoustoelastic and Third-Order Elastic Constants for Rail Steel," J. Acoust. Soc. Am., V. 60, Sept. 1976, pp. 741-744.
15. Ermolov, I. N. and Razygraev, N. P., and Sherbinskii, V. G., "Study of Shaping Process for Acoustic Field of Head-Type Wave in a Test Medium," Soviet J. NDT (translated from Defektoskopiya), V. 14, Nov. 1978, pp. 953-958.
16. Bray, D. E. and Leon-Salamanca, T., "Zero-Force Travel Time Parameters for Ultrasonic Head-Waves in Railroad Rail," Mats. Eval., V. 43, 1985, pp. 854-858, 863.
17. "DEBRO Ultrasonic Stress Meter," instructional manual, ELTRONA GmbH, Hamburg, Germany, 1990, p. 8.
18. Bray, D. E. and Stanley, R. K., "Nondestructive Evaluation, A Tool for Design, Manufacturing, and Service," McGraw-Hill Book Co., NY, 1989, p. 73.
19. Allen, D. R. and Sayers, C. M., "The Measurement of Residual Stress in Textured Steel Using an Ultrasonic Velocity Combinations Technique," Ultrasonics, V. 22, July, 1984, pp. 179-188.
20. Toda, H., Fukuoka, H., and Aoki, Y., "R-Value Acoustoelastic Analysis of Residual Stress in a Seam Welded Plate," Jap. J. Appl. Phys., V. 23, 1984, Suppl. 23-1, pp. 86-88.
21. Maxfield, B. W., Kuramoto, A., and Hulbert, J. K., "Evaluating EMAT Designs for Selected Applications," Mats. Eval., V. 45, 1987, pp. 1166-1183.

EXPERIENCES IN ULTRASONIC MEASUREMENTS OF STRESSES IN RAILS

Deputat J., Szelazek J., Adamski M.
Institute of Fundamental Technological Research
Swietokrzyska 21, 00-049 Warsaw, Poland

1. INTRODUCTION

The paper is continuation of our contribution presented to the Krakow International Conference in 1990 where the physical basis of ultrasonic stress measurements and some examples of its application for railroad industry were presented. Below we describe new results concerning stress evaluation in rails. Two problems are discussed - the influence of rail material texture on stress readings and possibilities of ultrasonic measurements of different origin stresses in rails both in steel mill and in the track.

Chapter two presents how the texture of rail material influences velocities of subsurface ultrasonic waves propagating along the rail in different locations around the rail profile. Next chapters describe applications of ultrasonic technique for measurements of various stresses created in the rail during its straightening, during track laying and due to train passages.

Ultrasonic stress measurements are based on elastoacoustic effect i.e. on the dependence of ultrasonic wave velocity on stress. In numerous experiments performed on various technical materials the linear dependence velocity - stress was observed for both compressive and tensile stresses, up to the plastic limit. Velocity changes caused by stress depend on wave type and on the orientation between direction of wave propagation, direction of stress and direction of wave polarization (particle motion) and the grade of material under test.
In practice of ultrasonic stress measurements not velocities but times of flights of ultrasonic pulses are measured. Time of flight Ts in stressed material is given as:

$$Ts = To / (1 - B * S)$$

where To - time measured in stress-free material
B - elastoacoustic constant [MPa^{-1}]
S - stress [MPa]

J.J. Kalker et al. (eds.), Rail Quality and Maintenance for Modern Railway Operation, 109–118.

The highest elastoacoustic constant is observed for longitudinal wave propagating parallel to the stress direction and therefore this wave is often used in ultrasonic stress measurements.
The value of elastoacoustic constant depends in some extent on material chemical composition and material heat treatment. Measurements performed on more than 100 samples cut from different rail parts show that elastoacoustic constant differs insignificantly over the whole cross section of the rail. It means that for texture ranges observed in the rails it is practically texture independent.

The velocity of ultrasonic waves is influenced not only by stress but also by numerous different factors which has to be taken into account in stress measurements. One of them, the very important one in case of rails - described in the next chapter - is material anisotropy due to texture (preferred grain orientation).
Next factors are temperature and chemical composition of the rail material. The temperature dependence of ultrasonic wave velocities can be easily measured and times of flights can be temperature corrected. Such corrections are specially important in measurements of stresses in rails in track where temperature differs in a wide range.
The chemical composition of rail material, various for different rail types and manufacturers, can be taken into account only by calibration on a given rail type.

Longitudinal stresses in rails are measured with probeheads generating subsurface longitudinal and transverse waves. Schema of such probehead is presented in Fig. 1. The advantage of subsurface waves is that they

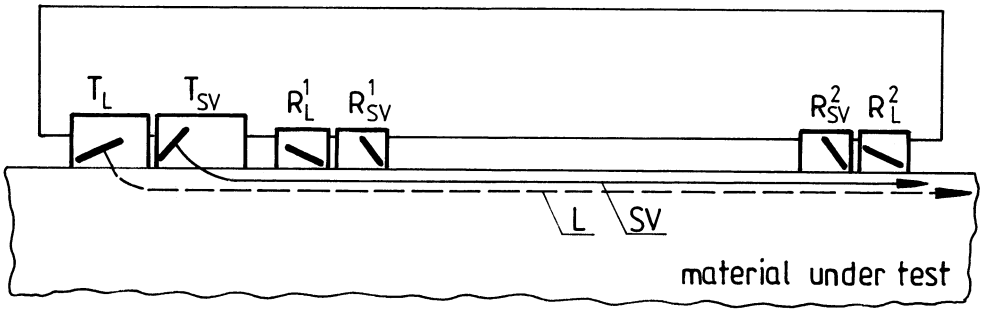


Fig. 1. Schema of the probehead for subsurface longitudinal (L) and transverse (SV) waves.
T_L, T_{SV}- transmitting transducers, R_L, R_{SV}- receiving transducers

can be generated and received at one surface on the object under test. Longitudinal wave is sensitive on both stress and texture. Transverse wave polarized in vertical direction (SV wave) is mostly sensitive on texture. This wave is also sensitive to stress parallel to the

polarization direction but stress component perpendicular to the rail surface, very close to the surface, can be neglected. The algorithm of stress calculation with subsurface longitudinal and transverse waves is described in [1].

Both waves propagate along one line. Measured times of flight are reverse proportional to the mean values of pulses velocities on distances between receiving transducers.
Subsurface wave gives information from a thin, subsurface layer of the material. It enables to measure stresses on flat surface without any special surface preparation. It is necessary only to remove debris from the surface in order to get proper acoustic coupling between probehead and the rail.

2. RAIL TEXTURE INFLUENCE ON ULTRASONIC WAVES VELOCITIES

The aim of experiments was to evaluate the influence of rail texture on accuracy of stress measurements performed around the rail profile. Material texture differs from massive rail head to the thin web or base what results in differences in ultrasonic pulse velocities. The measurements were performed on sample cut out from UIC-60 rail.
The sample was carefully stress relieving annealed to eliminate any velocity changes due to residual stresses. The efficiency of heat treatment was proved by measurements of distribution of velocities of ultrasonic waves along the sample. It can be assumed that rail texture gradient, if any, is very weak along the rail segment about 0.5 m long. Therefore any changes in the wave velocity observed along the sample could be related to remaining residual stresses with their maximum in the middle of the sample and zero values at its ends. During the test the same velocities over a whole sample length were observed what confirms that the sample is stress free.

Differences in times of flight measured on this sample in various locations around its profile which are result of texture differences only are presented in Fig. 2. Times of flights for both waves respond in similar manner to texture changes and their highest variations are observed on rail head side. Fig. 2 presents also values of residual longitudinal stresses calculated from measured times. It can be seen that for all measuring locations on the rail profile stress value is close to zero and stress value scatter is in the range +/- 10 MPa. This scatter denotes an error in stress evaluation with subsurface waves when the calibration (measurements of times of flight on stress free rail) was performed for one probehead location on the rail.

The above experiment proves that with subsurface longitudinal and transverse waves absolute values of residual stresses can be measured with technical accuracy around the rail profile. Such measurements allow to obtain more information concerning residual stress distribution in the straightened rail.

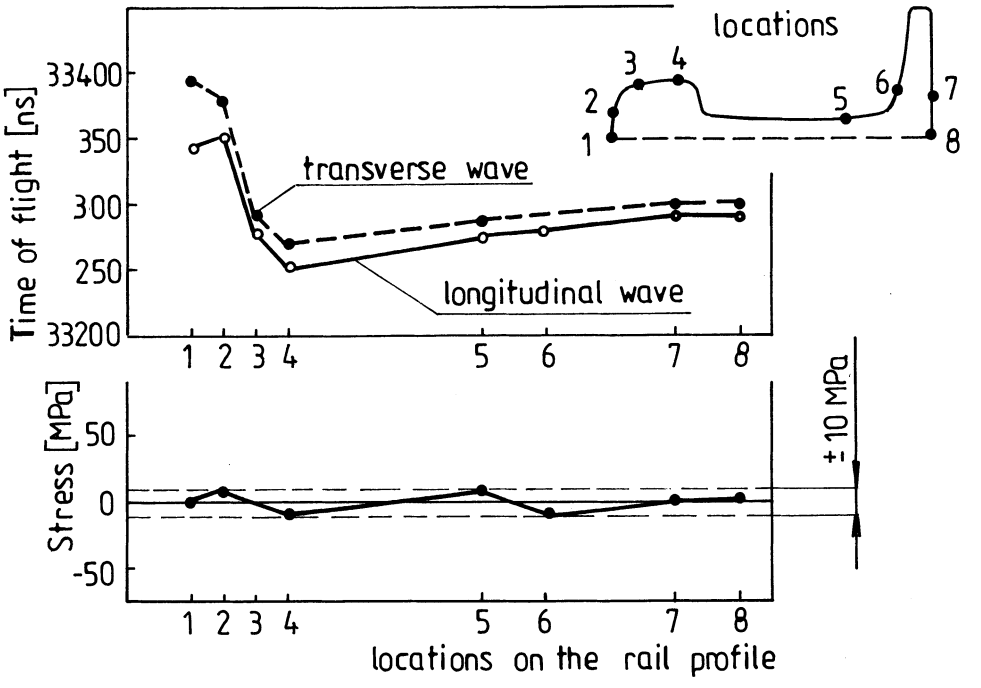


Fig. 2. Changes in times of flight of longitudinal and transverse waves as a result of rail texture, measured on stress free sample. Stress values calculated from measured times.

Texture changes along the rail are much lower than those observed from head to base therefore stress distribution along the rail can be also measured.

3. RESIDUAL STRESSES IN STRAIGHTENED RAIL

Fig. 3 presents an example of measurements performed around the profile of straightened rail with very high residual stresses. Tensile stresses are seen around the whole head and across base underside. The web and upper side of the base are in compression. The extreme compressive stress was measured in the middle of the web. Stresses are distributed symmetrically and practically the same values are observed on both rail sides.

With dashed line are presented the residual stresses measured on the same rail after annealing in 450 Centigrade during 3 hours. The aim of experiment was to observe if after partial annealing of straightened rail one can observe in it some regions where stresses are removed. Such regions could be used for the calibration. It can be seen that stresses

measured after partial annealing are reduced evenly about 60 % of their previous value and no zero stress regions are observed.

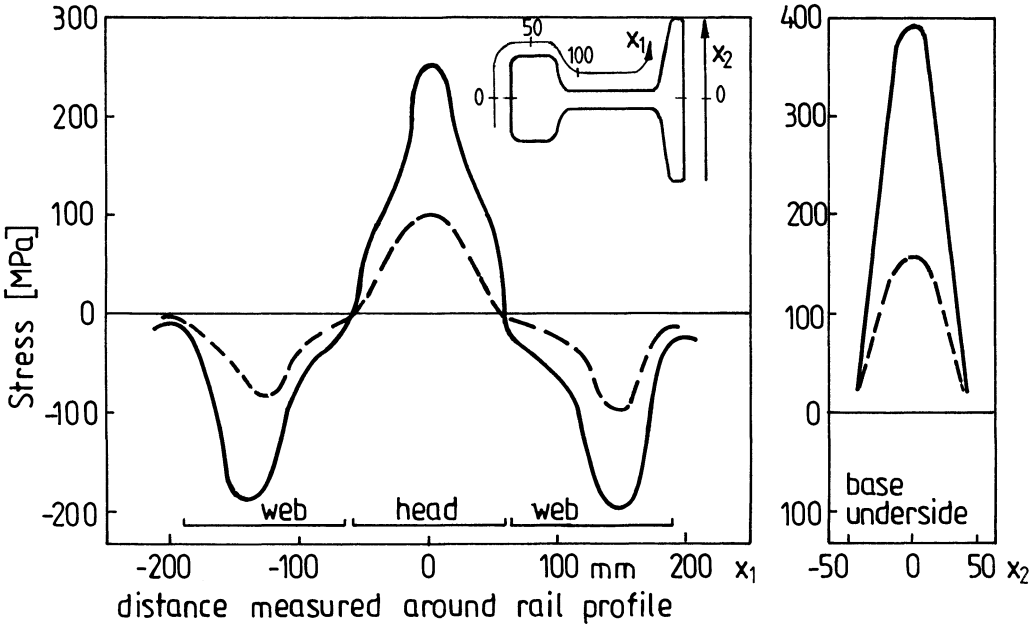


Fig.3. Residual stresses measured in 60 points around the profile of straightened rail. Dashed line presents stresses measured in the same rail but after partial annealing in 450 Centigrade in 3 hours.

Results presenting distribution of residual stresses along the rails straightened in different conditions and along the leading ends of rails, measured with ultrasonic technique are presented in [2].

4. STRESSES INTRODUCED INTO THE RAILS DURING TRACK LAYING

The aim of this measurements was to observe another kind of stresses which occur when the rail segment is positioned on ties and welded into the track. Such longitudinal stress sums up with residual stresses. Measurements were performed in field conditions on 400 m long CWR segment with the probehead for longitudinal subsurface wave only. Measured were both rails of the track, type S49.
Stresses which arise during rail laying are expected to be much smaller than residual stresses. It can be also assumed that these stresses are distributed evenly over the rail cross section and their gradients along the rail are weak. Therefore the distance between receiving probehead (see Fig. 1) and measured times of flight can be enlarged. Described

measurements were performed on new rails so the probehead could be attached to the rail head side not deformed by wheels yet.
The first measurement (calibration) were performed on the rails before they were positioned on ties. For each of 13 locations on both rails measured was time of flight and position of the probehead was marked on rail. It was assumed that the calibration was taken on the rails free of any longitudinal forces. In fact some forces resulting from rails transportation and unloading can exist.
The second measurement (stress measurement) was performed before the track was opened for traffic. The differences in times of flight measured during calibration and second measurement are proportional to longitudinal stresses introduced into the rails during track laying. To minimize the influence of thermal stresses on readings both calibration and stress measurement were performed in stable temperature conditions. The difference in rails temperature noticed during tests were below 5 Centigrade. The error of presented measurements, due to temperature changes, can be evaluated as +/- 5 MPa.

In measurements where only increments of stress are measured and calibration and stress readings are performed at the same locations on the rail, ultrasonic probehead has replaced a number of resistance strain gauges which were necessary to fix to the rail along the track. The difference is that to evaluate stress with resistance strain gauge it is necessary to know Young modulus and in case of ultrasonic technique - elastoacoustic constant for rail material.

Fig. 4 presents times of flight of subsurface longitudinal wave measured

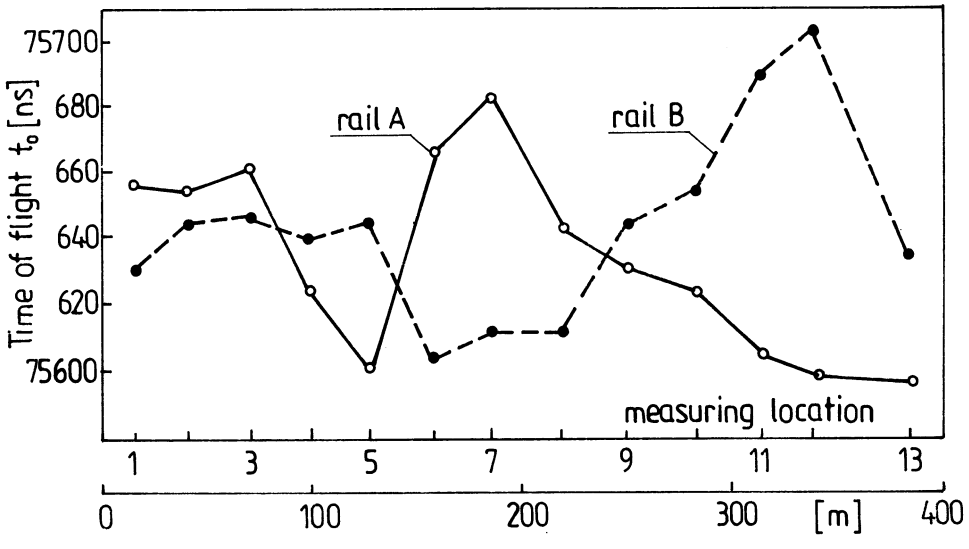


Fig.4. Times of flight of subsurface longitudinal wave propagating along the head side, measured in 26 locations on two 400 m long rails free of any external forces.

during calibration. It can be observed that measured times are different in different measuring locations. These differences can be result of variations in unknown residual stresses existing in the rail head and also of some unknown longitudinal forces introduced into the rails during transporting and unloading. It can be also seen that extreme values of times are observed for different locations on rail A and B.

Fig. 5 presents the distribution of longitudinal forces along the CWR segment before first train passed. Marked are also welds performed during track laying on both rails.
These stresses are calculated under assumption that the changes in times

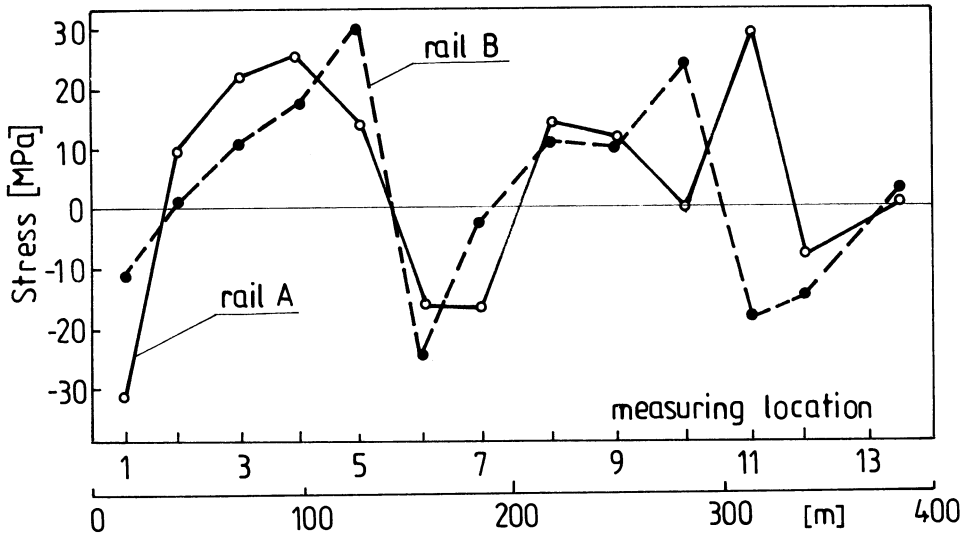


Fig.5. Distribution of longitudinal stresses introduced into the rails during track laying , welding and geometry regulation.

of flight are result of forces introduced into the rails during track laying and its geometry regulation. It can be seen that stresses are distributed along both rails of the track in similar manner. Tension is observed between locations 2 and 6 that again between locations 8 and 11. Also absolute values of stresses are similar for both rails. The explanation of such stress distribution can be displacements of the whole track when similar stresses are created in both rails.

Results of ultrasonic measurements of longitudinal stresses in CWR during destressing operation and rail temperature changes but performed with the probehead coupled to the rail web are described in [2]. Ultrasonic measurements presented there are compared with resistance strain gauges readings.

5. STRESSES DUE TO TRAIN PASSAGES

Passing trains exert loads on the rails in track. In result the material of rail head is plastically deformed and the state of stresses in the rail is changed. To measure such changes the same probehead as for residual stress evaluation was used.
Loads simulating rail work in curve were applied to the rails as shown in Fig. 6

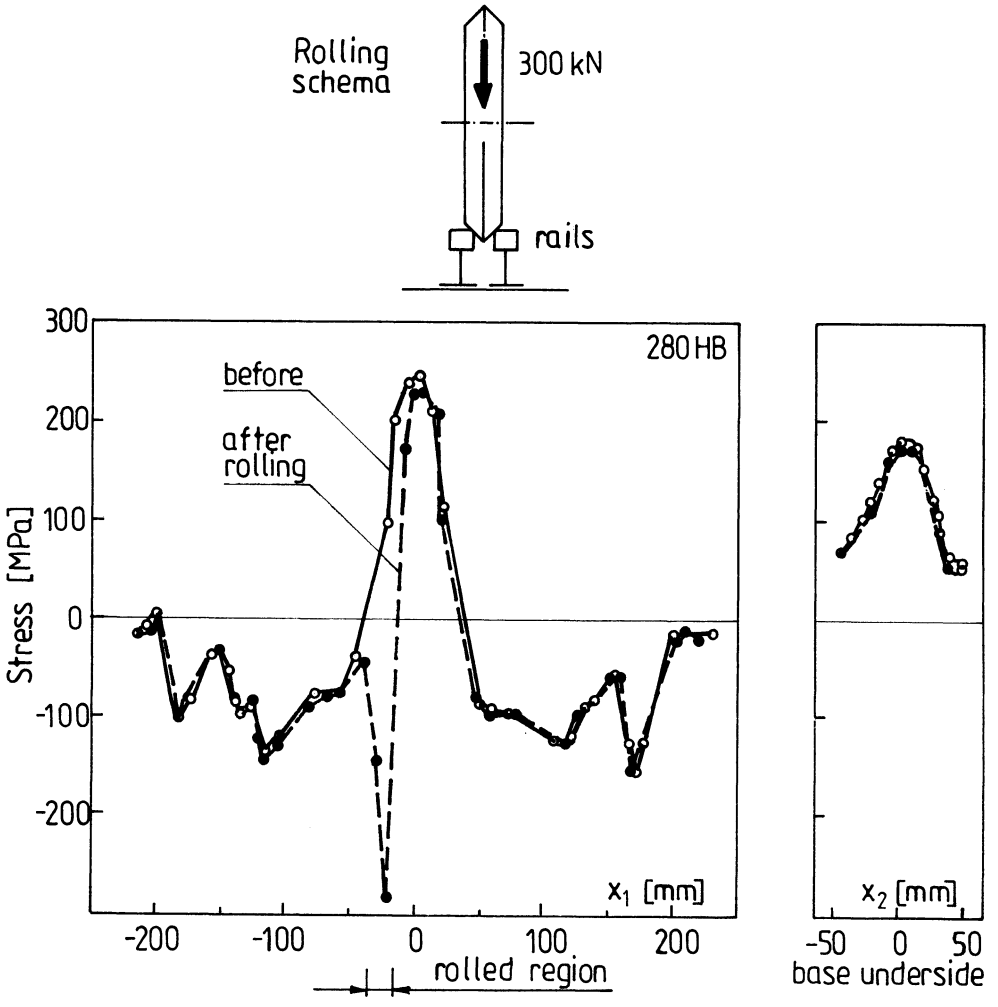


Fig. 6 . Schema of rails rolling in the stand and residual stresses measured around the 280 BH hardness rail profile before and after rolling.

To observe effects of rolling two rails with different hardness were chosen. Measurements were performed around the profile of two UIC-60 rails rolled at the bunch.
The diagram on Fig. 6 presents residual stresses measured around the rail profile before and after 1 million of load cycles on the rail with 270 HB hardness.
Fig. 7 presents results of similar measurements obtained on the rail with 370 HB hardness.
From both drawings it can be seen that the only changes in residual stresses were observed in rolled regions. In these regions, where tensile stresses due to straightening were observed in both rails prior rolling, compressive stresses are seen after test. Higher stress changes are observed for the rail with lower hardness.

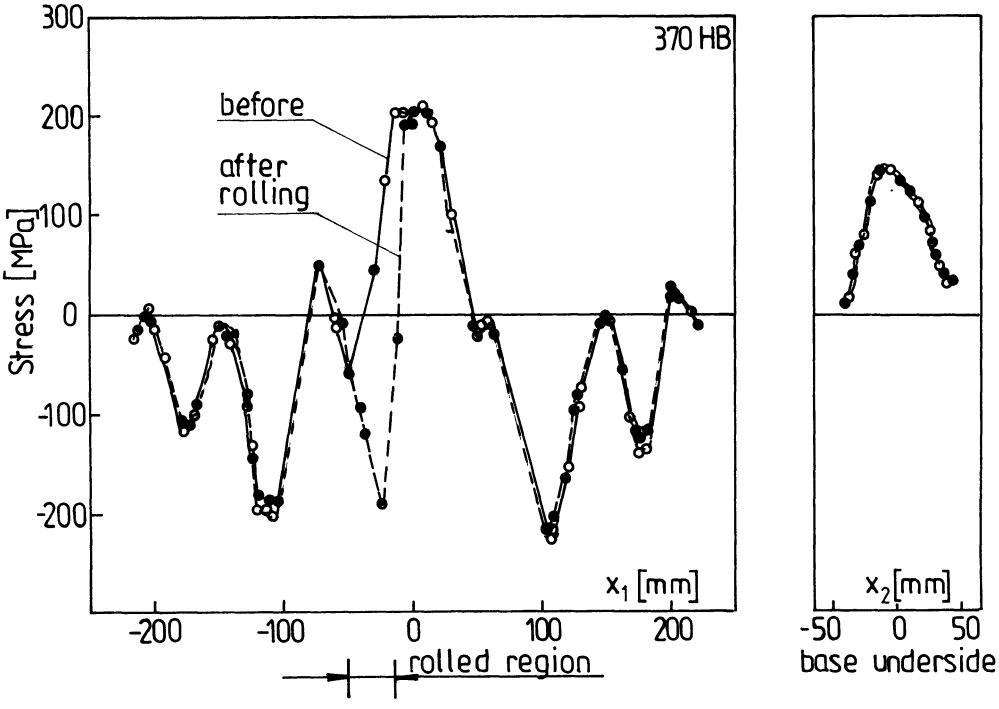


Fig. 7. Residual stresses measured around the 370 HB hardness rail profile before and after rolling.

Ultrasonic measurements of stresses were also performed after mechanical removing of 7 mm thick layer from rolled rail head. Readings obtained shows again tensile stress in the rail head.
The material texture in plastically deformed layer differs significantly from the rest of the rail. Therefore the accuracy of stress determination with ultrasonic technique in such a layer can be lower as compare with a new rail.

6. CONCLUSIONS

Results presented above were chosen to demonstrate how nondestructive, ultrasonic technique of stress measurements can be used to measure the absolute values or to monitor changes of stresses during rail life time from the steel mill to the track. Ultrasonic probehead can be easily positioned on the rail what allows to carry out measurements in field condition, even in the time lapse between passing trains.

Results obtained on rolled rails showed that residual stresses measured in web or base before and after rolling are the same. Therefore using probehead positioned on web it is possible to monitor longitudinal forces in CWR due to temperature changes (thermal stresses) or due to braking forces. During such measurements repeated for chosen location on the CWR both unknown residual stresses and material properties (chemical composition, texture) are constant and any changes of ultrasonic pulse velocity are result of stresses applied to the CWR.

Depending on distribution of stresses over rail cross section and along the rail different ultrasonic probeheads can be applied. Residual stresses can strongly change along the rail at the leading end. They also rapidly change around the rail profile. Therefore the volume of material between receiving transducers, over which stress value is averaged, should be as small as possible. Also material texture around the rail profile changes and these changes has to be taken into account during stress evaluation. In result probehead designed for residual stress measurements is comparatively short and generate two types of ultrasonic waves.
Thermal stresses or stresses due to track laying or braking forces are distributed evenly over the rail cross section. Also their gradients along the CWR are weak. Therefore they can be measured in one location on the rail profile (web for example) only and the distance over which their value is averaged can be long.

REFERENCES

[1]. Brokowski A., Deputat J.
Ultrasonic Measurements of Residual Stresses in Rails.
Proc. 11 WC NDT, Las Vegas, Taylor Publ. Co., Dallas, 1985, p.596

[2]. Utrata D.
Stress Measurements in Rails Using Ultrasonic-Based Nondestructive Methods and Strain Gage Holl-Drilling.
Association of American Railroads, Report R-779, June 1991

Rail Side Wear Problem On Chinese Railways

Zhiyun Shen

Southwest Jiaotong University Chengdu, Sichuan, P. R. China

ABSTRACT

The seriousness of the rail side wear problem on Chinese Railways has been described in this paper. Things account for the occurance of this problem and measures taken to solve it are discussed. Recommendations on improvements of the rail quality and rail maintenance are given as conclusions of this paper.

INTRODUCTION

Railways in China have known a history of more than 100 years. As the "main artery" of national economy, the railway has spotted itself as a vangard industry for the development of society and economy. Through technical remodelling of the existing railways and planned coustruction of new lines, the country's railway system grew up to route kilometerage of 57802 Km by the end of 1990, of which 53378 Km were state –owned and 4424 Km locally operated. In 1990, the railway accomplished 53.4% and 71.3% of the country's total passenger and freight turnover respectively. In this year, the traffic density on China's railways was 27.75 million tonne –kilometers per kilometer, ranking second in the railway world.

For the purpose of upgrading the technical condtions of railway facilities, the Chinese Railway has taken a range of measures of improvements, The double – tracked and electritied kilometerages have been raised to 13023 Km and 6940.8 Km, or 24.4% and 13.0% of the total respectively. Also the dieselized railway lines reached 16097 Km at the close of 1990. In the same time, rails of 60 Km / m and heavier deployed on trunk lines amounted to 15582.6 Km, constituting 23.1% of the total.

In the process of railway modernization, arised a problem of rail side wear on curves, the severeness of which was out of our expectation. At many intensively operated rail lines, new rails mounted in moderate curves were worn out in less than one year. For example, at the Shi –Tai Railway, only 225 Km in length, has to remove away rails in total length of 55 Km every year, which was only 4 Km in seventies. 98% of the removed rails were out of service due to side wear on curves.

In this paper, after a brieve survey of the rail wear problem, the mechanism of the rail wear on curves is analyzed and four methods of wear reduction are discussed, i.e.:

1, Rail quality improvements;
2, Rail lubrication;
3, Low wheel flange wear bogies;
4, Optimization of track parameters in rail maintenance.

J.J. Kalker et al. (eds.), Rail Quality and Maintenance for Modern Railway Operation, 119–125.

Recommendations on rail wear reduction are given as the conclusion of this paper.

RAIL SIDE WEAR ON CURVES

In recent years on Chinese Railways, the rail damages, mainly the severe side wear on curves, increases rapidly. Yearly requirements of steel to replace the worn out rails and wheels consist of 180 and 120 thousand tonne respectively. The rail life of the ordinary 60 Kg / m rail has been assigned as 600 MGT (Million Gross Tonne passage). But the real life of rails is only 100 ~ 135 MGT on curves of 450 ~ 750m radius, and about 300 MGT for curves of more than 1000m radius. In respect with the wear on curves, the change of 50 Kg / m rails to 60 Kg / m ones did not bring increase but decrease of rail life.

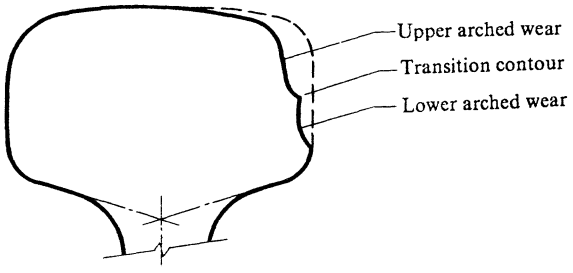


Fig.1 — Arch shaped rail side wear on Da –Qin Railway, 1991. [1]

Fig.1 shows an example of extraordinary rail side wear, occured on Dai –Qin railway [1], which was designed and constructed for modern operation of heavy freight trains for coal transportation. The curved track of 600 m radius with 100 mm superelevation was put into operation at the beginning of 1989. But by the end of this year, while the total passage was less than 40 MGT, the new 60 Km / m rails have been worn out, leaving a section shown in Fig.1. The replaced new 60 Kg / m rails were worn out again in less than one year.

The distinguishing features of this rail side wear is the two arcs with an overlapping contour at the middle. The reason which causes this phenomenon, may be the different angles of attack of the leading and trailing wheelsets of the freight bogies maunfactured specially for heavy hauling. In order to increase the load per meter track with the ordinary axle load of 21 tonne, short box cars have been manufactured for 10000 tonne coal train in Dai –Qin Railway. These cars have higher ratio of wheelset base to the distance between bogie centres in comparison with that of the ordinary freight cars. Under strongly enhanced tractional forces in the train, significant values of the angle of attack were resulted. In the same time, big differences between the attack angles of the leading and trailing wheelsets existed. In this consists the mechanism of formation of the rail side wear shown in Fig.1.

Increased rail side wear have been observed in the process of changing from steam locomotive to electrical and diesel ones. For moderate radius of curves, the rail wear rates were about 0.03 mm / MGT under steam engine traction. But, it increased to 0.14 ~ 0.15 mm / MGT after electrification and / or dieselization. The highest rate reached more than 0.22 mm / MGT, shortening the rail life to less than one year. As it was mentioned at the beginning of this paper, electrification and dieselization are the main trend in Chinese Railway modernization, the problem of rail wear reduction became one of the hot points of research. Some results will be given in this paper.

RAIL QUALITY IMPROVEMENTS

One of the measures to imporove rail quality is to use more heavier rails. The replace of less than 50 Kg / m rails to 60 Kg / m or 75 Kg / m rails greatly strengthened the track. But with respect to rail side wear, the heavier rails may bring no advantage, even neqative effects. It has been reported that rails of ordinary carbon steel without any heat treatment give higher wear rates than the 50 Kg / m ones.

In rail wear reduction, two directions have been considered on Chinese railways: to use high alloy steel and to apply heat treatment.

The constituents and tensile strength of new rails used on Chinese railways are listed in Tab. 1.

Tab.1 — Main types of rails to be used on Chinese Railways

Mark of Rail Steel	Alloy constituents(%)				Tensile Strength
	C	Mn	Si	V	
U74	0.67~0.78	0.70~1.00	0.13~0.28		~800 Mpa
U71Mn	0.65~0.77	1.10~1.50	0.15~0.40		883 Mpa
PD3	0.72~0.82	0.75~1.05	0.65~0.90	0.05~0.12	>900Mpa

Field observations of rail side wear of the high carbon steel, U71Mn and the steel of microvanadium alloy, PD3 have been carried out by Wu Hang Railway in 1988. Part of the results are given in Tab.2 [2].

Tab.2 — Results of Field Observation [2]

Curve radius	Rail Steel	Passage (MGT)	Rail Side Wear		Rail Life (months)
			max (mm)	Rate (mm / MGT)	
460 m	U71Mn	104	23	0.221	12
	PD3	223	22	0.099	25
500 m	U71Mn	80.4	15	0.187	15.6
	PD3	114.5	13	0.114	18.6

The results quoted in Tab.2 show that the microvanadium alloy rails give a side wear reduction of 55.2% –39% in comparison with that given by high carbon rails.

The most effective way to reduce the rail side wear is to harden the rail surface by quenching along the whole length. The obtained sorbitic pearlite metalographic structure pocesses very good triboresistant property and thus gives evident effect on wear reduction. This was approved in practice on most of the railways in China.

For example, a field test [3] will be given herewith, The ordinary carbon steel rail marked as U74 (see Tab.1) has been hardened by medium–frequency inductionthermy and then laid among rails of the same sort, but without heat treatment on a curve of 600 m radius. Fig.2 shows the distribution of hardness in the cross section of the U74 rail after quenching by medium–frequency induction hardening. After a passage of 10 thousand tonne, the side wear measured is 0.0675 mm / MGT, much lower than that of the untreated rails, which has reached to 0.12 mm / MGT.

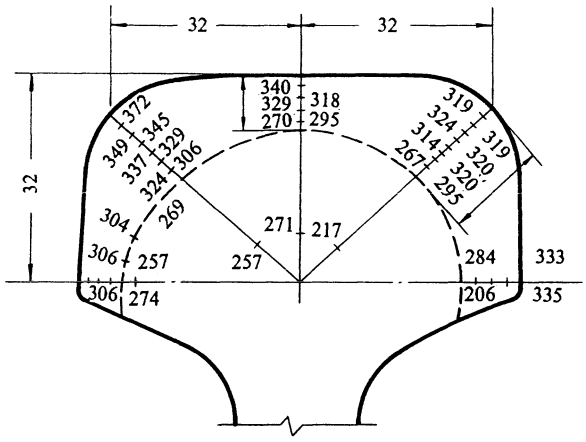


Fig.2 Hardness distribution of the U74 rail after quenching [3].

OPTIMIZATION OF TRACK PARAMETERS

Great potential of rail side wear reduction consists in rail maintenance, during which the track parameters, such as rail gauge, superelevation and cant of rails can be adjusted. Different values of these track parameters give different effects on side wear rates. On Chinese Railway numerous experiments under field conditions have been carried out to examine the influence on wear rates of the rail gauge, superelevation and rail cant. Almost the same conclusions have been obtained.

The rail gauge on curves has been assigned by Chinese Railway Standards to be widened according to the radius of the curve. But in practice, it has been observed that small gauge is advantageous to rail side wear reduction. The field observations, reported in 1989 [3], gave a good example. On the same curve of 600 m radius, two gauges, 1443 mm and 1433 mm, are deployed to stand the same tonne passage. After about one year's operation the side wear has been measured. The results showed that the narrow gauge gives a side wear reduction by a factor of 65.5% against the widened one.

In this test, 3 values of superelevation have been compared with respect to rail side wear rates. The equilibrium superelevation of this 600 m radius curve is 120 mm. Sections with 100 mm and 140 mm have been laid in both sides of the normal section. The measured results are given in Table 3.

Tab.3 — Rail side wear measured in sections with different superelevations [3].

Passed Tonnage (MGT)	Measured wear (mm)			Wear rates (mm / MGT)		
	h = (mm)			h = (mm)		
	100	120	140	100	120	140
93.50	3.86	4.45	4.05	0.041	0.048	0.043
119.84	4.46	4.92	4.47	0.037	0.041	0.040

The results show that to set up dificient superelevation gives positive effect on rail side wear reduction. Numerous tests of the same kind have been undertaken on other railways in China and the same conclusion has been obtained.

Tests on increased rail cant have been carried out also. According to the Chinese Railway Standards, the nominal value of rail cant is 1 / 40. To increase cant angle of the high rail, from 1 / 40 to 1 / 30 or 1 / 20, did not decrease, even increase the rail side wear. But shift this increase to the low rail, keeping the high rail cant as 1 / 40, an evident reduction of the side wear was observed [3]. After a passage of 26.34 MGT, the side wear of the high rail have been measured to be 0.24, 0.59 and 1.01 mm for low rail cant of 1 / 20, 1 / 30, 1 / 40 respectively.

RAIL LUBRICATION

Rail head side lubrication is an effective method to reduce intensive wear on tight curves. It is a wellknown fact to all of the railway countries and it has been reported in a vast of papers that a gain of 1.5 ~ 2.0 may be obtained purely by rail lubrication. Fig.3 is an example [4] showing by numerical calculations that the wear index, or the total work done at the wheel / rail coutact patches, may be reduced to about 1 / 2.5 of that without lubrication.

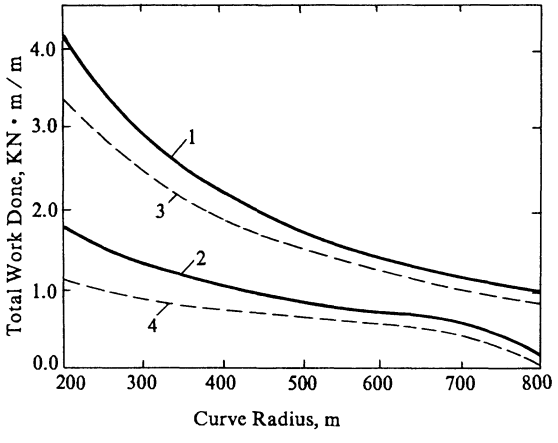


Fig.3 — Work done versus curve radius.
1. Total work, dry friction
2. Total work, lubricated
3. Flanging work, dry friction
4. Flanging work, lubricated [4]

On Chinese Railways curves with radius smaller than 800m are lubricated by track–side lubricator or lubricator carried by locomotive. Big research programs of rail lubrication development have been set up by Chinese Railway Ministry in recent years. Several types of lubricator, such as HB–1, developed by The Chinese Railway Academy, SPZ– by Shenyan Railway Research Institute, etc. have been put into daily service and positive effects on rail side wear,

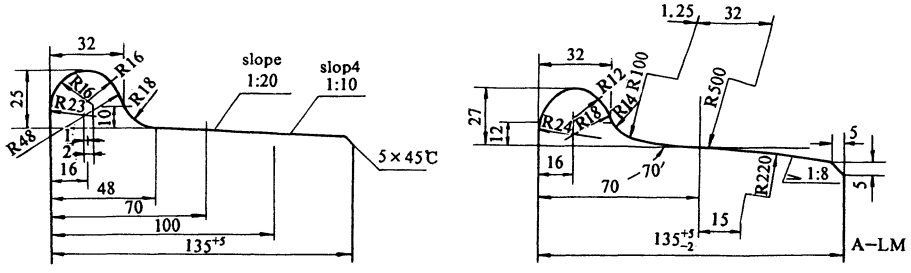
as well as wheel flange wear reduction have been resulted. The lubricant used is Calcium or Lithium Soap based grease containing certain percent of Graphite. The lubrication is so easy and effective that many railway sections of maintenance apply it to tight curves by hands, using common machine oil as lubricant. It is an urgent task on Chinese Railways to work out good rail or wheel lubricators and lubricants.

LOW WHEEL FLANGE WEAR BOGIES

The rail side wear is caused by the interaction between wheel and rail. In this tribosystem, the wheelset, as a base body, plays very important role in wheel / rail mechanism. Therefore, in solving the wear problem, it is very essential to improve the bogie design, to which much efforts are devoted in settling the research plan.

One of the basic research project is to determine the portion in rail side wear formation, how much from locomotives and how much from the freight cars. Metallic plates are inserted inbetween the wheel / rail contacts during the passing of locomotives, and are taken away when the freight cars passed by. Keep doing this for an enough volumn of passage and measure the rail side wear. Thus, the wear only from freight bogies are estimated at this isolated section. Comparing the measured results with those of the ordinary rails at the same site, a portion of 80% wear from freight cars has been evaluated. This study, although it was very expensive, revealed the great importance of development of low flange wear freight bogies.

The cheapest way of wheel flange wear reduction is to replace the conical wheel tread by a worn profile. The old wheel profile standard on Chinese Railway is shown in Fig.4,a, while the new stand, called LM profile for freight cars — in Fig.4,b. Experiences of applying this new standard of wheel profile indicated that the wheel tread wear may be reduced by a factor of 80%, and the wheel flange wear — by a factor of 50 ~ 30%.



(a) Conical profile (b) Worn profile, LM

Fig.4 — Old and New standards of the wheel profile of cars.

In rail side wear reduction, the most essential is to design new type of freight bogies without or with very scarce contact between the wheel flange and rail side surface. Various kinds of radial bogies, self–steering or forced steering, effectively decrease the angle of attack of the wheelsets by their radial orientation, and thus reduces the chance of wheel / rail contact. Therefore, radial bogie development is greatly supported by Chinese Rail.

As a successful example in this respect, may be mentioned the forced –steering freight bogie for Yunnan Railway. It was created in 1987 and was put into service in 1988. After about 4 year's service under very rigorous operation conditions, the wheel flange wear generated was very small, no more than 1 mm. The dynamic behaviour analyses have been reported in 1989

[5]. The total work done at the wheel / rail contact patches is given in Fig.5, which shows that with the optimum steering gain, the total work or the wear index, is only one seventh of that happened to conventional bogies (there the steering gain equals to zero). Observations under field conditions validated this conclusion.

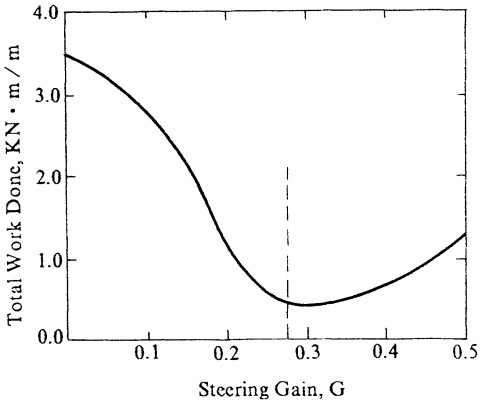


Fig.5 — Total Work done at w / R contact patches W_{tol} during steady state curving of the Forced Steering Freight Bogie [5].

CONCLUSIONS

Under modern operation of railways, the severe rail side wear problem arises, which merits a large share of attention in its solution. According to the experiences gained by Chinese Railway, the follwing measures are effective:

1. To improve the quality of rails mainly in two directions: to use high alloy steel, and to adopt heat treatment, such as surface quenching, etc.
2. To optimize the track parameters in track maintenance, such as relatively narrow rail gauge , deficient superelevation and small cant of the low rail.
3. To lubricate the rail side on curves or to lubricate the wheel flange by lubricants with triboresistant contaminants.
4. To improve the dynamical behaviour of the railway vehicles, especially the freight bogies. The worn profile of wheel tread should be used without exception. Radial bogies give big savings by effectively reducing the rail side wear.

REFERENCES

1. Zeng Shugu, Analysis of rail side wear on Da–Qin Railway, Chinese Academy of Railway Sciences, July, 1990.
2. Fan Tiehua, The reduction of rail side wear in tight curves, Wu–Hang Railway, Proceedings of Hangzhou Rail Lubrication Conference, Oct. 1990.
3. Lu Jiaofan, Rail lubrication and the comprechensive measures to reduce the rail side wear on Beijing –Shanghai Railway, Proceedings of Hangzhou Rail Lubrication Conferense, Oct. 1990.
4. Zhiyun Shen, J.K.Hedrick, The influence of rail lubrication on freight car wheel / rail wear rates, Proceedings of 9th IAVSD –Symposium, Linkoping, Sweden, 1985.
5. Zhiyun Shen, Junmo Yan, Jing Zeng and Jiangxin Liu, Dynamical behaviour of a Forced –Steering Three –Piece freight car track, Proceedings of 10th IAVSD –Symposium, Prague, 1987.

RAIL RECTIFICATION SPECIFICATIONS AND MODERN GRINDING STONE TECHNOLOGY

Jim Cooper, Jean-Claude Schaffner
Speno International S.A.

INTRODUCTION

Rail rectification treats in track the shape and the surface of the railhead, usually by grinding. The product is metres of finish ground rail. However, rail rectification has traditionally been remunerated on an equipment-presence basis and product specifications describing treated rail are not yet generalized. With the modern tendency toward production-related payment such specifications are now under discussion by railway administrations.

At the same time, a new generation of grinding equipment is coming into general use. The modern machines offer higher performance through their enhanced technical capacity. Also, their favourable influence on product costs renders rail rectification economically viable on a broader scale.

Key to the breakthrough is a new generation of grinding stones. Indeed, stone development has become acutely important. Recent years have seen the introduction of "aggressive" grinding stones with beneficial enhanced production. However, aggressive grinding is indeed aggressive. With metal removal rates up to six times greater than conventional grinding, the mechanism of metal removal changes. The high-production machines produce swarf-like particles rather than simple dust. The rail surface is clearly rougher. The question has been raised of possible detrimental effects to rail integrity.

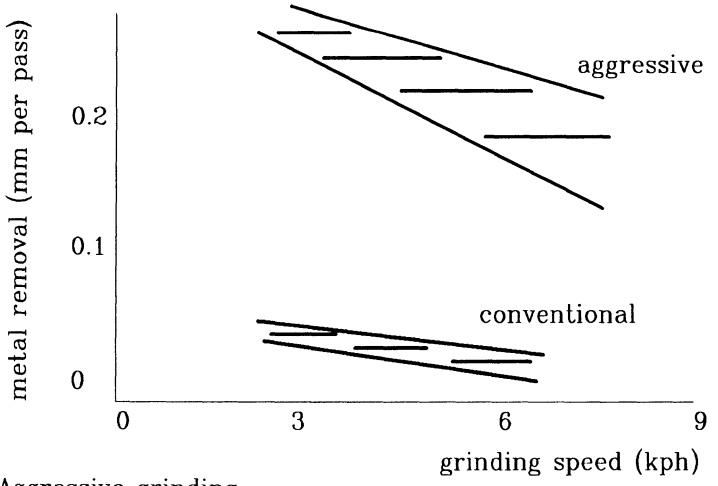


Fig 1 Aggressive grinding

J.J. Kalker et al. (eds.), Rail Quality and Maintenance for Modern Railway Operation, 127–136.

Originally deployed on heavy-haul lines, aggressive grinding has come under wider consideration. On European conventional lines, with higher train speeds, and rails lasting longer in service, engineers devote much attention to micro-crack initiation and propagation. The lighter axle loads may mean that the "disappearance" process of the grinding scratches under traffic is different, and the rail may be more vulnerable to micro-crack formation after aggressive grinding than heavy-haul experience shows.

There has been extensive reflection about grinding specifications with a view to introducing safeguards against excessive treatment of the rail. While much remains unknown, a generally accepted set of requirements for grinding rail is emerging. Meanwhile Speno International and its subsidiary ASI push forward with the development of stones that are ahead of today's increased demands.

Rectification
Rail rectification has particularities. First, the precision of the operation is unusual within the context of track maintenance. A result within a few tenths of a millimeter simultaneously in two planes is often the standard to be achieved.

Second, rail rectification has a multiple action. It can be deployed to treat the railhead:
- surface condition
- transverse profile
- longitudinal profile

These operations overlap in a complex manner. Assuring two or more operations requires a volume of work that is less than the cumulation of the individual tasks.

Production
Rectification in Europe is based upon multiple passes. Thus finish-ground rail is the sum of several intermediate runs. The number of passes to achieve the end-result depends upon:
- the crew: through its tactical skills (choices of angles and pressures, and speed, pass by pass)
- the machine: by the number and power of the grinding units, and specific constructional features

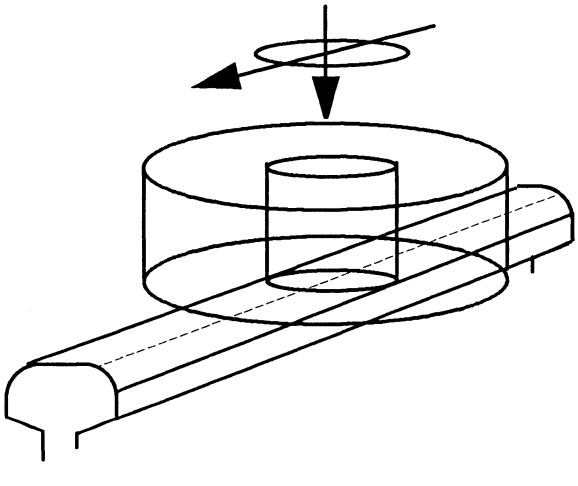


Fig 2 Grinding action

and last, but not least
- the grinding stones: by their cutting power, and service life

SPECIFICATIONS
This part of the paper is put forward with diffidence. Both authors contribute from the supplier's viewpoint. They may not be fully acquainted with the railways' attitudes.

Details of "specifications" are quoted as examples. These elements are not necessarily employed contractually by the railways. (BR, for example) In some cases they may have been replaced by later versions that the authors ignore.

The nature of specifications vary with the structure of the contract. They can be one or several of:
- presence-related: definition of the availability of means to effect the work
- performance-related: definition of the rate of work to be accomplished
- product-related: definition of what has to be produced

We are mainly interested here in the last case - the railway describes the finish-ground rail. However, all product-related specifications are situated within two constraints:
- the necessity for high production (to lower costs and to minimize track occupation)
- stone life should be at least a full shift (to avoid the time loss and the danger of changing in track)

A notion of desired production levels is given by:
BR: min. 1,200m/hr finished when removing 0.1mm wave formation
Sncf: min. 1,000m/hr finished when removing 0.4mm

Also for quality, the cost factor dominates. Rail rectification is like any production process - tighter tolerances mean higher production costs. Specifications that stretch currently available means will contribute to progress. Specifications that are frankly beyond those means can lead to expensive development work that increases product costs out of proportion to the technical advantages obtained.

Surface condition
Specific grinding for surface condition can be:
- preparative: cleaning millscale and construction damage from newly-laid rail to assure optimal initial conditions for commercial service
- preventive: removing fatigued metal before micro-cracking leads to more serious damage
- curative: recovering rail that has been hurt by external influences such as ballast imprints, foreign bodies pressed into wheels, cargo loss

Although grinding for surface condition is a specific application, grinding for *any* application affects the surface condition of the rail. Thus surface condition specifications concern virtually all grinding interventions. It is precisely here that the balance between rectification specifications and grinding stone technology is most critical.

At issue is the rail and the need to protect it from untoward material changes during the rectification process:
BR: Work shall be carried out...without adversely affecting the metallurgical structure of the finished rail or the surface finish.

Prima facie evidence of material damage is:
- discolouration
- excessive surface roughness

The first aspect is widely understood. The railways' site supervisors understand instinctively that colouring of the rail surface merits attention. However, explicit specifications in are not in general use. Two rare mentions:

BR : Light blue/straw colouring is acceptable

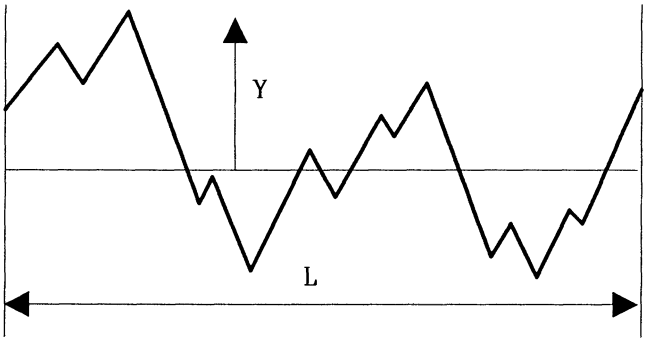
Sncf: The ground zones should present no trace of blueing at the end of the work

More precise wording presents the difficulty of quantifying an essentially visual effect. Nevertheless, it would be useful to know more about the meaning of colours on rails of varying qualities and compositions.

Since the introduction of aggressive grinding, surface roughness has become a major issue. First reactions were naturally prudent. With conventional grinding the scratch marks left during rectification are quickly re-absorbed under the influence of the passing wheels of the first few trains. The perceived danger with the increased surface roughness of aggressive grinding is that the relatively large-scale irregularities on the rail surface fold over during the absorption and initiate micro-cracks:

BR: Individual grinding stones should not leave transverse ridges on the railhead surface such that they fold over under traffic.

Fortunately, the measurement of surface roughness exists as a standard engineering procedure. The unit of measurement is the micron (one millionth part of a metre). Less clear, however, in the case of rail grinding is the correct direction of measurement. The scratch marks result from the locus of a point on the flat end surface of the grinding stone that has two components: rotational and advance. Thus the resulting ridges are neither longitudinal nor transversal. (Speno International measures surface roughness parallel to the longitudinal axis of the rail.)



$$R = 1/L \int_0^L |y(x)| dx$$

Fig 3 Roughness measurement

Only a few railways have expressed an opinion of suitable surface roughness after grinding in standard measurement terms:
BR: max. 7 microns at ten points across the rail
Sncf. max. 10 microns over the ground zone

These specifications present no practical problem for measurement. Apparatus are commercially available that indicate numerically the roughness as a centre-line average over a standard sampling length. They are small and of handy use. The roughness is generally stable over ground sections and sampling rate is not critical.

Transverse profile
There are several reasons for rectifying rail in the transverse plane:
- assuring vehicle behaviour
- controlling the distribution of internal stresses in the rail
- reducing rail wear
- relief for wheel profiles
- removing flowed metal that can obstruct track work processes

The product is a target rail transverse profile. Or more precisely, the objective is the reproduction of a certain shape over a selected zone on the surface of the railhead.
Thus a specification needs:
- definition of the shape
- tolerances
- the limits of the ends of the zone

Strangely enough, the definition of the shape poses problems. The matter is under consideration by railway authorities. For the time being reference is made to the manufacturing design profile (and the appropriate rail laying angle).

The shape - a curve of varying radius - is rendered by the series of straight facets of the individual grinding units. Thus, from the grinding process alone there is an unavoidable polygon-curve approximation.

A basic ride parameter - equivalent conicity - is very sensitive to the shape of the transverse profile. Calculating back from a desired conicity value and an assumed wheel profile to the corresponding rail profile shape indicates tight tolerances are necessary for the latter - in the order of tenths of a millimeter. Such accuracy is comparable with the polygon-curve approximation of a standard 32-stone machine. A call for tighter tolerances will have a quantum effect on costs as it implies extra passes or the deployment of bigger machines.

The ends of the grinding zone are mostly expressed as angles from the vertical of the axes of the grinding units, typical values being 70° toward the gauge side and 15° toward the field side (the flat grinding stone face being applied perpendicularly to these angles.)

Some examples of railways' wishes in connection with transverse profiles:
Shape:
Speno International is not aware of a target profile - other than a design profile - being currently specified, However, BR has stated its intention of specifying target profiles by rectangular cartesian coordinates.
Tolerances:
DB:

For line speeds	
>160kph target profile	*+0.3mm/-0.3mm*
140-160kph	*+0.5mm/-0.3mm*

<140kph service profile "improvement"

FS:
For line speeds >160kph target profile *+0.5mm/-0.3mm*

BV:
Without mention of line speed: target profile *+0.3mm/-0.3mm*

BR:
Without mention of line speed: target profile *+0.15mm/-0.15mm*

CFL:
Without mention of line speed: target profile *+0.5mm/-0.3mm*

Sncf:
Without mention of line speed: target profile *+0.3mm/-0.3mm*

Limits:

FS:
The zone is the area of contact between the wheel and the rail.

BV:
The zone is the part of the profile from the beginning (lowest point) of the 13mm radius on the gauge side to the beginning (highest point) of the 13mm radius on the field side

In addition, some railways perceive the need for safeguards, mainly concerning the subtended angle between two facets that may give rise to a ridge - considered bad for wheel wear, and possibly a fragile point on the rail.

Examples:

FS:
A regular finished profile without sharp edges.

Sncf:
Facet width shall not be larger than 5mm

In theory, checking on ground transverse profiles requires a continuous recording. Much effort is being deployed to achieve this result. In the meantime, two practical methods are being used. On Speno machines the operator has a visual aid showing for each rail the variance of 10 points on the current transverse profile from the appropriate target profile. (This indication is for checking that the pattern of angles and pressures being used is the best suited.) While the actual profile is retained by spot-checks with a manual recording device allowing a five times multiplication of the profile in analog or numerical form.

There remains the question of the appropriate sampling rate along the rail - a value determined by:

- the assumed rate of change of the transverse profile along the rail
- the perceived stability of the grinding process (that is accepting the notion that the accuracy stems from the control of the pattern of the angles and pressures)

Provisional guidelines call for checking the transverse profiles after grinding with the manual recorder at a frequency of:

DB:

- *once per grinding section, on lines with speeds >160kph*
- *once per shift, with speeds 140-160kph*
- *twice per week, with speeds <140kph*

Longitudinal profile
The purpose of rectifying rail in the longitudinal plane is to ensure a smooth running surface and thereby control unwanted disturbances such as impact loading, vibrations and noise emission. The geometrical shape sought is a "straight line". The variance from that line has two practical dimensions:
- severity: the wave depth (peak-to-peak)
- extension:the length of track over which the wave depth is representative

Within the general class of wave formation one can distinguish physical phenomena that merit separate specifications:
- corrugation (30- 80mm long)
- short waves (80- 250mm long)
- long waves (250-3,000mm long)

In practice, the specifications result from mixed considerations, with elements of both wishes and realities:
- how much the waves disturb railway operations
- how well the rectification process deals with them
- how well measurement systems describe them

The length of the waves plays a fundamental role in the grinding process. The usual grinding stones have an outside diameter of 250mm and they "bridge" over waves up to this length to give virtually the straight line that is sought. Above this wave length it is the construction of the machine, with its mechanical and operating tolerances, that determines the precision of the result.

Some examples of desired tolerances of residual wave depths in relation to wave lengths after grinding:

DB:
max. 0.02mm over 200mm
max. 0.30mm over 3,000mm

FS:
max. 0.02mm over 200mm
max. 0.30mm over 2,000mm

BV:
max. 0.02mm over 30- 80mm
max. 0.02mm over 80- 300mm
max. 0.30mm over 300-1,000mm
max. 0.30mm over 1,000-3,000mm

BR:
max. 0.010mm over 25- 90mm
max. 0.025mm over 90- 450mm
max. 0.050mm over 450- 750mm
max. 0.100mm over 750-1,050mm
max. 0.25mm over 1,050-1,400mm
max. 0.400mm over 1,400-1,700mm

Sncf:
max. depth less than 0.0001 times the corresponding wave length
over 100-1,000mm

Checking work to the precision of these specifications requires care. Also, the method of measurement is important. A differential measurement involves a geometrical transfer function with a selective zone of valid response.

GRINDING STONES
The grinding stone is the key to rail rectification: it determines performance. (Fundamental elements are grit size and binder tenacity, but several other ingredients play important roles.)

While grinding is a well-known practice, its use in track is particular. Specific safety and application methods must be respected.

Workshop grinding exists in two common forms:
- the heavy action of de-burring
- the precise action of polishing

The stones and the operating parameters differ widely.

Conventional rail rectification is between these extremes. It resulted from the adaptation to use in track of an uncritical process with average metal removal and average precision. The inhibitions to performance included:
- the fragility of the then-current stones
- the unsophisticated means of controlling the grinding pressure
- the undemanding action of the fixed-angle units
- the relatively low power of the grinding unit drive motors

A wave of parallel innovation has enabled rail rectification to move forward to a new level of capacity. The process has become closer to de-burring, but the end-product still requires a quasi-polishing finish. The problem is to reconcile both aspects by special means or by built-in flexibility of the stones.

Heavy-duty grinding removes metal by a complex action. There are elements of cutting, of tearing and of wearing. This condition results in part from:
- use of a mixture of various grit sizes
- differential cutting speeds (from 25m/sec to 50m/sec) across the rotating face of the stone

The metal debris varies widely in its characteristics. The surface finish reflects this fact. Discolouration and surface roughness of the rail can result from unsuitable stone characteristics. Such occurrences are not necessarily synonymous with heavy metal removal.

Stone design

The scheme for stone development (for a given drive motor power) is:

1. review of physical dimensions to clear obstacles in track (while respecting the maximum peripheral speed of 50m/sec imposed by European regulations)

and then a re-iterative process including:

2. choice of grit sizes and mixtures to give suitable surface finish
3. choice of binder, and compacting pressure, to give suitable stone life
4. choice of other ingredients to balance performance
5. bench tests of metal removal in relation to power absorbed

At this stage the objective is to balance stone life and metal removal and the evaluation parameter is:

$$\text{quotient "G"} = \frac{\text{volume of metal removed}}{\text{unit volume of stone consumed}}$$

The bench tests are valid for comparisons between prototype stones of the same series. They enable short-list selection of the pre-series stones for track testing. (On the other hand, comparisons between bench tests and track tests are not assured.)

The essential difficulties during this process are:
- the delicate balance between the opposing constraints of surface finish, stone life and metal removal
- the complexity and sensitivity of the ingredients and their interaction

- the difficulty of carrying out objective tests in the widely varying conditions in track

The current generation of grinding stones results from:
- new chemical ingredients (mainly stemming from zirconium)
- new physical aspects (specially shaped abrasive particles)
- new resin bonds
- new manufacturing process (high-pressure compaction)

A schematic comparison of the progress initially achieved in standard benchmark conditions:

	conventional stone	modern stone
surface finish	3-5 microns	10 microns
life	6 hours	5 hours
metal removal	0.05mm	0.20mm
quotient "G"	2-4	10-12

Since then progress has been made in stone life that renders the comparison much more favourable for the new technique.

Operation
The usual parameter for quantifying the grinding action in rail rectification is depth of cut, a value that needs understanding. In fact the rate of work is the removal of a volume of metal in a unit time. The width of the zone treated across the railhead and the advance speed of the machine must be taken into account for correct evaluation. There are also two geometrical effects:
- when grinding wave formation the first passes, which act upon the crests, remove a greater *depth* of metal
- when grinding the transverse profile the units on the gauge and field corners, which act upon a small-radius contact, remove a greater *depth* of metal

Thus defining performance by average metal removal requires careful specification of the circumstances.

The depth of cut of an individual stone depends upon:
- type of stone
- power of the drive unit motor
- area, and shape, of the contact rail-stone
- speed of advance

and the depth of cut of a machine with multiple stones depends upon:
- number of stones
- angular disposition of the stones
- order of cut

All performance considerations are subjected to the over-riding requirements of:
- respect of product specifications
- practical stone life

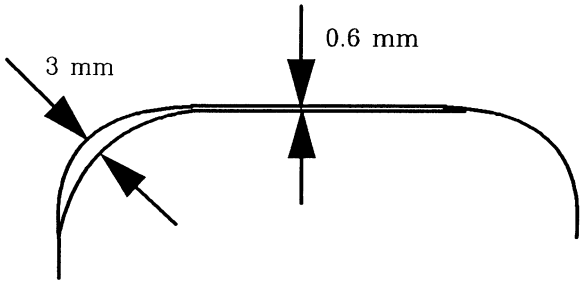
and within this scheme a distinction must be made between grinding on the gauge or field corner and grinding on the running surface. The railhead radius at the point of contact is sensitive to the extent that grinding on a 13mm radius versus a 300mm radius can:
- double the depth of cut
- double the surface roughness
- halve the stone life

To understand the new situation, the Sncf undertook in January 1991 a series of practical tests with a Speno International 32-stone machine fitted with modern stones. The trials centered on rail surface finish and hardness.

A partial illustration of results:
(measurements on running surface of roughness in microns, in conditions where six passes of the machine remove about 3mm on gauge corner and about 0.6mm on running surface) :

- before grinding
- after "roughest" intermediate pass aimed at removing metal
- after final pass intended to give acceptable finish
- after traffic smoothing (about 40,000 gross tonnes/day)):

example	*number* passes	*before* grinding	*"roughest"* grinding	*after* grinding	*after* 1 day
1.	6	"0"	15	2	2
2.	4	2	7	3	1
3.	5	1	18	5	4
4.	2	"0"	4	3	1



After six passes with a 32-stone machine

Fig 4 Typical metal removal

CONCLUSION

A broad comparison between current rail rectification specifications and grinding practice indicates that there is approximate equivalence. However, a more precise look at the characteristics directly related to stone technology - discolouration and surface roughness - shows that the limit of current acceptability may have been reached.

It may be possible to develop grinding stones of greater cutting power, but before going further suppliers need to know whether current constraints on rail treatment are definitive. While awaiting more information Speno International is concentrating on prolonging stone life for the same performance. On that score there is no doubt about the benefit for all.

IN TRACK QUALITY CONTROL AND ASSESSMENT OF RAILS

Coenraad Esveld

Consulting Engineer, The Netherlands

Summary

Rails are an important component of the total track structure. Irregularities and imperfections directly cause high dynamic forces which might be responsible for rapid deterioration and damage to the track. This paper discusses measuring systems presently in use to control rail geometry efficiently. Both high speed recording systems and static manual equipment are covered, while also attention is paid to monitoring of wheel imperfections via track mounted measuring systems.

1. Introduction.

When considering quality control of rails two main aspects can be distinguished: quality control at the rolling mill and quality control in track [1]. This paper primarily deals with the last aspect.

At the rolling mill there is normally an internal quality control system and an independent control system for acceptance of the rails. Most railways nowadays buy rails on certificate in which the supplier guarantees that the rails meet the specifications agreed upon. In Europe most rails are manufactured with an extreme tight tolerance for straightness: maximum peak deviation less than 0.4 mm relative to a 3 m base. The straightness is checked in a continuous measuring system and also the dimensions of the rail are checked mostly with optical equipment.

Via computer control the precise chemical composition is adhered to. With the present manufacturing techniques high cleanliness is achieved and the control on flakes with the aid of ultrasonic inspection is very effective.

Most railways, especially in Europe, almost exclusively lay continuous welded rail (CWR). For this purpose the rails are welded together either in a welding depot, with flash butt welds, or in situ with thermit welds or with flash butt welds made by a mobile welder. Weld straightness is an important aspect from the point of reducing the high wheel-rail impact forces.

J.J. Kalker et al. (eds.), Rail Quality and Maintenance for Modern Railway Operation, 137–147.

λ	PHENOMENON	MONITORING
1 - 30 cm	corrugations	axle box acc
1 - 100 cm	poor welds poor wheels	axle box acc rail accelerations
2 - 3 m	rolling defects	axle box acc/displ

Fig 1: Relevant wavebands for rail quality assessment.

In track the rails deteriorate as a result of the wheel-rail forces and the obvious remedial action is grinding: with phenomena like corrugations this action is focussed on the longitudinal direction, whereas with phenomena like shelling, headchecks and plastic deformation, due to inconsistent wheel-rail profiles, profile grinding is required to restore the appropriate cross sectional rail shape.

If the longitudinal rail geometry is considered in more detail different wave bands can be distinguished, which are characteristic for a particular phenomenon or a class of maintenance methods. Figure 1 presents a survey of these wavebands together with the maintenance processes available. Most of the modern track recording cars produce this, or part of this information in wavebands of similar dimensions.

With the aid of rail flaw detection, normally carried out by ultrasonic inspection systems, rail defects are located. This information is normally collected in a data base for further analysis aiming at determining when rail service life will expire.

When the causes of rail deterioration are studied these often happen to be correlated with the quality of the wheels passing over the track. It is therefore advised to also look at this aspect, with special emphasis on detecting damaged wheels in an early stage so that they can be taken out before ruining the track.

In the subsequent chapters a number of measuring tools for in track rail assessment will be discussed.

1. Measurement equipment for use on recording cars

1.1 Measurement of corrugations and poor welds.

It is extremely difficult to measure the geometry of corrugations and poor welds directly at high speed. Systems have been developed to do this indirectly via axle box accelerations. In fact such systems look at vehicle reactions, ie the response of the wheelset. A digital Corrugation Measuring System (CMS) was developed in the Netherlands by High Tech Automation [3]. This system produces per rail axle box accelerations at a fixed reference speed and geometry (wave depth).

The relationship between rail geometry and axle box acceleration is primarily governed by contributions from wheel, rail and Hertzian contact spring. As a first approximation this relationship may be considered as linear. Figure 2 shows this relationship. Obviously the wheel produces by far the greatest contribution in the frequency band up to about 50 Hz. The rail is mainly responsible for the behaviour in the 50 to 1000 Hz band and the Hertzian contact spring determines the behaviour above 1000 Hz.

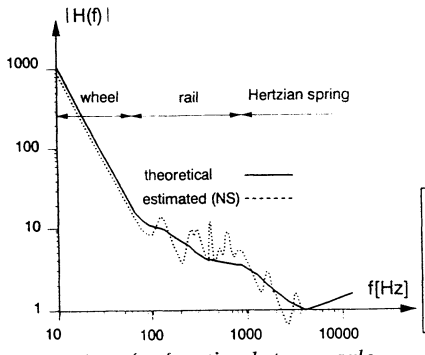


Fig 2: Transfer function between axle box acceleration and rail geometry.

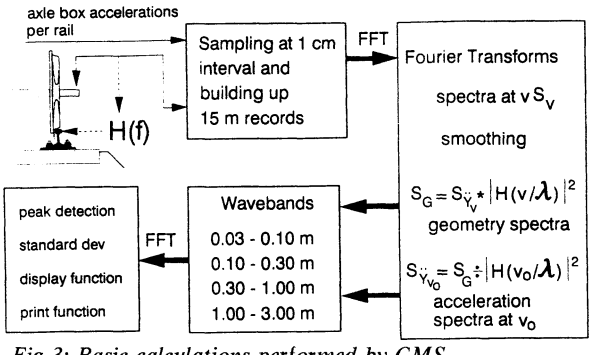


Fig 3: Basic calculations performed by CMS.

A special algorithm was developed to estimate the precise transfer function for a specific measuring wheelset from a number of repeatability runs at different running speeds. Preservation of energy, together with an optimization of the transfer function parameters according to the least square principle, form the basis of the estimation procedure. The modulus of the transfer function for the system running in The Netherlands is also displayed in figure 2.

The setup of the digital system is outlined schematically in figure 3. Left and right hand acceleration signal are sampled with an interval of 1 cm and overlapping blocks of 15 m are transformed to the spatial frequency domain. Psd functions are subsequently created and smoothed to satisfy statistical reliability. For the current speed the transfer values are determined and the squared values are multiplied by the corresponding acceleration spectra components to obtain geometry spectra. The acceleration spectra at the reference speed are obtained by dividing the geometry spectra by the squared transfer values at the reference speed v_0.

The spectra are split up into a number of user determined wavebands. The square root of the spectra, combined with the phase information, are used to back transform the signals to the space domain. The system checks whether the weld signals in the 0.3 - 1.0 m waveband exceed the user defined threshold levels.

Standard deviations are produced for both accelerations and displacements in the corrugation wavebands and for the displacements associated with the rail rolling process.

1.2 Rail profile measurement

In most of the track recording systems, especially in Europe, one important factor is lacking: the cross section of the rails. Knowledge of the rail profile may provide valuable additional information in the decision making process for rail profile grinding, aiming at restoring a suitable profile from the point of view of curving and contact stresses. Also track gauge and lateral rail wear can be monitored and, if the accuracy is sufficient, parameters like equivalent conicity and rolling line off-set can be produced for accommodating vehicle dynamics studies.

Several rail profile measuring systems are available on the market, like for instance RailScan and LiteSlice. Because of the extremely high requirements from the point of inspection speed, resolution and accuracy TNO's Institute of Applied Physics (TPD) in the Netherlands has proposed the development of a Rail Profile Measuring (RPM) system, as described in reference [5], for Netherlands Railways.

Compared to inertial systems RPM has the advantage that it is a static measurement system, which shows the rail's cross section with a constant resolution and accuracy. The longitudinal resolution increases with decreasing inspection speed.

In figure 4 an arrangement for the proposed RPM system is shown. The system consists of an illuminating system, an imaging system and a detector system. The illuminating system may consist of a (semi-conductor) laser with a collimator lens, to provide a light plane. The cross section of the light plane and the rail head is imaged onto the detector's surface. A special imaging arrangement features low distortion and large depth of view. This implies that no post processing is necessary to obtain rail profile data.

The user requirements, set for the development phase of the RPM system are:

- inspection speed up to 200 km/h;
- resolution/accuracy of rail profile change better than 0.1 mm;
- transversal sample distance (cross section) < 1 mm;
- longitudinal sample distance < 100 mm.

The PSD camera is very well suited to perform contactless and rapid 3D geometry measurements, for instance in a rail profile measuring system. The principle is based on the triangulation method. The cross-section of the light plane from an illumination system, e.g. a photo diode, and the object results in a thin light strip, which is imaged on the PSD-chip that consists of 96 photosensitive strips. Each of these strips generate an electrical current on the electrodes on the outer side of the strip. The exact position of each part of the profile is derived from the ratio of these two currents.

The main advantage of a camera built around the PSD-chip over a CCD camera, is the high speed with which measurements can be performed. It takes about 20 ms to read a frame from a CCD-chip, while it takes only 64 μs to scan the 96 strips of

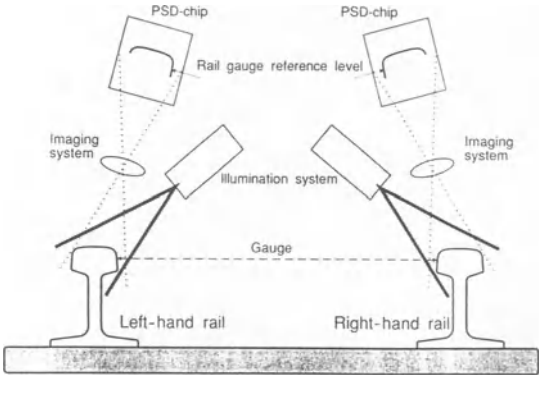


Fig 4: Rail profile and gauge measuring setup.

the PSD-chip and with that to register a complete picture. With a triangulation angle of 30° and a reduction of 4 times, the field shown on the chip is 10.8 (y) at 50.7 (z) mm.

The resolution depends on the adjustment of the optics, the triangulation angle and the amount of light which reaches the chip. A resolution between 0.03 and 0.1 mm in depth or z-direction is well achievable. The resolution in the y-direction typically is better than 0.03 mm. The resolution in the x-direction is determined by the movement of the camera with respect to the object.

1.3 Ultrasonic inspection

Most of the modern railways nowadays make use of ultrasonic inspection for rail flaw detection purposes. The state of the art system operated by Netherlands Railways is briefly described hereafter. All measurement data are recorded on an optical disk, enabling off-line trend analyses and optimization of measurement data classification. This, in turn, allows predictions to be made concerning:

- crack development, essential to decide when to replace a defected section;
- reduction in rail height, which gives a good estimate of the tonnage borne;

The maximum measuring speed is 50 km/h and the maximum distance which can be tested in one day is 200 km. Tests are carried out every 2 mm, regardless of speed. This provides the high precision needed for detecting cracks, particularly star cracks, as early as possible. In contrast to most other ultrasonic rail testing cars, the NS car can continue measuring through switches and crossings.

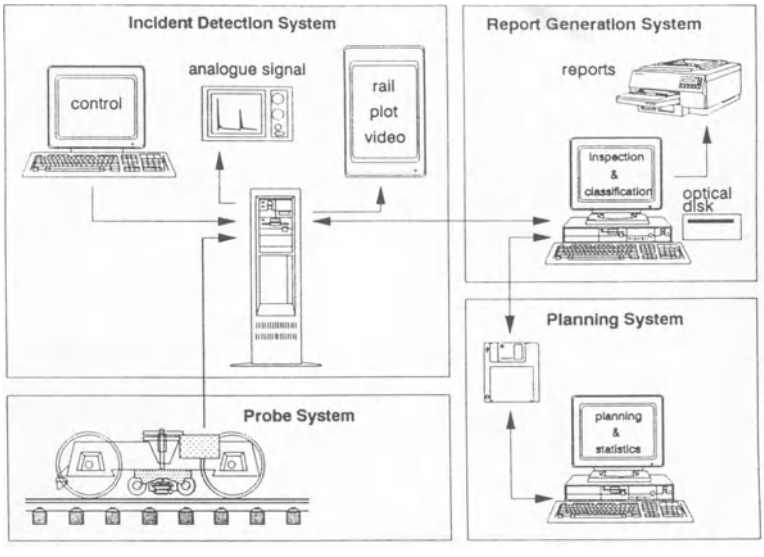


Fig 5: Setup of the NS ultrasonic measuring system.

The probe system has 0°, $+70^{\circ}$ and -70° probes for each rail, with the capability of investigating the head, web and central part of the foot. The probe shoe is of a special type, designed to maintain optimum contact with the rail even at high speeds. To improve the signal to noise ratio, the electronics for the probe are mounted remote from it, in a special cabinet inside the car.

The electronic system consists of the real time Incident Detection System (IDS) and the off line Report Generation System (RGS) as explained in figure 5. The IDS features are:

- menu driven operation;
- real-time display of analog probe signals on four dual channel oscilloscopes;
- real-time display of digitized results on a high-resolution monitor;
- high-resolution plot on fast digital plotter;
- storage of all measurement, configuration and administrative data on hard disk;
- automatic measurement start via detection of permanent magnets in the track;
- detection sensitivity standardized for depth;
- measuring area adjustable separately for each probe;
- automatic adjustment at changes of profile;

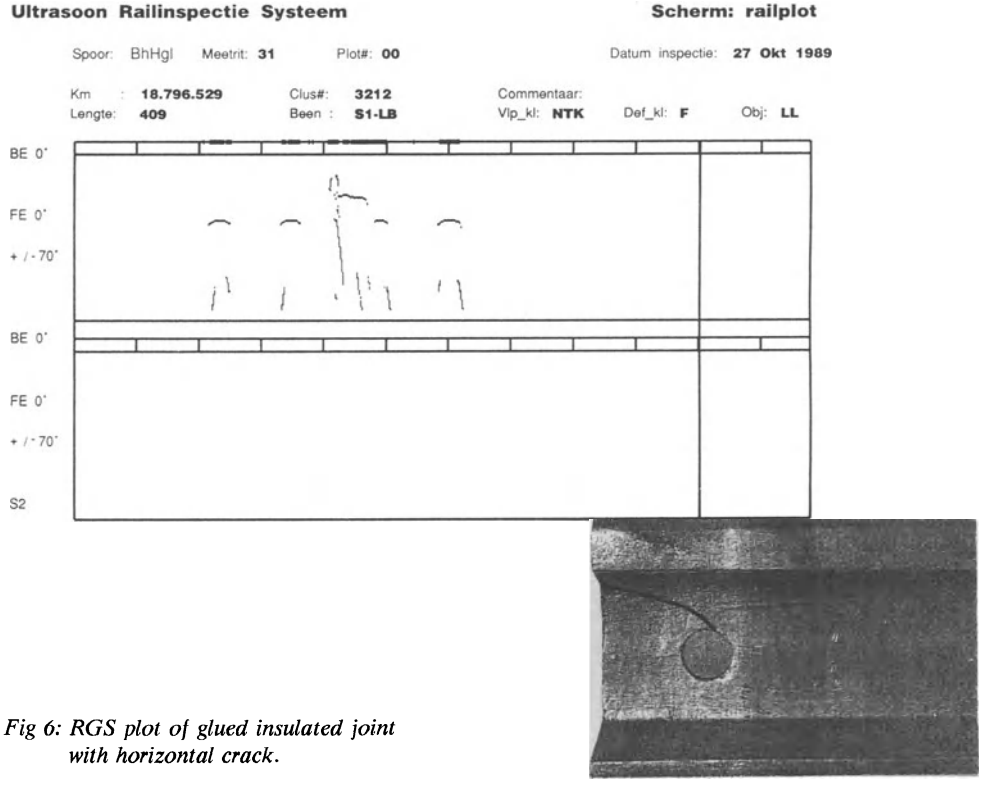


Fig 6: RGS plot of glued insulated joint with horizontal crack.

The Report Generation System (RGS) is based on a powerful IBM PS/2-90 PC, fitted with an 800 MB WORM drive for the storage of all measurement data. This runs an Oracle database with application-specific graphics functions. The main features are:

- menu driven operations;
- automatic classification of measurement data files;
- geometric transformation of probe signals;
- interactive classification on high-resolution colour monitor;
- consultation of historic data to aid interactive classification;
- graphic hard copy of defects on high-resolution plotter (figure 6);
- defect report with detailed location information, including distances to the nearest hectometre post and reference construction;
- rail hight tables, giving mean and standard deviation of the vertical wear; automatic reset each time profile changes;

2. Wheel monitoring

The mechanical exposure of tracks is increasing steadily due to higher axle loads on heavy haul railways, higher running speeds on European freight lines, introduction of high speed trains and implementation of more and more through-going trains.

High frequency wheel load components due to damaged wheels can be significantly higher than static loads. By detecting these wheels in an early stage via wheel flat detectors, the track engineer can - to some extent - protect his rails and sleepers.

The most well known principle of monitoring high impact loads is measuring vertical rail acceleration. The system described in [4] consists of 7 accelerometers per rail spaced at sleeper distance. The layout is sketched in figure 7. The length is sufficient to cover at least one wheel revolution.

The signal processing of the accelerometers is controlled by means of wheel sensors mounted in the track. These wheel sensors determine the position of each wheel, which information is used to switch on the appropriate accelerometer when a wheel passes.

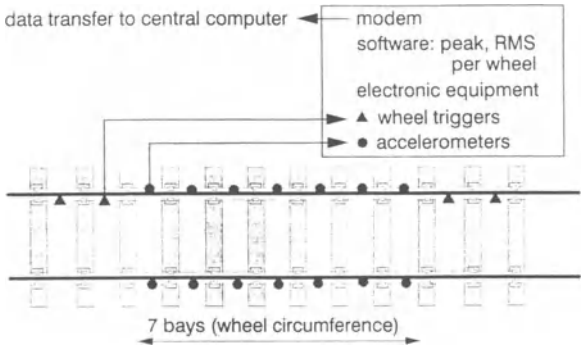


Fig 7: Layout of the Caltronic wheel monitoring system.

The system determines the maximum acceleration peak per wheel and the RMS-value to cover both single peaks and periodic forces. The information is stored in a memory and then transmitted via a telephone modem to a central computer for further monitoring purposes and off line statistical analyses.

A number of railways have initiated the use of wheel flat detectors. Caltronic installations have been installed at Danish railways DSB, the Finnish railways VR, the Swedish Rail Authority Banverket, and the Czechoslovak railways CSD. In the summer of 1991 the Caltronic system have been through a very comprehensive test programme at the Association of American Railroads (A.A.R.) testing facilities in Pueblo, Colorado.

The future need for heavier axle loads and faster trains will inevitably necessitate a more intense quality control of wheels. The well established "60 mm length" rule for wheel flats in Europe and the American "Rule 41A" will not be sufficient for this purpose and therefore have to be replaced by more precise descriptions required for assessment by means of automatic track mounted equipment.

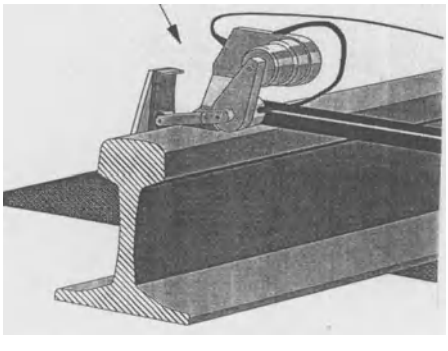
3. Manual equipment

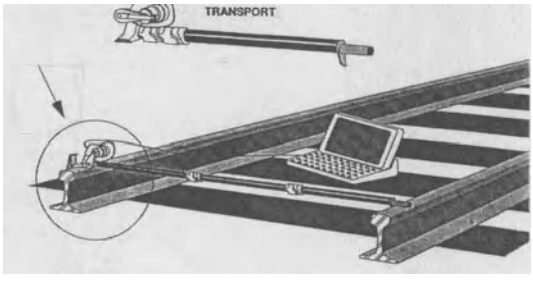
3.1 MINIPROF system

As the wheel and rail profiles constantly change by wear systematic control is useful to minimize maintenance costs. A recently developed system called MINIPROF, referred to in [2], will be briefly described here. The system consists of a notebook computer with dedicated electronics for interfacing and integration.

The Wheel-Unit is magnetically attached to the wheel. The Rail-Unit uses the opposite rail as a reference by a foldable fixture.

The sensing element, which is identical for both the wheel and rail unit, is a small magnetic wheel attached to the end arm, made by two joint extensions. When moving the arm the extensions are rotated. The rotatory method makes fast and precise measurements possible. The observed angles of the extensions are sent to the computer, where the actual profile is calculated and stored. A rechargeable battery provides approximately 3 hours of uninterrupted use. Figure 8 shows the measuring principle for rail profiles and figure 9 for wheel profiles.





The MINIPROF provides:

- geometric profile, available graphically in real time and as x, y and slope in ASCII format for further analysis;
- difference between reference profiles and actual profile. Material and time can be saved by optimising the new profile;
- wear patterns which ease diagnostics and help to explain excessive wear rates;
- Profile slope to determine and optimize dynamic interaction between wheel and rails;
- geometric characteristics for quality control, for instance for wheels Sd, qR and Sh; Some examples of wheel profile measurement are given in figure 10;

The specifications are as follows:

	wheel unit	*rail unit*
weight measuring unit:	0.6 kg	1.1 kg
weight notebook computer:	3 kg	
accuracy better than:	0.05 mm	
accuracy typically:	0.01 mm	
measuring rate:	10,000 points per second	
profile scanning:	3 sec	
profile processing:	2 sec (with co-processor)	
storage capacity on hard disk:	more than 10,000 profiles	
storage capacity on floppy disk:	more than 200 profiles	

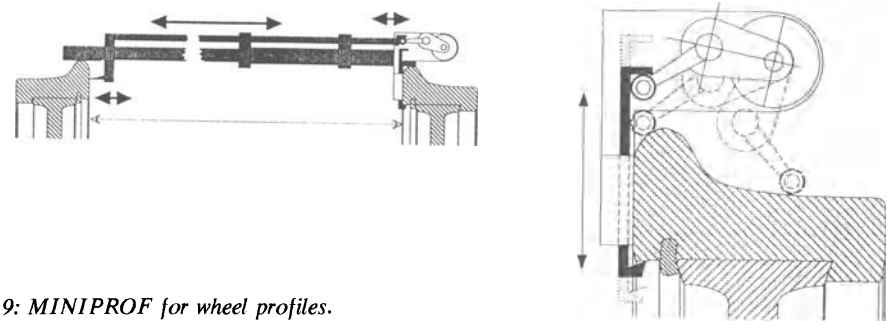


Fig 9: MINIPROF for wheel profiles.

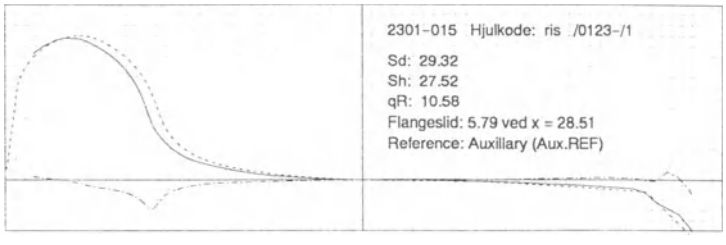


Fig 10: Wheel profiles recorded with MINIPROF.

3.2 Electronic measurement of rail geometry by RAILPROF

When considering longitudinal rail geometry this is normally measured continuously as discussed before. For manual recording the Cemafer apparatus is a commonly used tool, comprised of a 1.2 m reference bar relative to which a moveable, hand driven pen, records the off sets on paper. This tool is adequate for tests, but under practical conditions it is too slow and not sufficiently ruggedised.

This has led to the development of an electronic straight edge, further referred to as RAILPROF, designed as a rugged and easy to handle, yet accurate device, for use in the field and in the workshop. It records both vertical and lateral rail geometry over a length of 1.2 using one non-contact electronic sensors per measuring direction.

The recorded information is displayed on a LCD screen as graphs, or condensed information in the form of maximum versine, maximum step, maximum first derivative and standard deviation, in accordance with the selection of the user. All measurement data are stored in a memory to facilitate down loading in a central computer for post statistical analyses.

Typical applications of RAILPROF are:

- prior to thermit welding: monitoring and assessment of the rail end positions relative to each other to obtain the correct lateral and vertical geometry, including the overlift in vertical direction for compensating the shrinkage after welding;

- during and after weld grinding: recording of the rail geometry in order to meet the required standards, and to achieve effective quality control by both the welding troops and the management;

- in case of rail corrugations: assessment of vertical rail geometry to enable timely corrective grinding of the rail surface;

- in case op rail grinding: monitoring of the grinding result during the course of the process;

Relatively small irregularities in the rail running surface can already cause high impact loads, leading to accelerated deterioration of the track construction. Limiting of these irregularities to acceptable levels is an essential step in controlling the service life of the track and its components. Netherlands Railways, for instance, require for vertical weld geometry a versine on a 1.2 m basis of less than 0.2 mm and a step, defined as the differential displacement over 20 cm, of less than 0.1 mm. Corrugations, on the other hand, start to develop at values in the order of 0.05 mm peak to peak and then have to be ground off as soon as possible.

To be able to control the processes in conjunction with monitoring and grinding of rail corrugations and the production of welds, both on site and in the depot, the measuring equipment used should meet high performance specifications. Therefore RAILPROF is equipped with specially developed high accuracy sensors, which take the average value of an area of about 10 mm^2 to eliminate the roughness of the rail surface; these sensors have the additional advantage of not being influenced by rust and dirt on the rail.

With this concept RAILPROF achieves an overall accuracy which is better than 0.05 mm.

References

[1] Esveld C.: "Modern Railway Track", MRT-Productions, fax: +314180 16372, ISBN 90-800324-1-7.

[2] Gronskov L.: "MINIPROF a miniature tool for quick & neat measurement of profiles on train rails and wheels, Internal publication, December 1991.

[3] High Tech Automation: "Digital Corrugation Measuring System". HTA internal report, 1990.

[4] Danneskiold-Samsoe U.: "Improving Track Quality by Means of Wheel Flat Detection", Caltronic internal report, July 1991.

[5] Roos J. and F.G.W. van Trigt: "Rail Profile and Gauge Measurement", TNO/-TPD internal report, July 1991.

CONTACT PHENOMENA

COMPUTATIONAL CONTACT MECHANICS OF THE WHEEL-RAIL SYSTEM

by J.J. Kalker

Professor of Mathematics

Delft University of Technology

Abstract

In this paper a survey is given of contact mechanics as it is applied to the wheel-rail system. This means that we will concern ourselves with the elastic and plastic rolling contact of bodies made of identical materials. Moreover, these bodies may in vast majority be approximated by half-spaces.

After an introduction, the boundary conditions of contact are discussed. It is found that they are twofold, *viz.* the contact formation conditions and the frictional conditions. Numerical theories regarding elastic contact problems are treated, and also plastic rolling contact theory is considered.

Finally the impact of contact mechanics on some railway problems are discussed:

1. Simulation of vehicle motion;
2. Abrasive wear;
3. Crack inception and propagation;
4. Residual stress.

1. Introduction

Contact mechanics is the mechanics of solid bodies touching each other. Distinction may be made between rigid and deformable bodies: rigid bodies are treated by J.J. Moreau, eminent mechanicist and mathematician from Montpellier, France; the contact between deformable solids was initiated by Hertz (1882), and later by Cattaneo (1938) and Mindlin (1949-1955), and by K.L. Johnson (1958–the present).

In contact mechanics a special role is played by rolling contact, and indeed the wheel/rail contact falls in that category. The first to consider rolling contact was Carter (1926), actually in a railway setting. He considered two infinite cylinders rolling over each other with parallel axes, and made of the same linearly elastic material, *e.g.* steel on steel. As our present interest is the wheel-rail rolling contact which actually is steel on steel, we will omit the qualification "made from equal materials" from our considerations. — Carter solved his problem exactly, but it was 2D.

J.J. Kalker et al. (eds.), Rail Quality and Maintenance for Modern Railway Operation, 151–164.

Inspired by this, K.L. Johnson, later in cooperation with his assistant P.J. Vermeulen, approximately solved the rolling contact problem without conicity (1964). The purpose was again the railway application: in effect vehicle motion simulation.

Kalker started work to remove the restriction and the approximation; his results were laid down firstly in his thesis (1967), and later in his book (1990). As vehicle dynamics simulation requires the utmost of speed of calculation, he laid much emphasis on fast yet accurate calculations. So his work developed along two tracks: on the one hand, an exact contact theory was developed which took into account the complete theory of elasticity (CONTACT, see Kalker (1990)): this program was used to verify the various fast approximations of restricted classes of rolling contact that were constructed on the other hand:

1. **Table Book**: Developed by British Rail; a CONTACT-like program was used to construct the tables (ca. 1980);
2. **FASTSIM**: Approximate elastic equations are used;
3. **Linear Theory**: For small tangential forces, and small conicity;
4. **Explicit Formulae**: No (1968) or small (1990) conicity taken into account.
5. **Pascal's Method**: For non-Hertzian wheel-rail contacts.

In Section 2 of this paper, we will model the contact phenomena such as they occur in the ideally elastic contact of a perfectly smooth, elastic wheel rolling over a perfectly smooth elastic rail. In sec. 3, we describe the elastic theory that lies at the root of CONTACT; in sec. 4 we describe the theories 1-5. In sec. 5, we consider the plastic theory of rolling, and in Sec. 6 we discuss the impact of contact theories on

- Simulation of vehicle motion;
- Wear of surfaces;
- Crack inception and propagation;
- Residual stresses.

2. Modelling contact phenomena

Notation: Bold faced symbols designate vectors.

Contact mechanics is, in essence, a special boundary value problem of solid mechanics. Indeed, we must distinguish between contact formation and friction.

2.1 CONTACT FORMATION

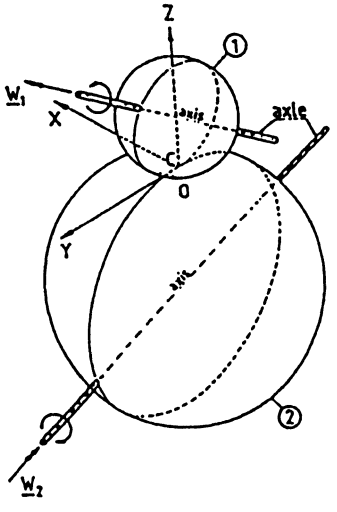


Fig. 1. Two bodies rolling over each other

Consider two bodies in contact, see Fig. 1. The bodies are numbered (1) and (2). The bodies touch each other in the contact area C; it is unknown *a priori*. We introduce the potential contact area P which must contain the contact area, but which is otherwise arbitrary; we assume it known. The area E ("exterior") is the difference of the pot.con. and the contact area. Let $\boldsymbol{x}$ be a point of body (1) and P; the distance of the bodies at $\boldsymbol{x}$ is denoted by $e(\boldsymbol{x})$. e is taken positive where the bodies are apart; zero where they touch; negative where they penetrate. The latter cannot happen:

$$e(\boldsymbol{x}) \geq 0 \text{ in } P, \; = 0 \text{ in } C, \; > 0 \text{ in } E \tag{1}$$

In addition, surface loads $p(x)$ act between the bodies at P; we define $\boldsymbol{p}(\boldsymbol{x})$ as the traction on the surface of body (1). A local coordinate system (x, y, z) is introduced with $\boldsymbol{x}$ as origin; x lies in the direction of rolling; z points normally into (1), and y completes the right-handed coordinate system. We decompose the traction into normal and tangential components:

$$\boldsymbol{p}(\boldsymbol{x}) = (\boldsymbol{p}_t, p_z) \tag{2}$$

and we have that the normal component is compressive in C, and vanishes in E:

$$p_z \; \geq 0 \text{ in } P, \; = 0 \text{ in } E, \; > \; 0 \text{ in } C. \tag{3}$$

These constitute the conditions of contact formation. In elasticity, $e(\boldsymbol{x})$ depends linearly on

$$u_z(\boldsymbol{x}) = u_{1z}(\boldsymbol{x}) - u_{2z}(\boldsymbol{x}) \tag{4}$$

where u_{az} is the z-component of the elastic displacement of body (a); indeed,

$$e(\boldsymbol{x}) = h(\boldsymbol{x}) + u_z(\boldsymbol{x}); \quad h(\boldsymbol{x}) : \text{Distance at } \boldsymbol{x} \text{ when the elastic deformation is omitted.} \tag{5}$$

2.2 DRY FRICTION

Essential for friction to be dry is that in C there exists a positive function of position called the traction bound $g > 0$, that g vanishes in E, and that it is such that

$$| \boldsymbol{p}_t(\boldsymbol{x}) | \leq g(\boldsymbol{x}) \quad \text{in } P: \ g > 0 \text{ in } C, \ = 0 \text{ in } E \tag{6}$$

The slip $s_t(x)$ is the tangential component of the velocity of body (1) over body (2) at x, in C. It plays an important role in dry friction, namely

$$\boldsymbol{s}_t(\boldsymbol{x}) \neq \mathbf{o} \Rightarrow \boldsymbol{p}_t(\boldsymbol{x}) = -(\boldsymbol{s}_t(\boldsymbol{x})/ | \boldsymbol{s}_t(\boldsymbol{x}) |) \, g(\boldsymbol{x}) \tag{7}$$

This means that the tangential traction is at the traction bound and opposes the slip, if the latter does not vanish. If it does vanish, one merely has (6). For anisotropy, (6) and (7) have to be modified. If this is taken into account, tribologists agree on the validity of (6) and (7).

The traction bound depends on several parameters. Although tribologists consider this an oversimplification, the traction bound is often written as

$$g(\boldsymbol{x}) = f p_z(\boldsymbol{x}), \quad f\text{: coefficient of friction} \tag{8}$$

Sometimes the coefficient of friction is regarded as being equal to two constants, f_{stat} and f_{kin}: *viz.* $\boldsymbol{s}_t = \mathbf{o}$: f_{stat}, otherwise f_{kin}. It is our experience with CONTACT that only when f_{stat} and f_{kin} differ considerably, a real effect is observed.

Like the distance $e(\boldsymbol{x})$, the slip $\boldsymbol{s}_t(\boldsymbol{x})$ may be expressed in the tangential component $\boldsymbol{u}_t$ of the *displacement difference* $\boldsymbol{u}$:

$$\boldsymbol{u} = \boldsymbol{u}_1 - \boldsymbol{u}_2, \quad \boldsymbol{u} = \text{displacement difference} \tag{9}$$

Indeed we have, after division by the rolling velocity V:

$$\boldsymbol{s}_t/V = \dot{\boldsymbol{x}}_t + \dot{\boldsymbol{u}}_t, \tag{10}$$

$$\boldsymbol{s}_t/V\text{: relative slip} \tag{11}$$

with

$$\begin{aligned} \dot{} &= d/dq \\ q &: \text{ distance traversed by body (1) over body (2)} \\ \dot{\boldsymbol{x}}_t &: \text{ Relative slip, when the elastic slip } \dot{\boldsymbol{u}}_t \text{ is omitted.} \end{aligned}$$

In numerical contact mechanics, equation (10) is discretized, with steps d of q. In this discretisation it should be observed that the differentiation ($\dot{}$) is particle fixed, *i.e.* one compares the displacement on one and the same particle at various instants of time. We multiply by d, and find

$$\boldsymbol{s}_t(\boldsymbol{x}, q)d/V = \dot{\boldsymbol{x}}_t(\boldsymbol{x}, q)d + \boldsymbol{u}_t(\boldsymbol{x}, q) - \boldsymbol{u}_t(\boldsymbol{x}, q-d),\ d: \text{ the discretization step.} \quad (12)$$

One can also write this in terms of coordinates rather than of particles. Let $\boldsymbol{y}$ be the surface coordinates fixed to the contact area of the point $\boldsymbol{x}$ at time q; when $\boldsymbol{i}$ is the unit vector in the direction of rolling, the point $\boldsymbol{x}$ was situated at $(\boldsymbol{y} + \boldsymbol{i}d)$ at the time *(q-d)*. So,

$$\boldsymbol{s}_t(\boldsymbol{y}, q)d/V = \dot{\boldsymbol{x}}_t(\boldsymbol{y}, q)d + \boldsymbol{u}_t(\boldsymbol{y}, q) - \boldsymbol{u}_t(\boldsymbol{y} + \boldsymbol{i}d, q-d), \quad (13)$$

$$\boldsymbol{i}: \text{ unit vector in the direction of rolling} \quad (14)$$

Note that this formula covers both transient and steady state rolling. Transient rolling is represented by eq.(12), where $\dot{\boldsymbol{x}}_t$ and $\boldsymbol{u}_t(\boldsymbol{y} + \boldsymbol{i}d, q-d)$ are known. In steady state rolling, there is no dependence on explicit q, so that the second arguments in (12) may be omitted:

$$\boldsymbol{s}_t(\boldsymbol{y})d/V = \dot{\boldsymbol{x}}_t(\boldsymbol{y})d + \boldsymbol{u}_t(\boldsymbol{y}) - \boldsymbol{u}_t(\boldsymbol{y} + \boldsymbol{i}d):\ \text{steady state rolling} \quad (15)$$

In addition, cases that are not rolling at all may be represented by (12). Indeed, gross sliding takes place when $|\ \dot{\boldsymbol{x}}_t d\ | \gg |\ \boldsymbol{u}_t(\boldsymbol{x}, q) - \boldsymbol{u}_t(\boldsymbol{x}, d-q)\ |$. Then the displacement may be neglected, and

$$\boldsymbol{s}_t(\boldsymbol{x}, q)/V = \dot{\boldsymbol{x}}_t(\boldsymbol{x}, q):\ \text{gross sliding} \quad (16)$$

We write (12), which is perfectly general, in an abbreviated way:

$$\boldsymbol{S} = \boldsymbol{X} + \boldsymbol{U} - \boldsymbol{U}' \quad (17)$$

$$
\begin{aligned}
S &= s_t(\boldsymbol{x}, q)d/V \text{ : shift} \\
X &= \dot{\boldsymbol{x}}_t(\boldsymbol{x}, q)/d \text{ : rigid shift} \\
U &= \boldsymbol{u}_t(\boldsymbol{x}, q) \text{ : present displacement} \\
U' &= \boldsymbol{u}_t(\boldsymbol{x}, q-d) \text{ : past displacement}
\end{aligned}
$$

3. The exact theory

In 3.1 we describe the contact formation, in sec. 3.2 rolling contact.

3.1 THE NORMAL CONTACT PROBLEM

During the late 1970's, the normal contact problem was attacked by a number of authors (Reusner (1977)), Nayak and Johnson (1978)) by dividing the candidate contact area into a number of strips, see Fig.2a,b.

Inside the contact area C, the distance $e(\boldsymbol{x})$ =0, so that

$$u_z(x,y) \;=\; -\,h(x,y) \tag{18}$$

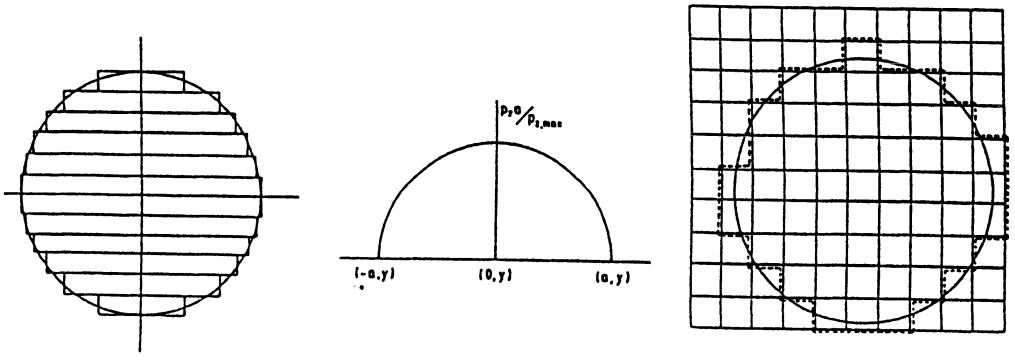


Fig.2. (a): A contact area divided into strips; (b): The pressure on the strips; (c): A contact area discretized in rectangles.

In many cases, the following expression holds for h,

$$h(x,y) \;=\; A(y)x^2 \;+\; D(y);\quad A, D \text{ : known functions of } y \tag{19}$$

Indeed, the distance h between two ellipsoids is of this type, and Hertz's semi-ellipsoidal pressure distribution is the solution. So we take the pressure semi-elliptical in each strip, see Fig. 2b, and we fit a (boundary of contact), and $p_{z,max}$ as well as we can to (17)-(18).

The representation of Fig. 2a,b has several disadvantages. The first is, that not all exact normal problems satisfy (18). The second is that the displacement all over the contact area due to the pressure in a strip is hard to calculate; the third is that the calculation as a whole may fail; and the fourth is that the validity of the approximation is open to doubt. So, new ways were sought in the early 1980's by Paul and Hashemi (1981) and by Kalker (1983), see Kalker (1990) for both methods.

These authors divided the contact area in small rectangles called elements, see Fig.2c. In Paul and Hashemi, moreover, the boundaries of contact are approximated in a special manner. The traction is taken constant over each rectangle. It was pointed out by Fichera already in 1964 that the conditions of contact formation can be embodied in a minimum principle, that as a consequence of convexity has a unique solution. It was shown by Duvaut and Lions (1972) that this holds also for friction, if the traction bound g is prescribed numerically. All this holds for elastic bodies; Duvaut and Lions prove their case also for viscoelastic bodies.

Based on this minimum principle Kalker (1983; see 1990) proposed an algorithm that solves the discretised contact formation in a finite number of steps (typically 4) where each step consists essentially of the solution of n linear equations with n unknowns, where n is the number of elements in the candidate contact area.

Kalker solved these equations with the aid of Gauss Elimination (GE) (1983); as the coefficient matrix is positive definite, Carneiro (1987) and Vollebregt (1992) used Gauss-Seidel (GS) instead; Vollebregt, in addition, uses a MultiGrid method (MG) in conjunction with GS. Both used a vector computer. As GE's timing is proportional to n^3 and GS's to n^2, while MG yields a constant acceleration factor, this yields a spectacular acceleration. Carneiro went as far as using 1000 elements in C. The algorithm is formulated in a plausible manner in the pressure p_z, the deformed distance e, and the discretized contact area C. No explicit reference is made to the functional that is minimized: yet minimized it is.

The algorithm seems simple, but appearances deceive. It is perfectly robust.

3.2 THE TANGENTIAL PROBLEM OF FRICTION

The algorithm can be adapted to the frictional problem, where it works equally well, although no proof is offered in this case. The equations at the heart of each iteration are non-linear; Kalker solved them by linearization (Newton-Raphson (NR)) and GE. Pascal (1992) showed that under extreme conditions N-R may converge slowly, or even diverge. Jäger (to be published) and Vollebregt each made attempts to apply GS to the problem. Jäger abandoned Kalker's algorithm, which results in non-robustness. Vollebregt could retain Kalker's algorithm by a neat trick; yet he also cannot claim complete reliability.

When the bodies have different elastic constants, the minimization problems do not exist. By a trick, *viz.* by prescribing the traction bound g, one regains the minimization problems; g is then adjusted by iteration (fixed point process).

4. The Fast Rolling Contact Theories

We will treat the theories in the order they were enumerated in the Introduction.

4.1 TABLE BOOK

It was stated before that the exact theoretic algorithms are quite slow. This induced British Rail Research (*ca.*1980) to use such a program to extensively tabulate the rolling contact laws with the aid of CONTACT's predecessor DUVOROL. The table has to satisfy certain demands, *viz.* that entries are easily extracted from it, and that the entries are close enough to allow easy interpolation. DUVOROL is confined to steady state rolling with a Hertzian normal pressure, and it is completely reliable. BRR made about $10^4 - 10^5$ entries, confined itself to the Poisson ratio of steel $\nu = 0.28$, and to elliptical contact areas. Needless to say, the elastic constants were equal.

Nowadays one is not satisfied with Hertzian contact areas. I propose to generate a Table Book in a few hours or days for *each* combination of wheel and rail. Parameters are: the lateral position y of the wheel with respect to the rail, the normal force F_z on the wheel, and the longitudinal and lateral creepage at a reference point: four parameters in all, leading to a Table Book of about 10^4 entries. The spin follows from the local conicity that is known from the profile, and Poisson's ratio and Young's modulus are of course likewise known.

It has appeared , however, that difficulties arise with CONTACT in flanging. But instead of CONTACT, Pascal's multi-elliptical contacts together with FASTSIM may be used to generate the Table Book.

4.2 FASTSIM

FASTSIM is a fast simulation routine to calculate frictional contact problems on the basis of the Simplified Theory of Rolling Contact. In this simplified theory, the *displacement difference* $\boldsymbol{u}_t(\boldsymbol{x}, q)$ is assumed to depend on the traction $\boldsymbol{p}_t(\boldsymbol{x}, q)$ alone, and not on the tangential traction at other points $\boldsymbol{x}$ and times q, as distinct from classical elasticity where the *strain* at $\boldsymbol{x}$,q depends on the stress on the same point alone. In fact, we have in simplified theory

$$\boldsymbol{P} = L\boldsymbol{U}, \quad L\text{: a constant called } \textit{flexibility}. \quad \boldsymbol{P} = \boldsymbol{p}_t(\boldsymbol{x}, q), \quad \boldsymbol{P}' = \boldsymbol{p}_t(\boldsymbol{x}, q\text{-}d) \tag{20}$$

Then, (16) becomes

$$\boldsymbol{S} = \boldsymbol{X} + L(\boldsymbol{P} - \boldsymbol{P}') \tag{21}$$

In the so-called *area of adhesion* $\boldsymbol{S} = \boldsymbol{o}$. Assuming $(\boldsymbol{x}, q)$ to be in the area of adhesion, we calculate the corresponding traction $\boldsymbol{P}$, which we call $\boldsymbol{P}_H$. $\boldsymbol{P}'$ and $\boldsymbol{X}$ are assumed known:

$$\mathbf{o} = \boldsymbol{X} + L(\boldsymbol{P}_H - \boldsymbol{P}) \quad \Rightarrow \quad \boldsymbol{P}_H = (\boldsymbol{P}' - \boldsymbol{X})/L \tag{22}$$

1. Assume that

$$| \boldsymbol{P}_H | \leq fP_z, \quad P_z : \text{ normal pressure, } f\text{: coefficient of friction}$$

Then we can show that $\boldsymbol{x}$ lies in the area of adhesion, and that $\boldsymbol{P} = \boldsymbol{P}_H$.

2. Otherwise, $\boldsymbol{x}$ is in the area of slip where $\boldsymbol{S} \neq \mathbf{o}$, and we can show that $\boldsymbol{P} = \boldsymbol{P}_H f P_z / | \boldsymbol{P}_H |$.

In order to calculate $\boldsymbol{P}$ in the entire contact area, we start with an $\boldsymbol{y}$ outside the contact area. There $\boldsymbol{P} = \mathbf{o}$. we set

$$\boldsymbol{P}' = \boldsymbol{P}; \quad \boldsymbol{y} = \boldsymbol{y} - \boldsymbol{id}; \quad \text{calculate } \boldsymbol{P}; \text{ set } \boldsymbol{P}' = \boldsymbol{P}, \text{ etc.} \tag{23}$$

Outside the contact $\boldsymbol{P}$ will be zero, as $P_z = 0$. Once inside the contact, $\boldsymbol{P}$ starts to build up.

The determination of the flexibility constitutes a problem. We solve it with the aid of the linear theory, see below (Sec. 4.3).

Only steady state rolling can be treated with simplified theory with acceptable accuracy, see Kalker (1990), and that only when the contact area is elliptic: see Sec. 4.3. So, in particular, simplified theory cannot be used in the normal contact problem, and we have to use a complete theory as described in Sec. 3.1.

4.3 THE LINEAR THEORY

The basic assumption of the linear theory is that we assume the tangential traction to be smaller than the traction bound fP_z in most of the contact area, so that we may neglect the condition "$| \boldsymbol{P} | \leq fP_z$" altogether. This means that we have adhesion everywhere in the contact, or, in other terms, that $\boldsymbol{P} = \boldsymbol{P}_H = (\boldsymbol{P}' - \boldsymbol{X})/L$, when simplified theory is adopted. Now the linear theory can also be implemented on the basis of the full theory of elasticity, where the local strain depends on the local stress. A comparison of these theories yields various values of the flexibility L. A weighted sum of these values was proposed by Kalker (1990) as a final value. *At present, the linear theory can only be calculated with any accuracy when the contact area is elliptic.*

4.4 EXPLICIT FORMULAE

Nowadays the best explicit formulae are based on the theory of Vermeulen-Johnson (1964), which to that end are slightly modified. This modification is necessary because V-J leads to a linear theory which is slightly wrong. Correction with the aid of the linear theory, see Sec. 4.4, leads to a very close fit with FASTSIM, CONTACT, and Johnson's experiments (1958, 1964).

The first formula of this type was given by Kalker (1968); it can handle arbitrary longitudinal and lateral creepage. The most advanced formula is due to Shen, Hedrick and Elkins (1984), which handles longitudinal and lateral creepage and *small* spin.

4.5 PASCAL'S THEORY

Recently, Pascal(1992) proposed to approximate the non-Hertzian contact area in wheel-rail contact by by the union of several not necessarily disjunct ellipses, which he finds in an automatic manner. The tangential traction distribution is found by FASTSIM. A remarkably good agreement is found with CONTACT.

5 Plastic Rolling Contact

Up to now, we have concerned ourselves exclusively with elastic contact, and it was seen that there is a wealth of theories, computer programmes and explicit formulae in that area. The situation in plasticity is less favourable.

The reason is that plasticity, and especially elastoplasticity, is essentially different from and more difficult than elasticity. Plasticity is essentially nonlinear (**unilateral**), so that superposition is only possible to a limited extent. Only a handful of papers on plastic rolling contact of the type that is relevant for the wheel–rail problem are mentioned by Johnson in his authoritative book (1985), Ch. 9.

Residual stresses are produced in rails by cyclic loading, and the question arises whether the rail shakes down, and whether the rail suffers damage or not. When shakedown occurs, the plastic deformation due to cyclic loading stops eventually. Shakedown in rolling was first investigated by Johnson (1962), and by Johnson and Jefferies in 1963. In subsequent years these considerations were refined, and the latest developments were reported on by Bower and Johnson in (1991) and in this Session (1993).

In my opinion the last word has not been spoken on contact plasticity theory. It sems to me that a consistent use of unilateral theory both in the contact leg (see Kalker (1990) Ch. 4) and in the plasticity leg in the framework of a 3D FEM holds much promise.

6 Applications

We will now consider the contact mechanical aspects of some wheel rail applications.

6.1 SIMULATION OF VEHICLE MOTION

Towards the end of the 1950's the investigators Johnson and de Pater interested themselves in rail vehicle motion simulation. Especially de Pater saw great possibilities in the then novel computers, and he became one of the pioneers in this field in the Netherlands. As the most important forces on a rail vehicle are those the rail exerts on the wheel, a powerful impulse was given to contact mechanics. This resulted, on Johnson's side, in the theory of Vermeulen and Johnson (1964), and in various theories on plastic rolling contact discussed in Sec. 5, which, by the way, pertain more to the theory and practice of rails — subject of this conference — than on vehicle dynamics.

De Pater delegated the elastic contact mechanics to me. As in vehicle motion simulation hundreds of thousands of contact problems must be solved without much loss of accuracy, a *quest of speed* came into existence which led to a great number of rolling contact theories, approximate and accurate, fast and slow, see Kalker (1990) Ch. 2.

In my opinion the best bet for wheel-rail contact is found in a non-Hertzian table book, generated for a single pairing of wheel-rail profiles, either by CONTACT-Vollebregt or by the non-Hertzian multi-elliptic contacts of Pascal.

An important role will be played by the problem of flanging, about which we will be told by Hanson and Keer in this conference (1993). There are two types of flanging contact, viz. two-point contact, and a single contact smoothing itself about the corner of the rail.

Two-point contact can often be treated as two independent contact areas, one on the tread and one on the flange, see Piotrowski (1982). "Smoothing" contact, that is a continuous contact from the tread to the flange, is very difficult as the half-space approximation fails: we look forward to the paper by Hanson and Keer (1993).

6.2 ABRASIVE WEAR

By abrasive wear we mean wear that is proportional to the local frictional work. Distinction is made between regular wear and irregular wear.

Regular wear leads to an evolution of the profiles of wheel and rail; irregular wear leads to unroundness of the wheel and to corrugation of wheel and rail. These problems have in common that hundreds of thousands of contact problems have to be solved, which requires the utmost of the robustness and speed of the rolling contact codes. Table book and explicit formula solutions are ruled out as the profiles change (regular wear), *resp.* the rolling contact is transient (irregular wear). In addition, the local frictional work is needed in regular wear. Remaining are FASTSIM (regular wear) and CONTACT (irregular wear). Regular wear has been considered by Kalker (1987,1991), and by Pearce and Sherrat (1991). Irregular wear has been the focus of attention of Knothe (TU Berlin; see *e.g.* Knothe and Ripke (1989)) and of Frederick (1987).

In irregular wear, an important role is played by the response of the track to vibration. In this conference this will be treated by Grassie (1993).

6.3 CRACK INCEPTION AND PROPAGATION

Cracks constitute problems in self-contact, and can therefore be attacked by CONTACT-type algorithms. This has been done by Allan Bower; work on the same lines was done by Marie-Christine Dubourg. Dr. Dubourg will report on this theory at the present conference (1993). Bower and Dubourg consider cracks in a half-space. On the other hand, Prof. Stupnicki used an FE code to simulate cracks in much the same way as Bower and Dubourg. Needless to say that Stupnicki's program is

relatively slow, which results in his using the absolute minimum of elements at the crack, but he has the advantage over Bower and Dubourg that the geometry of wheel and rail can be very realistic. It will be interesting to compare the results of Dubourg and Stupnicki.

The contact mechanics at the outer surface need be given only sketchily. What might be of interest is a program to calculate the elastic field inside the half-space (CONTACT), but that may be severely deformed by the presence of cracks and other iregularities.

6.4 RESIDUAL STRESS

We considered residual stress in Sec. 5, in the context of shakedown. For the calculation, only a low accuracy of the surface loading is required. However, the elastic field inside the rail seems to be required with rather good accuracy.

Conclusion

In this paper we have tried to give an overview of wheel-rail rolling contact theory with applications.
We started from the formulation of the contact as a boundary value problem.
Then we gave some idea of the computer code that implements this as accurately as possible, together with the experimental results supporting it.
Finally we mentioned a number of wheel-rail problems in which contact mechanics plays a major role.

References

A.F. Bower, K.L. Johnson (1991): Plastic flow and shakedown of the rail surface in repeated wheel-rail contact, Wear **144**, p.1–18

A.F. Bower, K.L. Johnson (1993): Analysis of plastic flow under contact loading, In: Rail Quality and Maintenance for Future Railway Operation. Ed. J.J. Kalker, D.F. Cannon, O. Orringer, Kluwer Academic Publishes, Dordrecht/Boston/London.

A. Carneiro Esteves (1987): Résolution du contact élastique entre deux corps rugueux. INSA Lyon, Thèse 87, ISAL 0047.

F.C. Carter (1926): On the action of a locomotive driving wheel. Proc.Roy.Soc.London **A112**, p. 151–157.

C. Cattaneo (1938): Accademia dei Lincei, Roma, Fis. Mat.

M.-C. Dubourg, J.J. Kalker (1993): Crack behaviour under rolling contact fatigue. In: Rail Quality and Maintenance for Modern Railway Operation. Ed. J.J. Kalker, D.F. Cannon, O. Orringer, Kluwer Academic Publishers, Dordrecht/Boston/London.

J. Duvaut, J.-L. Lions (1972): Les Inéquations en Mécanique et en Physique. Dunod, Paris. Also in English translation.

G. Fichera (1964): Problemi elastostatici con vincoli unilaterale: il probleme di Signorini con ambigue condizioni al contorno. Mem.Acc.Naz.Lincei, Ser. 8,7 p.91–140.

C.O. Frederick (1987): Theory of rail corrugation. In: Contact Mechanics and Wear of Rail/Wheel Systems II, Ed. G.M.L. Gladwell, H. Ghonem, J. Kalousek. U. of Waterloo Press p. 181-212.

S.L. Grassie (1993): Dynamic modelling of the track. In: Rail Quality and Maintenance for Modern Railway Operation. Ed. J.J. Kalker, D.F. Cannon, O. Orringer, Kluwer Academic Publishers, Dordrecht/Boston/London.

M.T. Hanson, L.M. Keer (1993): Analytical treatment of flange contact. In: Rail Quality and Maintenance in Modern Railway Operation, Ed. J.J. Kalker, D.F. Cannon, O. Orringer, Kluwer Academic Publishers, Dordrecht/Boston/London.

H. Hertz (1882): Über die Berührung fester elastischer Körper. J.Reine, Angewandte Math. (Crelle) **92**, p.156–171.

J. Jäger (1992) Elastic Impact with Friction. Thesis Delft.

K.L. Johnson (1958): The effect of a tangential contact force upon the rolling motion of an elastic sphere on a plane. J.Appl.Mech. **25** p. 339-346.

K.L. Johnson (1962): A shakedown limit in rolling contact. In: 4th US National Congress of Applied Mechanics, Berkeley, ASME.

K.L. Johnson, J.A. Jefferies (1963): Plastic flow and residual stress in rolling and sliding contact. In: Proc.Inst.Mech.Engrs. Symposium on Rolling Contact Fatigue, London p.50.

K.L. Johnson (1985): Contact Mechanics. Cambridge UP.

J.J. Kalker (1967): On the rolling contact of two elastic bodies in the presence of dry friction. Thesis Delft.

J.J. Kalker (1968): The tangential force transmitted by two elastic bodies rolling over each other with pure creepage. Wear **11** p. 421–429.

J.J. Kalker (1983): Two algorithms for the contact problem in elastostatics. In: Contact Mechanics and Wear of the Rail/Wheel System. Ed. J. Kalousek, R.V. Dukkipatti, G.M.L. Gladwell. U. of Waterloo Press, p. 101-120.

J.J. Kalker (1990): Three-Dimensional Elastic Bodies in Rolling Contact. Kluwer Academic Publishers, Dordrecht/Boston/London.

K. Knothe, B. Ripke (1989): The effect of the parameters of wheelset, track and running conditions on the growth rate of rail corrugations. In: The Dynamics of Vehicles on Roads and Tracks, Ed. R. Anderson, Swets and Zeitlinger, Lisse, p. 345–356.

R.D. Mindlin: Compliance of elastic bodies in contact. J.Appl.Mech. **16** p. 259-268.

L. Nayak, K.L. Johnson (1978): Pressure between elastic bodies having a slender area of contact and arbitrary profiles. Int.J.Mech.Sci. p.237–247.

J.-P. Pascal (1992): About multicontact Hertzian hypothesis and equivalent conicity in the case of S 1002 and UIC 60 analytical wheel/rail profiles. Report INRETS–LTN, Arcueil, France.

B. Paul, J. Hashemi (1981): Contact pressures in closely conforming elastic bodies. J.Appl. Mech. **48**, p. 543–548.

T.G. Pearce, N.D. Sherratt (1991): Prediction of wheel profile wear. Wear **144** p. 343–351.

J. Piotrowski (1982): A theory of wheel set forces for two-point contact between wheel and rail. Veh.Sys.Dyn. **11** p. 591–605.

H. Reusner (1977): Druckflächenbelastung und Oberflächenverschiebung in Wälzkontakt von Rotationskörpern. Thesis Karlsruhe.

Shen Z.Y., J.K. Hedrick, J.A. Elkins (1984): A comparison of alternative creep–force models for rail vehicle analysis. In Dynamics of Vehicles. Ed. J.K. Hedrick. Swets and Zeitlinger, Lisse, p. 591–605.

P.J. Vermeulen, K.L. Johnson (1964): Contact of non-spherical bodies transmitting tangential forces. J.Appl.Mech. **31** p. 338–340.

E.A.H. Vollebregt (1992): A fast solver for normal and tangential contact mechanics in the elastic half-space. Ir. Thesis Delft.

DYNAMIC MODELS OF THE TRACK AND THEIR USES

Dr Stuart L Grassie
Department of Mechanical Engineering
Imperial College
LONDON
UK SW7 2BX

SUMMARY

Mathematical models developed for the track's response to dynamic excitation in the frequency range 5Hz-5kHz are reviewed, particularly with regard to their use in practice. Experimental data strongly suggest that for vertical excitation the most significant deficiency of a model of a beam on a uniform, continuous support is that it does not satisfactorily represent behaviour at the so-called "pinned-pinned" resonance. The severity of this deficiency has yet to be quantified satisfactorily. The errors incurred in neglecting loss of contact are also relatively poorly explored. A variety of measurements in track indicate that typical ballast stiffnesses are in the range 30-80MN/m per railseat; the stiffnesses of 5mm and 10mm elastomeric railpads are typically about 250MN/m and 100MN/m respectively. To model the lateral response satisfactorily it is necessary to respresent both bending and torsion of the rail at lower frequencies, and bending of the web for frequencies of about 1500Hz or more. Track longitudinally behaves as a strong viscous dashpot in parallel with an elastic spring.

1 INTRODUCTION

Mathematical models of the track's dynamic behaviour are particularly valuable to the railway engineer because they enable phenomena to be explored which cannot easily be measured, and effects of changes to the vehicle/track system to be examined without making costly and perhaps damaging modifications to track. The concern of this paper is with dynamic models of the track for excitation at frequencies from about 5Hz to 5kHz, with particular emphasis upon vertical excitation. These models are needed to explore physical phenomena such as corrugation and wheel/rail noise.

The mathematical models of track which have been developed for these purposes vary in complexity from analytical models, through those of Ahlbeck et al with several degrees of freedom e.g. [1], to the finite element models of Thompson [2,3] and Knothe

J.J. Kalker et al. (eds.), Rail Quality and Maintenance for Modern Railway Operation, 165–183.

et al [4,5]. Different models of track for vertical, lateral and longitudinal excitation are discussed in Sections 2 to 4 respectively of this paper, particularly in relation to their uses and limitations.

In the lower part of the frequency range of interest the rail moves significantly on the ballast and on the more or less resilient pad which may exist between rail and sleeper. It is accordingly necessary to model these components and to have values for the appropriate parameters. Relatively few values have been been published, and those found in [6] from a very limited number of tests in track are often used. Techniques for finding appropriate parameters from experiments in track are discussed in Section 5 and a range of values for vertical excitation is presented.

2 MODELS FOR VERTICAL DYNAMICS

2.1 Infinite beam on a uniform, continuous support

The most familiar model of track, and that used for static track design, is that of an infinite beam, representing the rail, on an elastic foundation representing the ballast. This model is also most useful dynamically, and it is accordingly appropriate to quote some standard results from refs [6-8]. The static deflection at a distance x from a load P applied to an infinite rail of flexural rigidity EI on a continuous, uniform support of stiffness β per unit length of track (the "track modulus") is

$$Y(x) = \frac{P}{\left(64\beta^3 EI\right)^{0.25}} \left(\sin|kx| + \cos|kx| \right) e^{-|kx|} \qquad (1)$$

where $k = \left(\frac{\beta}{4EI} \right)^{0.25}$

The bending moment and shear force in the rail are found by differentiation.

The deflection $Y(x)e^{j\omega t}$ of an Euler beam of mass m per unit length on a support of dynamic stiffness $\beta^*(\omega)$ under a dynamic point force $Pe^{j\omega t}$ is calculated from eqn (1) simply by substituting

$$\beta = \beta^* - m\omega^2 \qquad (2)$$

For example, many characteristics of typical track with sleepers are evident from a support comprising 3 layers: the railpads, the sleepers and the ballast (Fig 1). If these have stiffness β_p, mass m and stiffness β_b respectively per unit length, then

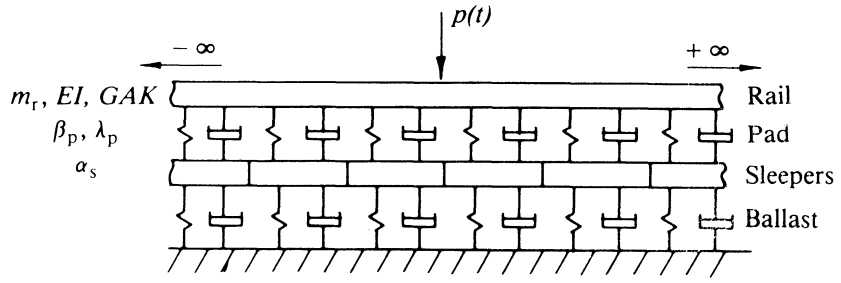


Fig 1 Model of track on a continuous, three layer support

$$\beta^* = \frac{\beta_p \beta_b}{(\beta_p + \beta_b)} \frac{(1 - \frac{m\omega^2}{\beta_b})}{(1 - \frac{m\omega^2}{(\beta_p + \beta_b)})} \tag{3}$$

The receptance Y(x)/P (which is, in general, complex) is calculated from eqns (1)-(3). In practice a dynamic contact force is caused when the wheel rolls over irregularities on the rail or wheel. If the irregularity is assumed to be sinusoidal and of amplitude Δ, the dynamic contact force P is given by [6]

$$\frac{P}{k_H \Delta} = \frac{-1}{(1 + k_H(\alpha_r + \alpha_w))} \tag{4}$$

where k_H is the linearised contact stiffness, α_r is the direct receptance of the track, found from eqns (1)-(3), and

$$\alpha_w = \frac{-1}{M_w \omega^2} \tag{5}$$

is the receptance of the wheel, which is modelled as a rigid body of mass M_w. This is typically 350kg for a passenger coach and 1000-2000kg for a locomotive.

Equations (1)-(5) can be programmed quickly on a microcomputer using a good scientific software package, and give a satisfactory estimate of the dynamic contact force for frequencies up to about 500Hz, provided that wheel/rail contact is not lost. Dynamic loads caused by discrete irregularities with relatively low frequency components can be found using a Fourier transform method in which eqn (4), or the equivalent for the response of interest, is taken as the transfer function relating the Fourier transform of the response to the Fourier transform of the irregularity. This technique has been validated tentatively by fair correlation of calculation with measured contact forces and sleeper moments for different railpads and vehicle speeds [9].

British Rail have have used the model of an Euler beam on an elastic foundation to calculate loads arising from a wheelflat and from dipped welds [10,11]. For dipped

welds, which contain essentially long wavelength components, good correlation of calculated and measured contact forces and stresses was obtained [11]. Newton and Clark [10] obtained satisfactory correlation of calculated and measured data for wheelflats for speeds up to about 80km/h, above which their model is deficient primarily because they neglect the railpads which decouple the rail from the sleepers, thus reducing the high frequency contact force. They show that shear deformation and rotatory inertia significantly affect the rail at high frequencies, and their neglect overestimates the dynamic contact force. When the railpads, shear and rotatory inertia are neglected, high frequency components of the contact force may misleadingly be associated with a "contact resonance" of the rail and sleepers moving approximately in antiphase to the wheel on the contact stiffness, whereas such components are caused essentially by sleeper resonances. Analysis equivalent to that of eqns(1)-(3) for a Timoshenko beam representing the rail on a 3 layer damped elastic support is given in ref [6].

A limitation of the uniform support model is that it does not show effects of the so-called "pinned-pinned" resonance which occurs when the wavelength of bending waves is twice the sleeper spacing. The frequency at which this occurs is approximately

$$f = \frac{1}{2\pi}\left(\frac{\pi}{L}\right)^2 \sqrt{\frac{EI}{m}} \qquad (6)$$

where m is the rail mass per unit length, EI is its flexural rigidity and L the sleeper spacing. The frequency calculated from eqn(6) for 56kg/m rail and L = 0.7m is 950Hz, whereas in fact the effects of shear and rotatory inertia reduce the pinned-pinned resonant frequency to about 750Hz in this case. Models which can be used to calculate the track response when the pinned-pinned resonance is significant are considered in the following Section.

The complexity and nature of the support model clearly influence its utility and validity. Representation of sleepers as a uniform, continuous layer of elemental beams spanning the rails is a useful means of studying damage which occurs from excitation of flexural modes of vibration [12,13]. The non-uniformity of typical sleepers appears to have a relatively small effect on the dynamic moment. Railpads and ballast are commonly modelled as damped, elastic components. Values for the stiffness and damping, and the validity of different damping models, are discussed in Section 3.

Although the model of track with a 3 layer support presented in [6] appears to be the standard reference in this area, it transpires that a substantially similar model with the rail as an Euler rather than a Timoshenko beam was developed at a slightly earlier date [14]. Recently the same group has developed a track model with a 5 layer support to calculate forces on the ballast [15]. The support comprises 3 layers of distributed viscoelasticity representing the railpads, ballast or resilient sleeper support, and the subgrade, with two intervening layers of distributed mass representing the sleepers and the ballast.

The model of track as a beam on a continuous support has been particularly useful for calculation of the 'P2' force arising from resonance at about 25-50Hz of a vehicle's unsprung mass on the track stiffness and examination of associated problems. For example, it was demonstrated [11] that a reduction in unsprung mass of locomotives would be necessary to prevent cracking at boltholes on track with bolted joints, as a result of which British Rail's early Class 86 locomotives were fitted with resilient wheelsets. Excitation of this mode also either gives rise to or exacerbates at least three types of corrugation [16-19]. The model and Fourier transform technique for calculating the response to discrete irregularities has also been used to help design track components to withstand both the quasi-static loads associated with vehicle dynamics and the higher frequency loads arising from rail and wheel irregularities [9].

The rail was modelled as a uniform beam on a continuous support in early work on wheel/rail noise [20], and should be adequate for calculation of vertical dynamic loads provided the rail is modelled as a Timoshenko beam. A more sophisticated rail model is now used to model distortion of the rail cross section under lateral loads in the acoustic frequency range [2,3].

2.2 Finite element models

The finite element (FE) models of track which have been developed in the last decade have provided a powerful means of examining cross-sectional deformation of the rail and of including the discrete nature of the track support. Deformation of the rail, and particularly of the web, occurs primarily under lateral excitation and is discussed in Section 3. However, it is clear from measurements of bending strains in track that the rail does not always deform vertically as a single beam: a clear high frequency "spike" is measured, which cannot be calculated using the simple beam models described in Section 2.1. This appears instead to arise from the railhead moving as a beam on the elastic foundation of the web.

Thompson [2,3] has developed an FE model of the rail to study wheel/rail noise. In his model the rail is divided into elements 10mm long in which the railhead is a beam and the web and foot (where most of the deformation occurs), comprise elemental plates. The rail support is assumed to be uniform and continuous, and the model track is of infinite length. The model is validated most directly by the satisfactory correlation obtained between calculated receptances and several published sets of measurements, and less directly by comparing measured noise levels with those calculated from a model in which the track is one component. The possibility of obtaining good correlation with receptance data is limited by there being up to 10dB difference between some sets of published data. It is clear from Thompson's calculated receptances [3] that a single Timoshenko beam is a satisfactory model for vertical vibration of the rail up to at least 4kHz. Resonance of the railhead relative to the foot, which gives rise to the aforementioned "spike" on measurements of bending strain, occurs at about 5kHz for Thompson's 56kg/m rail ; whereas Bender and Remington found a corresponding frequency of 3-4kHz in experiments on lighter AREA 100 rail [20].

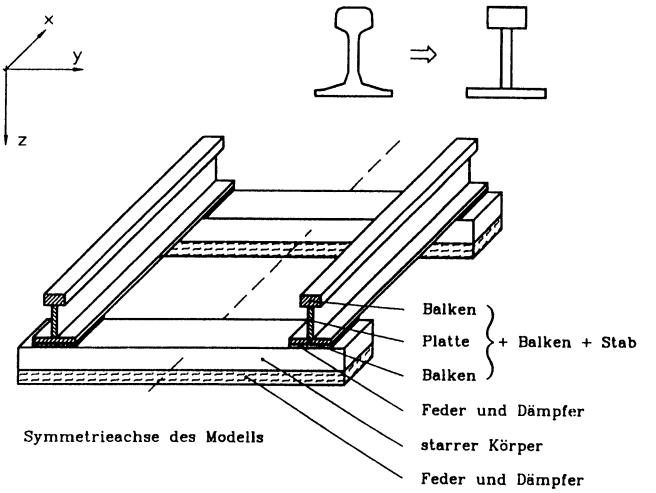


Fig 2 Finite element model of track on discrete sleepers (from [4])

Several rail and track models have been developed by Knothe et al e.g.[4,5,21], in order to study the development of rail corrugation. They conclude that it is sufficient to model the rail as two Timoshenko beams representing the head and foot separated by an elastic element representing the web [21] for both lateral and vertical excitation up to about 2.5kHz. The frequency range in which the simpler single beam model of Section 2.1 is valid for vertical excitation alone is not examined. The more complicated rail models which are proposed for higher frequencies are tested by comparison of the calculated dispersion relationships with the experimental data of Scholl [22]. The pinned-pinned resonance which occurs because of the discrete support of track at sleepers (Section 2.1) significantly affects the track's dynamic behaviour at a frequency of about 1kHz, which is in the region excited by typical corrugation on high speed main line track. The discrete support has accordingly been included in these track models, with the sleeper either as a rigid body or as a Timoshenko beam spanning the rails [4] (Fig 2). Railpads and ballast are included as viscously damped elastic elements. The infinite nature of the track is included using a technique of "harmonic continuation" [5,21]. The dynamic response of this model track, whose receptance varies between sleepers, is found by the Fourier Transform technique described in [5]. The calculated receptance correlates well with that measured and has been used to find support parameters [38], as described in Section 5.

At the opposite extreme of frequency from that of noise, one of the most prominent characteristics in ground-borne vibration is the "sleeper-passing" frequency [23]. It has been shown by Auersch e.g. [24,25], that a track model with discrete sleepers is required to predict these effects.

2.3 Modal analysis

In their work on wheelflats [10] and corrugation [26] British Rail developed a track model which includes a finite length of track (e.g. 20 "sleeper spans" in ref [10]) with

discrete supports at sleepers, from which the response to a moving load is calculated by modal analysis and forward integration. The importance is demonstrated of the railpad (or the equivalent local resilience in timber sleepered track) in attenuating the contact force and of shear deformation and rotatory inertia in calculating behaviour of the rail [10], but the latter cannot easily be included in their model.

BR's model includes loss of contact, in contrast to the linear models described in Sections 2.1 and 2.2; in later work the correct contact position is also calculated [27]. Calculations suggest that the contact force following a period of "free flight" may be underestimated by as much as 25% if it is assumed that the contact point is always under the wheel centre. However, it is not yet clear to what extent the contact force is under- or overestimated by a linear model in which contact is not lost.

It is difficult in such a finite model of track satisfactorily to represent high frequency behaviour, essentially because travelling waves rather than standing waves exist in the rail. Also in BR's calculations it is assumed, for convenience, that damping is constant in each of the many modes. This may overestimate the influence of some modes while underestimating that of others, particularly the relatively ill-damped pinned-pinned mode [6].

2.4 Other models

The work on dynamic behaviour of track undertaken by Ahlbeck, Harrison and their colleagues is distinguished by a concern to derive practical limits and correlations with minimum effort. The rail is accordingly represented as an effective mass and the track overall as a system with relatively few degrees of freedom e.g. 11 and 18 in refs [1] and [28] respectively. By comparison, 126 modes of BR's 20 span track model [10] were found in the frequency range 80-3000Hz. In a study of the dynamic loads caused by coal trains, good correlation was obtained between calculated and measured peak wheel/rail forces for two wheel profiles and for speeds of 8-88km/h [28]. Calculations indicated that the dynamic contact force caused by short wavelength irregularities would be reduced using more resilient railpads, whereas the converse would be true for long wavelength defects.

Such models have been useful in examining essentially the track's lower frequency behaviour, but it is questionable whether they can be used reliably to predict high frequency behaviour because the asymptotic behaviour of the rail is not that of a rigid body . Their use as a general predictive tool is also limited by the apparent absence of a routine way of calculating effective modal parameters.

Restriction of the effect of the discrete support at sleepers to a narrow frequency range around the "pinned-pinned" resonance was revealed by a model of the track as an infinite, discretely supported Euler beam [6]. The transfer matrix method used there to model the infinite rail is substantially similar to the technique used by Knothe et al [5,21].

3 LATERAL DYNAMICS

Track behaves differently to lateral excitation at the railhead than to vertical excitation in two particular respects: torsion of the rail, as well as bending, is significant because the force does not pass through the shear centre, and deformation of the cross section itself is significant at higher frequencies because the web is like a relatively slender beam with the railhead and foot at either end. Although these effects were examined somewhat qualitatively in [29], a more thorough treatment of their significance has awaited development of finite element rail models such as those of Thompson [3] and Ripke and Knothe [4] in which deformation of the rail's cross section is explicitly examined (Section 2.2). With the insight which these finite element models provide it is accordingly now possible to propose how the track may most appropriately be modelled in different frequency ranges, and perhaps to reconcile apparently conflicting observations.

At relatively high frequencies the rail cross section is distorted significantly: a mode in which the railhead and foot move in antiphase on the web's flexural stiffness occurs at about 1800Hz [3,4]. However, the lack of significant distortion at lower frequencies suggests that a substantially simpler model may be used to represent the rail. Thomson [3] has obtained satisfactory correlation of his calculations with a few sets of experimental data up to their limit of about 1.5kHz, and states that three significant features are present in the track's response: resonance of the rail and sleepers laterally on the ballast at about 80Hz, lateral vibration of the rail on the railpad at about 150Hz and a well damped torsional resonance on the railpad at about 350Hz. Representation of the pinned-pinned mode at about 350Hz requires a model with discrete sleepers e.g. [4]. This model has been used by Fingberg [30] to study wheel squeal.

The satisfactory correlation obtained by Thompson suggests that an appropriate track model to about 1.5kHz should include the rail's response as an infinite beam in bending and torsion, on appropriate continuous, distributed supports. The frequency of torsional resonance on the support would be determined primarily by the resistance of the railpad to roll of the rail. By varying this roll stiffness it might be possible to obtain satisfactory correlation of calculation with the data of [29] for track on both timber and concrete sleepers. Such a model was developed in [31] but may have been abandoned prematurely. "Beam" like behaviour, such as that associated with track on timber sleepers in [29], could be retained to a relatively high frequency by having a railpad with high roll stiffness, whereas the higher lateral receptance of track on concrete sleepers may be obtained from a pad with low roll stiffness.

Regardless of the model's detail, the rail is not only much more "lively" laterally than it is vertically, but it is also in general more lively than the wheel. Measurements made of the lateral receptance of a wheel and of a rail individually and together [32] show that the rail behaves like a weak spring and damper attached to the wheel.

4 LONGITUDINAL DYNAMICS

Interest in the longitudinal dynamic response of track arises primarily because longitudinal vibration is excited by variations in longitudinal traction. These are caused, for example, by stick-slip oscillations of the type which give rise to wheel squeal and may be associated with development of corrugations.

An extremely simple model of the track which gives acceptable correlation of calculated and measured direct receptances is of a spring and dashpot in parallel. The spring is of stiffness

$$k = \sqrt{4\beta AE} \tag{7}$$

where β (which is typically 25MPa) is the support stiffness per unit length, A is the cross-sectional area of the rail and E is its Young's modulus. The dashpot has strength

$$c = 2m_r \sqrt{\frac{E}{\rho}} \tag{8}$$

where $m_r = \rho A$ is the mass per unit length of the rail [33,34].

More complicated track models which include the sleepers, give slightly better correlation with experimental data and can be used to find the response away from the driving point are described in [33].

5 PARAMETERS FOR MATHEMATICAL MODELS

Parameters are required for the various track components in order to use the models described in Sections 2-4. Only vertical excitation is considered here; some parameters for lateral and longitudinal excitation can be found in refs [29] and [33]. Parameters for the rails are well defined: the mass per unit length and the flexural rigidity for the Euler beam model, plus the additional shear constant κ for a Timoshenko beam model: a value of $\kappa = 0.34$ is appropriate for BS113A (56kg/m) rail [10]. For the FE models of Section 2.2 the rail is divided into well defined beam or plate elements.

The geometry, spacing and mass of sleepers are also well defined, but an effective flexural rigidity is required if they are represented as the uniform beams of refs [12,13]. The satisfactory correlation of measured track response with that calculated for the first 3 modes indicates that the effective flexural rigidity for a typical non-uniform sleeper is approximately that at the centre. A shear constant of 0.83, corresponding to a rectangular section, is appropriate.

The parameters which are relatively ill-defined are those for the railpads and the ballast. Techniques for finding appropriate values of these are discussed in Section 5.1 and several values are presented in Section 5.2.

5.1 Ballast and railpad parameters

The majority of techniques for identifying parameters involve fitting transfer functions calculated from reliable mathematical models to those measured. Clearly the railpad and ballast parameters which are of greatest interest are those pertinent to the track's response under traffic, and it is accordingly desirable to obtain appropriate values from experiments involving a vehicle of the type of interest with track of the type of interest. In such tests the track's response is measured to an irregularity on wheel or rail: the latter is preferable as its position along the track is known. Experiments of this type are expensive and alternative methods which are more routine and do not require a vehicle are desirable. Examples of both types of tests are described below. The relevance of parameters found from the latter tests to calculation of the track's response under traffic is discussed in Section 5.2.

Excellent examples of controlled experiments in track under traffic are those undertaken by British Rail to study effects of dipped welds [11], wheelflats [10] and corrugation [26]. In the latter two cases an irregularity of known profile was ground into the rail and measurements were then taken on the track as a test train ran over the instrumented site. Instrumentation has included transducers to measure shear strain on the rail web and deflection of the sleeper. Parameters for the ballast and railpad were found by fitting a calculated response e.g. the amplitude of sleeper displacement as a function of train speed, to that measured (Fig 3). The ballast stiffness influences behaviour primarily at frequencies below 100Hz, whereas the railpad is most significant around the first sleeper resonance (at about 200Hz in Fig 3). Ballast is the principal source of damping. Data from BR's corrugation experiment were also used in [12] to calculate support parameters.

Field tests have been undertaken in which railpad and ballast parameters were found from the measured vibration of rail and sleeper caused by wheel irregularities on a test train [9,35]. The stiffening of resilient railpads with increasing static wheel load was clearly demonstrated: for the 10mm railpads tested, railseat bending moments were consistently attenuated (compared to the moments with non-resilient HDPE or EVA railpads) by about 75% under a 30kN wheel load but by only 50% under a 110kN wheel load [35].

Values of ballast and railpad stiffness and damping found from several experiments with test trains are shown in Table 1. These values are expressed for a single railseat as it seems reasonable to expect, for example, that the effective ballast stiffness would be approximately doubled if there were twice as many sleepers. Values reported for unit length of track have been converted by assuming a sleeper spacing of 0.7m.

A technique which has gained wide acceptance is to find the support parameters by fitting a calculated value of the track receptance to that measured on unloaded track. This method was used in refs [6], [13] and [18] to find support parameters, and is also now used by British Rail [36]. Measurements can be made particularly quickly using

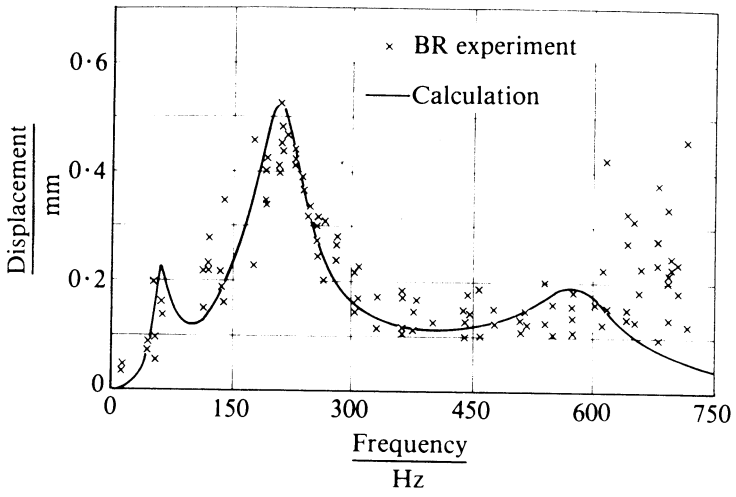


Fig 3 Sleeper end displacement, test train (from [12])

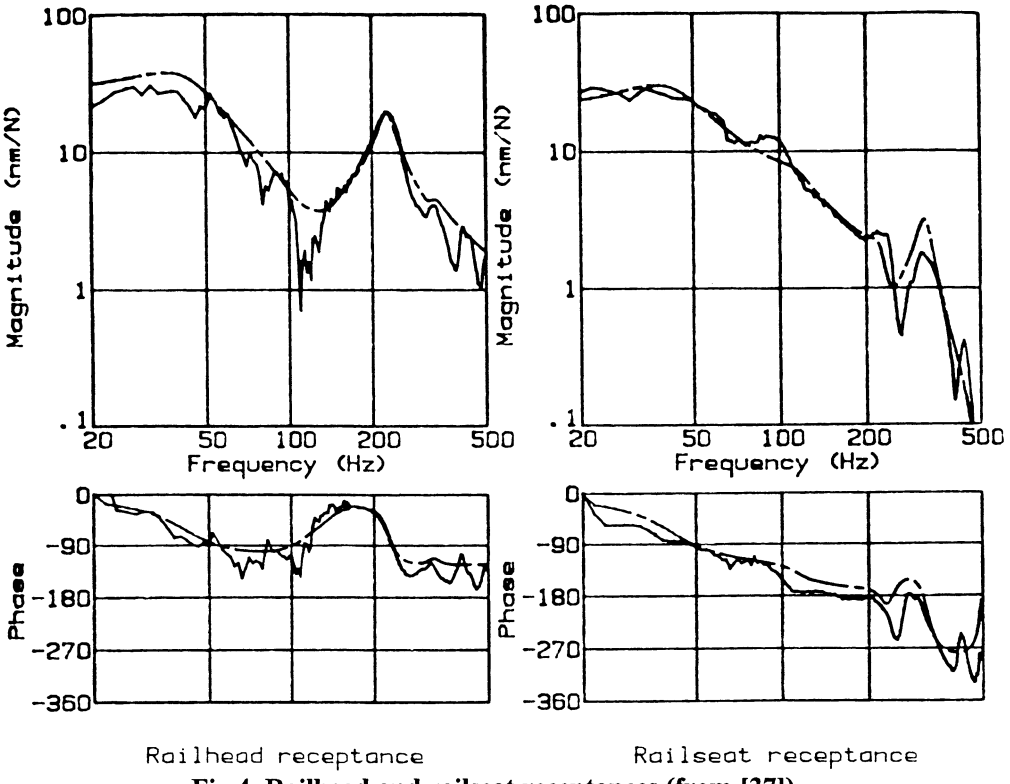


Fig 4 Railhead and railseat receptances (from [37])

calculate the transfer function. It is prudent to obtain some corroboration of the fit of calculation and measurement, such as by fitting two sets of data. An example is shown in Fig 4 where correlation is established for both the direct receptance at the railhead and the cross receptance at the railseat [37]. Values of ballast and railpad parameters found from the receptance on unloaded track are shown in Table 2. An essentially identical technique has been suggested by Kopke [43] for finding areas where gas pipelines are poorly supported.

ballast		railpads			
stiffness (MN/m)	damping (kNs/m)	stiffness (MN/m)	damping (kNs/m)	ref	notes
50	51	250	26	12	2
31.6	21.8	200		10	5
46.6		250		26	2
64.6	40.8			11	
35	35	70	7	9,35	3
35	35	350	70	9,35	4

Notes

1 All data are for ballasted track with concrete sleepers unless stated otherwise, and are for a single sleeper end
2 5mm thick railpads
3 10mm thick railpads
4 HDPE railpads
5 timber sleepers

Table 1 Ballast and railpad parameters: measurements with test train

Track parameters can also be found from the response of a vehicle passing over a length of rail of known profile. If $S_x(\omega)$ and $S_y(\omega)$ are the power spectra of the railhead profile and the vehicle response respectively, the transfer function is simply

$$H(\omega) = \sqrt{\left|\frac{S_y(\omega)}{S_x(\omega)}\right|} \qquad (9)$$

The measurements of this type shown in Fig 5 were made up to a frequency of about 60Hz during calibration of a railhead profile system using axlebox accelerometers [17]. The profiles of several 200m lengths of rail were measured using a self-propelled profilometer which runs along the rail at walking speed: this was developed at Cambridge University in the 1970s. The corresponding axlebox accelerations were recorded on the measuring coach at speeds of about 80km/h for several runs over the same site. The transfer functions shown in Fig 5 are for several runs over the same site.

The illustrative calculations are for an Euler beam on a uniform, continuous support (Section 2.1) with effective stiffness and viscous damping constants of 70MN/m and 70kNs/m respectively per metre length of track. Fair correlation is also found in ref [6] of measured data with the calculated transfer function of contact force per unit railhead irregularity for a few frequencies in the range 300-800Hz. Although this is a powerful technique in principle, a practical difficulty with such measurements is that the wheel may not run along the line measured by the profilometer.

ballast		railpad			
stiffness (MN/m)	damping (kNs/m)	stiffness (MN/m)	damping (kNs/m)	ref	notes
180	82	280	63	6	2,7
27.8	16.6			6	4
70	30	225	28	6	2,6
72	132	280	50	13	2,5
100	72	300	45	38	2,8
		20-100	2-12	39	3
30.7	(0.2)	221	(0.2)	18	2,9,10
41.6	(0.3)			18	10
12	30	76	5	37	3
55	100	76	5	37	3
65	150	76	5	37	3

Notes

1 All data are for ballasted track with concrete sleepers unless stated otherwise, and are for a single sleeper end
2 5mm thick railpads
3 10mm thick railpads
4 timber sleepers
5 frozen ballast
6 post-tamping
7 pre-tamping
8 light rail
9 sleepers with resilient "boots"
10 damping is viscoelastic with the specified loss factor (see eqn (11))

Table 2 Ballast and railpad parameters: measurements on unloaded track

Lyon [36] reports results of low frequency tests undertaken for ORE Committee D117 to obtain track stiffness and damping. A known load was applied to a section of track carrying an equivalent unsprung mass and a static load . Although details of the test method are unclear, equivalent ballast stiffnesses per sleeper end can be calculated by

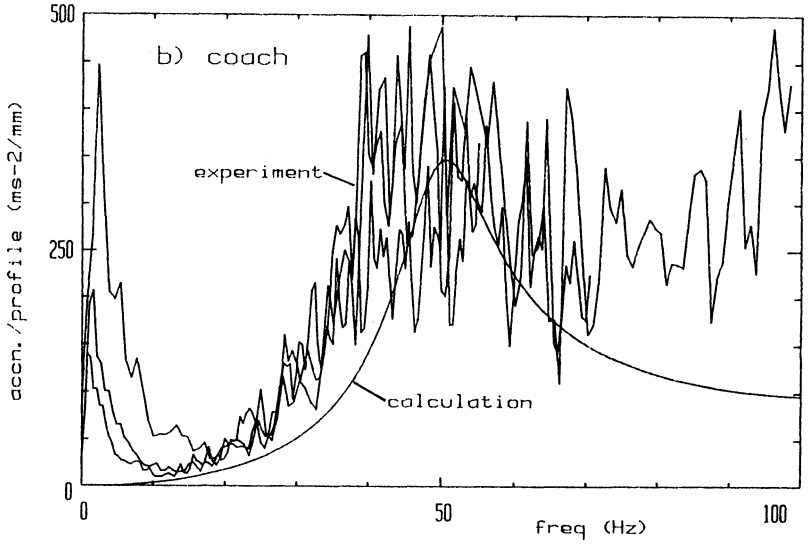


Fig 5 Transfer function of axlebox acceleration to railhead profile (from [17])

assuming that the deflections quoted are for track with a sleeper spacing of 0.7m. Values thus calculated are a secant stiffness of 78.4MN/m for a static load of 100kN and a quasi-static stiffness of 172MN/m for a load of +/- 50kN about a mean load of 150kN, which are high compared with values in Tables 1 and 2.

Stiffness and damping values have been found for railpads in laboratory experiments. Kanamori [40] reports measurements of the static railpad stiffness in the range 40-60MN/m and a dynamic stiffness at 5Hz of 1.2-1.5 times the static stiffness. For one type of pad the dynamic stiffness was 80MN/m at 5Hz, 150MN/m at 1kHz and 200MN/m at 2.5kHz. The loss factor ranged from 0.2 at 5Hz to 0.6 at 1kHz. Stiffnesses of several resilient railpads measured in a laboratory resonance test correlated well with, but were consistently greater than, those found from receptance measurements in track [39]. In this case consistently good correlation was obtained between the static stiffnesses of lightly damped resilient railpads and the stiffnesses found in the resonance test. Poor correlation was found of corresponding values of the (relatively low) railpad damping. Such techniques provide a convenient means of comparing different railpads.

Sato et al [15] report measurements of ballast loss coefficients of 5 and 15 for concrete and timber sleepered track respectively, which are substantially greater than damping values found in other tests reported here.

5.2 Discussion

The majority of effective ballast stiffnesses reported here for vertical excitation are in the range 30-80MN/m per railseat. There are insufficient data to decide whether there is any consistent difference between values found with a test train and those found from receptance measurements on unloaded track. Lower values may be more typical of

recently tamped track and higher values of well consolidated track. These stiffnesses are about 50% greater than static values commonly used in track design. For example, Raymond [41] suggests that an optimum track modulus is in the range 35-70MPa (i.e. 25-50MN/m per railseat for a typical sleeper spacing of 0.7m), while Stewart [42] has measured a variation in track modulus from about 21MPa with 250mm depth of ballast to 48-55MPa with 750mm depth of ballast. Stewart's findings were that track modulus consistently increased with depth of ballast and that there was little difference in modulus before and after tamping, whereas ballast stiffnesses measured by Sato et al [15] decreased with depth of ballast.

Data are shown in Tables 1 and 2 for two thicknesses of more or less resilient railpads. The 5mm thick railpads, some of which are rubber and some a composite material, consistently have a stiffness of 200-300MN/m, whereas the stiffness of the 10mm thick pads is consistently about 70MN/m under the fastening load or a relatively light wheel load. A more representative stiffness under a static wheel load of about 100kN would be 125MN/m. HDPE railpads would typically be at least as stiff as the 350MN/m shown in Table 1.

Damping in the railpad and ballast has commonly been represented as viscous i.e. the dynamic stiffness at frequency ω is

$$\beta^* = \beta(1 + j\frac{\omega\lambda}{\beta}) \tag{10}$$

where λ is the effective viscous damping constant. However, viscoelastic damping is a better physical representation for typical resilient railpads i.e.

$$\beta^* = \beta(1 + j\eta) \tag{11}$$

where η is the loss factor. Kanamori's range of 0.2-0.6 for the loss factor [40] should cover the majority of railpads in current use. However, there is evidence that a more lightly damped railpad is less likely to degrade from dissipation of thermal energy [39]. The mechanism of damping in ballast is not yet well understood, and both viscous and viscoelastic models are used. The values of viscous damping in Tables 1 and 2 are equivalent to loss factors of 0.3-0.6 for the frequency range 50-100Hz. Although higher values would be obtained from correlation of data at higher frequencies, it would be surprising if values as large as those reported in [15] were obtained as the mode in which the unsprung mass moves on the ballast would then be attenuated more than is suggested by most experimental data.

6 CONCLUSIONS

Many mathematical models of railway track of considerable complexity have appeared in the literature, but for a very large number of applications the track can be modelled satisfactorily as a beam on a uniform, continuous support. For vertical excitation at frequencies up to about 400Hz the rail is represented satisfactorily as an Euler beam. This gives rise to a particularly simple analytical expression for the track receptance. For frequencies of about 400-4000Hz the rail can be represented as a single Timoshenko beam, while a more sophisticated model is required for higher frequencies.

If detail of vibration of the sleepers is not of interest, these can be modelled as rigid bodies, but should otherwise be modelled as beams spanning the track. Railpads and ballast are conveniently represented as viscoelastic layers. For railpads the loss factor varies from 0.2 or less for lightly damped materials to about 0.6. The majority of measurements of ballast damping correspond to loss factors of 0.3-0.6 in the frequency range 50-100Hz.

The only significant deficiency of a model of track with a continuous support is that it fails to represent the so-called pinned-pinned resonance which occurs when the length of bending waves is twice the sleeper spacing. If the effects of this are significant the discrete support at sleepers must be modelled explicitly. The method of harmonic continuation used by Knothe et al e.g. [4,5] is a means of doing this while maintaining the physical representation of the rail as an infinite beam which is lost in other methods.

When a rail is excited laterally its behaviour is more complicated than when excited vertically. At frequencies above about 1500Hz a rail model is required which includes distortion of the cross section. At lower frequencies a model which allows both bending and torsion of the rail should be adequate but this has not been demonstrated conclusively. When the track is excited longitudinally it behaves at the driving point essentially as a spring and dashpot in parallel.

There are several interesting areas for further work and questions to be answered. For example, existing models could be used to quantify the extent to which the contact force may be under- or over-estimated by assuming that the support is continuous rather than discrete. Loss of contact is neglected in most existing, linear models, but it is not known to what extent calculations are in error which assume that contact is not lost and that tension accordingly exists in the "contact spring". Railpad and ballast characteristics are at present commonly found from receptance measurements on unloaded track, yet these appear to underestimate both stiffness and damping. It would be desirable to know in what circumstances these parameters give an adequate estimate of dynamic loading and whether they can be modified in a consistent manner to calculate loading more accurately.

A broad area for research is in the prediction and reduction of degradation of the track support. For example, damping in the support is desirable insofar as it reduces dynamic

loads, but may be obtained at the cost of damage to the support, as indeed appears to be the case with highly damped railpads. Development of appropriate and well validated models may indicate to what extent damping can satisfactorily be obtained in practice.

7 ACKNOWLEDGEMENTS

The author is grateful to Prof Knothe of TU Berlin, Prof Sato of NKH, Japan, and his colleagues at Pandrol International Ltd, particularly for giving him access to reference material which he would not otherwise have seen.

8 REFERENCES

1 Ahlbeck D.R. and Hadden J.A., "Measurement and prediction of impact loads from worn railroad wheel and rail surface profiles", Jnl. of Engng. for Industry, 1985, 107, 197-205

2 Thompson D.J., Workshop on Rolling Noise Generation, T.U. Berlin 1989

3 Thompson D.J., "Theoretical modelling of wheel-rail noise generation", Proc. of Instn. of Mech. Engrs., 1992, 205F, 137-149

4 Ripke B. and Knothe K., "Die unendlich lange Schiene auf diskreten Schwellen bei harmonischer Einzellasterregung", VDI Fortschritt-Bericht Nr 155, Dusseldorf, 1991

5 Hempelmann K., Hiss F., Knothe K. and Ripke B., "The formation of wear patterns on the rail tread", Wear, 1991, 144, 179-198

6 Grassie S.L., Gregory R.W., Harrison D. and Johnson K.L., "The dynamic response of railway track to high frequency vertical excitation", Jnl. of Mech. Engng. Sci., 1982, 24, 77-90

7 Hetenyi M., "Beams on elastic foundations", University of Michigan Press, Ann Arbor, 1946

8 Fryba L., "Vibration of solids and structures under moving loads", Noordhof Publishing Co., Groningen, 1972

9 Grassie S.L., "A contribution to dynamic design of railway track", Vehicle System Dynamics, 1991, 20, 195-209

10 Newton S.G. and Clark R.A., "An investigation into the dynamic effects on the track of wheelflats on railway vehicles", Jnl. of Mech. Engng. Sci., 1979, 21, 287-297

11 Jenkins H.H. et al, "The effects of track and vehicle parameters on wheel/rail vertical dynamic forces", Railway Engng. Jnl., 1974, 3, 2-16

12 Grassie S.L. and Cox S.J., "The dynamic response of railway track with flexible sleepers to high frequency vertical excitation", Proc. of Instn. of Mech. Engrs., 1984. 198D, 117-124

13 Grassie S.L. and Cox S.J., "The dynamic response of railway track with unsupported sleepers", Proc. of Instn. of Mech. Engrs., 1985. 199D, 123-135

14 Sato Y., "Study on high frequency vibrations in track operated with high speed trains", JNR Quarterly Reports, 1977, 18,109-114

15 Sato Y., Odaka T. and Takai H., "Theoretical analyses on vibration of ballasted track", RTRI Research Report, 1987, no 1347 (in Japanese)

16 Mair R.I., "Natural frequency of rail track and its relationship to rail corrugation", Civil Engng Trans., The Instn. of Engineers Australia, 1977, paper 3494

17 Grassie S.L., "Corrugation on Australian National: cause, measurement and rectification", 4th Intnl. Heavy Haul Railway Conf., Brisbane, 1989

18 Tassilly E. and Vincent N., "A linear model for the corrugation of rails", Jnl. of Sound and Vibn., 1991,150,25-45

19 Tassilly E. and Vincent N., "Rail corrugations: analytical model and field tests", Wear, 1991, 140, 163-178

20 Bender E.K. and Remington P.J., "The influence of rails on train noise", Jnl. of Sound and Vibration, 1974, 37, 321-334

21 Knothe K., Willner K. and Strzyzabowski Z., "Rail vibrations in the high frequency range", submitted for publication in Jnl. of Sound and Vibration, 1992

22 Scholl W., "Two theoretical models for wave propagation in rails", 2nd Intnl. Conf. on Recent Advances in Structural Dynamics, Southampton, 1984

23 Dawn T., "Ground vibrations from heavy freight trains", Jnl. of Sound and Vibration, 1983, 87, 351-356

24 Auersch L., "Zur parametererregung des rad-schiene-systems: berechnung der fahrzeug-fahrweg-untergrund-dynamik und experimentelle verifikation am hockgeschwindegkeitszug Intercity Experimental", Ingenieur Archiv, 1990, 60, 141-156

25 Auersch L, "Structural dynamics and soil-structure interaction induced by railway traffic", Dynamics of Structures, 1989, Karlovy Vay, Czechoslovakia

26 Clark R.A., Dean P.A., Elkins J.A. and Newton S.G., "An investigation into the dynamic effects of railway vehicles running on corrugated rails", Jnl. of Mech. Engng. Sci., 1982, 24, 65-76

27 Tunna J., "Wheel/rail forces due to wheel irregularities", 9th Intnl. Wheelset Congress, Montreal, 1988

28 Harrison H.D., Bethune A.E. and Ahlbeck D.R., "Comparison of measured and predicted loads on 100 ton coal gondolas", 4th Intnl. Heavy Haul Railway Conf., Brisbane, 1989

29 Grassie S.L., Gregory R.W., and Johnson K.L., "The dynamic response of railway track to high frequency lateral excitation", Jnl. of Mech. Engng. Sci., 1982, 24, 91-95

30 Fingberg U., "A model of wheel/rail squealing noise", Jnl of Sound and Vibn., 1990, 143, 365-377

31 Grassie S.L., "The corrugation of railway track", PhD thesis, University of Cambridge, 1980

32 Grassie S.L., "Vibration on Docklands Light Railway", ref GEC-88-1.R3, 1988

33 Grassie S.L., Gregory R.W. and Johnson K.L., "The dynamic response of railway track to high frequency longitudinal excitation", Jnl. of Mech. Engng. Sci., 1982, 24, 97-102

34 Grassie S.L., "Dynamic modelling of railway track and wheelsets", 2nd Intnl. Conf. on Recent Advances in Structural Dynamics, Southampton, 1984

35 Grassie S.L., "Allowable impact loading and wheel irregularities", 8th Intnl. Rail Track Conf., Sydney, 1990

36 Lyon D., "Dynamic measurements in the research and development of rail vehicles", Vehicle System Dynamics, 1987, 16, 149-166
37 Rhodes D. and Cox S.J., "CNR tangent tie test", report no 1100, Pandrol Intnl. Ltd., London, 1985
38 Ripke B. and Knothe K., "Anpassung der Modellparameter eines Gleismodells an gemessene Gleisrezeptanzen", T.U. Berlin 1992
39 Grassie S.L., "Resilient railpads: their dynamic behaviour in the laboratory and on track", Proc. of Instn. of Mech. Engrs., 1989, 203F, 25-32
40 Kanamori T., "A few experiments on elasticity of rubber pad", JNR Quarterly Reports, 1980, 21, no 4
41 Raymond G.P., "Analysis of track support and determination of track modulus", Transport Research Record, no 1022, 1985, 80-90
42 Stewart H.E., "Measurement and prediction of vertical track modulus", Transport Research Record, no 1022, 1985, 65-71
43 Kopke U.G., "Condition monitoring of buried gas pipes using a vibrating pig", Ph.D. thesis, University of Cambridge, 1992

Rail Service Load Environment by Measurement and Modeling

Donald R. Ahlbeck, Battelle, Columbus, Ohio USA

1. Background

A key factor in all aspects of rail vehicle and track performance, but particularly in the service life and reliability of the rail, is the wheel/rail load environment. The steel rail is subjected to certainly one of the more severe mechanical environments known to man. In many ways it is analogous to the roller bearing in a severe impact application. The steel rail is overloaded and overstressed by increasingly heavy locomotives and freight cars. In curves the rail must resist the triaxial loads of the nonsteering solid-axle wheelset, Figure 1. It can also be subjected to severe dynamic load conditions even under the lighter axle loadings of passenger and rail transit vehicles.

Load-induced rail damage can take many forms. Repetitive high loads under unit trains can induce plastic flow at the rail surface, as well as subsurface fatigue failures. Cracks can be initiated at metallurgical anomalies, and rail flaws such as the transverse defect can grow under the service load regime. Curving wheelsets can produce extensive flange face and gauge corner wear on the high rail, as well as plastic deformation on both high and low rail running surfaces. Impact loads at rail running surface geometry errors, such as rail joints, engine burns, and poorly-ground welds, cause surface batter and deformation, as well as subsurface head and web damage such as bolt hole cracks. Finally, the oscillatory dynamic phenomena that cause rail surface corrugations can produce plastic deformation, abrasion and wear, and even thermal damage to the running surface.

In the context of flaw growth and rail fatigue life, the rail service load characterization is best defined in a statistical sense. To this end, sophisticated load measurement instrumentation and extensive data acquisition experiments have been employed. In the context of wheel/rail dynamics and the damage initiation processes, transient load events can be of even greater interest. Both measurement and computer modeling techniques have been developed to explore these aspects of the rail load environment.

J.J. Kalker et al. (eds.), Rail Quality and Maintenance for Modern Railway Operation, 185–202.

2. Load Measurements

2.1 Measurement Techniques

Measurement of dynamic loads on the various track components has long been of interest to railroad administrations and track designers. Measurement of dynamic forces directly at the wheel/rail interface has been the ultimate goal. The contact patch, however, is for all practical purposes inaccessible to measurements, unless some unforeseen future technique makes this possible. At this time, one must measure from a distance and sort out the intervening uncertainty as best one can.

Techniques used to measure wheel/rail loads fall into one of two categories: loads measured at a specific location in the track (wayside measurements), or loads measured by a specific wheelset (on-board measurements). The two methods provide an orthogonal view of the service load environment. Wayside load data pertain only to the instrumented bit of rail or track, but accumulate these measurements under a wide range of passing cars and locomotives. Wheelset load data accumulate over a wide variety of track conditions and geometry errors, but pertain only to the instrumented wheelset and its specific vehicle. Combinations of both are needed to extrapolate the load environment to the greater universe of vehicles and track.

The development of the strain gage in the late 1930s was the key factor in wheel/rail load measurement technology. Almost all successful load measurement techniques are strain-gaged based. Some accelerometer-based techniques have been used. Unless the accelerometer is mounted to a rather stout structural mass, however, the relationship between acceleration and force cannot be easily defined.

The art of load measurements with strain-gaged wheelsets has been developed and refined over the past thirty years[1]. A number of major research programs have used this technique. Strain gages are mounted in arrays on either the wheel plate[2] or wheel spokes[3] to measure vertical, lateral, and even longitudinal forces. The use of the wheel as a transducer does have certain limitations. Short-duration impact loads at the wheel/rail interface (the traditional "P1" force) are attenuated by the relatively massive wheel rim. Structural modes of vibration in the rim, spokes, or plate are excited by impact loads. Because of these vibrations, the wheel-mounted strain gage signals are distorted and must be low-pass filtered below 200 Hz, and much of the wheel/rail impact load data is lost.

The art of load measurements using the strain-gaged rail has also been advanced in the past thirty years[1]. Vertical wheel/rail loads are measured with the standard ORE-design strain gage pattern on the rail web between crossties (sleepers). These gages measure wheel loads between the chevron gage patterns (i.e., within the crib) through shear effects. As a transducer, the ORE gage pattern is remarkably linear and free from crosstalk due to orthogonal loads, rail bending, or contact patch position. It has a trapezoidal "influence zone" within which a sample of the passing wheel load is captured, Figure 2. Since the intervening rail head mass is small relative to the wheel rim, the rail-based transducer has a useful frequency range of 1000 Hz or more.

Measurement of lateral load by rail-mounted strain gage patterns has had a less satisfactory history. ORE experiments have, in the past, used vertically-oriented gages on the rail web located over the sleeper to measure lateral load-induced bending. This arrangement suffers a rather large crosstalk error due to the position of the vertical load on the rail running surface. A "base chevron" pattern was therefore devised[1] to circumvent these problems. This pattern is reasonably linear up to a lateral to vertical load ratio (L/V) of about 0.3 and is reasonably immune to crosstalk within its influence zone.

Measurement of rail dynamic longitudinal loads at the track can currently be accomplished only by major structural modifications to the rail. Railroad administrations frown upon this approach. Rail web-mounted gage patterns that can measure thermally-induced longitudinal rail loads[4] suffer, unfortunately, from large crosstalk errors from passing vertical wheel loads. We look to the ingenious to solve this measurement problem with a physically and economically practical transducer.

2.2 Measured Rail Service Loads

Several large-scale measurement programs were undertaken during the late 1970s to characterize wheel/rail loads in the revenue train service environment[5,6,7]. Examples of the statistical occurrence of vertical and lateral loads on the rail are given in Figures 3 and 4, respectively. The curves, on a normal probability plot, represent at any point the percentage of passing wheel loads exceeding that particular load level. These plots[5] were developed from seven randomly-chosen measurement sites on tangent wood-tie track with continuous welded rail (CWR). Traffic on this single-track line consisted of mixed freight, including unit grain and coal trains and intermodal (piggyback trailer and container) trains, predominantly in the 75 to 120 km/h speed range. From Figure 3, we see that the median vertical wheel load of this traffic mix is about 80 kN.

The statistical description of rail loads in Figures 3 and 4 can change substantially from month to month (with seasonal harvest traffic, for example) and from year to year as traffic density and types of loads change. The predicted vertical load spectrum under unit train operation (a mine railroad, for example) is also shown in Figure 3. The sharp knee in the curve near the 50 percent (median) level shows that most of the axles are under either fully-loaded cars (and locomotives) from the mine or empty cars returning to the mine.

In the lateral load exceedance curve, Figure 4, a rather sharp discontinuity is seen below the 15 percent probability level. The rapidly increasing lateral loads at the lower probability levels reflect the speed-dependent truck hunting phenomena. This type of instability drives some wheelsets into periodic hard flange contact at higher speeds. At lower speeds, this behavior subsides and a more Gaussian distribution of lateral creep forces (on this tangent track) results.

2.3 Extreme-Value Loads

Perhaps of greater interest to the track engineer and rail metallurgist is the frequency of occurrence of extreme-amplitude loads. The development and commercialization over the past eight years of the Wheel Impact Load Detector (WILD)[8,9] has provided us with an

invaluable source of data on these low-probability events. By storing measured load maxima as counts in "load bins", the WILD can develop statistically significant data bases under weeks or months of traffic.

Examples of extreme-value vertical wheel/rail loads are shown in Figure 5. Two different populations are illustrated here. First (the top two curves) are the peak (per wheel) values, each representing the maximum value per passing wheel from a group of load measurement sites. A point on this curve represents the probability that a passing wheel will exceed that load level at some point on the rail, within a wheel diameter. With the current design of the WILD (in terms of the number of circuits and the sleeper spacing), about 80 to 90 percent of the passing wheel diameter is inspected. The second population (the bottom two curves) is the "event", the probability of the load on a given small piece of rail (about 20 cm long, Figure 2) exceeding that particular load level. Traffic on two different railroads is compared in Figure 5. Traffic on Railroad C is mixed freight and loaded coal trains, while traffic on Railroad F is more predominantly intermodal trains.

The extreme-value vertical load statistics can vary substantially from week to week, as shown in Figure 6. This variation can result from one particular unit train with marginal wheel conditions passing by one week, but not the next. Impact loads are usually speed dependent: in Week 3 of Figure 6, a speed restriction was enforced for track maintenance during part of the week. This is seen in the train speed distributions of Figure 7. Impact load amplitudes are also dependent on track modulus. The load statistics on wood-tie track (75 to 120 km/h freight traffic) are more in line with the lower-speed (30 to 100 km/h) traffic on concrete tie track in Figure 6.

2.4 Curved Track Loads

Measurements of wheel/rail loads on curved track have for the most part been concentrated in particular test programs on the behavior of specific rail vehicles. Analytical studies (see for example the paper by Newland[10]) have defined the lateral and longitudinal loads in terms of steady-state curving behavior. With the solid-axle wheelset in a two-axle bogie arrangement, the leading wheelset assumes some angle of attack relative to the curve radius. On smaller radius curves, flange contact occurs on the high rail. The vector (net) combination of lateral creep and flanging forces on the high rail, plus lateral creep force on the low rail, result in a track gauge-spreading load. Except in extremely short-radius curves, the trailing wheelset assumes a near-zero angle of attack and develops low lateral and longitudinal creep force levels.

Steady-state curving in the real world is, however, an idealized average condition. Lateral and vertical loads under a vehicle in a curve will oscillate about the steady-state values due to curve entry dynamics, track geometry errors, or kinematic hunting behavior. An example is shown in Figure 8 for the leading bogie of a heavy coal-hauling freight car on a 5-degree (349 m radius) curve on 11.9-meter bolted-joint rail (BJR) track with staggered rail joints. Plots show the variation in wheel/rail loads by distance from the high rail joint. Even on CWR, the effects of geometry errors at welds will cause similar (but attenuated) variations in load.

3. Analytical Modeling of Rail Loads

Analytical models of the wheel/rail dynamic system provide a means for predicting the load environment at the point of greatest interest -- here, the wheel/rail contact patch. This location is inaccessible to current practical measurement techniques and its loads are somewhat obscured at "remote" measurement sites on rail or wheel. Therefore the analytical approach can be a "best bet" for both a qualitative and quantitative understanding of the contact patch loading.

Analytical techniques used by researchers have ranged from closed-form solutions to lumped- and distributed-parameter mathematical models to very complicated finite-element models. The analyst must steer a course between simplicity (and possibly missing important dynamic phenomena) and too much complication, which can obscure important relationships and increase computation time and costs unnecessarily.

3.1 Curving Loads

The closed-form solution methods have worked well for studying steady-state curving load behavior[10,11]. Examples of predicted lateral and longitudinal loads from this type of analysis for our old favorite, the three-piece freight car bogie, are given in Figures 9 and 10. The model SSCRV2 uses an iterative matrix solution to solve a set of nonlinear algebraic equations and to satisfy wheel flanging and wheel/rail creep limit conditions. In this example, a 5-degree (349 m radius) curve on branchline track was simulated.

In Figure 9, the net lateral force of the lead outer wheel on the high rail is the sum of the positive (outwardly directed) lateral flange force and negative (toward the center of curvature) creep force. This is the lateral load measured at the rail (Figure 8). The flange force, which cannot be measured directly, produces the greater damage to curved rail through rail gauge-face wear, which is controlled by lubrication and (on transit systems) by guard rails. The negative (inward) lateral creep force of the leading inner wheel on the low rail can result in running surface damage, rail rollover, and wheel drop derailments if sufficiently large. Lateral creep forces on the trailing wheels can be positive or negative in direction and, for this combination of vehicle and track characteristics and wheel/rail geometry, are quite small.

Longitudinal loads on the rails under the four wheels are shown in Figure 10 for this specific curving situation. Positive loads (as plotted) force the rail away from the direction of vehicle forward motion. Additional loads due to torques on the wheelsets (such as those due to braking) have not been included in this particular analytical solution.

Dynamic models are needed, however, to predict the actual varying load behavior illustrated in Figure 8. In a time domain model, the nonlinear wheel/rail creep and flanging relationships are used directly in the force calculations. The model may be "started" from tangent track passing through an entry curve (spiral) and into the body of the curve, or it may be started from analytically calculated steady-state conditions in the body of the curve itself.

3.2 Wheel Impact Loads

Impact loads due to wheel tread or rail running surface geometry errors can be a major factor in rail service life. The loads at rail geometry errors occur with some statistical distribution of the peak amplitudes, but always near the same point on the rail. Damage accumulates relatively quickly. Impact loads from wheel flats occur with some small, but measurable probability at any given point on the rail. A rail with an existing mature crack can break under these low-probability loads.

Analytical models have been used to explore the effects of these impact loads on rail, track structure, and wheel bearing life[12,13,14]. The rail itself has been modeled to varying degrees of complexity -- as, for example, a Timoshenko beam on an elastic foundation, continuous or discrete[15,16]. Even a relatively simple, but nonlinear rail and track structure model can provide good correlation with measured impact loads under measured wheel tread geometries[17], Figure 11. In this model, the rail effective mass has been approximated by standard beam-on-elastic-foundation (BOEF) relationships. Nonlinear Hertzian wheel/rail contact and rail/tie support stiffnesses are used, plus the transverse bending modes of the concrete or wood crosstie.

Using this type of model, some interesting differences in rail contact patch forces can be predicted between concrete and wood-tie track. These are shown in Figure 12, where the solid curve plots the rail/tie reaction force, which approximates the measured force of the WILD system, and the dotted curve plots the wheel/rail contact force. On rail stiffly coupled to the more massive (350 kg) concrete ties, the two force time histories are nearly identical. On rail loosely coupled to the more resilient and lighter (102 kg) wood ties, there is much higher frequency content in the force at the contact patch.

3.3 Rail Corrugations

One of the problems that has plagued railroad and rail transit administrations for over 100 years is that of rail corrugations. These periodic waves on the rail running surface generate excessive noise, shorten rail life (as well as the life of track and vehicle components), and increase maintenance costs. Mathematical models of varying degees of complexity have addressed this problem in an effort to define the initiation and growth phenomena[18,19,20]. In a recent study of corrugations on the Baltimore Metro[21], a nonlinear model of the wheelset and track components was used to explore the magnitude and phase relationships of the triaxial loads at the wheel/rail contact patch under known rail corrugation-producing conditions. Examples of load time-histories are shown in Figures 13 and 14. These instantaneous load relationships can be used in complex stress analyses to predict rail surface and subsurface Hertzian contact and shear stresses. The key point in this study is the realization that the wheelset, rail, and track structure act in concert as a complex and total dynamic system. In this particular case, the rail gets much the worst of the situation.

4. Summary

Measurements and mathematical modeling of track loads have provided a better understanding of the nature of wheel/rail loads. Rail loads have been defined quantitatively for specific vehicle and track combinations and for representative traffic and operating conditions. In the design of railway track, both the statistical and temporal load regimes must be addressed by the civil engineer. Both track and equipment designers must view the wheel/rail interface as part of a complex dynamic system, not as a convenient boundary to one or the other's responsibilities.

References

1 P. Tong and R. Greif (Ed.), *International Conference on Wheel/Rail Load and Displacement Measurement Techniques*, January 19-20 1981, DOT-TSC-UMTA-82-3, Sept. 1981.

2 M. R. Johnson, Development and Use of an Instrumented Wheelset for the Measurement of Wheel/Rail Interaction Forces, op. cit., Paper 3, pp. 3-1 to 3-17.

3 A. R. Pocklington, The B.R. Load Measuring Wheel, op. cit., Paper 4, pp. 4-1 to 4-29.

4 D. R. Ahlbeck and H. D. Harrison, Monitoring Rail Neutral Temperature and the Long-Term Stability of Track, *Rail Transportation - 1989*, ASME RTD-Vol.3, pp. 135-140.

5 D. R. Ahlbeck, H. D. Harrison, J. M. Tuten, and M. R. Johnson, *Measurements of Wheel/Rail Loads on Class 5 Track*, Report No. FRA/ORD-80/19, February 1980.

6 H. Harrison and M. Moody, Correlation Analysis of Concrete Cross Tie Track Performance, *Proceedings, Second International Heavy Haul Railway Conference*, 1982, Paper 82-HH-39, pp. 425-431.

7 H. D. Harrison, et al, *Correlation of Concrete Tie Track Performance in Revenue Service and at the Facility for Accelerated Service Testing*, FRA/ORD-84/02.1, August 1984.

8 J. M. Tuten and H. D. Harrison, *Design, Validation and Application of a Monitoring Device for Measuring Dynamic Wheel/Rail Loads*, ASME Paper 84-WA/RT-10, Dec. 1984.

9 H.D. Harrison, Second generation wayside impact detector, *Railway Technology International '90*, pp. 197-201.

10 D. E. Newland, Steering a Flexible Railway Truck on Curved Track, *Journal of Engineering for Industry, Trans. ASME*, August 1969, pp. 909-918.

11 J. A. Elkins and O. K. Ahmad, *Development and Applications of a Steady-State Curving Program*, ASME Paper No. 85-WA/RT-15, Dec. 1985.

12 S. G. Newton and R. A. Clark, An Investigation into the Dynamic Effects on the Track of Wheelflats on Railway Vehicles, *Journal Mechanical Engineering Science*, IMechE 1979 Vol 21 No 4, pp. 287-297.

13 D. R. Ahlbeck and J. A. Hadden, Measurement and Prediction of Impact Loads from Worn Railroad Wheel and Rail Surface Profiles, *ASME Journal of Engineering for Industry*, 107, May 1985.

14 S. Williams, D. R. Ahlbeck, and H. D. Harrison, *Railroad Bearing Performance Under the Wheel Impact Load Environment*, ASME Paper No. 87-WA/RT-1, Dec. 1987.

15 S. L. Grassie and S. J. Cox, The dynamic response of railway track with flexible sleeper to high frequency vertical excitation, *Proc Instn Mech Engrs Vol 198D No 7, IMechE 1984*, pp. 117-124.

16 S. L. Grassie and S. J. Cox, The dynamic response of railway track with unsupported sleepers, *Proc Instn Mech Engrs Vol 199 No 2D, IMechE 1985*, pp. 123-135.

17 H. D. Harrison, A. E. Bethune, and D. R. Ahlbeck, Comparison of Measured and Predicted Impact Loads on 100 Ton Coal Gondolas, *Proc. The 4th International Heavy Haul Conference 1989*, The Institution of Engineers, Australia.

18 R. A. Clark and P. Foster, On the mechanics of rail corrugation formation, 8th IAVSD Conf. Proc., Cambridge, MA, 1983, Swets and Zeitlinger, Lisse, pp. 75-85.

19 R. A. Clark, G. A. Scott, and W. Poole, Short Wave Corrugations - An Explanation Based on Stick-Slip Vibrations, *Applied Mechanics Rail Transportation Symposium - 1988*, AMD-Vol. 96, RTD-Vol. 2, ASME 1988.

20 E, Tassilly and N. Vincent, A Linear Model for the Corrugation of Rails, *Journal; of Sound and Vibration (1991) 150(1)*, pp. 25-45.

21 D. R. Ahlbeck and L. E. Daniels, Investigation of rail corrugations on the Baltimore Metro, *Wear 144*(1991), pp. 197-210.

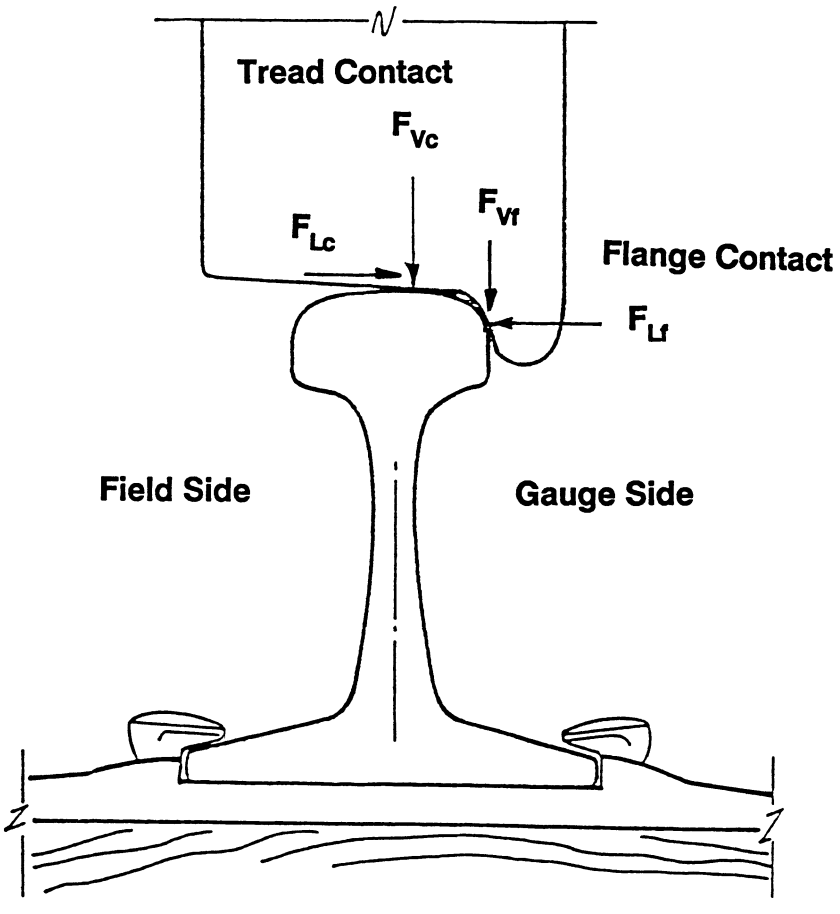


Figure 1. Loads at the Wheel/Rail Interface.

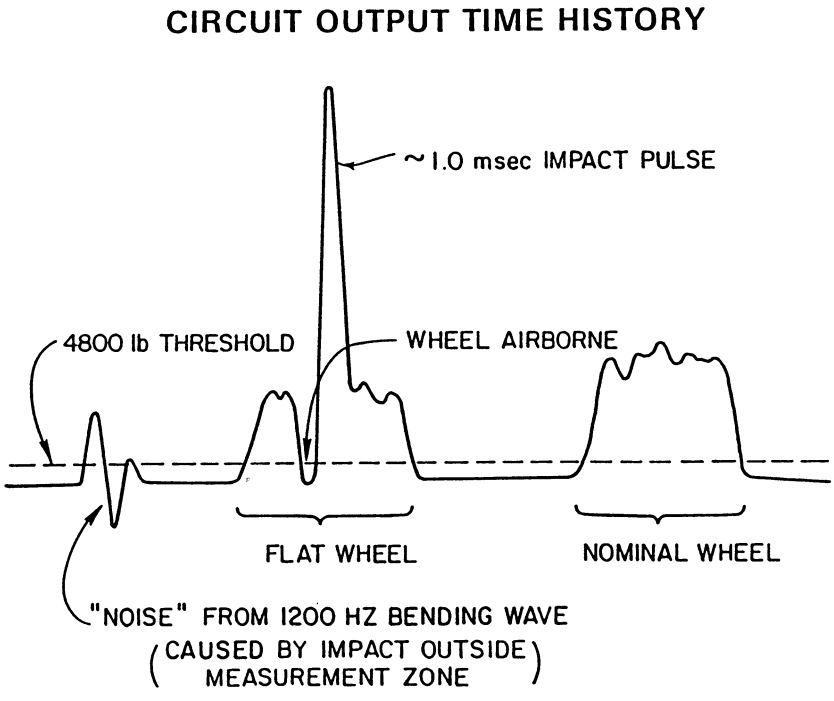


Figure 2. Typical Vertical Load Circuit Output.

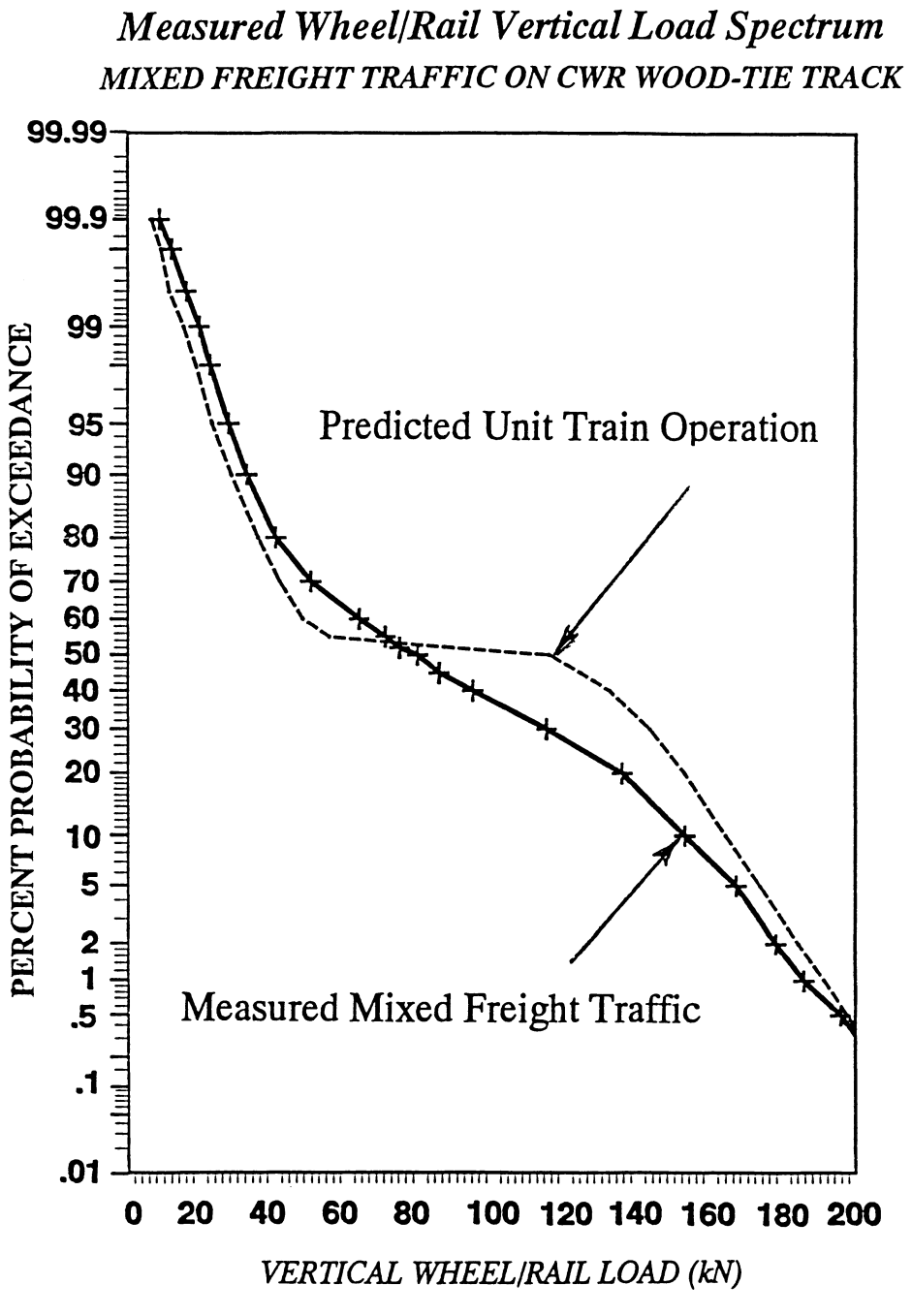


Figure 3. Example of Rail Vertical Service Load Environment.

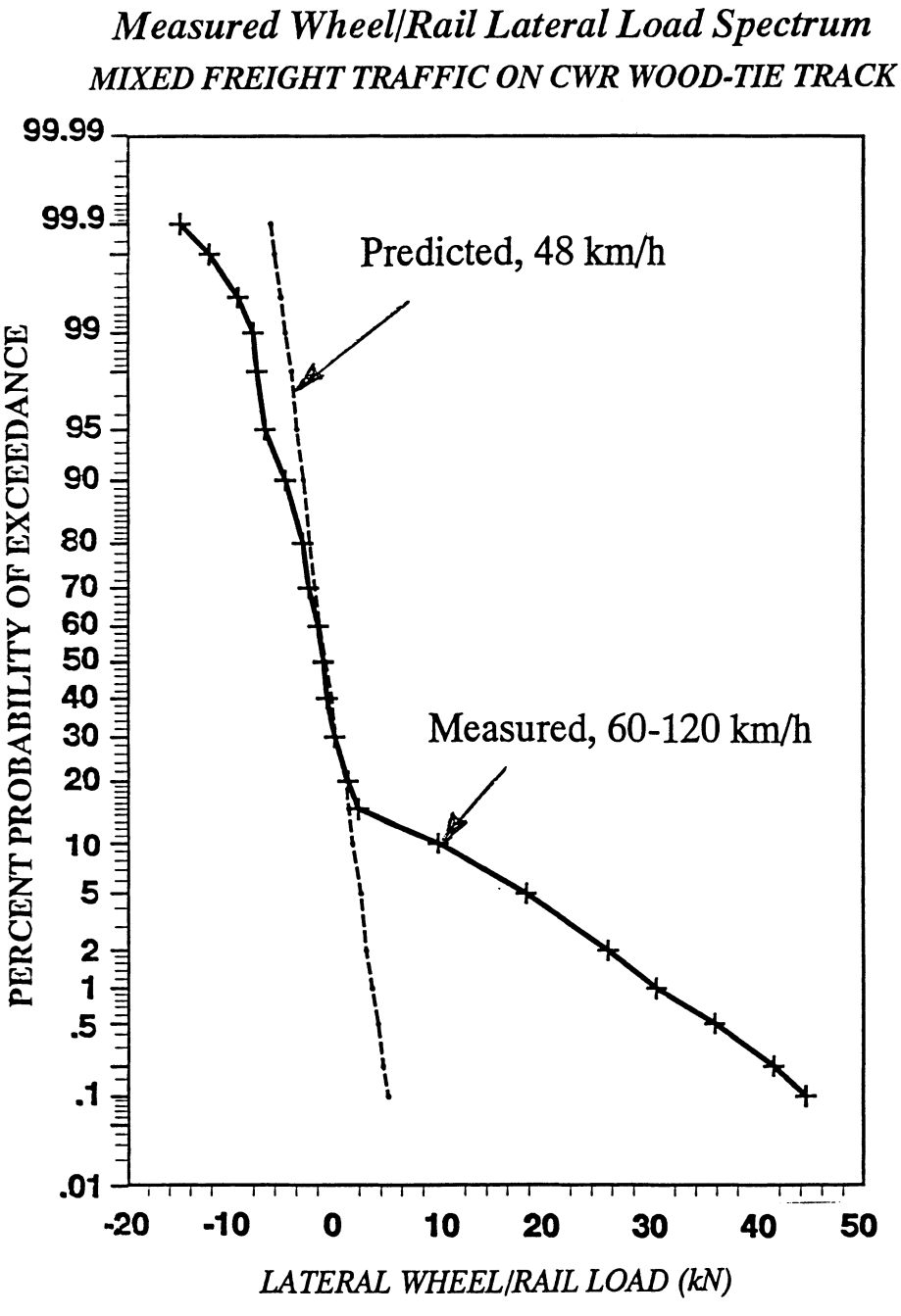


Figure 4. Example of Rail Lateral Service Load Environment.

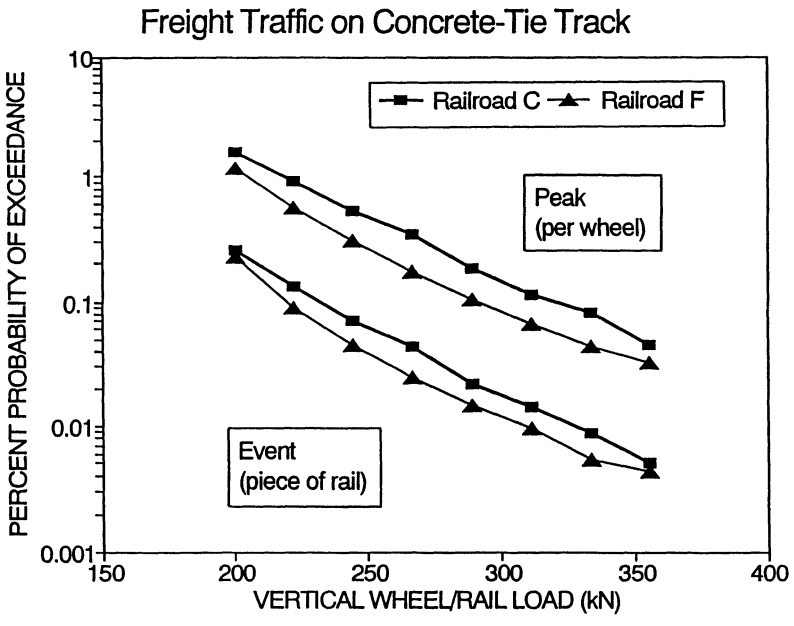


Figure 5. Extreme-Value Vertical Load Environment for Wheel and Rail.

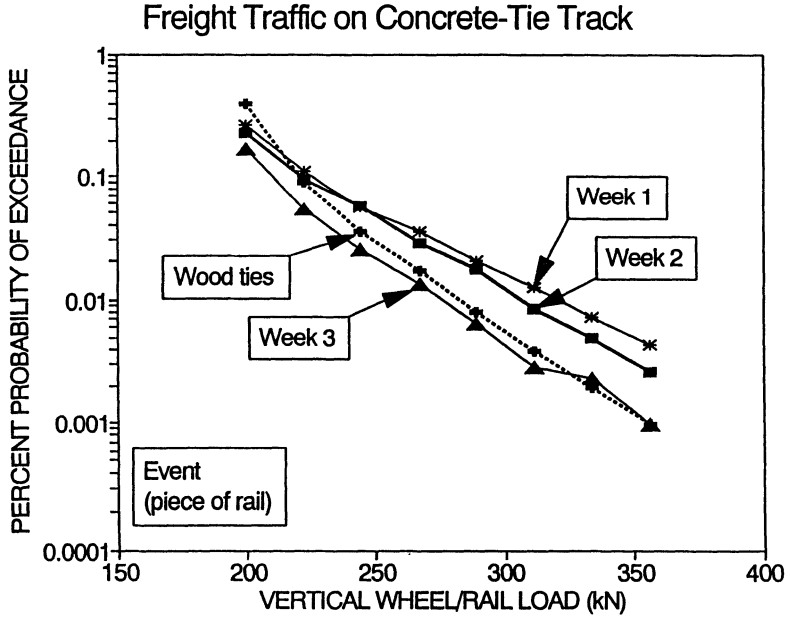


Figure 6. Variability of Extreme-Value Vertical Load Environment.

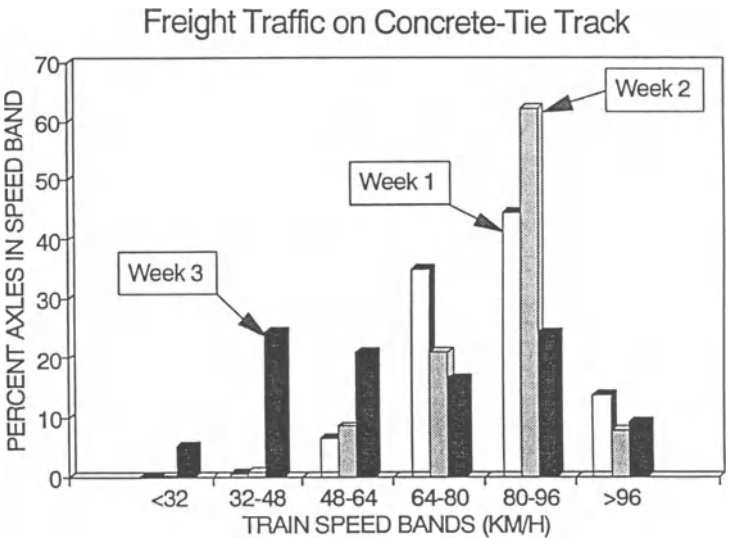


Figure 7. Speed Distribution of Wheelsets for Three Weeks of Traffic.

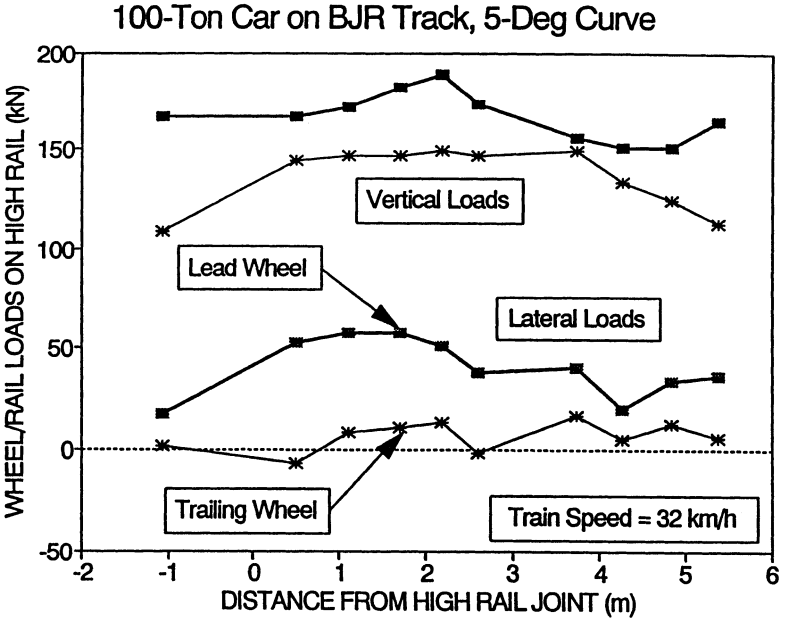


Figure 8. Loads on High Rail of Curve Under Heavy Freight Car.

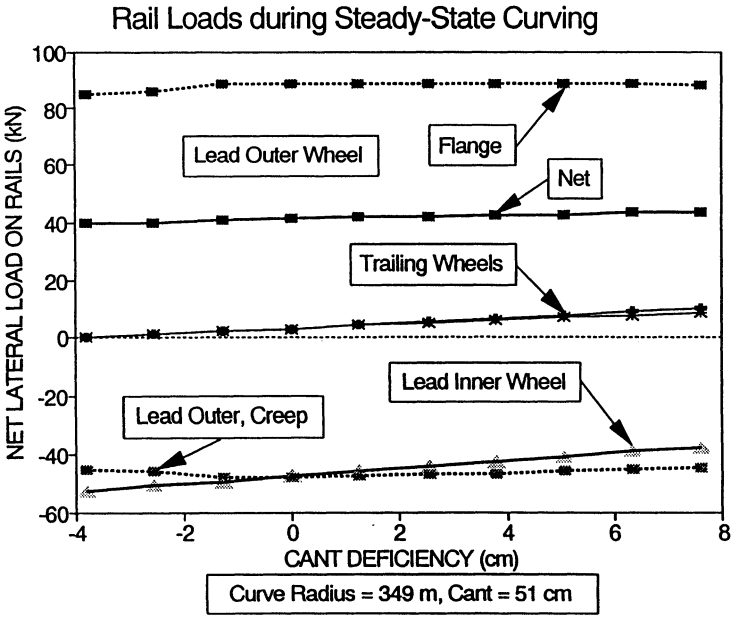


Figure 9. Predicted Lateral Loads on Curved-Track Rail.

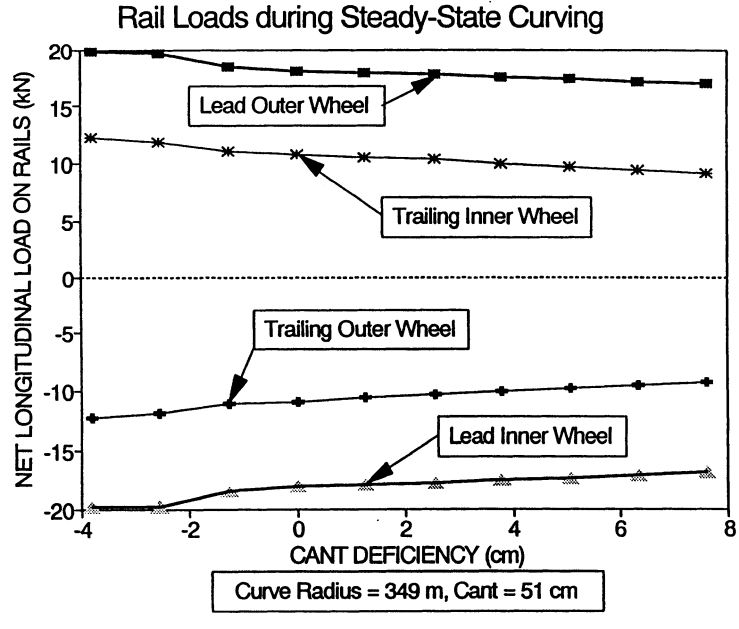
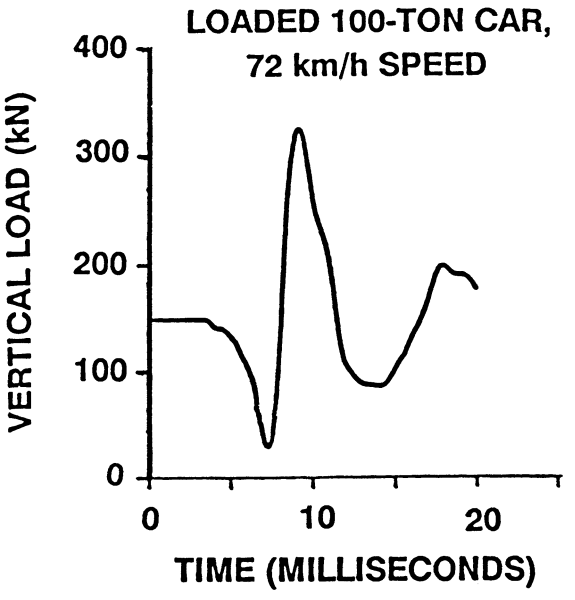
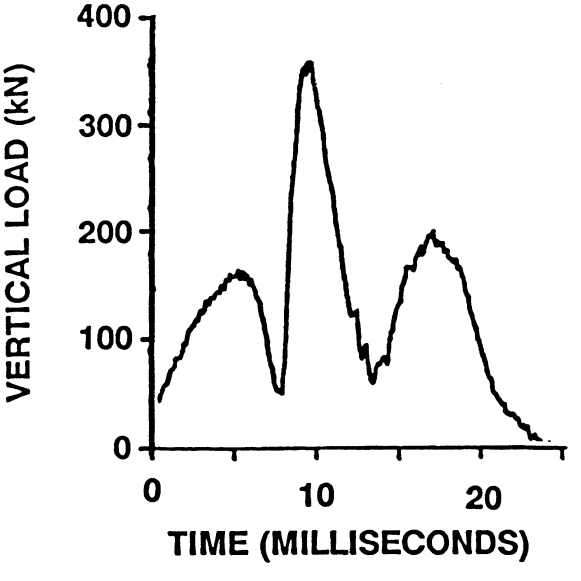


Figure 10. Predicted Longitudinal Loads on Curved-Track Rail.

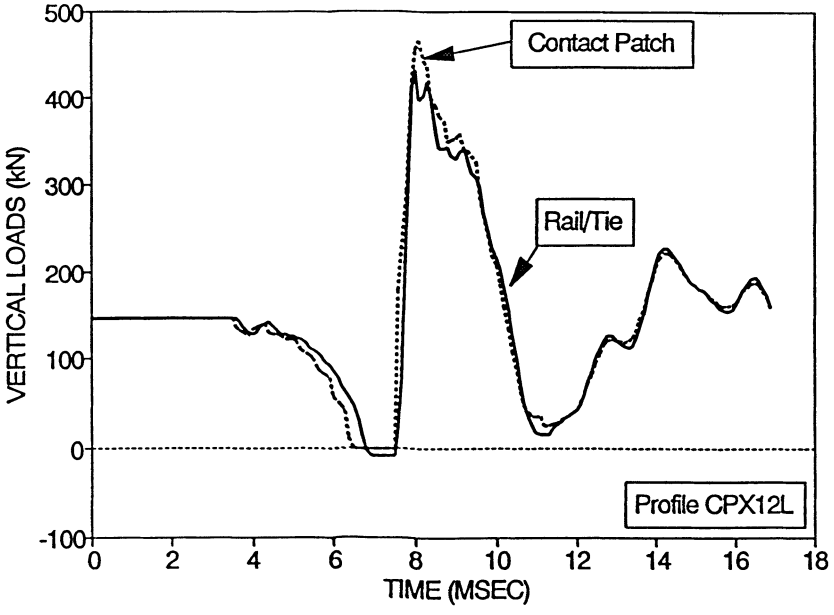


a. Predicted Load (Program IMPWHL)

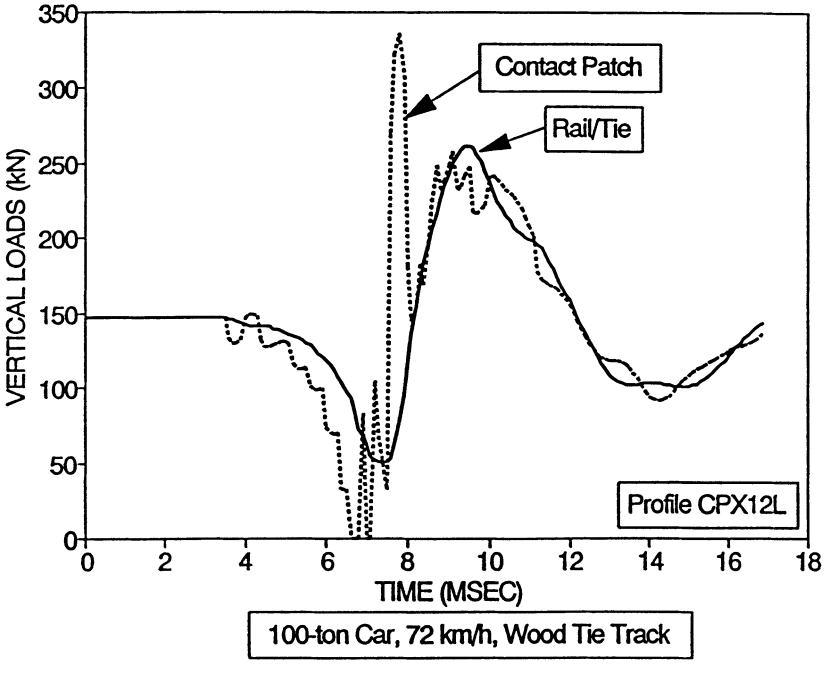


b. Measured Load (CP Rail Tests)

Figure 11. Comparison of Predicted and Measured Impact Loads.



a. Impact Loads on Concrete-Tie Track



b. Impact Loads on Wood-Tie Track

Figure 12. Impact Loads on Different Track Types.

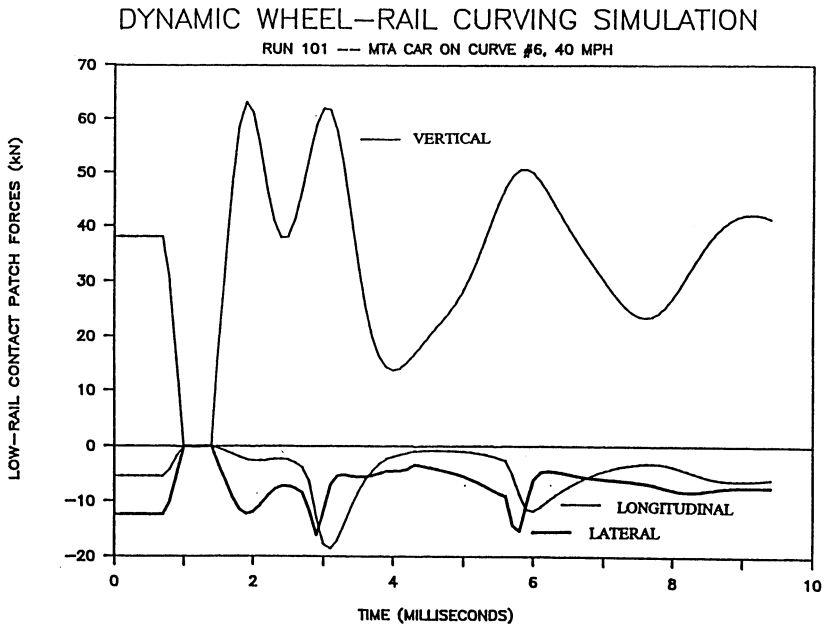


Figure 13. Predicted Corrugation-Initiating Loads.

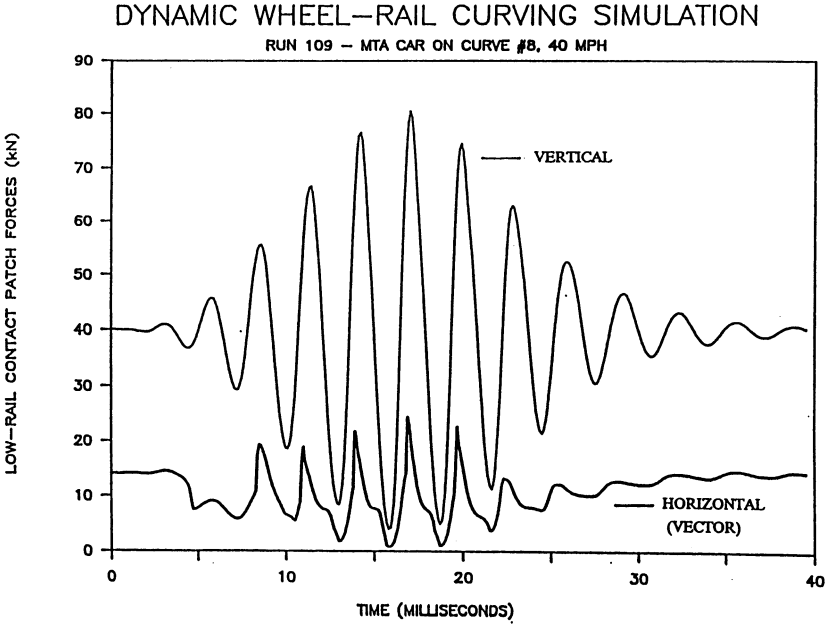


Figure 14. Predicted Wheel/Rail Loads on Corrugated Rail.

Analytical and Numerical Treatment of Flange Contact

M. T. Hanson
Dept. of Engrg. Mechs.
University of Kentucky
Lexington, KY 40506-0046

L. M. Keer
Dept. of Civil Engrg.
McCormick School of Engrg.
and Applied Science
Northwestern University
Evanston, IL 60208-3130

Abstract

This paper addresses contacts which occur in rail-wheel systems. In particular the edge effect caused by the gauge face of the rail is examined. The methods developed by the authors in a previous paper for contact analysis of a quarter space geometry are presently extended to investigate frictionless contacts near the edge of a wedge with varying wedge angles. Numerical results illustrate the effect of wedge angle on contact stress and contact area geometry. The loss in stiffness caused by proximity of the contact region to the gauge face of the rail is also quantified. To simulate side loading of the gauge face by the wheel flange, a modification of the analysis is made to accomodate frictionless contact on both faces of the wedge. Numerical results are given for the contact stress when symmetric contact is assumed.

1. Introduction

In rail-wheel contact the evaluation of the contact stresses and the internal stress fields in the rail head are extremely important for understanding and predicting failure mechanisms. The stress field is quite complex since in addition to non-linear (plastic) deformation which may occur under heavy loading, the presence of the gauge face of the rail head may significantly alter the contact stress and internal stress fields in the rail head from a simplified half space model. A traction free gauge face will cause an edge effect on the stress fields. In some instances the wheel flange will exert a loading on the gauge face, thus compounding the edge effect.

It is the purpose of this study to examine the edge effect caused by the gauge face of the rail. In a previous paper [1], the authors addressed this issue by developing a stress analysis capability for a three-dimensional incompressible wedge under normal loading on one face while the other face was traction free.

J.J. Kalker et al. (eds.), Rail Quality and Maintenance for Modern Railway Operation, 203–214.

It was shown for a right angle wedge (quarter space) that proximity to the edge significantly affects contact stress and contact area geometry as compared to similar half space results.

The present analysis will further investigate edge phenomena by considering frictionless contact near the edge of a wedge with varying wedge angle. Thus the effect of wedge angle on contact stresses and contact area geometry will be examined. Furthermore, the loss in stiffness as a function of distance from the edge will be obtained. As noted above, the wheel flange may exert a loading on the gauge face thus altering the contact and subsurface stress fields. To address this point, the analysis in [1] is modified to account for normal loading on both faces of the wedge. The solution for point normal loading on each face will be derived. The point force Green's function is then used to obtain integral equations for the contact stress on both the upper surface of the rail head and the gauge face of the rail head. These integral equations are solved using the methods developed in [1]. Numerical results for contact stress and contact area geometry are given for various wedge angles and for contact on the gauge face of the rail head.

2. Frictionless Contact for a Three-Dimensional Wedge

In this section frictionless contact on the upper face of a three dimensional wedge is examined. The geometry and coordinate system is shown in Figure 1. The z axis is taken along the edge of the wedge while the polar coordinate r is measured from the edge and the angle θ is measured from the middle surface. The wedge faces are located as $\theta=\pm\alpha$. Rigid spherical indentation is considered and the indenter tip is located a distance r_c from the edge. The coordinates r^* and z^* are attached to the indenter tip. Denoting $v(r,\alpha,z)$ as the normal displacement of the upper surface the integral equation derived in [1] is given as

$$v(r,\alpha,z) = -\frac{1}{4\pi}\iint_{\Omega}\frac{\bar{p}(r_o,z_o)\,dr_o dz_o}{\sqrt{(r-r_o)^2+(z-z_o)^2}} + \frac{1}{4\pi}\iint_{\Omega}\bar{p}(r_o,z_o)\left(\frac{1}{\sqrt{rr_o}}\int_o^{\infty}\frac{F(\tau)}{G(\tau)}P_{i\tau-\frac{1}{2}}\left(\frac{r^2+r_o^2+(z-z_o)^2}{2rr_o}\right)d\tau\right)dr_o dz_o, \tag{1}$$

where $\bar{p}(r_o,z_o)$ is the contact pressure divided by the shear modulus. The functions $F(\tau)$ and $G(\tau)$ are given in [1] as

$$G(\tau) = \tau^2\sin^2(2\alpha) - \sinh^2(2\alpha\tau) \tag{2}$$

$$F(\tau) = G(\tau) + \tanh(\tau\pi)\{\tau\sin(2\alpha)\cos(2\alpha) + \cosh(2\alpha\tau)\sinh(2\alpha\tau)\} \tag{3}$$

In equation (1) $P_{i\tau-\frac{1}{2}}(s)$ is a Legendre function of the first kind which can be defined by an integral representation given in equation (30) of [1]. The numerical procedure is based on a piecewise constant approximation to $\bar{p}(r_o, z_o)$ with an iterative approach to determine the contact area. The details are given in [1].

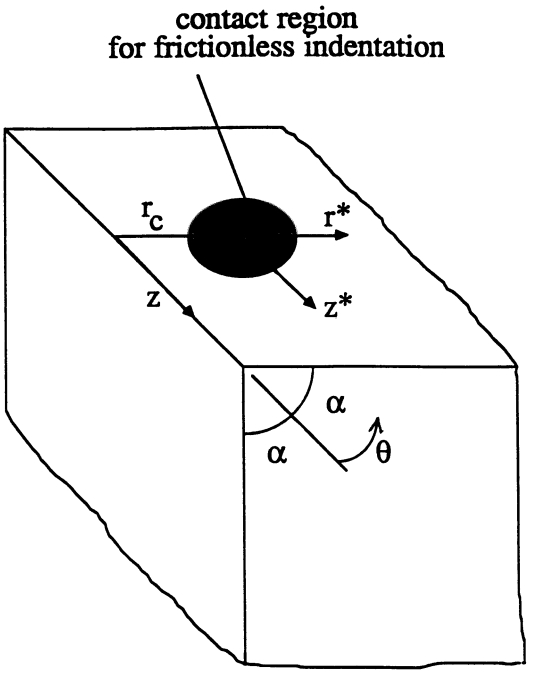


Figure 1. Geometry and coordinate system for frictionless contact on one face.

For spherical contact the displacement in the contact region is given as

$$v(r, \alpha, z) = -\delta + \frac{1}{2R}\{(r-r_c)^2 + z^2\} \qquad (4)$$

where R is the indenter radius and δ is the rigid body displacement of the indenter. Since the contact area will not in general be circular, there is not a convenient length scale useful for non-dimensionalization. Thus here as in [1] δ and R are taken in dimensional form although the actual units (mm, inches, etc.) need not be directly specified. All subsequent quantities will be understood to have dimensions consistent with those chosen for δ and R. In [1] the quantities $\delta=\pi/8$ and $1/2R=\pi/16$ were used and the reduction in contact area and pressure

were studied for α=45° (quarter space) as r_c was decreased. The effect of different indentation depths and wedge angle effects are examined here. The distance r_c is taken as ∞,3.0,1.5,1.0,0.5,0.0 where $r_c=\infty$ implies the half space results while r_c=0 places the indenter tip on the wedge tip.

The first example considers a wedge angle of 63 degrees. For an indentation depth of $\delta=\pi/8$, Figure 2 displays the contact area as r_c varies. The contact area remains circular (Hertzian) for $r_c \geq 1.5$ while for r_c=0.5,0.0 contact is made with the wedge tip. The pressure through the center of contact (z^*=0) is shown in Figure

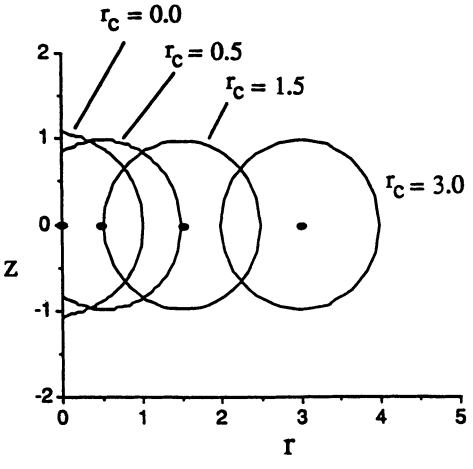


Figure 2. Contact area variation versus distance from the edge for a 63 degree wedge.

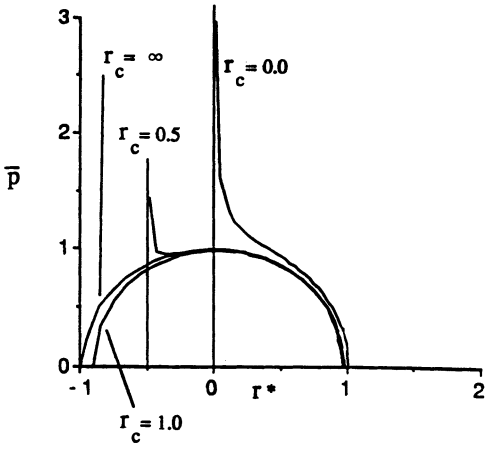


Figure 3. Contact stress through the center of contact for a 63 degree wedge.

3. Again for $r_c \geq 1.5$ the contact stress is Hertzian while a slight deviation occurs for $r_c=1.0$. When $r_c=0.5,0.0$ contact extends to the wedge tip and the contact stress is singular along the edge. The coefficient of the singularity varies along the wedge tip and increases as r_c decreases. This result is in contrast to a 45 degree wedge where it was shown in [1] that for $r_c=0.0$ the stress is not only non-singular but in fact vanishes along the edge.

The second example considers a wedge angle less than 45 degrees, namely 30 degrees. Figure 4 displays the contact area for various r_c values. In this case the contact area decreases much faster as r_c is decreased and hence one must go further away from the edge to recover the Hertzian results. It is interesting to note that as r_c is decreased the center of contact area shifts off the indenter tip in a direction away from the edge. In fact, for $r_c=0.0$ the indenter tip is over the edge but contact does not extend to the edge, i.e. some separation exists between the wedge tip and indenter tip. Figure 5 and 6 show the contact stress through the center of contact for various r_c values. As r_c is decreased, the contact stress falls rapidly, consistent with the decrease in contact area.

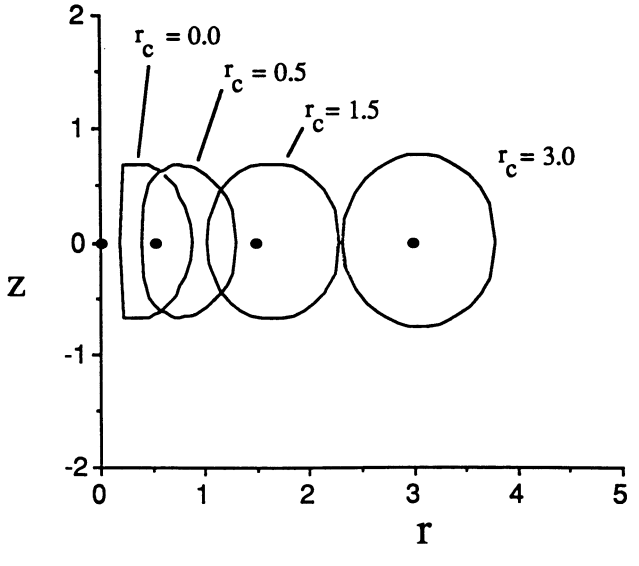


Figure 4. Contact area variation versus distance from the edge for a 30 degree wedge.

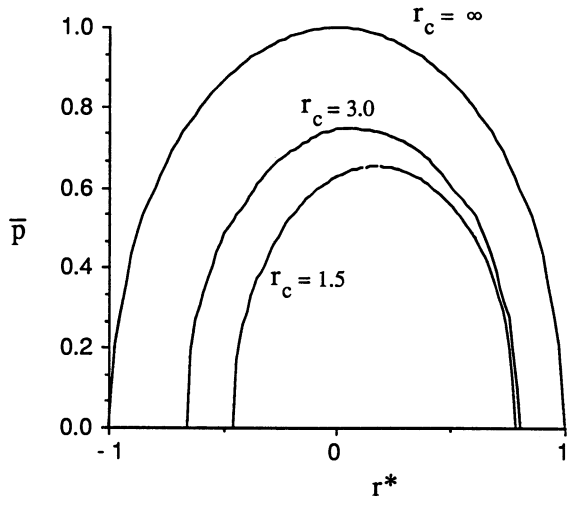


Figure 5. Contact stress through the center of contact for a 30 degree wedge.

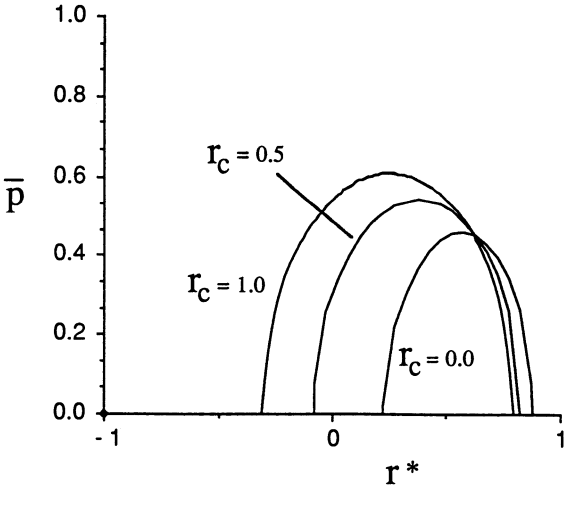


Figure 6. Contact stress through the center of contact for a 30 degree wedge.

In the present study wedge angles of 63 degrees and 30 degrees have been considered while a 45 degree wedge was used in [1]. In all cases the following trends in the results hold. As the wedge angle is decreased from 90 degrees (half space) the contact stress and contact area decrease for a given value of r_c, the distance from the edge to indenter tip. For a given wedge angle, the contact stress and contact area decrease as r_c is decreased. The smaller the wedge angle the more pronounced the decrease in these quantities. For wedge angles less than 45 degrees the edge will displace downward to cause a loss of contact along the edge.

Since the contact area and stress decrease as the contact approaches the edge for a given indentation depth, one may conclude that the edge effect leads to an increase in compliance of the body. To illustrate this, Figure 7 displays the contact force divided by the shear modulus versus r_c for an indentation depth of $\delta=\pi/8$ and various wedge angles. As $r_c \to \infty$ all three curves will approach the half space result. As r_c decreases, all wedge angles give an increased compliance with the smaller wedge angles giving a more pronounced effect. To further study this behavior the relation between indentation force and depth was investigated. The results for a 45 degree wedge (quarter space) are shown in Figure 8. For a given r_c value the displacement was incrementally increased and the force calculated. For all values $r_c>0.0$, the force deflection relation is initially Hertzian. As δ is increased the force increases but at a rate less than the Hertzian value, again displaying the loss in stiffness caused by the edge.

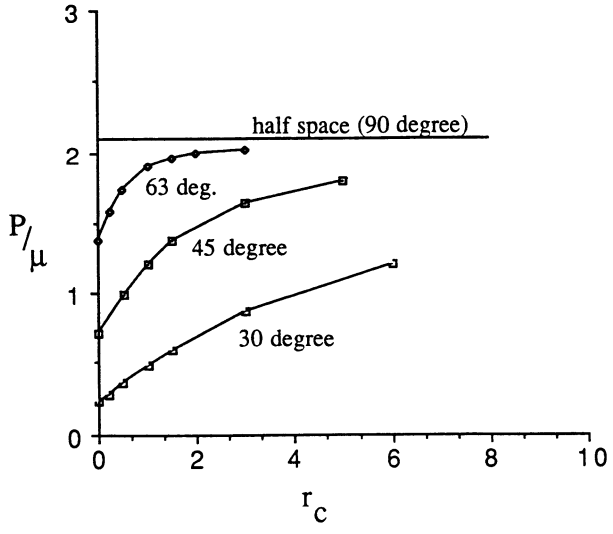


Figure 7. Contact force versus distance from the edge for a 45 degree wedge (quarter space).

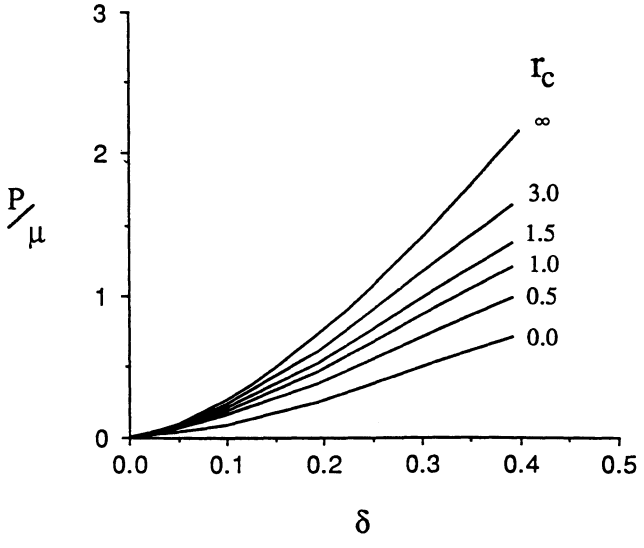


Figure 8. Force deflection relation for spherical indentation of a quarter space.

3. Simulated Contact on the Gauge Face

In this section the effect of contact on both the upper surface and gauge face of the rail is considered. The analysis in [1] was modified to include point normal compressive forces on both faces of the wedge. Denoting P_1 as the force on $\theta=+\alpha$ at $r=r_o$, $z=z_o$ and P_2 on $\theta=-\alpha$ at $r=r_o'$, $z=z_o'$, the expressions for the normal displacement on $\theta=\pm\alpha$ become

$$v(r,\alpha,z) = -\frac{P_1}{4\pi\mu}\frac{1}{\sqrt{(r-r_o)^2+(z-z_o)^2}} - \frac{P_2}{4\pi\mu}\frac{1}{\sqrt{(r+r_o')^2+(z-z_o')^2}}$$
$$+ \frac{P_1}{4\pi\mu\sqrt{rr_o}}\int_o^\infty \frac{F(\tau)}{G(\tau)} P_{i\tau-\frac{1}{2}}\left(\frac{r^2+r_o^2+(z-z_o)^2}{2rr_o}\right)d\tau$$
$$- \frac{P_2}{4\pi\mu\sqrt{rr_o'}}\int_o^\infty \frac{H(\tau)}{G(\tau)\cosh(\pi\tau)} P_{i\tau-\frac{1}{2}}\left(\frac{r^2+r_o'^2+(z-z_o')^2}{2rr_o'}\right)d\tau \qquad (5)$$

$$v(r,-\alpha,z) = \frac{P_1}{4\pi\mu}\frac{1}{\sqrt{(r+r_o)^2+(z-z_o)^2}} + \frac{P_2}{4\pi\mu}\frac{1}{\sqrt{(r-r_o')^2+(z-z_o')^2}}$$
$$+ \frac{P_1}{4\pi\mu\sqrt{rr_o}}\int_o^\infty \frac{H(\tau)}{G(\tau)\cosh\pi\tau} P_{i\tau-\frac{1}{2}}\left(\frac{r^2+r_o^2+(z-z_o)^2}{2rr_o}\right)d\tau$$
$$- \frac{P_2}{4\pi\mu\sqrt{rr_o'}}\int_o^\infty \frac{F(\tau)}{G(\tau)} P_{i\tau-\frac{1}{2}}\left(\frac{r^2+r_o'^2+(z-z_o')^2}{2rr_o'}\right)d\tau$$

The function $H(\tau)$ is given as

$$H(\tau) = -G(\tau) + \sinh(\pi\tau)\{\tau\sin(2\alpha)\cosh(2\alpha\tau) + \cos(2\alpha)\sinh(2\alpha\tau)\} \qquad (6)$$

To simulate contact on both faces the forces P_1 and P_2 are distributed with pressures $p_1(r_o,z_o)dr_odz_o$ and $p_2(r_o', z_o')\,dr_o'dz_o'$ over regions Ω_1 and Ω_2 to arrive at an integral equation over the contact region on each face. The solution of these coupled integral equations is beyond the scope of this paper. To investigate contact on both faces and stay within present capabilities the contact on each face is assumed identical, Figure 9. Therefore the midplane of the wedge is one of symmetry and the integral equations on each face become identical. The resulting equation is given as

$$v(r,\alpha,z) = -\frac{1}{4\pi}\int\int_{\Omega_1}\frac{\bar{p}_1(r_o,z_o)\,dr_odz_o}{\sqrt{(r-r_o)^2+(z-z_o)^2}} - \frac{1}{4\pi}\int\int_{\Omega_1}\frac{\bar{p}_1(r_o,z_o)\,dr_odz_o}{\sqrt{(r+r_o)^2+(z-z_o)^2}}$$
$$+\frac{1}{4\pi}\int\int_{\Omega_1}\bar{p}_1(r_o,z_o)\left(\frac{1}{\sqrt{rr_o}}\int_o^\infty\frac{1}{G(\tau)}\left[F(\tau)-\frac{H(\tau)}{\cosh(\pi\tau)}\right]\right.$$
$$\left.P_{i\tau-\frac{1}{2}}\left(\frac{r^2+r_o^2+(z-z_o)^2}{2rr_o}\right)d\tau\right)dr_odz_o. \qquad (7)$$

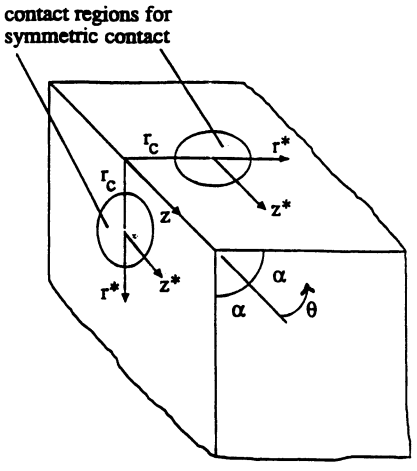


Figure 9. Symmetric contact on both faces.

This equation is similar to equation (1) for contact on a single face and hence the methods developed in [1] can be directly applied.

To illustrate the effect of contact on both faces spherical indentation will be used and hence $v(r,\alpha,z)$ is given by equation (4). Again $\delta=\pi/8$ and $1/2R=\pi/16$ are taken for comparison purposes.

Figure 10 displays the results for a wedge angle of 90 degrees (half space). The contact stress through the center of contact is plotted for $r_c=\infty$ (Hertzian) and $r_c=2.5$. The value of 2.5 is approximately the indenter radius and thus represents the minimum closeness physically allowable. It is apparent that the interaction

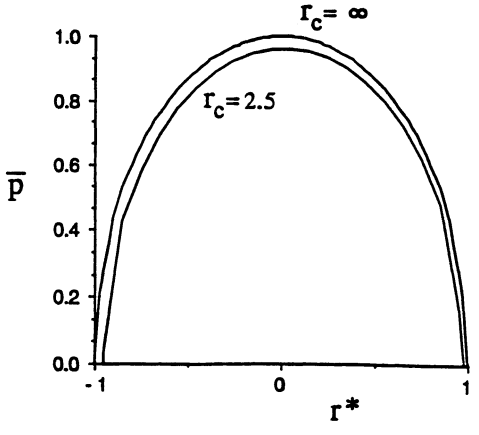


Figure 10. Contact stress for symmetric indentation of a half-space.

reduces the contact stress under each indenter. This is easy to explain since each indenter causes a downward surface displacement. Since the indentation depth of each indenter is specified, less indentation pressure is required because of the other indenters presence. The contact force for r_c=2.5 is P/μ=1.862 which is 11% less than the Hertzian value.

Figure 11 displays the results for a 45 degree wedge (quarter space). In this case as the indenters are brought closer to the edge the indentation pressure increases. This behavior is opposite of the half space results above. The explanation follows from the fact that the midplane of the quarter space is a symmetry plane and thus undergoes no displacement normal to itself. The symmetric contact gives a stiffening effect and thus requires more indentation pressure and force to achieve a specified displacement. The contact force has the values P/μ=2.096, 2.624, 2.982 for r_c=∞, 3.0, 2.0 respectively. The value r_c=2.0 was the minumum used for this double contact analysis since smaller values led to significant difficulties in numerically evaluating the infinite integrals for the integral equation (7). The increase in force is 25% for r_c=3.0 and 42% for r_c=2.0 as compared to the Hertzian value.

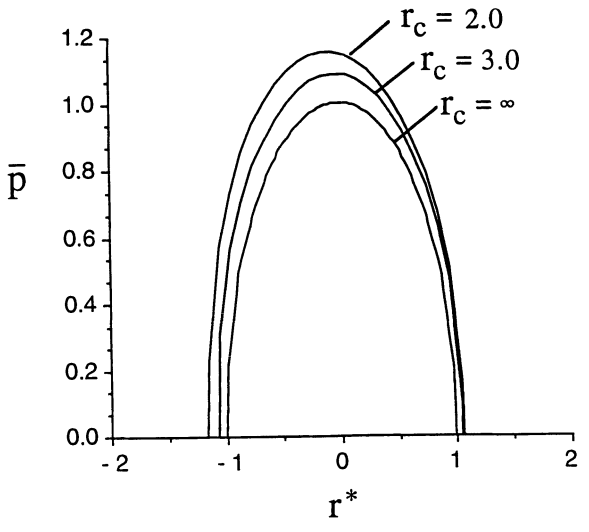


Figure 11. Contact stress for symmetric indentation of a quarter space.

4. Conclusions

It has been shown that edge effects can play a dominant role in stress analysis for frictionless contacts. For contacts on one face of a three-dimensional wedge it was shown that contact stress and contact area are significantly decreased as a spherical indenter is brought closer to the edge. Thus the edge effect is to cause a reduction in stiffness for the contact. The stiffness loss increases as the wedge angle is decreased. For a fixed value of r_c, stiffness is initially Hertzian but decreases as the indentation depth is increased. Furthermore, for wedge angles greater than 45 degrees, the contact stress becomes singular if the edge comes into contact whereas angles less than 45 degrees prevent the edge from contacting the indenter tip.

Symmetric contact on both faces causes a different behavior. For wedge angles of 90 degrees, the interaction causes a reduction in force for a specified indentation depth. As the wedge angle decreases, the effect of symmetry is to increase the stiffness of contact and give an increased force. Although only a 45 degree wedge was used, it is anticipated that this stiffness increase will be more pronounced for smaller wedge angles.

Acknowledgement

The authors are grateful for partial support from the Association of American Railroads.

References

[1] M. T. Hanson and L. M. Keer, "Analysis of Edge Effects on Rail-Wheel Contact," Wear, Vol. 144, pp. 39-55 (1991).

ON AN APPROACH TO PREDICTION OF CONTACT LOADING OF RAIL FOR TWO-POINT CONTACT BETWEEN WHEEL AND RAIL

J. Piotrowski
Warsaw University of Technology, Poland

Abstract

Two-point contact interaction between the wheel of a railway vehicle and rail takes place in many instances. It is important for the investigation of rail strength and fatigue to determine the contact loading occurring for two-point contact interaction. An approach to prediction of contact loading has been presented, which employs the ellipticization procedure of non-Hertzian contact between wheel and rail for the description of contact in vehicle-track model and reverts to non-Hertzian contact problems to calculate detailed surface stress distributions after interaction forces and creepages have been predicted.

Introduction

Interaction between rail vehicle and track takes place in the contact zones between wheels and rails and has a form of surface stresses distributed over the contact regions. At some distance from the contact zone the distributed stresses have the same effect as the concentrated forces calculated as the integrals of surface stresses over the contact region. These forces are referred to as the interaction forces and it is usual for vehicle dynamicists to think of interaction between vehicle and track in terms of the interaction forces. The description of interaction forces for vehicle/track models requires solving some geometrical problem and contact mechanical problems. Usually, such a description is based upon the geometrical data for profiles of rolling surfaces of wheels and rails. These data are collected with the use of measuring machines which have been developed by several railways and which allow to acquire the data describing geometry of both profiles (left and right) and the relative position of the two profiles. We can classify wheel/rail contact, as detected by numerous measurements and contact calculations into several characteristic modes shown in Fig.1.

J.J. Kalker et al. (eds.), Rail Quality and Maintenance for Modern Railway Operation, 215–224.

Fig.1

The two-point contact is shown in Fig.1 b,c,e. The situation b. frequently occurs on curves between guiding (front outer) wheel and high rail. The other form of two-point contact on curves appears between groove rail and wheel on tramway lines. This contact has different properties from b. but no comment will be made on it in this paper. Another form of two-point contact, Fig.1c frequently appears in many situations on curved and on nominally straight track. Two-point contact influences interaction between vehicle and track and has a profound impact on stresses in wheel and rail. In order to calculate a stress field coming from vehicle/track interaction it is necessary to determine these forces. For two-point contact we have to rely almost completely on predictions as standard measuring technique (load measuring wheelsets) can measure only the total forces along principal wheelset axes. Therefore we will outline an approach to mathematical modeling of interaction for two- point contact and will give some examples of contact surface stress distributions coming from two- point contact interaction. The question of the influence of these stresses on strength and fatigue is not a topical point of this paper and could be more competently answered by the authors of Residual Stress and Crack Propagation sections of this Conference.

Contact geometry

Having the data for profiles collected, the first problem to solve is purely geometrical one. We need to know where the two profiles of wheels meet the profiles of rails if the wheelset axis is perpendicular to the track centre line and is shifted in lateral by a distance y relative to track centre line. But before commencing such calculations we smooth the measured data and describe the profiles by some continuous functions as the raw data usually have a form of the co-ordinates of a number of points on profiles.

The geometrical problem at hand is a two dimensional one and is solved iteratively by a number of programs.

Notation:
input: y lateral displacement of wheelset relative to track centre line
output: θ roll, γ_ℓ, γ_r contact angles, r_ℓ, r_r actual rolling radii.

After the calculations have been done for a set of y: $[y_{max}, y_{min}]$ we sort the results of the output in the form of a table with y as an entry. This is a a geometrical one for yaw angle $\psi = 0$. The two-point contact leads to singularities in geometrical quantities of the form of jumps in rolling radii difference and in the contact angles.

For some applications, especially for curving predictions, we are interested in the geometry when the yaw angle $\psi \neq 0$. There are basically two ways of calculating geometrical quantities for $\psi \neq 0$. The first is approximate and is valid when the profiles for $\psi = 0$ and an arbitrary y are not closely conforming. This method employs local approximation of surfaces near the contact point for $\psi = 0$ by regular surfaces.

The simplest result by this method was obtained by Bödecker as early as in 1895, in the form of a closed formula for "cone-line" approximation. The only effect of yaw by this approximation is that the contact point on the wheel shifts forward/backward by a distance

$$s_\alpha = \pm r_\alpha \psi \tan \gamma_\alpha, \quad (\alpha = \ell, r).$$

The other method, which can be used without any qualification is based on so-called "visible profiles" of the wheelset. We yaw the wheelset by the angle ψ and we calculate the visible profiles of the wheels as seen along the direction of the track centre line. Thus, the three dimensional geometrical problem is reduced to two dimensions and can be solved in the same way as the problem for yaw angle $\psi = 0$. With this method we can detect two-point contact for $\psi \neq 0$ when the profiles conform, even if we have one-point contact for $\psi = 0$.

Now, the geometry is described by θ - roll angle, n_α $(\alpha = \ell, r)$ - unit normal vectors, r_ℓ, r_r - rolling radii, s_ℓ, s_r -longitudinal shifts of contact points Again, we can arrange the results for various yaw angles ψ in a table with two entries: y and ψ. We have a two dimensional table.

The contact problems

Under load, the point of geometrical contact grows to a contact area. The shape and size of the contact area and the pressure distribution is important in order to describe accurately the interaction forces in a mathematical model. The tangential (creep) forces depend on shape, size and pressure distribution and are described by theory of rolling contact. So, to describe rolling of wheel over rail in a mathematical model we need the solutions of two problems of contact mechanics:

1. Normal contact problem (contact area, pressure distribution)
2. Tangential contact problem (calculation of creep forces)

At the present time, these problems are usually solved under the assumption that the elasticity of bodies in contact is approximated with the elastic half-spaces. When the bodies are made of the same material then they are elastically symmetric and the two problems are not coupled. Solutions to both contact problems are at present available for most situations of interest of vehicle system dynamics and the most advanced system providing solutions is CONTACT by Kalker [2].

However, for vehicle system dynamics, the situation is greatly complicated by the fact that the normal load at contact and the creepages

are not known beforehand. These are exactly the quantities which we want to predict. One way out of this difficulty is to employ the iterative calculations, starting with a guesstimate on normal loads and creepages. Now, the contact problems have to be solved in each cycle of the iterative calculations for each wheel of the vehicle model. Since geometry usually leads to a non-Hertzian contact, the solutions of contact problems are time consuming and costly. Therefore, the description must be simplified in order to reduce the cpu time and cost to an acceptable level. One way of doing this is to exclude contact problems from iteration of the vehicle/track model.

Exclusion of normal contact problem from iteration of vehicle/track model by equivalent Hertzian ellipse.

Suppose that for prescribed normal load N the contact patch has been calculated with a non-Hertzian method. We know the shape, size and pressure distribution, see Fig.1 as an example.

Now we replace this contact patch with an equivalent ellipse and attribute the properties of Hertzian contact to the ellipticized contact. We infer from properties of the rolling contact theory by Kalker that the most important parameters of contact are those describing its size and shape. The results of simplified theory by Kalker are not strongly sensitive to the distribution of contact pressure if the resulting normal load is the same as for half-ellipsoidal pressure.

For rail vehicle/track interaction models, the main feature of the ellipticized contact should be that the tangential creep forces calculated for equivalent ellipse do not differ drastically from those resulting from non Hertzian rolling. Therefore we hope that the ellipticization retaining size and general shape of contact patch will be reasonable for approximation of the tangential contact problem (problem 2.) with the solution for Hertzian rolling, which is much easier to obtain. The following criterion for ellipticization seems well suited for this purpose.

Let the equivalent ellipse be described by requirements:

$$\frac{a}{b} = \frac{L}{W}, \qquad \pi ab = A, \quad A \text{ - area of contact patch}$$

It follows that the semi-axes of the equivalent ellipse are equal to

$$a=\sqrt{AL/(\pi W)}, \quad b=\sqrt{AW/(\pi L)}$$

The other parameter of the Hertzian contact is its characteristic length ρ. We calculate it with a geometry such that the contact ellipse with semi-axes a,b forms under load N. This geometry has the characteristic length equal to

$$\rho = \frac{4E}{3(1-\sigma^2)N}\left(\frac{A}{\pi n_a n_b}\right)^{3/2}$$

E- Young's modulus, σ- Poisson's ratio, n_a, n_b- tabulated coefficients, functions of ellipse eccentricity, see for example [8].

Now we attribute the Hertzian properties to ellipticized contact. We know from Hertz that the semi-axes a,b are proportional to normal load in power 1/3, so if we want to know semi- axes at load N' we factorize a,b following formulae of Hertz

$$a' = a(N'/N)^{1/3} \qquad b' = b(N'/N)^{1/3}$$

but the shape of the contact ellipse does not change with load, so $\rho' = \rho$. The factorization is usually valid if N' does not drastically differ from load N used for non-Hertzian calculations. When it occurs we have to extend non-Hertzian calculations, see section on multi- dimensional functions. For two- point contact (see Fig.1.b,c) we ellipticize independently each of the two contact zones.

The contact functions

After calculating non-Hertzian normal contact for the same set of y as for the geometrical table, we calculate the a,b of ellipticized contact and merge the results in a common table with entry y and parameter N. For normal load N', not drastically different from N, we are going to use factorization. We define the contact functions by interpolation over the table. If linear interpolation is used then contact functions are linear, piece-wise functions of displacement y. For some applications, especially for two-point contact, it is necessary to extend the functions to more dimensions.

Exclusion of tangential contact problem from iteration of vehicle/track model

After contact between wheel and rail has been replaced with an ellipse the creep forces can be calculated for Hertzian rolling. One way of excluding tangential contact problem from iteration of vehicle/track model is to pretabulate the results for creep forces with the use of some exact method for tangential contact problem. For a regular figure such as ellipse, we parameterize the results using Kalker's parameterization. The normalized tangential forces are functions of creepage parameters and ratio a/b. The tangential contact problem is not solved in each iteration of the vehicle/track model. Instead, we call pretabulated forces. Other approaches using fast tangential force generators for Hertzian rolling include FASTSIM by Kalker, a method of Shen, Hedrick and Elkins or Johnson and Vermeulen.

Use of an approximate method for normal contact

The primary aim of the proposed use of non-Hertzian normal calculations is to determine parameters of equivalent ellipses. However, an equivalent ellipse is a departure from the real situation and does not require exact calculations for normal contact which are both time consuming and costly.

It is obvious that the ellipticization will be valid only for a "quasi-Hertzian" contact and much can be gained on time and expense using less general methods and calculating only quasi-Hertzian contact. The two known methods are due to Reusner [9] and Hung Le The & Knothe [3]. Both methods assume that the pressure distribution in strips in one main direction is semi- elliptical. But even these methods seem to be too slow for the task of preparing the full description of wheel/rail contact for vehicle/track models with the use of multi- dimensional contact functions.

Therefore, to illustrate the proposed approach, we have opted for an even less general but faster method for quasi- Hertzian contact. The idea of such a method is due to Gostling [1]. It differs in assumptions from Reusner's and Hung Le The & Knothe's in that the shape of semi-elliptical pressure is the same in each strip.

Multi-dimensional contact functions

The Hertzian factorization is valid when the load N' is close to load N used for contact calculations. In practice, we can usually use it when $0.7 < (N'/N) < 1.3$. In some situations, especially for curving with two-point contact, the load distribution over two contact zones may lead to a high normal load in the one zone and a low load in the other. In this case we need more normal contact calculations for a set of normal loads. For each load of the set we will have a table of results and the contact functions will have an additional independent variable - the normal load. The other independent variables arise from track widening (narrowing) and rotation of cross-sections of rails. This extension is necessary if we want to describe the contact conditions on flexible track.

Singularities of "rigid" contact geometry

It is common for both measured and nominal profiles that we find singularities in calculated geometrical quantities. They have a form of a sharp increase (jumps) in rolling radii difference and in contact angles. When such a jump is encountered it indicates that a special contact zone would form, as geometrically, we have the two- humped bodies in contact. In fact we will have a two-point contact of the form shown in Fig.1 c. The jumps in geometry lead to difficulties in mathematical models and a method to smooth the singularities has been proposed by Pascal & Sauvage [10]. We propose an approach different but compatible

with ellipticization [4]. With this approach we use properties of equivalent ellipses at jump to predict the range of y near jump where two-point contact takes place. Then, we exclude non-Hertzian results for this range and we calculate the contact functions from remaining results (two non- simultaneous ellipses). *In this way two contact ellipses with centres at points on different rolling radii and with different contact angles are replaced by a single ellipse whose centre is at a point where rolling radius and contact angle have some intermediate values.* A refinement of this prediction is possible, taking into account more details of actual geometry. "Rigid" and smoothed (elastic) contact geometry parameters near the jump for refined prediction are shown on example in Fig.2 and Fig.3.

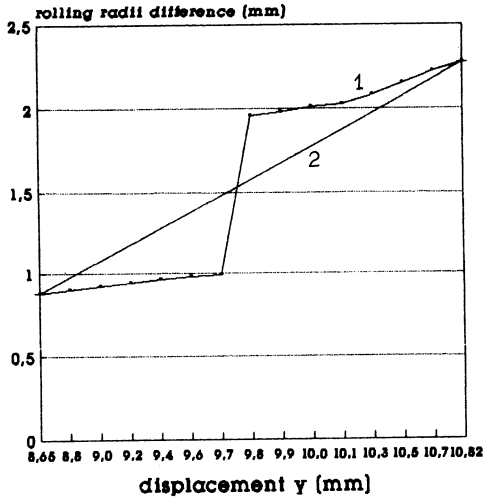


Fig.2 1-"rigid" 2-smoothed (elastic)

Contact loading of a high rail in curve

To describe wheel/rail in a mathematical model we use ellipticized contact (contact functions) allowing to exclude both contact mechanical problems from iteration of vehicle/track interaction model. For genuine two-point contact (as in Fig.1a) we do not smooth geometrical quantities as the distribution of forces and frictional work over two contact zones is technically important. For two-point contact occurring on tread (as in Fig. 1b) we smooth geometrical quantities by replacing two contact zones with one equivalent ellipse.

Using this description we predict with a mathematical model the interaction forces, creepages and co-ordinates describing positions of various elements of the system. Then, we revert to the contact mechanical problem of normal contact and, if necessary, to the tangential problem. To solve contact mechanical problems we can use exa-

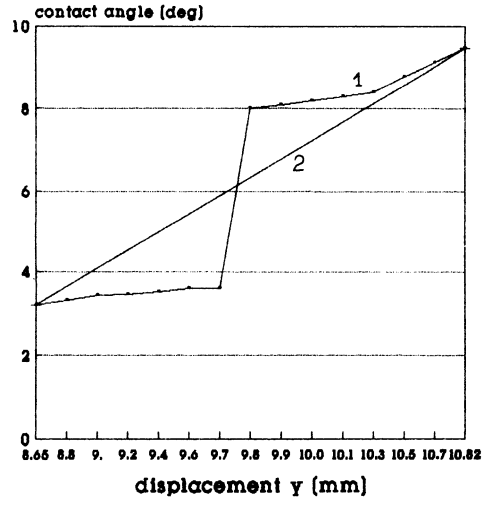


Fig.3. 1-"rigid" 2-smoothed (elastic)

ct methods (CONTACT by Kalker) or some approximate methods, depending on the aim of calculations.

We will illustrate this approach on two examples of quasi- static curving predictions calculated with the curving programs by the author. The main assumption of the model is that the two genuine contact zones are elastically independent so that the load at one contact zone does not influence surface deflections near the other contact zone. This assumption is valid when the distance between two contact zones is higher than the width of the broader contact patch, see also [7].
After calculating normal load between the guiding (front outer) wheel and a high rail, we calculate the contact patches and normal pressure distribution with the approximate non-Hertzian method which was used for ellipticization. Other examples of implementation of this approach to modeling of contact forces can be found in [5,6].

The first example is for two-axle bogie locomotive running with cant deficiency 0 deg on R=1000 m curve. The main dimensions of the vehicle are given in Fig.4. The contact functions are calculated using measured data from poor quality track. Predicted normal load for guiding wheel/high rail is 94.9 kN. The leading wheelset is shifted by -19.05 mm

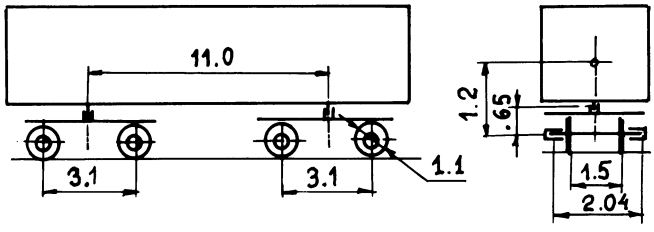


Fig.4 Main dimensions of BoBo locomotive

relative to track centre line, the contact angle is 14.36 deg. Now we calculate normal contact problem for wheelset at -19.05 mm, with a norm-

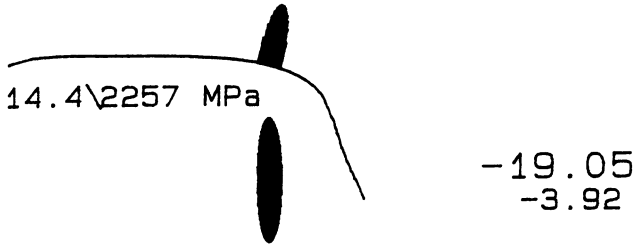


Fig.5. Contact patch and pressure profile on high rail

al load of 94.9 kN. The result is shown in Fig.5. We have one-point contact interaction. The contact patch is nearly elliptical and the profile of pressure distribution is almost semi-elliptical.

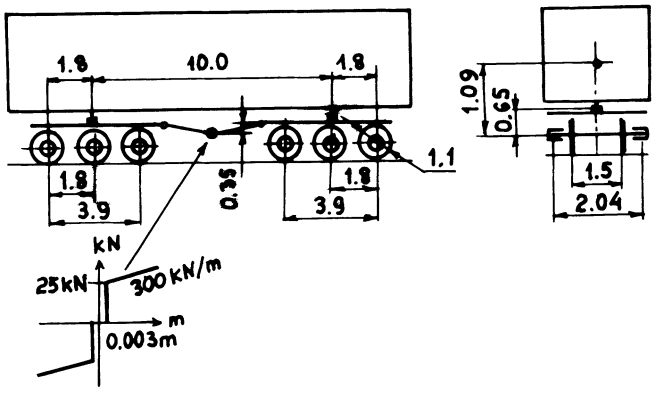


Fig.6 Main dimensions of CoCo loco with transversal coupler

The second example is for a three-axle bogie locomotive with transversal, inter-bogie coupler running with -1.0 deg cant deficiency on R=580 m curve. The main dimensions of locomotive are given in Fig. 6. together with the characteristic of the coupler. The measured data for profiles have been used to calculate the contact functions. Two-point contact on a high rail occurs when the vertical distance between geometrical contact points is 6.84 mm and the contact angle on the rail side is 38.7 deg. This is a "genuine" two-point contact.

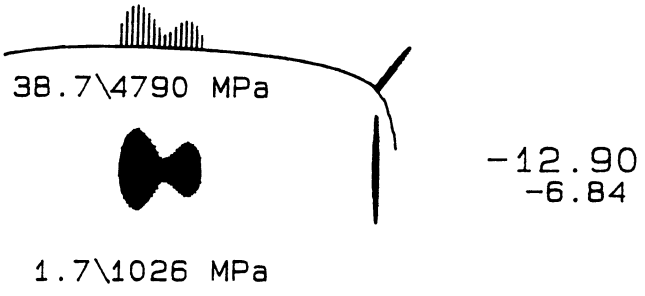


Fig.7 Contact patches and pressure profiles on high rail

Normal loads for guiding wheel from curving prediction are: 34 kN between rail gauge corner side and flange and 62.8 kN between tread and rail head. The contact patches and pressure profiles are shown in Fig.7. Between rail flange and rail side we have an almost elliptical but highly concentrated contact. The rail here has rather sharp corner. Between rail head and wheel tread the contact is less concentrated. In fact we have here another two-point contact which for vehicle/track modeling has been smoothed by replacing it with one equivalent ellipse according to method given in [4].

Conclusions

The author hopes that despite of using of the procedure of ellipticization for the description of contact in vehicle/track model, the discrepancy between predicted and actual interaction forces will be small. The final results of interest for investigation of rail strength and fatigue are the contact patch shape and size and detailed distribution of surface stresses. The quality of this result will depend on the methods used for normal and tangential contact problems. At present, we have a number of methods available starting with the exact methods by Kalker (CONTACT) to less general but nevertheless technically useful approximate methods.

References

1. Gostling R.J., Piotrowski J., Rose K.A.: Comparison of two theoretical methods of calculating contact area between measured wheelsets and track. BR Tech. Memo. TM DA 31, 1981.
2. Kalker J.J.: Three-Dimensional Elastic Bodies in Rolling Contact. Kluwer Academic Publishers, Dordrecht/Boston/London 1990, 316 pp.
3. Knothe K., Le The, Hung: A contribution to calculation of contact stress distribution between two elastic bodies of revolution with non-elliptical contact area. Comp. & Struct. Vol. 10, No. 6, 1984.
4. Piotrowski, J.: Calculation of equivalent Hertzian ellipses with an approximate non-Hertzian method for normal contact problem. 3. Reutlinger und Berlin Arbeitstagung "Mehrkörperdynamik in der Praxis". 18-20th May 1992, Berlin.
5. Piotrowski J.: A theory of wheelset forces for two-point contact between wheel and rail. VSD 11 (1982) pp. 69-87.
6. Piotrowski,J: Contact loading of a high rail in curves. Physical simulation method to investigate shelling. VSD 17(1988) pp. 57-79.
7. Piotrowski, J. Kalker, J.J.: The elastic cross-influence between two quasi-Hertzian contact zones. VSD 17(1988) pp. 337-355.
8. Prochnost, Ustojchivost, Kolebanija. A handbook in three volumes. Edited by Birger I.A. and Panovko J.G. Publ. House "Mashinostrojenije", Moscow 1968. (In Russian)
9. Reusner H.: Druckflachbelastung und Oberflachenverschibung in Waltzkontakt von Rotationskörper. SKF Schweinfurt 1977.
10. Pascal, J.P. and Sauvage, G.: New method for reducing multi- contact wheel/rail problem to one equivalent contact patch. 12th IAVSD Symposium. Lyon, France Aug.26-30 1991.

Experimental Tribo-Analysis of Rail/Wheel Interface

J. Kalousek
National Research Council, Canada

Abstract

Usage of modern experimental techniques and instruments can significantly enhance the understanding of wheel-rail rolling-sliding phenomena. Two novel experimental techniques are described. The first technique utilizes a long distance microscope to directly observe the wheel or rail surfaces in the field. The second technique involves an accurate replication of macro- and microscopic features of worn surfaces and permits their examination by a variety of laboratory instruments. A short pitch rail corrugation study is used to demonstrate the general usefulness of both techniques.

1.0 Introduction

Tribologists all over the world rely heavily on experimental observations to learn about processes involved in wear, friction or lubrication. The success of laboratory experiments often depends on observational procedures. In particular, in experiments in which a third body is involved a lot depends on how the specimens are handled between the end of the wear process or test and the time of the observations. Prof. M. Godet, for example, encourages us not to disengage the contacting pairs till they have been transported from the test rig room to an instrument stage or its chamber(1). Some researchers even go as far as to build their wear apparatuses inside the scanning electron microscopes.

In modern railway operations the understanding of wear, adhesion, lubrication, contact fatigue or plastic flow on wheels and rails can be greatly enhanced by novel and bold experimental methods and techniques. In addition to the dilemma of deciding between laboratory or field tests, the tribologist must carefully weigh the difficulties involved in taking the instruments to the field or bringing the specimens from the field to the laboratory. Even when all possible precautions are taken during the removal of a rail specimen from the track, it usually arrives at the laboratory with its wear band surface features disturbed, if not substantially changed or completely obliterated.

J.J. Kalker et al. (eds.), Rail Quality and Maintenance for Modern Railway Operation, 225–238.

This paper describes two novel experimental techniques adopted at NRC. The first technique attempts to bring a delicate instrument - an optical microscope - into the field. The second is concerned with the delivery of undisturbed surface features, with high macro- and micro-dimensional accuracy, from the field into the laboratory. The latter uses fast setting acrylic resin to replicate the surfaces. Although these techniques are equally applicable to wheel or rail surfaces, our further discussions are limited to rail surfaces only for the sake of brevity. In particular, the usefulness and limitations of both techniques are demonstrated on an example involving an investigation into the causes of short pitch corrugation formation. All the field microscope observations and acrylic replicas originate from Vancouver's SkyTrain mass transit system.

2.0 Direct microscopic observations of rail surfaces.

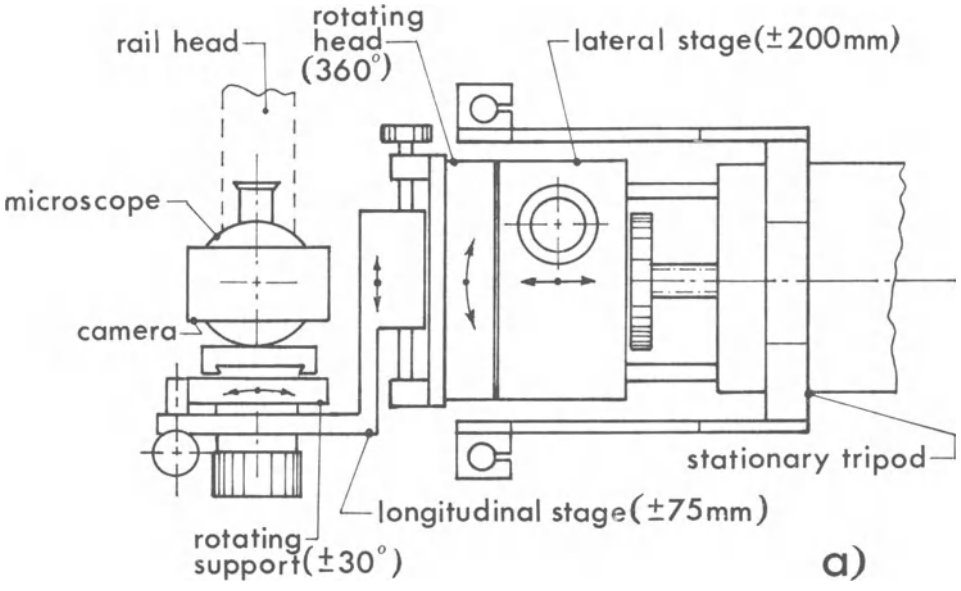
2.1 Description of the field microscope and its support.

Conventional optical microscopes are designed solely for laboratory use. Most of these microscopes have heavy cast iron bodies which eliminate vibrations. The weight involved makes it difficult, for example, to turn a metallurgical microscope upside down and move it with high precision over the surface of the rail. The changes in the vertical displacement needed for focussing would be particularly difficult to control.

A "Questar" long distance microscope is not too heavy and can therefore be utilized for field work. It is based on the design of a telescope but its optics are modified so it can focus at a much shorter distance than infinity. Its field of view is less than 250μm, its resolution is 1.1μm and its working range is from 15 to 35cm. The magnification varies from 50x to 250x depending on the lenses used and the distance from the target. The focussing is internal; this eliminates the need for high precision vertical movement in its support mechanism.

To "walk" the curved surface of the rail or wheel, the microscope must remain perpendicular to the surface at any point of observation. The design of a support stand to permit observation of either wheel or rail surfaces, which at the working distance of 35cm are approximately perpendicular to each other, has posed quite a challenge. Figure 1a shows how all the necessary motions are achieved. The rotating head can place the microscope into vertical or horizontal positions to observe the rail or wheel surface respectively. In addition to longitudinal and lateral motions, a ±30° rotational support stage allows compensation for the curvature of the wheel or rail profiles.

The overall view of the experimental setup, which includes a light source with fibre optic light guides, is shown in figure 1b. As with all other microscopes, any minute vibration during the exposure time reduces the clarity of the pictures. Making use of a long exposure time of about one second, and of a camera body with a partial fixed mirror and electronic shutter is highly recommended.



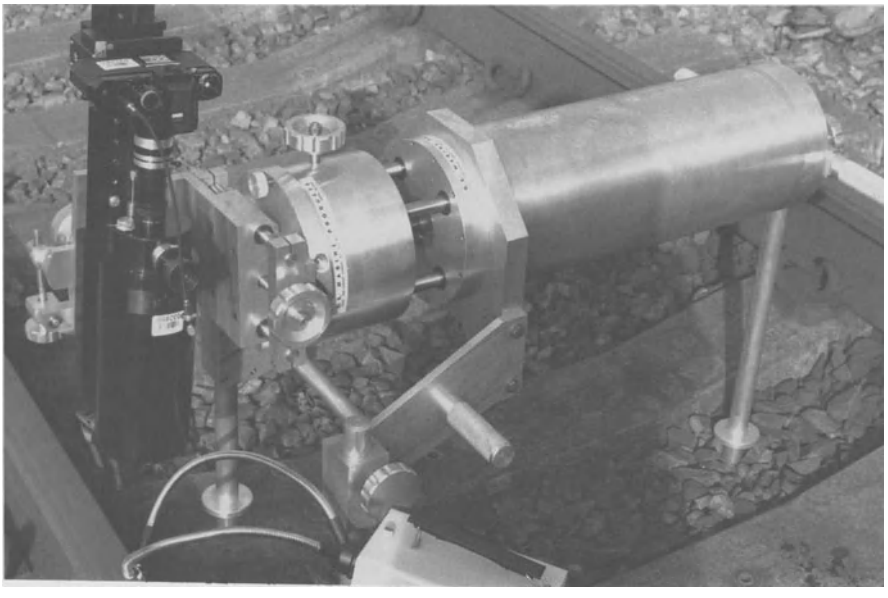


Figure 1: Field microscope with support stand; a) schematic sketch - plan view, b) photo of the complete instrumental set-up.

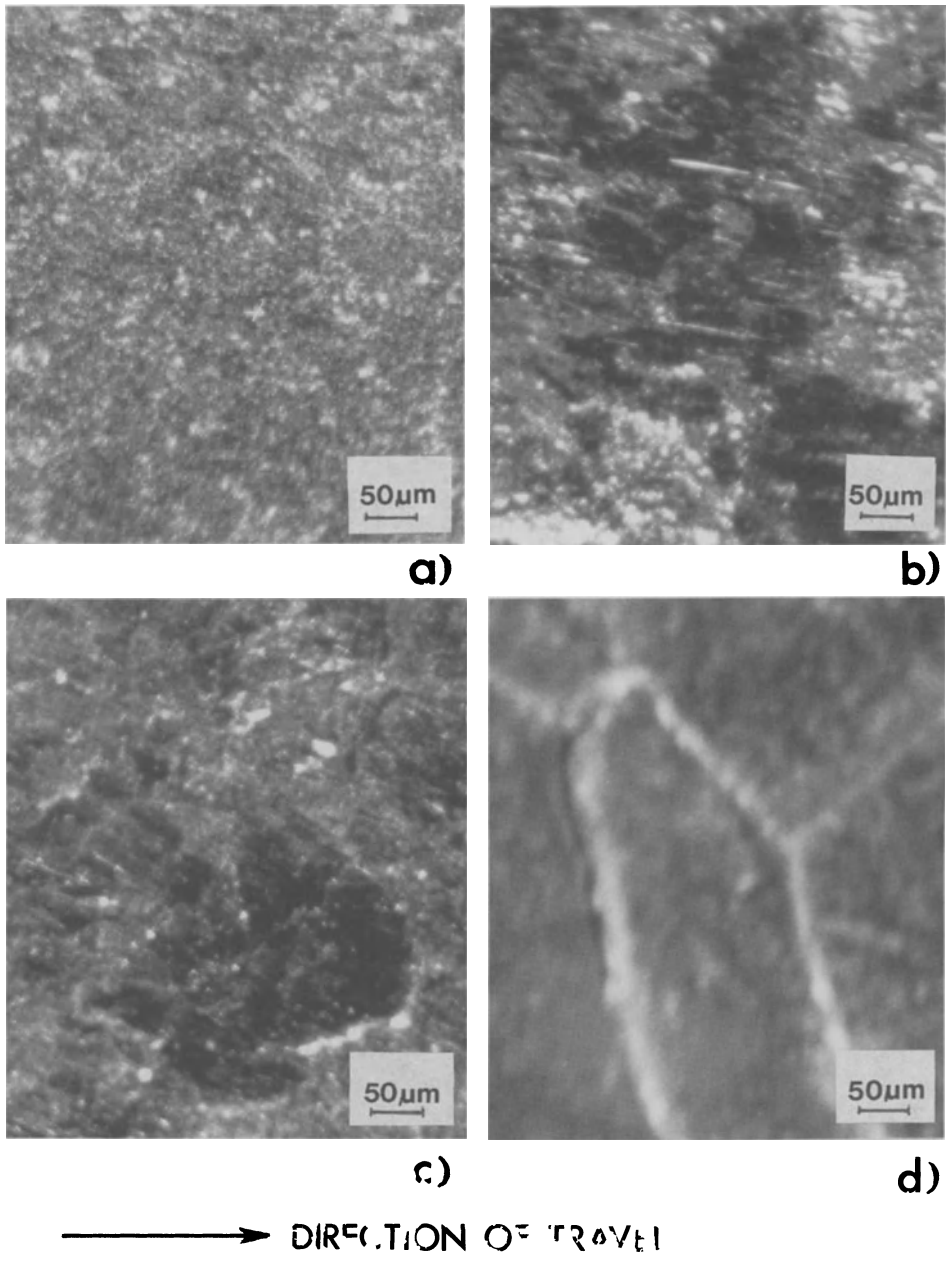


Figure 2: Field optical micrographs of a) uncorrugated rail and b), c) and d) corrugated rail. b) and c) valley at the field side and center of the rail respectively, d) peak at the center of the rail.

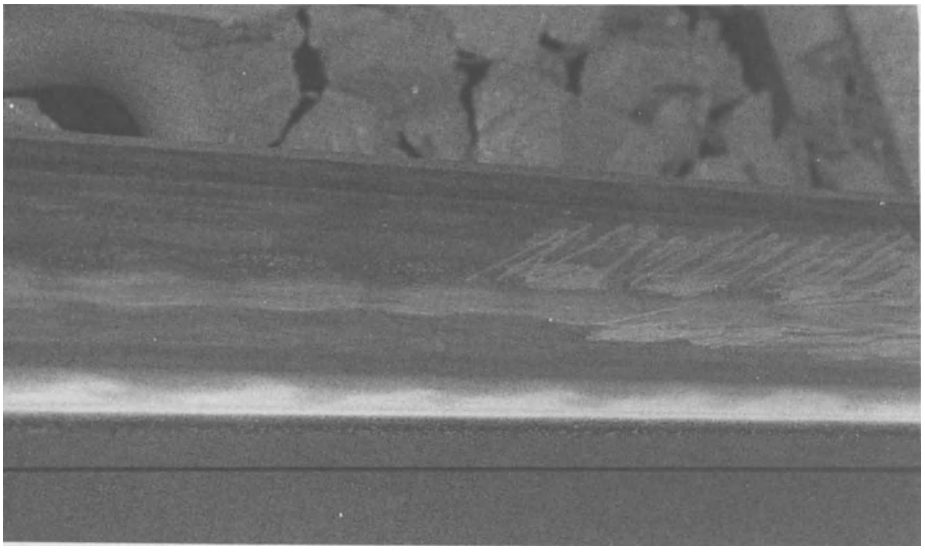
2.2 Observation of corrugated rail surface

Looking through the long distance microscope convinces any observer that the entire wear band is covered by fine powder-like debris. The consistency of this debris, however, is not the same on uncorrugated and corrugated rail. Some of the observed changes in the debris coverage are illustrated in figure 2. Figure 2a shows a rail surface which is not corrugated. The layer of debris seems to be relatively undisturbed, and to be of uniform texture and possibly of uniform thickness as well.

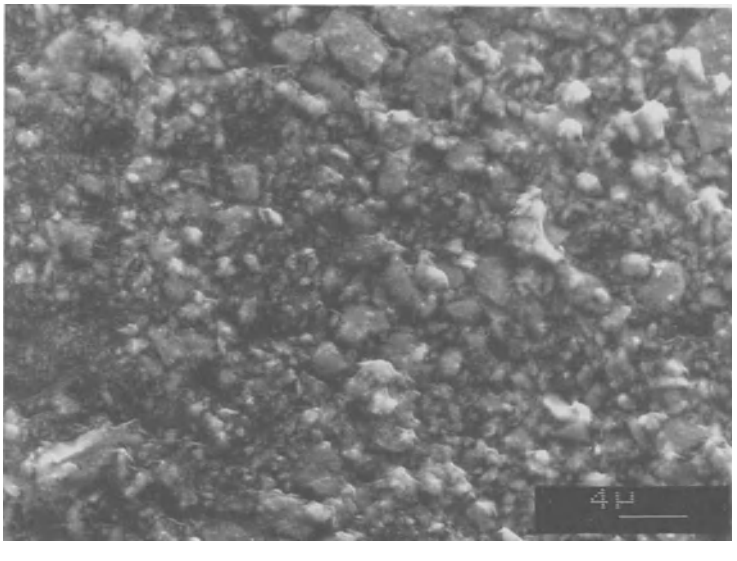
The situation is different when the rail is corrugated. Examples of the surface texture of a mildly corrugated high rail located in a 100m radius curve are shown in figure 2b,c,d. The appearance of the debris layer changes from location to location. At the corrugation peak the layer appears uniform but is cracked as shown in figure 2d. This implies, that the layer is not smeared or disturbed by passing wheels at least for the period needed for cracks to develop. Elsewhere, within the corrugated wear band, the debris is disturbed. At the field side of rail, the debris is smeared in a nearly longitudinal direction as shown in figure 2b. In the middle of the corrugation valley, whole areas of debris layer are missing. An example of such an area is the dark spot in figure 2c. Because this dark spot has a sharp edge and a smeared edge on its right and left hand sides respectively, the debris originally located in this area has been displaced to the left. These observations suggest that the debris is gradually moved over the rail surface by the passing wheels from the corrugation valleys to the corrugation peaks.

The microscopic observation showed that the debris is not too firmly attached to rail surface. Analysis of the replicas described later (see figure 8) showed that debris can be compacted into layers several times thicker than its mean size. It, therefore, should not be surprising that some of it can be scraped off as shown in figure 3a. Lacking image analysis and spectroscopic chemical analysis equipment it was not possible to determine the exact particle size distribution or chemical composition of this debris. Nevertheless, it can be seen from figure 3b that particle size ranges from a fraction of a micrometer to several micrometers. It has been shown by a number of researchers that the debris contains several kinds of iron oxides; however, its bulk is magnetite, Fe_3O_4 (2,3).

Observation of surface debris in the field has shown that some of it is always loose. This strongly supports the assumption that it is mobile and can thus intermix under the action of passing wheels. Therefore, by introducing other solid lubricants into this layer, such as by film transfer from a stick, the shear strength, thickness and frictional properties of the iron oxide debris layer can be modified.



a)



b)

Figure 3: a) Photo of corrugated rail with portion of debris layer scraped off
b) scanning electron micrograph of the scraped debris.

3.0 Observations of replicated surfaces in the laboratory.

3.1 Description of replication technique

Taking an acrylic resin casting or replica is simple and takes about half an hour to complete. In addition to fast curing acrylic resin (dental acrylic is cheap and easily available) one needs aluminum strips of various sizes, plasticine, a propane torch, a utility knife, a hammer, solvent and rags, a scriber and a track level bar. The track level bar, which has two straight edges at each end, can be made from any suitable material. It is used to inscribe a line parallel to the track level on both sides of each replica.

To produce a high quality replica of good dimensional accuracy that is free of bubbles it is recommended to heat the rail with a propane torch to about 40°C. When the rail is too hot, at a temperature of 50°C or more, the replica becomes dimensionally distorted. To avoid unwanted scratches on the replicated surface, cleaning of the wear band is not recommended. The excess dirt, rust or grease on sides of the rail, however, should be cleaned with solvent so that the plasticine will stick at these locations. Using aluminum strips and plasticine, a "pan" is constructed around the selected portion of the wear band.

Mixing of the resin powder with the hardener produces a viscous liquid which is poured into the pan to a level of about 1cm above the highest point of the rail. Following the exothermic reaction and the hardening of the replica, the plasticine and aluminum strips are removed. Any flared out plastic at the vicinity of the replicated surface is cut off with the utility knife. The hardened and cooled down replica is scribed with track level lines and marked for identification. Tapping lightly on the sides of the replica with the hammer "peels off" the replica from the rail surface.

3.2 Evaluation of corrugated rail from replica

Photographs of replicas taken from unground and ground corrugated rails are shown in figure 4a,b. Using a variety of laboratory instruments a wealth of knowledge can be derived from such replicas. This includes:

a) transverse rail head profile and its orientation with respect to track level

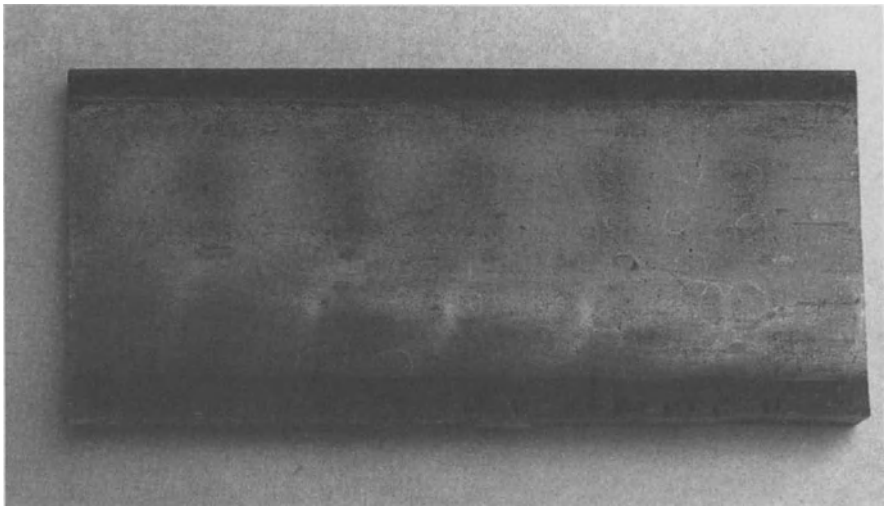
b) longitudinal topography of corrugated rail surface

c) mean magnitude and direction of slip at locations where slip marks are observed

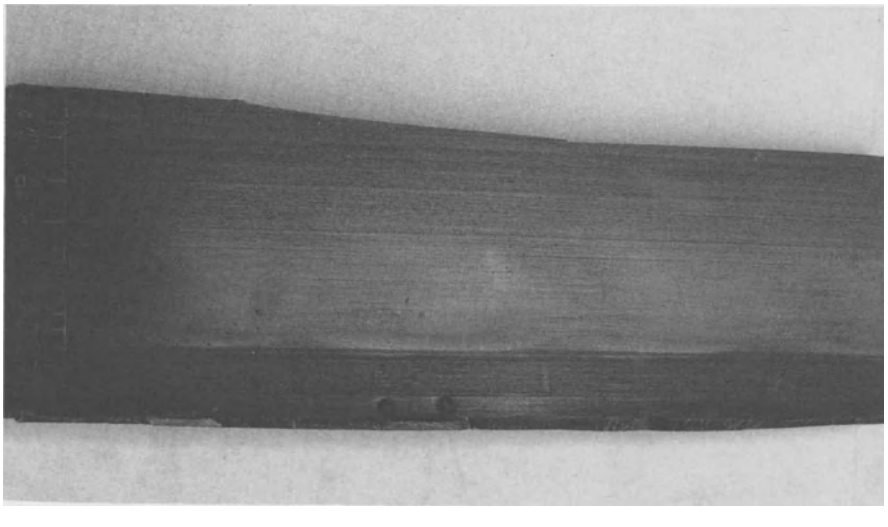
d) surface roughness at any location

e) depth of the debris layer at locations where a portion of such a layer has delaminated.

The information in a) and b) is obtained with a surface profilometer and the one in c) with an optical microscope. A surface roughness measuring instrument yields d) and if it has computer-aided surface mapping equipment, e) is obtained.



a)



b)

Figure 4: Photos of the surface replica taken from
a) unground and corrugated
b) ground and corrugated rail

3.2.1 Using surface profilometer

Surface profilometers are usually mechatronic, stylus type instruments. As the stylus is dragged over the surface of the replica the x-y coordinates of the stylus tip are recorded electronically or can be plotted by a plotter at several stepwise magnifications from 1x to 200x. The accuracy of our "Mitutoyo" profilometer is 0.01mm in both directions.

Many replicas of corrugated wheels and rails were taken from the Vancouver's SkyTrain mass transit system. Numerous superpositions of wheel and rail transverse profiles which were magnified ten times, revealed that these profiles are tightly conformal to each other. In some instances, a minute gap is observed at the centre as shown in figure 5. In this figure, the size of the gap is exaggerated in order to avoid the loss of definition when reducing from the large to the small scale. Depending on the exact value of such a gap, the wheel could roll over the rail either supported simultaneously by two distinct contact areas at the sides of the wear band or by a single contact area across the entire wear band with a non-Hertzian pressure distribution.

Figure 6 shows the macroscopic topography of the replica shown in figure 4a. The replica was taken from a fully corrugated low rail in a 60m radius curve in the overnight storage yard. The transverse rail profile and points at which the longitudinal traces were taken are shown at the extreme left of the figure. The depth of the corrugation troughs at the gauge side of the wear band are up to 130μm deep while at the field side they are approximately 25μm; Therefore, there is a nearly five fold difference in the corrugation depth at these two locations. Also noteworthy is the 90° phase shift observed between the peaks of the traces on the gauge and field sides, i.e. traces 1 to 7.

3.2.2 Using optical microscope with polarized light.

The right hand side of figure 6 shows a map of the microscopic slip vectors. These vectors represent the average length and direction of scratch marks observed with an optical microscope under polarized light. None of these scratch marks would be visible if observed with white light or with a scanning electron microscope. Four examples of the scratch marks are shown in figure 7. The first two, figure 7a,b, originate from the vicinity of the corrugation peak, and the remaining two from the corrugation valleys. The locations of these micrographs are cross-referenced by letters a to d in figure 6.

The scratch marks located within the area of the traces 1 to 4, which are deep, point to the left. The marks in the vicinity of traces 5 to 7, which are shallow, point in the opposite direction. Bearing in mind the conicity of the rail wheel interface, the directions of the observed scratch marks are consistent with the distribution of slip due to the spin in the contact ellipse which spans the entire width of the wear band. In a pure spin situation, the rail surface moves more slowly at the gauge half and more quickly at the field half of the contact area with respect to the wheel surface. Thus the observed difference in the depth of corrugations is consistent with the known fact that in combined rolling and sliding contact the wear rate is greater on the slower moving surface (4).

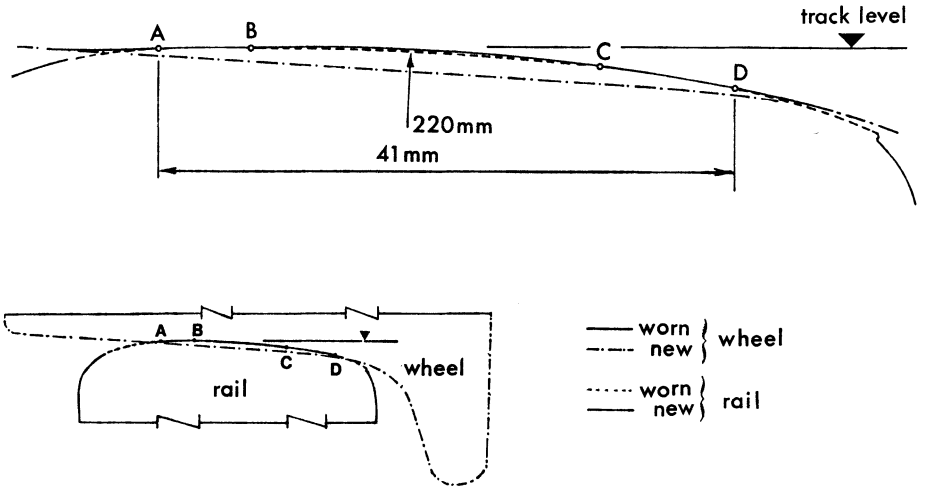


Figure 5: Superposition of worn wheel profile to rail profile taken from unground corrugated rail

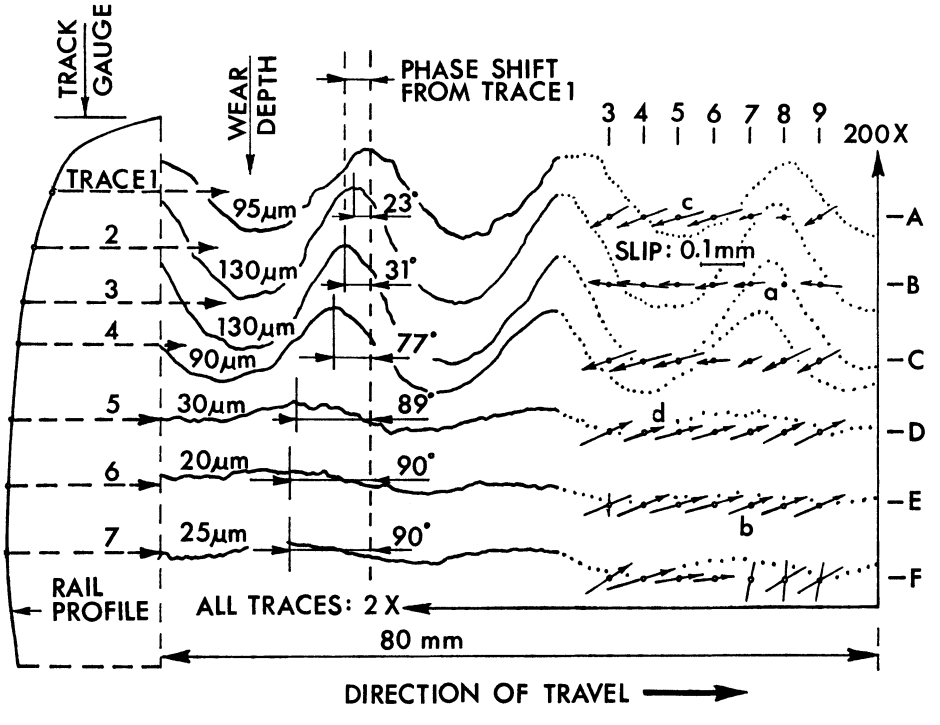


Figure 6: Topography of fully developed corrugations with a map of slip vectors.

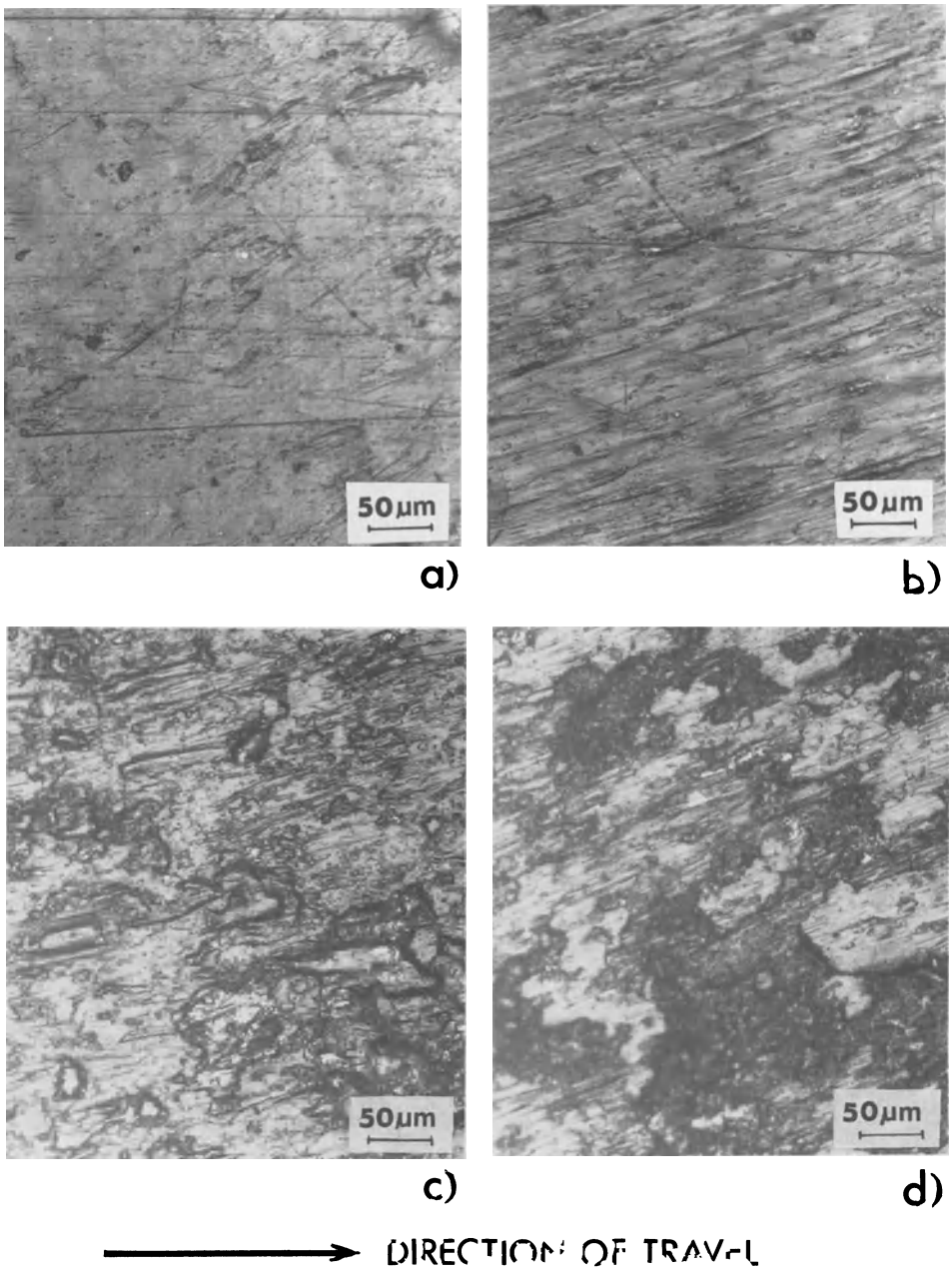


Figure 7: Sample of optical micrographs used in construction of slip vector map shown in Figure 6; Micrographs a), b), c) and d) were taken at positions 8B, 7E, 5A and 4D respectively.

3.2.3 Using computer-aided surface mapping

When searching the surface of a replica for scratch marks with the optical microscope, one occasionally finds, a microscopic protrusion with a flat top. These protrusions are only found in the field half of the wear band. One such protrusion can be seen in the lower right hand corner of a computer aided surface map shown in figure 8a. Figure 8b shows the optical micrograph of the same protrusion with the focus at its top, where distinct longitudinal scratch marks can be seen. No scratch marks are present at the base of the protrusion, as can be seen in figure 8c. The height of this particular protrusion is about 17μm. The height of other protrusions varied from 5μm to about 30μm in the vicinity of corrugation valley and peak respectively.

The protrusion on the replica is actually a depression in the surface it replicated. The depression or hole can be seen as such by viewing the surface map upside down. These holes are most likely formed by small segments of the cracked surface layer which adhere to the wheel and are lifted up. Therefore, the actual rail surface is at the bottom of these holes. Since scratch marks are also visible at the bottom, it can be assumed that the entire layer slips over the surface to which it is attached. This complements the observations made by the field optical microscope and shows that in addition to debris being moved within the layer, the entire layer can be displaced and incrementally move with respect to the rail surface.

4.0 Concluding remarks

The field microscope and surface replication experimental techniques are extremely useful in helping to understand the causes of numerous wheel/rail problems associated with wear, lubrication, rail spalling, wheel shelling and others. In the studies of short pitch corrugations on the Vancouver's SkyTrain mass transit system, these techniques contributed to the solution of the short pitch corrugation and the associated noise problem in a number of areas.

Firstly, the observation of the debris layer and its thickness by the above techniques led to the assumption of debris mobility and of the intermixing of the debris within the layer. These observations inspired the development of the high positive friction (HPF) solid lubricant (5). This lubricant reduces friction when it is too high, but maintains an adequate friction for the braking and adhesion purposes particularly under wet conditions. It also modifies the friction characteristic of the debris from negative to positive to alleviate the excitation of frictionally induced quasi-harmonic (roll-slip) oscillations.

The high accuracy superpositions of ten times magnified profiles, made possible by the replication technique, identified close conformity of the worn wheel and rail profiles. Close conformity gives high effective conicity and can cause dynamic instability of vehicles (6). Spin creepage in closely conformal contact yields strong equal and opposite slip and, under dry conditions, high and opposite friction forces between gauge and field sides of the wheel/rail interface. Elimination of close conformity and associated reduction in spin creepage can be used to eliminate the short pitch corrugations and reduce noise (7) wherever close conformity can be identified as a source of these problems.

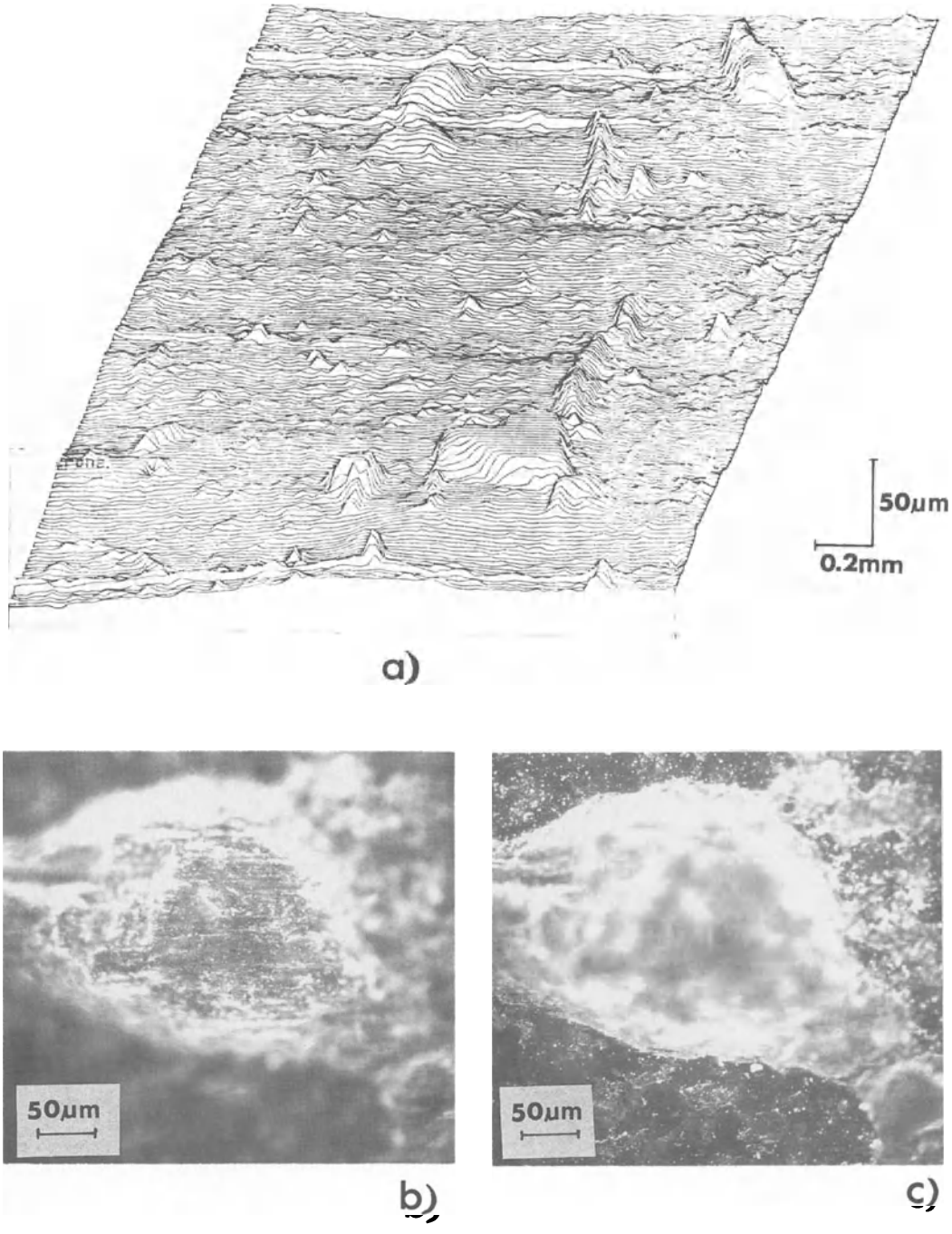


Figure 8: Delaminated surface layer
a) at the right-hand lower corner of computer aided surface map
b) optical micrograph focussed at the bottom of the layer (i.e. top of the replica)
c) optical micrograph focussed at the top of the layer.

Some of the findings these techniques can yield, however, may not always be immediately understood. For example, the significance of the 90° phase shift between gauge and field side corrugation traces and that of the unequal thickness of the debris layer are not clear. It is hoped that the significance of such observed but insufficiently understood phenomena can be explained by theoretical analysis.

References

1. Godet, M.: "The Third-Body Approach: A Mechanical View of Wear", WEAR, 100(1984), 437-452.

2. Broster, M., Prichard, C. and Smith, D.A.: "Wheel/Rail Adhesion: Its Relation to Rail Contamination on British Railway", WEAR, 29(1974), 309-321.

3. Beagley, T.M.: "The Rheological Properties of Solid Rail Contaminants and Their Effect on Wheel/Rail Adhesion", Proc. Instn. Mech. Engrs., 190, 39(1976) 420-428.

4. Milne, A.A. and McDonald, S.D. "Institute of Mechanical Engineers Conference on Lubrication and Wear", London, 1957, 740-749.

5. Holdcroft, J. and Townsend, .F, " The Development of Tribological Solutions to Alleviate Problems, Associated With Wear, Corrugation Formation, and Noise Generation on Railway System Track", Proc. Lubrication on the Railway Interface, Canfield Campus, Monash University, Australia, 1991, 136-173.

6. Wickens, A.H.: "The dynamic stability of railway vehicle wheelset and bogies having profiled wheels" Int. J. Solid Struct., 1(1965), 319-341.

7. Kalousek, J. and Johnson, K.L., " An Investigation of Short Pitch Wheel & Rail Corrugations on the Vancouver Mass Transit System, to appear in Proc. Inst. Mech. Engrs., London, Part F, Jl. of Rail and Rapid Transit, vol. 206.

Acknowledgements

The cooperation and support in the form of labour and equipment as well as financial support by the British Columbia Rapid Transit Company (Skytrain) and its staff is greatfully acknowledged. The author is most indebted to Mr. Johan G. Vigen for his input into the design of the field microscope stand.

Shakedown, Residual Stress and Plastic Flow in Repeated Wheel–Rail Contact

A.F. Bower† and K.L. Johnson*

†Brown University, Division of Engineering, Providence, RI 02912 U.S.A.
*Cambridge University Engineering Department, Cambridge, CB2 1PZ U.K.

Abstract

We review the results of recent theoretical investigations of plastic flow caused by repeated rolling and sliding contact. A hardening law is descibed which is capable of predicting the behaviour of rail steel under the cycles of stress caused by contact loading. The hardening law is used to calculate the loads required to cause repeated plastic flow under both line and point contacts. The calculation also leads to an estimate of the magnitude of the residual stresses induced by the deformation, which is compared with experimental measurements. Finally, approximate methods are used to analyse the plastic flow which occurs when the loads exceed the shakedown limit.

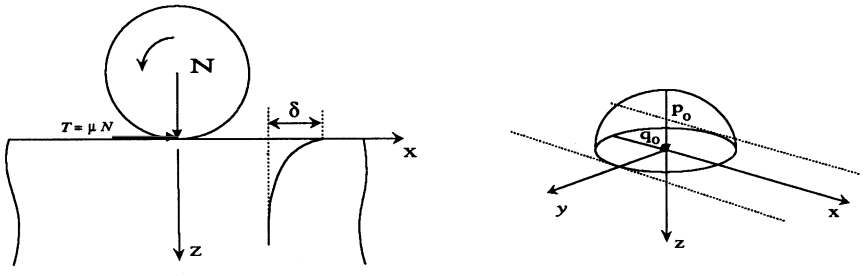
1. Introduction

Railway lines are among the few structures that are routinely loaded above the elastic limit during service. In heavy haul lines, the wheel loads can be so severe that the rails fail due to excessive plastic deformation. It is unusual for high–speed lines to fail by plastic flow. Nevertheless, a close examination of track taken from high speed lines almost invariably reveals some evidence of plastic deformation near the surface of the rail head. The most common type of deformation is illustrated in Figure 1(a): a thin layer of material near the surface is sheared relative to the bulk material. The deformed zone is typically 1–2mm deep, and the strains in the surface layer can exceed 100%. The deformation is driven by the high normal and tangential wheel loads acting on the surface of the rail, and is an example of incremental plastic flow, or "ratchetting:" each successive pass of a wheel over the track causes a small increment in the plastic strains, which continue to accumulate during many thousands of cycles. This plastic flow has a number of consequences. It introduces residual stresses in the rail, which may affect the growth of fatigue cracks. In addition, there is experimental evidence to suggest that the deformation may be responsible for sliding wear, and may also initiate rolling contact fatigue cracks.

Modelling the deformation caused by contact loading is an important step towards developing models capable of predicting the mechanisms of failure in railway track. In this paper, we summarize the results of recent theoretical investigations of plastic flow caused by rolling and sliding contact. The calculations have shown that the plastic deformation is strongly influenced by the strain hardening behaviour of the deforming material. We begin, therefore, by describing a nonlinear kinematic hardening law, which is capable of modelling the behaviour of rail steel under the cycles of stress associated with contact loading. The hardening law is

J.J. Kalker et al. (eds.), Rail Quality and Maintenance for Modern Railway Operation, 239–249.

used as a basis for a theory of elastic/plastic rolling and sliding contact. The loads required to cause repeated plastic flow are found. This calculation also leads to an estimate of the residual stresses introduced by the deformation, which are compared with experimental measurements of residual stresses in a rail head. Finally, we investigate the nature of the deformation which occurs when the loads exceed the critical value required to cause plastic deformation.



(a) Line contact (b) Point contact

Fig. 1: Notation and sign convention for contact geometry.

For simplicity, we will consider only the contact geometries illustrated in Fig. 1: the contact area may be a line, such as occurs between a cylinder and a half–plane, or a circle, which occurs when a sphere is pressed into a flat surface. In addition, we will assume that full slip occurs between the contacting surfaces, and that the tractive force either acts in the direction of motion of the load, such as occurs due to braking, or else is opposite to the direction of motion (this corresponds to the driving wheel of a locomotive). The effects of partial slip and lateral traction will be neglected: they are discussed in more detail in Bower and Johnson (1991).

2. Elastic Limit

It is instructive to begin by reviewing the elastic stress field caused by contact loading. In this case, the contact pressure distribution is Hertzian, so that the normal and tangential tractions p and q acting on the area of contact are

$$p = p_0\sqrt{1-(x/a)^2} \qquad q = \mu p_0\sqrt{1-(x/a)^2} \tag{1}$$

for line contact, or

$$p = p_0\sqrt{1-(x/a)^2-(y/a)^2} \qquad q = \mu p_0\sqrt{1-(x/a)^2-(y/a)^2} \tag{2}$$

for point contact, where μ is the friction coefficient, a is the radius of the contact area and p_0 is the peak Hertzian pressure. The critical value of p_0 required to cause yield may be calculated using expressions for the subsurface stresses given in Johnson (1985) and Hamilton (1983). The elastic limit is shown as a function of friction coefficient in Fig. 3. The results demonstrate an important feature of

contact loading: if the coefficient of friction $\mu < 0.3$, the critically stressed material lies below the surface. Under high tractive loads ($\mu > 0.3$), yield first occurs at the surface.

If the elastic limit is exceeded, some plastic deformation must take place under the first few passes of the load. However, repeated plastic flow will not necessarily occur, for three reaons. Firstly, residual stresses are introduced by the deformation which act so as to prevent further plastic flow. Secondly, the material may strain harden. Finally, in the case of a point contact, a groove may be formed in the surface which tends to attenuate the contact stresses. If the load lies below the so called 'shakedown limit,' these processes may prevent plastic deformation, so that in the steady state the material is subjected to an elastic cycle of strain. If the load exceeds the shakedown limit, repeated plastic flow will occur.

Shakedown due to the formation of a groove has been investigated by Kapoor and Johnson (1992). However, the effects of grooving are significant only when rolling takes place repeatedly along the same track. This is not the case in wheel–rail contacts and in consequence the wheels and rails develop fairly standard non–conforming worn profiles. Provided these profiles are used in the calculation of the contact stress, the effects of grooving can be neglected.

To account for the effects of strain hardening, it is necessary to choose a hardening law which is capable of predicting the response of the deforming material to the stress cycles associated with rolling and sliding contact. The merits of various possible choices are discussed by Bower and Johnson (1989). In our analysis, we have adopted a simple nonlinear kinematic hardening law. The advantages of this model are that it can predict the accumulation of plastic flow observed in cyclic tension/compression and torsion tests on rail steel, and is straightforward to use. The hardening law is outlined briefly in the next section.

3. A cyclic hardening law for rail steel

The hardening law described here is similar to simple linear kinematic hardening. We assume that the yield stress for the material is defined by a von Mises yield surface which is free to move in stress space but cannot change shape or size. The yield criterion is thus given by

$$F = \left\{ \frac{1}{2}(S_{ij} - X_{ij})(S_{ij} - X_{ij}) \right\}^{1/2} - k_e = 0 \tag{3}$$

where k_e is the initial yield stress of the material in shear, S_{ij} are the components of the deviatoric stress, $S_{ij} = \sigma_{ij} - \delta_{ij}\sigma_{kk}/3$, and X_{ij} are the components of the centre of the yield locus in stress space. Plastic flow follows the normality rule

$$d\varepsilon_{ij} = d\lambda \frac{\partial F}{\partial S_{ij}} = \frac{\sqrt{3}}{2} d\lambda \frac{S_{ij} - X_{ij}}{k_e} \tag{4}$$

where $d\varepsilon_{ij}$ is an increment in plastic strain and $d\lambda$ is the magnitude of the plastic strain increment

$$d\lambda = \left\{ \frac{2}{3} d\varepsilon_{ij} d\varepsilon_{ij} \right\}^{1/2} \tag{5}$$

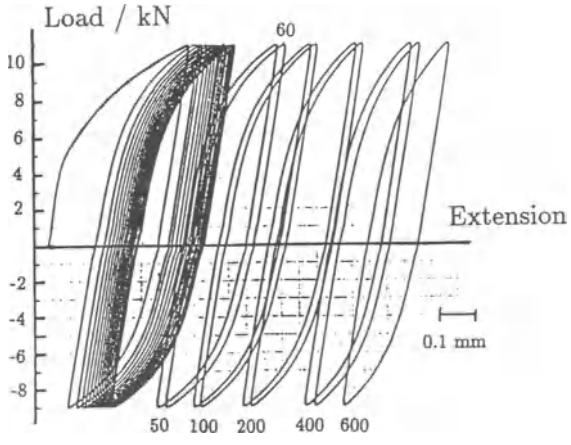


Fig. 2: The behaviour of of rail steel under load controlled tension / compression.

The hardening characteristics of the material are determined by a relationship between the plastic strains and the motion of the yield surface. Here we assume that

$$dX_{ij} = \frac{2}{3} c d\varepsilon_{ij} - \gamma_1 (X_{ij} - Y_{ij}) d\lambda \tag{6}$$

and

$$dY_{ij} = \gamma_2 (X_{ij} - Y_{ij}) d\lambda \tag{7}$$

where c, γ_1 and γ_2 are material constants.

The constants c, γ_1 and γ_2 in this model determine the behaviour of the material under cyclic loading. The parameter c governs the initial slope of the plastic stress–strain curve, while γ_1 and γ_2 control the ratchetting behaviour. For the particular case $\gamma_2 = 0$, the material accumulates strain at a steady rate. If both γ_2 and γ_1 are nonzero, the ratchetting rate decreases with continuing cycling and eventually settles to a closed cycle of strain.

The parameters in the hardening law were determined from the results of the tension/compression test shown in Fig. 2. The figure shows the response of a solid cylindrical specimen of rail steel, initial length 38mm and diameter 4.76mm to a load controlled cycle of tension and compression with a non–zero mean stress. It can be seen that the material accumulates strain in the direction of mean stress, but the ratchetting rate decreases with time. The material constants were chosen so as to predict correctly the yield stress, the amplitude of the cyclic plastic strain, and the total accumulated strain after 100 and 400 cycles. The measured values of the parameters are given in Table 1.

We have compared the predictions of this hardening law with experimental measurements of the behaviour of BS11A rail steel under cycles intended to be representative of contact loading. The predictions of the model have been shown to predict accurately the behaviour of the material under both the proportional

Shear modulus G	80.7 GNm^{-2}
Yield stress in shear k_e	231 MNm^{-2}
Hardening rate c	33.9 GNm^{-2}
Feedback rate γ_1	8.3
Feedback rate γ_2	0.41

Table 1: Hardening parameters for rail steel

cycles of stress associated with subsurface plastic deformation, and also the non-proportional cycle which occurs when material at the surface is deformed by high tractive loads (Bower, 1989; Bower and Johnson, 1991).

4. Shakedown Limits and Residual Stresses

It is straightforward to calculate the loads required to cause repeated plastic deformation in a kinematically hardening material. Since the yield locus is free to move in stress space, the solid can always strain harden to an elastic state if the stress cycle can be enclosed within the yield locus. Repeated plastic flow takes place if the amplitude of the stress cycle exceeds twice the initial yield stress of the material. A graphical method of calculating the shakedown limit has been presented by Johnson (1990).

Shakedown limits for a kinematically hardening solid under line and point contact loading have been calculated by Hearle (1984) and by Johnson (1987). The results are shown in Fig. 3. The load is specified by the ratio of the peak Hertzian stress to the initial yield stress of the material in shear p_0/k_e, and by the traction coefficient μ. The figure shows the behaviour of the material under the contact for any given load. If the load lies below the shakedown limit, then in the steady–state, the material will always follow an elastic cycle of strain. If the load exceeds the shakedown limit, then repeated plastic deformation will take place, and the nature of the deformation is determined by the magnitude of the tractive load acting on the surface. Under low tractive loads, the deforming material is confined below the surface, so that the maximum plastic strain occurs at a depth of about $0.5a$ below the surface. If high tractive loads act on the solid, then the critically stressed material lies at the surface, so that a thin layer of material at the surface is deformed. Under particularly high loads, both surface and subsurface material may be deformed, but this situation rarely occurs in practice and will not be considered further here.

It is instructive to compare the loads acting on a typical rail head with the shakedown limits shown in Fig. 3. On heavy haul lines, the magnitude of the peak Hertzian contact pressure is of the order of 1300 MNm^{-2} (Bhargava et al 1987). The coefficient of friction may vary between $\mu = 0.18$ if the rail head is lubricated, to $\mu = 0.6$ on a dry rail. Using the value of k_e given in Table 1 shows that $p_0/k_e \approx 5.6$ in this case. The contact pressure on a heavy haul line is thus sufficient to cause repeated subsurface plastic deformation in the rail, and under high tractive loads it may cause surface deformation as well. It is perhaps not surprising that subsurface initiated contact fatigue cracks and excessive plastic flow are a cause of failure in heavy haul lines.

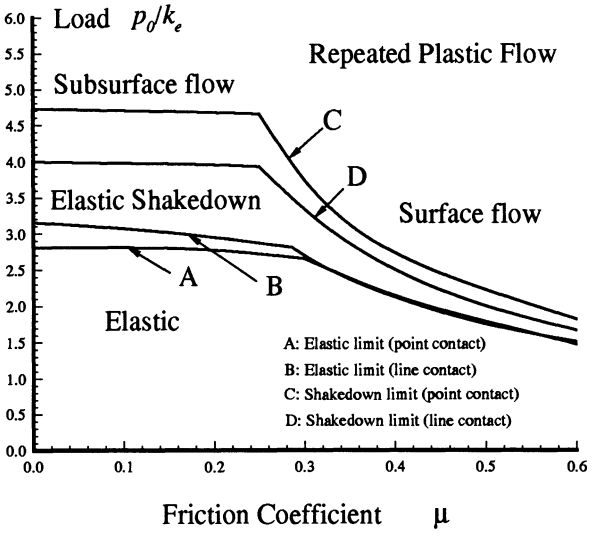


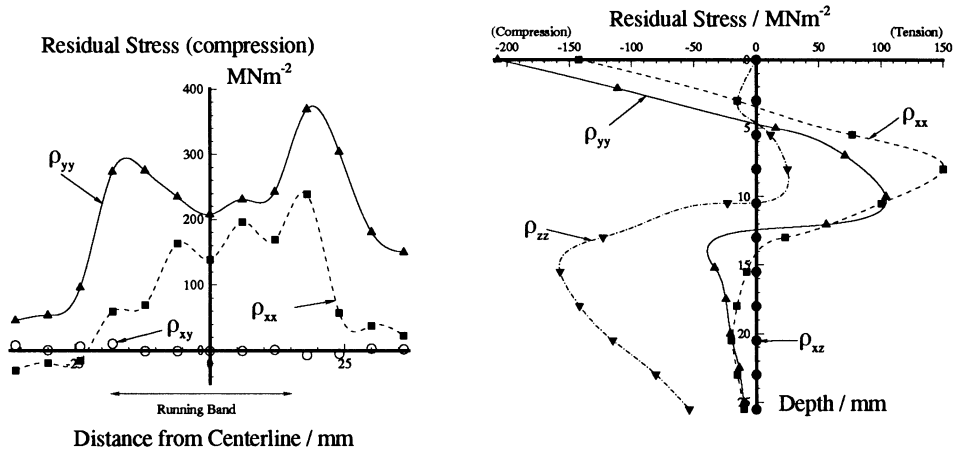
Fig. 3: Elastic and shakedown limits under point and line contact loading.

The contact pressure on high speed lines is less severe: $p_0 \approx 900$ MNm^{-2} is a representative value, which gives $p_0/k_e \approx 4.2$. Provided the friction coefficient $\mu < 0.28$, this is below the shakedown limit. Repeated subsurface flow is not likely to occur in high speed lines. However, under dry conditions the friction coefficient can exceed this value, which results in repeated plastic flow at the surface of the rail head. This may go some way towards explaining the predominance of surface initiated rolling contact fatigue cracks in high speed lines.

The shakedown calculations also give estimates of the residual stresses introduced by contact loading. In a kinematically hardening solid, shakedown occurs due to the combined effects of strain hardening and residual stresses. The combination of residual stress and hardening at shakedown is not unique, but the two contributions may be separated if we assume that residual stresses in the solid are identical to those in an elastic–perfectly plastic solid at shakedown. The reasoning leading to this assumption is that the magnitude of the plastic strains required to set up the residual stress distribution are smaller than the strains required to produce the equivalent amount of strain hardening.

The residual stresses at shakedown in an elastic–perfectly plastic solid subjected to line contact are given in Johnson (1985). In this case, only two components of residual stress ρ_{xx} and ρ_{yy} can be set up by the deformation. Under conditions of pure rolling, the magnitudes of the residual stresses at the critically stressed depth $z = 0.5a$ beneath the surface are $\rho_{xx} = -0.536k_e$, $\rho_{yy} = -0.852k_e$, where k_e is the yield stress in shear of the solid. In a rail head, this corresponds to compressive longitudinal and lateral stresses of the order -120 MNm^{-2} and -200 MNm^{-2}, respectively.

More recently, Hearle (1984) and Ponter, Hearle and Johnson (1985) have investigated shakedown of an elastic perfectly plastic material subjected to a point



(a) Variation of stress across the rail surface

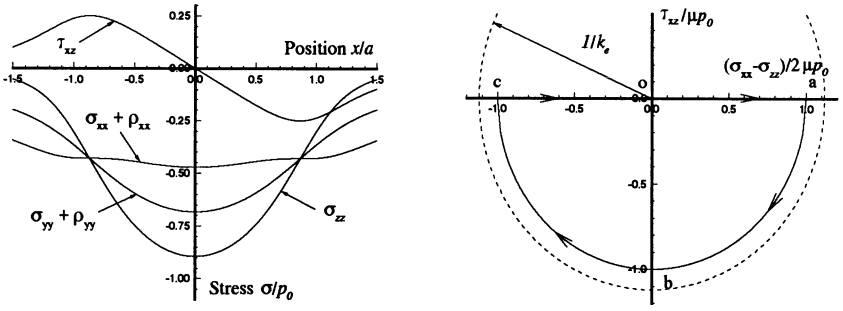
(b) Variation of stress down the center–line

Fig 4: Measured residual stress distribution in a rail head.

contact. They show that in general five components of residual stress ρ_{xx}, ρ_{yy}, ρ_{zz}, ρ_{xy} and ρ_{zx} can be set up in the solid. The longitudinal and lateral stresses predicted by these analyses are similar to those predicted for line contacts. In addition, Hearle (1984) suggests that a compressive vertical stress will be set up under the load: the predicted magnitude of the vertical stress $\rho_{zz} \approx -0.1k_e$, giving $\rho_{zz} \approx -20\text{MNm}^{-2}$ in rail steel. Finally, the analyses predict that two components of residual shear stress ρ_{zx} and ρ_{xy} are likely to be developed by the deformation caused by a point contact. Both calculations predict that ρ_{xy} is greatest on the surface, on either side of the running band ($z = 0, y = \pm a$), while ρ_{zx} is a maximum just below the surface, on the axis of symmetry $y = 0$. The predicted magnitudes of the shear stresses differ somewhat: Hearle (1984) gives a conservative estimate, while Ponter et al give an upper bound. Taking an average of the two sets of values suggests that $\rho_{xy} \approx 0.4k_e$ (giving about 90MNm^{-2} in rail steel), while $\rho_{zx} \approx 0.15k_e$ ($35\ \text{MNm}^{-2}$ in rail steel).

We have compared these predictions with experimental measurements of residual stresses near the surface of a rail head. The experiments are described in detail in Bower and Cheeswright (1987). Our results are summarized in Fig. 4: Fig 4a shows the variation of residual stress ρ_{xx}, ρ_{yy} and ρ_{xy} across the surface of the rail head, while Fig. 4b shows the variation of ρ_{xx}, ρ_{yy}, ρ_{zz} and ρ_{zx} with depth down the center line of the rail.

There are two main causes of residual stresses in rails. The stresses in the top 5–10 mm of the rail head are caused by the plastic deformation that occurs under the repeated passage of wheels over the track, while the stresses below this depth are set up during manufacture by the straightening process. It is evident that the magnitude of the longitudinal, lateral and vertical stresses in the top 5mm of the rail head are in good agreement with theoretical predictions. However, the residual



(a) Subsurface stress distribution (b) Surface stress trajectory

Fig. 5: Stress states at the shakedown limit under line contact.

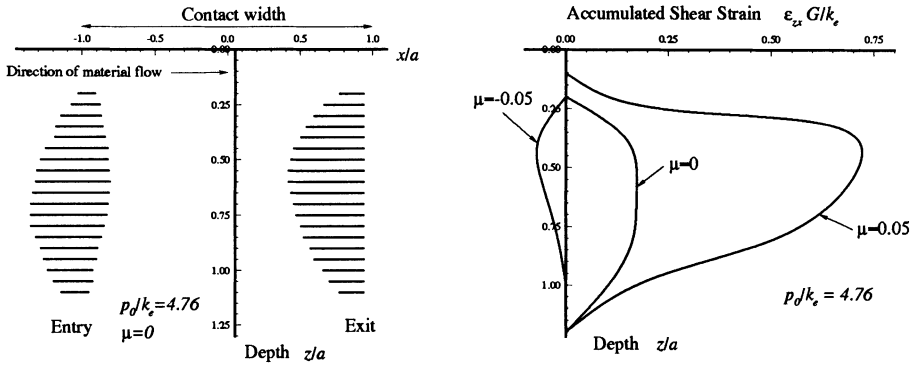
shear stresses ρ_{zx} and ρ_{xy} are much smaller than the theory predicts. The most likely explanation for this discrepancy is that the theoretical calculations assume that the load always passes down the same track over the surface. This is not the case in railway track: the load tends to move from side to side over the rail head. The random lateral motion of the load would be expected to attenuate the residual shear stresses.

5. Plastic flow above the shakedown limit

If the load exceeds the shakedown limit, it is of interest to estimate the magnitude of the cyclic and accumulated plastic strains. To date, only the deformation caused by a line contact has been investigated in detail. Since different methods of analysis must be used to investigate the subsurface deformation which occurs under low tractive loads and the surface flow caused by high tractive loads, they will be considered separately.

A. Subsurface flow

A good guide to the nature of the deformation which occurs under low tractive loads can be obtained by considering the state of stress under a pure rolling contact at shakedown. The stresses at the critically stressed depth $z = 0.5a$ under a pure rolling line contact are plotted in Fig. 5a. At shakedown, the orthogonal shear stress τ_{zx} just reaches yield at two points under the surface: $x = \pm 0.875x/a$. Residual stresses ρ_{xx} and ρ_{yy} develop so that $\sigma_{xx} + \rho_{xx} = \sigma_{yy} + \rho_{yy} = \sigma_{zz}$ at these two points. If the load is increased by a small amount above the shakedown limit, it follows from the flow rule in eq. (4) that the predominant component of plastic strain is the orthogonal shear ε_{zx} parallel to the surface. This deformation can be modelled by using distributions of dislocations with Burgers vector parallel to the free surface to describe the shear deformation, (Hearle and Johnson 1987, Bower and Johnson 1989). The two plastic zones under the contact are divided into equally spaced bands of dislocations, as illustrated in Fig. 5a. The distribution of dislocations in each band is chosen so that when the stress field due to the dislocations is added to the stresses due to the applied load, the stresses inside the bands are equal to the yield stress of the material in shear. As a result of the dislocation distribution, the material above each band is displaced tangentially



(a) Plastic zones (b) Total accumulated strain

Fig. 6: Predicted subsurface deformation under line contact.

relative to that below it, thus approximating the shear deformation in the two plastic zones. The shear strain in the plastic zone ahead of the contact is negative, while that behind it is positive.

The interaction between the two plastic zones gives rise to a small accumulated plastic strain below the surface due to each pass of the load. In rail steel, the rate of accumulation of strain progressively decreases, until in the steady state, the material is subjected to a closed cycle of strain. This steady state is well represented by the linear kinematic hardening calculations of Bhargava et al (1987), so the dislocation model has only been used to calculate the total accumulated strain. The results are shown in Fig. 6b, which shows the variation of total accumulated strain as a function of depth below the surface for three values of the friction coefficient μ. The strains are specified in dimensionless form as $\varepsilon_{zx}G/k_e$, where G is the elastic shear modulus. It is evident that the accumulated strain is of the order of the elastic strain caused by the contact load. The strain is therefore too small to be measured. This prediction is supported by experiments reported by Bhargava et al (1987), who measured the deformation of markers in rail steel disks under conditions of pure rolling. They were unable to detect any accumulated displacement of the markers even at loads twice the shakedown limit.

The implication of these results for railway practice is that on heavy haul lines, there may be a region $\approx$ 10mm deep below the surface of the rail which is plastically deformed by every wheel passing over the track. The material does not accumulate a unidirectional strain, so the plastic flow may not be perceptible by merely sectioning the rail and etching it. Nevertheless, the closed cycle of plastic strain may lead to failure by low cycle fatigue.

B. Surface flow

The deformation caused by a sliding contact with traction coefficient $\mu > 0.25$ differs significantly from the subsurface flow which occurs under low tractive loads. As before, a qualitative guide to the nature of the deformation can be obtained by examining the elastic stress state at shakedown. The critically stressed material

now lies at the surface. The history of stress for an element of material at the surface as it passes under a line contact is illustrated in Fig. 5b, which shows the trajectory of shear stress τ_{xz} as a function of the deviatoric stress in the horizontal direction $(\sigma_{xx} - \sigma_{zz})/2$. Results are shown for a driving traction. The trajectory follows the sequence (o,a,b,c,o): while the material is distant from the contact, it is stress free (at o). As the element of material approaches the leading edge of the contact, it is subjected to a tensile stress, (at a). As it passes under the contact, it follows the sequence of tension, followed by shear (at b), followed by compression at the trailing edge of the contact (point c in Fig. 5b).

For reference, the yield locus has also been plotted as a dashed line on Fig. 5b. It is evident that when the load reaches a magnitude $p_0/k_e = 1/\mu$, the material under the entire width of the contact reaches yield simultaneously. Now, imagine that the load is increased by a small increment, so that a thin layer at the surface exceeds yield. We have used an approximate method to estimate the resulting plastic deformation. Provided the thickness of the layer is small compared with the contact width, the distribution of contact pressure and traction remains approximately hertzian. Then the *elastic* stress field, which is in equilibrium with these tractions, provides a good approximation to the true stress field. The history of strain can then be deduced by integrating the constitutive equations (3–7). In addition, we assume that the accumulation of longitudinal and lateral strain, ε_{xx} and ε_{yy} is prevented by small residual stresses σ_{xx} and σ_{yy}, which are found by iteration.

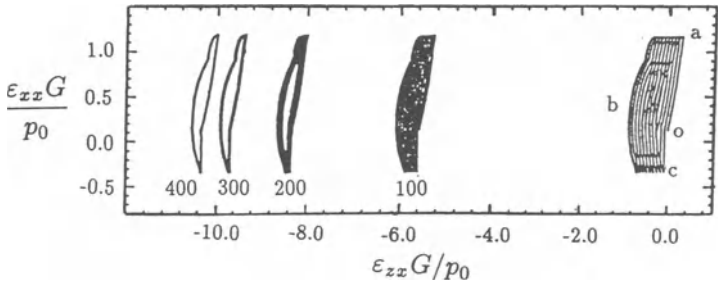


Fig. 7: Predicted cycles of strain at $z = 0$ under a sliding line contact.

The predicted cycles of strain at the surface are shown in Fig. 7. The strains are again shown in dimensionless form as $\varepsilon_{ij}G/p_0$, where G is the elastic shear modulus and p_0 is the peak Hertzian pressure. Points oac on this figure correspond to those shown on the stress cycles in Fig. 5b. The material accumulates a shear strain as a result of the loading, but the ratchetting rate progressively decreases as the deformation continues. Eventually, the deformation settles to a closed cycle of strain.

By computing similar cycles of strain for elements of material under the surface, it is possible to calculate the total accumulated strain as a function of depth below the surface. Integrating these strains allows the total displacement of the surface δ before ratchetting ceases to be estimated. For a load $\mu p_0/k_e = 1.5$, we find a total accumulated displacement $\delta G/p_0 a = 3$.

The near surface plastic flow caused by a sliding contact has also been measured experimentally, using a disk machine (Bower and Johnson, 1989). By inserting markers into a rail steel disk, we were able to measure the total accumulated displacement of material at the surface. The results were in qualitative agreement with the theory: the material was observed to accumulate a strain, but the ratchetting rate progressively decreased, until after about 50 000 cycles, no further displacement of the markers was observed. However, the total accumulated displacement in our experiments was of the order $\delta G/p_0 a = 200$ at a load $\mu p_0/k_e = 1.4$. This is almost two orders of magnitude greater than our predictions. We believe that the discrepancy is caused by our choice of the parameter γ_2 in the hardening law for rail steel. The value for γ_2 given in Table 1 was determined from the results of a tension/compression test, which was limited to about 600 cycles by fatigue failure. This value predicts accurately the initial ratchetting rate in the disk machine experiments, but it is not possible to extrapolate to 50 000 cycles using a value of γ_2 based on a test of only 600 cycles. Nevertheless, it is likely that the theory predicts accurately the cyclic plastic strain.

6. References

Bhargava, V., Hahn, G.T., Ham, G., Kulkarni, S. and Rubin, C.A., "Influence of kinematic hardening on rolling contact deformation," *Proc. Symp. on Contact Mechanics and Wear of Rail/Wheel Systems II, Kingston, R.I. July 1986*, University of Waterloo Press, pp. 133–146

Bower, A.F. and Cheesewright, P.R., (1987), "Measurement of residual shear stresses near the surface of a rail head," *Cambridge University Engineering Department Report CUED/C-Mech/Tr.41*

Bower, A.F., (1989), "Cyclic hardening properties of hard-drawn copper and rail steel," *J. Mech. Phys. Solids*, **37** (4) pp.455–470

Bower, A.F. and Johnson, K.L., (1989), "The influence of strain hardening on cumulative plastic deformation in rolling and sliding contact," *J. Mech. Phys. Solids*, **37** (4) pp.471–493

Bower, A.F. and Johnson, K.L., (1991), "Plastic flow and shakedown of the rail surface in repeated wheel–rail contact," *Wear*, **144**, pp.1–18.

Hamilton, G.M., (1983), "Explicit equations for the stresses beneath a sliding, spherical contact," *Proc. I. Mech. E.*, Part C, **197**, p.53

Hearle, A.D. and Johnson, K.L., (1987), "Cumulative plastic flow in rolling and sliding line contact," *A.S.M.E. Journal of Applied Mechanics*, **54**, pp. 1–5.

Hearle, A.D., (1984), "Deformation, shakedown and fatigue in rolling contact," Ph.D dissertation, University of Cambridge.

Johnson, K.L., (1985), "Contact mechanics," Cambridge University Press.

Johnson, K.L., (1987), "Plastic flow, residual stress and shakedown in rolling contact," *Proc. Symp. on Contact Mechanics and Wear of Rail/Wheel Systems II, Kingston, R.I. July 1986*, University of Waterloo Press, pp. 83–97

Johnson, K.L., (1990), "A graphical approach to shakedown in rolling contact," *Proc. Conf. on Applied Stress Analysis, Nottingham, August, 1990*, T.H. Hyde and E. Ollerton (eds.), Elsevier, pp.263–274

Kapoor, A., and Johnson, K.L., (1992), "Effect of changes in contact geometry on shakedown of surfaces in rolling/sliding contact," *Int. J. Mech. Sci*, **34** (3) p.223

Ponter, A.R.S., Hearle, A.D. and Johnson, K.L., (1985), "Application of the kinematical shakedown theorem to rolling and sliding point contacts," *J. Mech. Phys. Solids*, **33** (4) p.339

RESIDUAL STRESSES

HOW RESIDUAL STRESSES CAN AFFECT RAIL PERFORMANCE

Oscar Orringer
U.S. Department of Transportation
Volpe National Transportation Systems Center
Cambridge, MA 02142, U.S.A.

Abstract

The effects of rail residual stress on rail performance are discussed. Methods for measuring or calculating residual stress are summarized and compared. Based on current knowledge, it appears that better control of straightening stress in the mill and wheel contact stress (via grinding) in service may improve rail performance.

Introduction

Iron and steel have been used to make rails for roughly 150 years, yet only in the last 15 years has attention been focused on the effects of residual stresses on rail performance. What is meant by "residual" stress? Why has residual stress apparently become important only after more than a century of railway experience? Answers to these questions can be found in the theory of solid mechanics and in the history of rail/wheel system development.

What levels of residual stress exist in modern rail? How do those levels affect rail performance? Can they be changed, for better or worse, by changes in production methods, maintenance practices, or railway operations? Answers to these questions require experiment, analysis, and enlightened re-examination of field experience. Various aspects of these topics are developed in depth by the other contributors to this conference.

This paper offers a framework for general understanding and summarizes the present state of knowledge about residual stresses in rails. It is written mainly from the author's perspective of the North American rail research program sponsored by the U.S. Federal Railroad Administration.

Residual stress and its consequences

The simple example of a nut retaining a sleeve surrounding a bolt (Figure 1) illustrates the concept of residual stress. Tightening the nut tensions the bolt shank and compresses the sleeve. (To further simplify the example, we may assume that these stresses are elastic.) The tension and compression persist after the external nut-tightening torque has been removed, as long as the position of the nut is not disturbed. Also, since no external forces now act on the assembly, the total tensile force (product of tension and bolt shank cross section area) is exactly counter-balanced by the total compressive force (compression times sleeve cross section area). Persistence and self-equilibrium are the two defining characteristics of a residual stress field.

J.J. Kalker et al. (eds.), Rail Quality and Maintenance for Modern Railway Operation, 253–271.

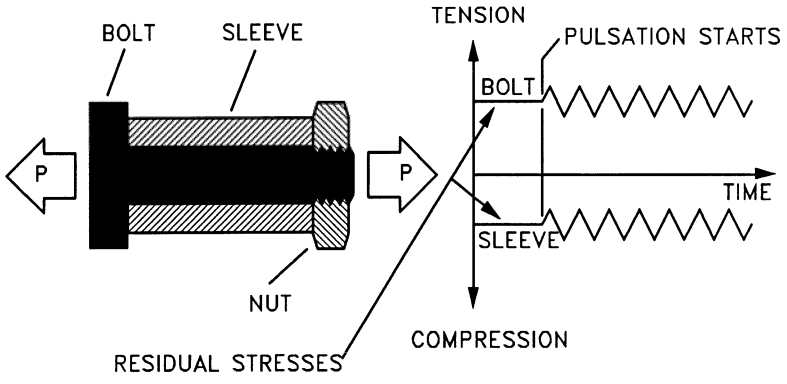


Figure 1. Effect of residual stress in a mechanical assembly.

Now consider what may happen if an external pulsating tensile force (*P-P*) is applied to the bolt. This force is distributed to the bolt shank and sleeve as pulsating tension which combines with the residual stresses to produce the stress-time histories shown at the right of Figure 1. If the pulsating force is applied long enough, fatigue cracks may develop in either the sleeve or the bolt. A crack most likely occurs in the bolt first, since the pulsating and residual tensions add. This crack will also tend to grow more quickly with the residual tension present than it would do otherwise. Conversely, a crack in the sleeve would require much longer to develop and would most likely not grow in the presence of the residual compression.

The pulsating force might also cause wear by inducing minute sliding motions between the bolt and sleeve. If wear did occur, however, it would have at most only a weak dependence on the magnitude of the residual stress in the assembly.

Now assume that we are given the assembly with no information about the thread pitch, friction coefficient, or the tightening torque. What can be done to estimate the residual stress magnitude? On one hand, we might apply a strain gauge to the outside surface of the sleeve and measure the change in strain when the nut is removed. (An equal strain of opposite sign would represent the state before removing the nut.) On the other hand, we might leave the nut tight and expose the sleeve surface to X-rays to determine the Bragg diffraction angle, and hence the strain as indicated by the atomic lattice spacing. Applying Hooke's law in either case would then provide a direct estimate of the residual compression, at least in the sense of an average value over that part of the sleeve surface where the measurement was made. We might assume further that the residual stresses are uniformly distributed in both the sleeve and bolt shank, and if the cross section areas were known, we could then apply the self-equilibrium principle to infer the residual tension magnitude in the bolt.

The above speculations illustrate the following limitations of experimental and analytical methods for estimating residual stress. Destructive measurements require confidence that total relief is achieved. Both kinds of measurement are localized and, strictly speaking, represent only averaged strains. Experimentally determined stresses

are, therefore, also localized, averaged, and inferred based on the assumption that relief obeys Hooke's law. The principle of self-equilibrating forces is a powerful tool for analysis, but it cannot be applied without making assumptions about the distribution of stresses in a continuum. These limitations may seem trivial for the bolt example, but as discussed next, they pose formidable challenges in practical situations.

Residual stresses in rails

The external forces responsible for residual stresses in rails come from rolling contact in the straightener during production as well as under wheels in service. Local plastic deformation (cold work) supplies the mechanism to establish the stresses, and the counter-balancing is done within the rail itself. The effect can be understood by thinking of a thin layer of material, immediately under the contact zone, as if it were pastry dough being rolled out on a table. The rail remains a continuum, however, and the stretch of the cold-worked layer is resisted by the bulk of the rail. The stresses decrease but do not disappear, when the external force is removed, because the permanent stretch in the cold-worked layer cannot be recovered in the elastic unloading process.

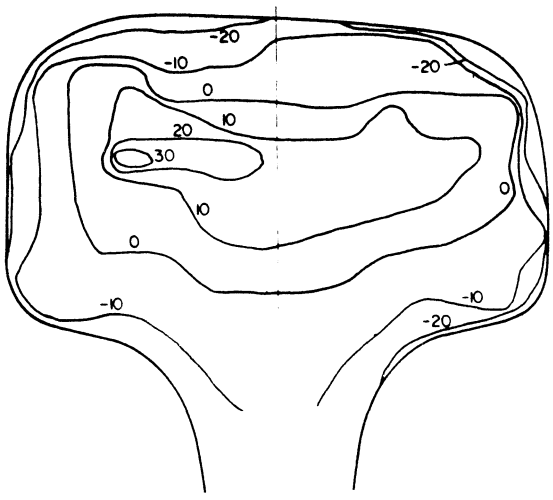
The effects of production and service forces are quite different. On one hand, the roller straightener applies relatively high contact forces across both the head and base of the rail. These parts are thus stretched laterally and shortened lengthwise [1], leading to a residual stress distribution of axial tension in the head and base, balanced by axial compression in the web [2]. On the other hand, a wheel applies light to moderate contact forces across a relatively narrow zone[1] on the head. The surrounding material prevents lateral deformation[2] in this case. The cold-worked layer stretches axially, acquiring residual axial compression balanced by tension in the central region of the head.

Figure 2 illustrates the nature of the residual stresses caused by wheel contact. The stress contours shown here were developed from destructive measurements of strain and length changes as the sample was sectioned [3]. Two important features appear in these plots. First, an internal region of 1/3 to 1/2 the rail head cross section is in a state of triaxial tension. This tension, when combined with the pulsating triaxial compression from passing wheels, provides the bias to promote fatigue crack development and propagation. Second, the residual stress patterns are nonuniform and asymmetric. Thus, one cannot directly apply the self-equilibrium principle, as was done in the bolt example to infer the complete stress distribution from a measurement made only (for example) on the surface.

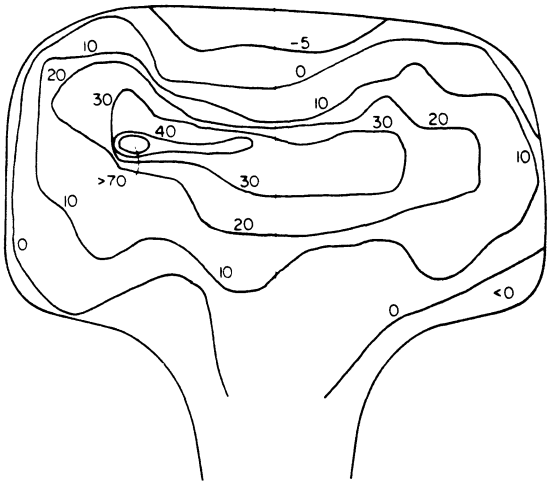
---

1 The lateral position of the zone depends on the wheel profile. Since a rail in service encounters wheels with profiles ranging from new to heavily worn, the plastic deformation is spread among many overlapping zones which generally cover an area from the gauge corner to halfway between the rail crown and field-side face.

2 A transitional case appears to occur when older rail stock of relatively low strength is subjected to extremely high axle loads. Lateral metal flow is often reported in such cases, suggesting the possibility that axial compression might have been established in the rail head. However, there have been no attempts to measure residual stresses in such rails.

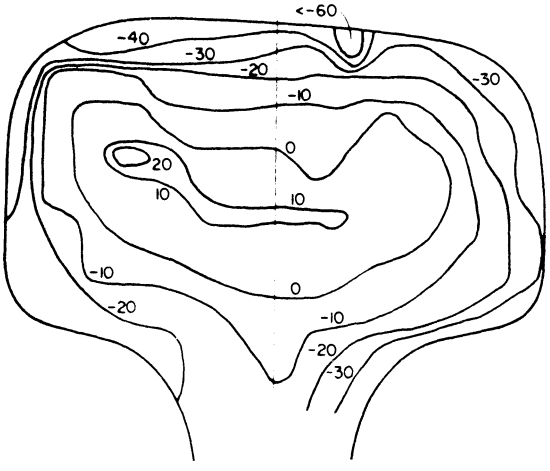


(a) axial stress



(b) maximum principal stress in cross section

Figure 2. Typical residual stresses in 132 RE rail after service under North American heavy haul freight traffic. (Reproduced from [3]; stress contours in ksi; 1 ksi ≅ 6.9 MPa.)



(c) minimum principal stress in cross section

Figure 2. (concluded)

Three other features distinguish the practical problem of measuring rail residual stress from the bolt example. First, the distribution details differ from one rail to another [4], and lengthwise measurements made on individual rails exhibit axial variations as well [4,5]. Second, large stress gradients require a high density of measurements, each of which should cover as small an area as possible in order to minimize averaging bias error. Third, using Hooke's law to infer stress from strain may produce considerable error in regions which have undergone plastic deformation.[3]

Historical perspective

Even the earliest railway operations must have produced rail residual stresses. For more than a century, however, rails either wore to removal limits before the residual stresses produced recognizable effects, or rail failures were attributed to manufacturing imperfections. Improvements in rail-making technology [2] have by now almost eliminated these imperfections, leaving rolling contact fatigue as the only remaining significant source of rail failure. Development of high-hardness alloys and increasing

---

3 Plastic deformation produces atomic lattice defects which broaden the Bragg diffraction peaks, thus introducing errors in X-ray strain measurements. (Similar texture effects on ultrasonic wave speeds may occur.) The stress-strain curve for unloading after plastic deformation is often nonlinear below 20 to 30 percent of the maximum stress attained (Bauschinger effect). Thus, both destructive and nondestructive strain measurements in plastically deformed material are subject to measurement or interpretation error.

use of lubrication have doubled rail wear life over the past 15 years. Wear remains the principal life-limiting factor, but many more rails now than in the past stay in service long enough for fatigue cracks to form and propagate.

Few comparisons of rail wear and rolling-contact fatigue life have been published because it is quite difficult to make unbiased estimates. Figure 3 summarizes the comparisons made between 1979 and 1985 for North American freight service [6 - 8]. The wear life plots (solid lines) are based on projections of measured average wear rates and include tangent track (straight line) [6] as well as curves [8]. The fatigue life plots represent calculations for tangent track by means of the Palmgren-Miner rule. Reliance upon a theory was necessary because fatigue life measurements could not be derived from the available data.[4] The earliest model was based on the axial components of rail bending, contact, and thermal[5] stress (dashed straight line) [6], the life estimate being modified by wear of the running surface (head height loss) at an average rate. A later model (triangle, circle, and square symbols) [7] was based on octahedral shear and hydrostatic stresses from combined bending, contact, thermal, and residual stresses. The results represent life estimates for three different depths below the rail head.[6]

The perceptions of those who made the life estimates are perhaps as important as the numbers. The author of the first comparison (the two straight lines in Figure 3) suggested that fatigue might become a life-limiting factor at a critical axle load just slightly below the presently allowed 33-ton maximum for interchange service [6]. The authors of the second comparison found fatigue life to be an elusive quantity, strongly dependent on the material fatigue curve, the rail head wear rate, and the depth assumed for crack development [7]. They noted that extrapolation of the calculations (e.g., as shown for the 6 mm depth by the dotted line in Figure 3) would lead to quite different predictions for the critical axle load.[7]

Figure 3 reflects the general situation on the North American railways at the beginning of their transition from mixed freight to widespread unit train operations. Fatigue was expected to become the life-limiting factor as average axle loads increased, and as wear life was extended by lubrication and improvements of rail hardness. The forecast was soon confirmed by experiments at the Transportation Test Center and by field experience on heavy haul freight lines [9].

Changing the assumed residual stress magnitude did not appear to have much effect on fatigue life [7]. Residual stress was also judged to be a secondary variable in the

---

4 Reports of detection covering all types of rail defects requiring expeditious remedial action. These reports cover transverse defects but not shelling or gauge corner spalling.

5 Caused by restraint of continuous welded rail against thermal expansion or contraction.

6 This model was limited to points directly under the center of contact, and the residual stresses were approximated by an arbitrarily chosen mean axial tension. The effect of wear on fatigue was also approximated by displacing the contact stress field downward into the rail head at a rate consistent with measurements of average head height loss.

7 These fatigue life estimates are somewhat too pessimistic because they have been based on first percentile failure data from laboratory tests of rail steel. One may question whether a permanent way engineer would consider development of cracks in just one percent of the rail stock to be the fatigue life limit.

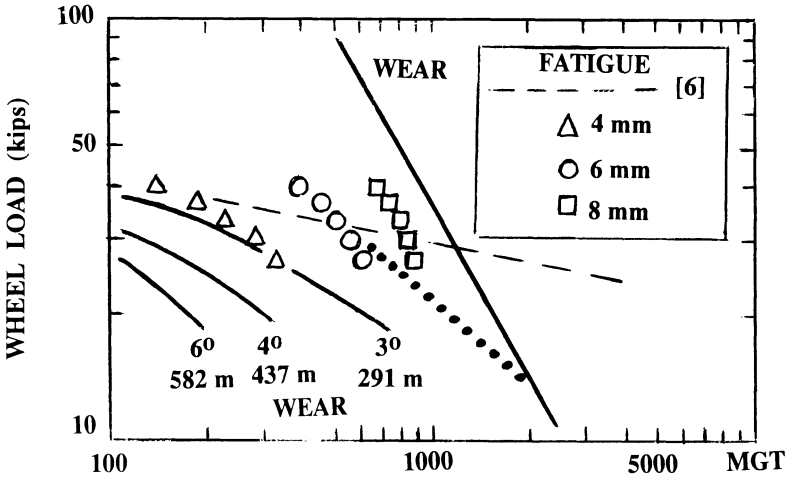


Figure 3. Rail wear and rolling-contact fatigue life estimates. (Based on North American freight service circa 1975-1980. For wear life on curves, track curvature is given in terms of AREA "degree" and radius in meters.)

first theoretical investigation of detail fracture propagation [10]. However, later correlation of full scale tests, stress measurements, and rail failure reports with a similar model showed that residual axial tension in the rail head has a strong effect on detail fracture propagation life (Figure 4) [11]. The effect on propagation rate is greatest when the crack is small[8]; thus, a small increase of residual tension may dramatically shorten the propagation life. Subsurface shells [12, 13], transverse defects [13], and vertical split heads[9] have also been treated as propagating cracks in models which may include residual stress.

Wheel/rail contact loads are mainly responsible for the residual stress effects just discussed. Manufacturing stress became a topic of equal importance in the early 1980s, as the result of an investigation of a rail failure which caused the derailment of a passenger train [14]. Experiments [15] and analyses [1, 14] have suggested that the roller-straightening residual stress field, described earlier, can drive a web fracture through an entire rail length if the stress magnitude exceeds a critical level. (Fabrication stress can also be beneficial, e.g., as in the prolongation of mechanical joint life by applying cold work to the bolt holes.[10])

---

8 The crack relieves the residual tension as it grows.

9 R.A. Mayville et al., Arthur D. Little, Inc., Cambridge, MA, work in progress.

10 See the contribution to these proceedings by L. Reid.

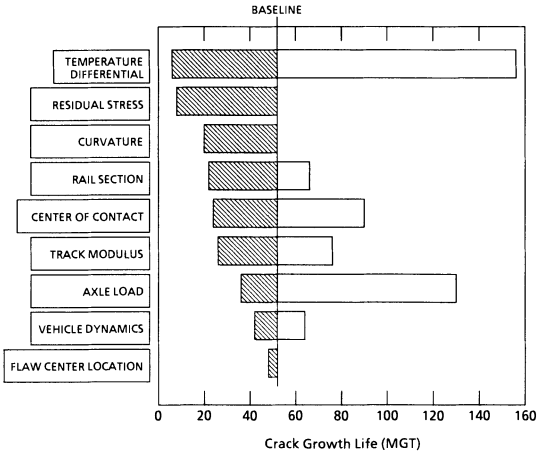


Figure 4. Effect of several variables on detail fracture safe life.

Measurement of residual stress

Destructive examination (DE) with strain gauges has been a widely accepted method of measuring both surface and internal rail residual stress for many years [16 - 19]. Following the first exploratory investigations, DE procedures based on surface measurements were developed for evaluation of new rail quality (roller-straightening stress magnitude).[11] Concerns about subsurface shelling and the propagation of transverse defects later led to development of a DE procedure for mapping the internal stress field in three dimensions [4]. For the rail head alone, this requires up to 360 gauges arranged in rosettes on a grid of 120 squares on a transverse slice, which is cut into cubes to relieve the residual strains.[12] In addition, a 0.5-m long section must be cut into rods, following the same grid pattern, and each rod length measured to provide data for the calculation of axial strain relief.[13]

The great cost per stress map has prompted research on cheaper methods and procedures, including X-ray [21] and neutron [22] diffraction, ultrasonics [5], and moire

---

11 Typically, eight sites around the rail profile are instrumented, and small pyramids containing the gauges are trepanned to relieve the residual strains. Used rails are also surveyed by this procedure to evaluate the general effect of wheel contact.

12 The web and base may also be instrumented if information about the entire roller-straightening stress field is sought. However, most of the rails examined with this procedure had not been roller straightened, so attention was focused on the rail head.

13 Based on earlier work by Meier [20]; see [3] or [4] for detailed description of the entire procedure.

interferometry [23], as well as other alternates for partial surveys [3,24]. Neutron diffraction and moire now appear to be the best alternates from the viewpoint of resolution and accuracy (Table 1). Ultrasonics does not appear to have the resolution needed for stress mapping but is beginning to be considered as a nondestructive examination (NDE) method for monitoring roller-straightening stress on the production line [15].[14] Ultrasonics is also suitable for monitoring continuous welded rail, where a large beam area can be used to average out the residual stress and thus indicate thermal stress levels which might approach track buckling limits [25].

Table 1. Comparison of stress mapping methods.

Method	Resolution (mm)	Bias error (MPa)	Random error (MPa)
Strain gauge [4]	5	35 to 50	± 24
Neutron diffraction [22]	2	not reported	± 25
Moire [23]	1	not reported	± 10

The stress mapping procedure remains subject to certain interpretation errors, as well as the errors associated with specific measurement methods. The resulting stress maps fail to satisfy the self-equilibrium principle [1]. This dilemma can only be solved by means of independent theoretical analysis, as discussed in the next section.

Measurement errors

Bias error from averaging over finite resolution length is a fundamental property of all methods. DE and NDE methods are also subject to bias errors of different nature when applied to plastically deformed regions (Figure 5).

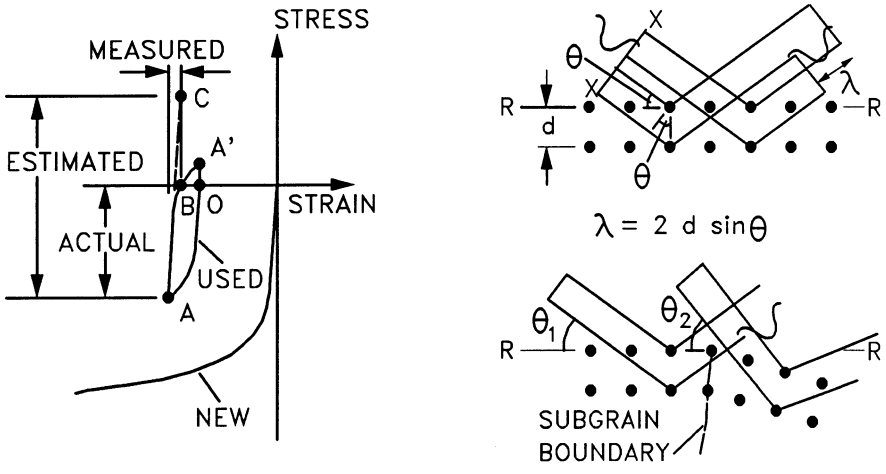
Figure 5(a) illustrates the Bauschinger effect, which biases the DE methods. The stress-strain diagrams, shown here to approximate scale, represent a typical location in the rail head plastic zone. After wheel contacts have developed the residual stresses (represented by point A), each wheel passage causes a cycle $ABA'OA$. The strain relief measured after cutting includes a nonlinear segment in the unloading curve AB, and the residual stress magnitude estimated by Hooke's law (AC) is incorrect.

The diffraction methods are based on Bragg's law, illustrated in the upper diagram of Figure 5(b). Here d is the distance between the strained atomic lattice planes, λ

---

14 DE procedures based on initiating a fracture or slow cutting through the web have been used to evaluate rail alloys and roller-straightening practices [1,15] but are not suitable for production quality control. Neutron diffraction may become a competitive quality control method if a safe portable neutron source can be developed.

the beam wavelength, and θ the beam incidence angle with respect to a specimen reference axis *R-R*. The incident plane wave *X-X* is split into two reflected wave fronts, and *d* is measured by finding the angle θ at which the two fronts reinforce. In the lower diagram, plastic deformation has created atomic-scale lattice faults which collect to form subgrain boundaries, across which the lattice planes tilt. The reflected waves then partially reinforce over a range of nominal incidence angles, and *d* can only be estimated from the broadened diffraction peak.



(a) Bauschinger effect *(b) peak broadening*

Figure 5. Errors caused by plastic deformation.

Cutting can be the worst source of error for the DE methods, since the process itself plastically deforms material near the site of strain relief measurement. The original stress-mapping procedure specified slow-feed cutting with an air-cooled bandsaw, the only equipment practicable for cutting the long-section rods. In the initial trials of moire [23], which involved only transverse slices, an oil-cooled low-speed carbide circular saw produced better results, but plastic effects were still observed as far as 1 mm in from the edges of the typical 4 × 4-mm surface region left after cutting. The effects were reduced further by cutting with an oil-cooled high-speed diamond saw.

Interpretation errors

Without a procedure much more elaborate than the one now used to map internal stresses, one is forced to interpret the data based on the assumption that the residual stresses are axially uniform, at least over the 0.5-m sample length. Although axial uniformity is a physically plausible hypothesis, strain gauge [4] and ultrasonic [5,15] surveys have indicated the presence of small fluctuations along the rail, even in regions

distant from the mill ends. This implies the presence of residual shear stresses with axial-vertical and axial-lateral orientations: stresses for which there are currently neither measurement nor analysis provisions.

The last source of error to be considered here arises from plotting. The usual practice is to plot level contours (e.g., Figure 2) as the best aid to visualizing the stress map. Such plots are based on treatment of discrete data as points which define a smooth topographic surface. This is a poor substitute for enforcement of the self-equilibrium principle, especially in regions where the apparent stress gradient is large compared with the measurement spacing.

Improvements based on moire

Moire interferometry is an optical diffraction method for measuring the surface displacements on gratings typically ruled to the order of 1000 lines/mm (see [23] for details). Surface strains are obtained by measuring both cartesian components of the displacement field and then numerically differentiating the data. Since only the grating must be attached to the surface of a body, moire has an obvious advantage over the tedious mounting and wiring of hundreds of strain gauges. Therefore, it may be possible to improve the stress-mapping procedure by making two modifications, which are now under investigation.

First, it is practical to apply moire DE routinely to more than one slice from one rail sample. In the simplest procedure, transverse and oblique slices may be oriented as shown in Figure 6, with Mohr's circle applied to transform the oblique-slice data to rail coordinates.[15] For the configuration shown, the transform provides the axial normal stress and axial-vertical shear. One is still forced to assume axial uniformity when interpreting the results for corresponding points as stresses in a common cross section, but at least the effective sample length is reduced.

Second, heat treatment at 400 to 500 ^{o}C is under consideration as a substitute for cutting to relieve the residual stresses which remain in rail slices. This has not been an option in the past because moire gratings are usually replicated on thin aluminum films, which are then bonded to slice surfaces. In a recent experiment, however, gratings of 1000 lines/mm have been etched directly onto specially prepared surfaces of rail slices. Figure 7 illustrates the quality of the fringes obtained from one of these gratings with stress relieved by high-speed cutting.[16] It should be possible to obtain full-field data of equal quality if the slice stresses are relieved by heat treatment in an inert atmosphere.

---

15 This idea was first proposed by R. Czarnek and J. Orkisz during the April 1990 conference on "Residual Stress in Rails", held in Cracow.

16 Work in progress by F.P. Chiang, State University of New York at Stony Brook. This work will also include a trial of the dual slice procedure.

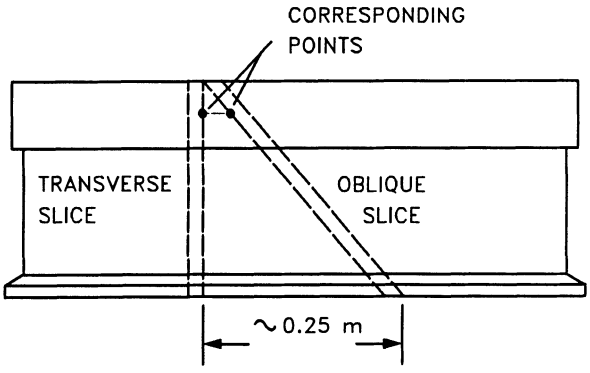
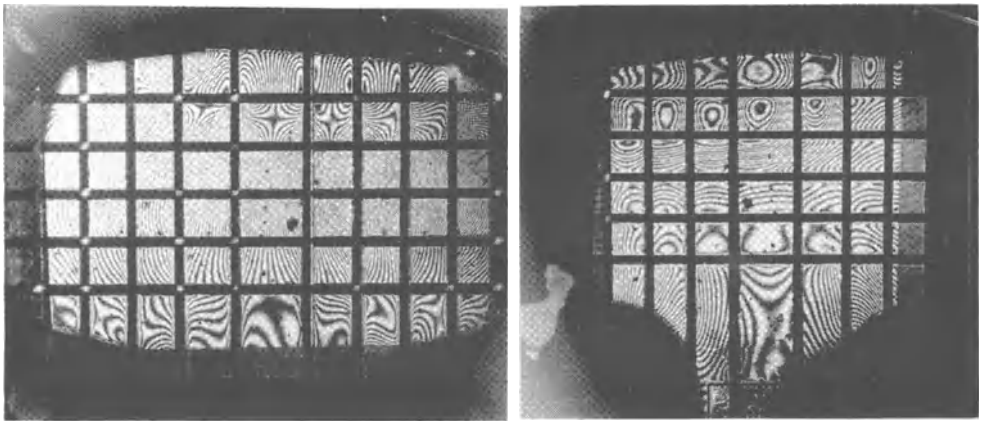


Figure 6. Czarnek-Orkisz dual slice procedure.



(a) lateral *(b) vertical*

Figure 7. Cartesian displacement fringes after cutting a rail slice.[16]

Theoretical estimation of residual stress

No practical measurement scheme can furnish a complete picture of the residual stress field in a rail. Theoretical methods are thus essential to the quantitative understanding of relations between applied loads and residual stresses. The rôle of theory is the same as for the bolt example in the introduction. To complete the picture, one must use the

self-equilibrium principle to extrapolate the stress field from discrete measurements. The complex shape of the rail cross section and the three-dimensional character of its stress fields requires numerical analysis, which can at most approximate the actual residual stress field.

All numerical analysis schemes require geometrical approximation of the actual body. Analyzing a three-dimensional model of a rail is thus a formidable task, even if purely elastic behavior is assumed. Consequently, most numerical analyses of rail stress are done with two-dimensional models.

The load sources have already been noted for the complexity of their effects on the character of the actual residual stress field. Roller straightening is a precisely specified sequence of less than ten loads. Conversely, a rail in service carries millions of wheel loads; their magnitudes can only be estimated as probability distributions, and their sequences cannot be predicted. These differences also affect the application of theory.

Residual stresses caused by roller straightening

The best approach to problems involving a few precisely specified loads is to calculate the body's response to the specified load-time history. Conventional numerical analysis methods are well suited for this approach. It is only necessary to divide the history into a moderate number of small steps (typically 20 to 50 per load cycle). The response can then be estimated in steps using Hooke's law and corrected with a model of the nonlinear material behavior based on the accumulated stresses and strains. Actual material properties such as strain hardening and the Bauschinger effect (Figure 5) can be approximated in the material model, if desired.

Alternative geometrical models for stepwise analysis of roller straightening have been devised by Wineman [26]. A three-dimensional model was found to require so much computing effort that the analysis had to be limited to passage through only one roller. The stresses calculated for such cases failed to approximate the effects of the full straightener but were useful for vetting the results of a simpler two-dimensional model. The latter model represents the rail in its vertical center plane, with all effects averaged across the profile width, and has been used for stepwise calculations simulating passage through several rollers.

Figure 8 illustrates the results obtained by Wineman from a two-dimensional simulation of an experiment reported by Schweitzer et al. [27], in which the axial residual stress distribution was measured at a cross section which had passed through 7 out of 9 rollers in the straightener. The theoretical results approximate the experimental observations, suggesting that the model is suitable for comparison of different roller-straightening practices. However, this model cannot provide a three-dimensional stress map like that shown in Figure 2.

Residual stresses caused by wheel contact

The conventional theoretical approach is totally impractical for the problem of estimating residual stresses caused by wheel contact. Even when a rail is loaded repeatedly along the same center-of-contact path in a controlled experiment, some 2×10^5 to 1×10^6 wheel passes are typically required to fully develop the residual stress field [28]. Still more are required in service, where the center of contact wanders across

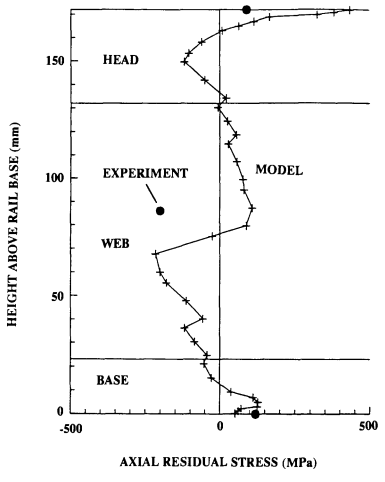


Figure 8. Comparison of model prediction with measurements.

the running surface as a rail encounters wheels with different profiles. In practical terms, full development of a rail residual stress field may thus require on the order of 10 to 100 million gross tons of traffic.[17]

Despite the large number of loads which must be considered and the lack of precise knowledge about their magnitudes, the problem of wheel contact can be analyzed by means of a theoretical approach based on the concept of a so-called shakedown state [29]. The residual stresses in a body are said to have reached a shakedown state if, after some number of loads, the body experiences no further plastic deformation under continued application of the same range of load magnitudes. Although the theory cannot be used to predict when such a state will be reached, it can be used to estimate the shakedown residual stress distribution from the greatest expected load magnitude, without regard to any specific load sequence. The shakedown state is estimated by searching, among the possible (i.e., self-equilibrating) stresses, for those which also: (1) do not exceed the yield limit either alone or when combined with the applied stresses; and (2) minimize a potential function of elastic energy stored in the body and

17 The range exhibited in the cited experiments reflects differences between the rates at which different stress components develop. The range quoted for service is an estimate which also includes the effects of contact wandering and a range in average wheel loads.

work done upon the body by external forces.[18]

Although a rail cannot be said to reach a true shakedown state in service, experimental observations do suggest that the plastically deformed region eventually enters a stable cycle, like that illustrated by the loop *ABA'OA* in Figure 5(a). The idealized shakedown state may be viewed as an approximation of the actual residual stress (the point *A* in the figure). The shakedown theory is also limited to an idealized material model: the so-called elastic - perfectly plastic (E-PP) characteristic, which is regulated by a single strength parameter called a flow stress.[19]

The shakedown theory has been applied to the present problem based on simplified numerical models of the rail and its service environment. Axial uniformity is assumed for the residual stress field, as has been done to interpret experiments. This reduces the rail model to a typical cross section. The effects of the wheel loads are represented by applied stresses in the typical section, which is assumed to lie directly under the center of contact. Bending stresses are calculated from a model of the rail as a beam on a continuous elastic foundation [11]. Contact stresses are calculated from a bi-parabolic pressure distribution over a rectangular patch, with peak pressure and patch dimensions based on the Hertz contact theory [30].[20]

The behavior of the model was investigated by varying the geometrical representation of the rail cross section and the ad hoc rules for sizing the contact patch [30]. The sensitivity of residual stress to wheel load was studied by varying the greatest load magnitude from 90 to 222 kN, with 147 kN considered as a nominal value.[21] In order to reduce the number of variables and allow for comparison with the available experimental data, a 132 RE section with a flow stress of 483 MPa (70 ksi)[22] was modeled, and the lateral position of the center of contact was fixed.

Attention was focused on the predictions for axial residual tension because of its effect on the propagation of detail fractures (Figure 4). The tension predicted for nominal wheel loading was much lower than the experimental observations, e.g. Figure 2(a).[23] In a later study, with the load magnitude and peak contact pressure fixed at

---

18 The potential is used as the objective function for the search. The correspondence between its minimum and the correct solution is a fundamental theorem of solid mechanics.

19 E-PP material either follows Hooke's law or, if it is plastically deforming, has a fixed octahedral shear stress equal to the assumed flow stress. Thus, actual properties like strain hardening and the Bauschinger effect cannot be directly modeled.

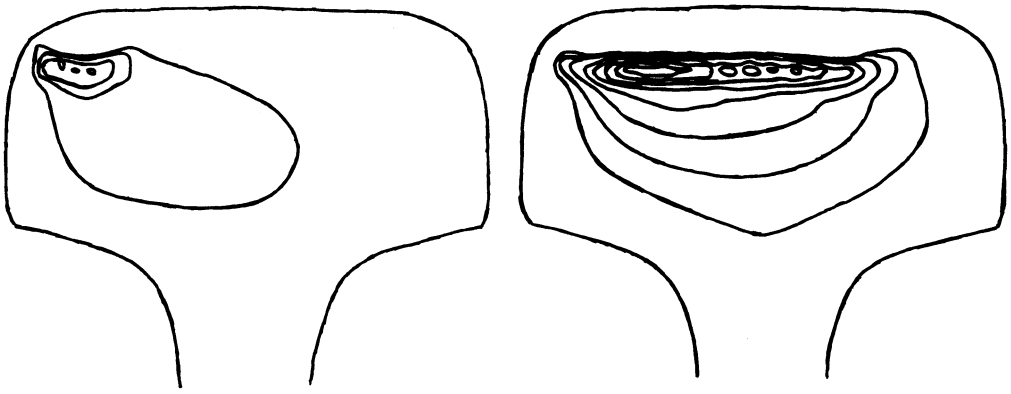
20 The shakedown approach requires elastic representations of the applied stresses. Therefore, these calculations are routine once the model assumptions have been specified.

21 The largest static wheel load presently allowed for freight cars in North American interchange service is 147 kN.

22 The nominal yield strength of the 280 BHN rails from which the North American stress maps have been obtained.

23 This result was not unexpected, since dynamic wheel loads routinely exceed the nominal value, and peak dynamic loads well above 222 kN have been measured on North American railway lines. The actual residual stresses may also be increased by tangent tractions (creep and friction forces), which have been omitted from the calculations to date.

their nominal values, the contact wandering phenomenon was simulated and was found to triple the axial residual tension relative to loading along a fixed rolling path (Figure 9).[24] Therefore, it appears that realistic stress maps may be obtained from the shakedown model, once realistic specifications are established for dynamic load magnitudes, tangent tractions, and contact pressure as a function of lateral center-of-contact position.



(a) center of contact fixed near gauge corner

(b) center of contact wandering from near gauge corner to crown

Figure 9. Effect of contact wandering on axial residual tension. (Wheel load = 147 kN; stress contour interval = 1 ksi; outermost contour = 0 ksi; 1 ksi ≅ 6.9 MPa.)

Enhancement of results

The idea of searching among possible states for those stresses which satisfy the yield constraints and minimize the mechanical potential [29] has also been generalized to formulate a powerful concept for combining theoretical and experimental stress analysis [31]. Additional constraints are incorporated by requiring the computed results not to deviate from measurements at discrete points by more than a specified tolerance.[25] The objective function is also extended to be a weighted sum of the

---

24 M. Hołowiński, Cracow Institute of Technology, unpublished work at the Volpe National Transportation Systems Center, 1991.

25 Based on a rational estimate of the expected measurement error.

mechanical potential and a geometrical expression for contour smoothness. Varying the weights furnishes a continuous spectrum of estimates with bases ranging from purely theoretical to purely empirical.

The generalized search method can potentially make three important contributions. The first and most obvious is for analyses of measurements with good resolution and coverage, where a small influence from theory may produce maps of self-equilibrating stresses which best fit the data. Second, the method seems able to project from regions where measurements are available into regions where the data is missing.[26] Such a feature would be useful for rational interpretation of experiments in which some of the instrumentation has failed, or as a means of reducing test costs by permitting the elimination of sensors in non-critical regions.[27] Third, the method may be used to enhance the theoretical model by training its parameters in the presence of experimental results. For example, heuristic rules for assigning higher flow stress values in the plastic volume may thus be developed as a means to approximate strain hardening.

Concluding remarks

Research and field observations in the past 15 years have shown that residual stress is an important factor affecting rail performance in modern railway service. Residual stresses caused by wheel contact contribute to economic life limits (gauge corner spalling and subsurface shelling) when lubrication extends wear life and increased axle loads reduce fatigue life. Residual stresses caused by service and/or by roller straightening establish conditions (crack propagation and fracture) which can lead to rail failure.

It now appears possible to evaluate different grinding practices based on the effects their wheel/rail contact patterns may have upon residual stresses in the rail head. Making rail stronger also reduces these stresses, but residual stresses increase when a rail of higher strength is straightened to specified dimensional limits. Increasing strength is known to be an effective measure for extending fatigue life, at least where the limit is associated with subsurface shelling. Extending rail life limits without risking unacceptable rates of rail failure thus demands better control of roller-straightening stress as rails are made stronger. Either mechanical or thermal procedures may be used to provide the necessary control. Preserving fatigue life under greater axle loads requires the same combination of measures. The research on these questions should lead to quantitative conclusions in the next 5 years.

---

26 Based on a numerical test with an idealized one-dimensional problem.

27 One possible application is the analysis of residual stress caused by roller straightening, for which the available measurements are from surveys of surface stress. Roller-straightening stresses appear to be as close to axially uniform as residual stresses from service. One is thus led to consider the mechanical potential from the two-dimensional shakedown model as a part of the objective function, although there is no evidence that roller-straightening residual stresses approach a shakedown state.

References

[1] Wineman, S.J. and McClintock, F.A., "Residual stresses and web fracture in roller-straightened rail", in: *Residual Stress in Rails - Vol. II: Theoretical and Numerical Analyses* (O. Orringer, J. Orkisz, and Z. Świderski, eds.), Kluwer Academic Publishers, The Netherlands, 1-22, 1992.

[2] Esveld, C., *Modern Railway Track*, MRT-Productions, Duisburg, Germany, 1989.

[3] Cundiff, C.H. and Rice, R.C., "Comparative evaluation of several alternative methods for measuring rail residual stress", in: *Residual Stress in Rails - Vol. I: Field Experience and Test Results*, op.cit. 121-142.

[4] Groom, J.J., "Determination of residual stresses in rails", Battelle Columbus Laboratories, Columbus, OH, report no. DOT/FRA/ORD-83/05, May 1983.

[5] Deputat, J. et al., "Experiences in ultrasonic measurement of rail residual stresses", in: *Residual Stress in Rails - Vol. I*, op.cit., 169-183.

[6] Zarembski, A.M., "Effect of rail section and traffic on rail fatigue life", *American Railway Engineering Association Bulletin 673, vol. 80*, 514-527 (1979).

[7] Perlman, A.B., Jeong, D.Y., Orringer, O., and McConnell, D.P., "Rail flaw growth investigations", *American Railway Engineering Association Bulletin 688, vol. 83*, 536-550 (1982).

[8] Zarembski, A.M., "Changing face of rail replacement", *Railway Track and Structures*, 17 (January 1985).

[9] Reiff, R.P., "Report from FAST - recent test experiences", *Transportation Research Record 1174*, Transportation Research Board, National Research Council, Washington, DC, 81-93 (1988).

[10] Besuner, P.M., "Fracture mechanics analysis of rails with shell-initiated transverse cracks", *Rail Steels - Developments, Processing, and Use, ASTM STP 644* (D.H. Stone and G.G. Knupp, eds.), 303-329 (1978).

[11] Orringer, O. et al., "Crack propagation life of detail fractures in rails", Volpe National Transportation Systems Center, report no. DOT/FRA/ORD-88/13, October 1988.

[12] Farris, T.N., Xu, Y., and Keer, L.M., "Some factors influencing the transition from shelling to detail fracture", in: *Residual Stress in Rails - Vol. II*, op.cit., 23-44.

[13] Olzak, M., Stupnicki, J., and Wòjcik, R., "Analysis of crack front propagation in contact", in: *Residual Stress in Rails - Vol. II*, op.cit., 45-62.

[14] Orringer, O. and Tong, P., "Investigation of catastrophic failure of a premium-alloy railroad rail", *Fracture Problems in the Transportation Industry* (P. Tong and O. Orringer, eds.), American Society of Civil Engineers, New York, 62-79 (1985).

[15] Steele, R.K. et al., "Catastrophic web cracking of railroad rail", in: *Residual Stress in Rails - Vol. I*, op.cit., 1-19.

[16] Yasojima, Y. and Machii, K., "Residual stresses in the rail", *Permanent Way 8*, Japan (1965).

[17] Konyukhov, A.D., Reikart, V.A., and Kaportsev, V.N., "Comparison of two methods for assessing residual stresses in rails", *Zavodskaya Laboratoriya 39(1)*, 87-89 (1973).

[18] Deroche, R.Y. et al., "Stress releasing and straightening of rail by stretching", *Proc. Second International Heavy Haul Railway Conference*, Colorado Springs, paper no. 82-HH-17 (1982).

[19] Radomski, R., "Residual stress measurements at rail surface and inside rail head", in: *Residual Stress in Rails - Vol. I*, op.cit., 205-214.

[20] Meier, H., "Eigenspannungen in Eisenbahnschienen", *Zeitschrift Vereines deutscher Ingenieure 81(12)*, 362-363 (1937).

[21] Schilling, C.G. and Blake, G.T., "Measurement of triaxial residual stresses in railroad rails", Association of American Railroads, Chicago Technical Center, report no. R-477, 1981.

[22] Webster, G.A. et al., "Neutron diffraction determination of residual stress patterns in railway rails", in: *Residual Stress in Rails - Vol. I*, op.cit., 143-152.

[23] Czarnek, R., Lee, J., and Lin, S.-Y., "Moire interferometry and its potential for application to residual stress measurements in rails", in: *Residual Stress in Rails - Vol. I*, op.cit., 153-167.

[24] Bijak-Żochowski, M., "Investigation of residual stress by the penetration method", in: *Residual Stress in Rails - Vol. I*, op.cit., 185-203.

[25] Utrata, D., "Experience with two ultrasonic-based measurement techniques for residual stress determination in railroad rails", *American Railway Engineering Association Bulletin 733, vol. 92*, 392-398 (1991).

[26] Wineman, S.J., "Residual stresses and web fracture in roller-straightened rail", PhD thesis, Department of Mechanical Engineering, MIT, Cambridge, MA, June 1991.

[27] Schweitzer, R., Flügge, J., and Heller, W., "Einflüsse auf das Bruchverhalten von Schienen", *Stahl und Eisen 105*, 1451-1456 (1985).

[28] Świderski, Z. and Wójtowicz, A., "Plans and progress of controlled experiments on rail residual stress using the EMS-60 machine", in: *Residual Stress in Rails - Vol. I*, op.cit., 57-66.

[29] Orkisz, J., "Prediction of actual residual stresses by constrained minimization of energy", in: *Residual Stress in Rails - Vol. II*, op.cit., 101-124.

[30] Perlman, A.B. and Gordon, J.E., "Application of the constrained minimization method to the prediction of residual stresses in actual rail sections", in: *Residual Stress in Rails - Vol. II*, op.cit., 151-177.

[31] Karmowski, W., Magiera, J., and Orkisz, J., "Enhancement of experimental results by constrained minimization", in: *Residual Stress in Rails - Vol. II*, op.cit., 207-217.

PREDICTION OF RESIDUAL STRESS IN RAILS: PRACTICAL BENEFITS FROM THEORETICAL APPROACH

J. Orkisz and M. Hołowiński
Cracow Institute of Technology
Cracow, Republic of Poland

Abstract

Residual stresses constitute a significant part of the total stresses in rails and thus play an important rôle in rail fatigue phenomena. An effective theoretical-numerical approach has been recently proposed for residual stress analysis. It is shown here that it is a valuable tool for practical engineering applications. The method is used to evaluate how the residual stress distribution is influenced by: yield strength, load magnitude and sequence, presence of initial stresses, and lateral wheel wandering. Practical conclusions are presented, based on a variety of numerical tests.

Introduction

Residual stresses constitute a significant part of the total stresses in rails and thus play an important role in rail fatigue phenomena [1]. Theoretical evaluation of the true residual stress state in a railroad rail is a difficult task. However, an effective theoretical-numerical approach based on concepts of the shakedown theory has been recently proposed [2,3]. It results in solution of the following constrained optimization problem:

> *find self-equilibrated stresses that minimize the total complementary energy of a rail, and together with the live stresses caused by the current loads (vehicle wheels, restrained thermal expansion of CWR, etc.) follow the yield constraint (i.e., the effective total stress is not above the material flow stress)*

This approach takes into account only the most important factors of the phenomena and is valid for an elastic - perfectly plastic rail material, quasi-static cyclic loadings of deterministic nature, and small strains. As opposed to such simplified assumptions, the real service conditions involve dynamic effects, repetitive non-cyclic loadings of rather stochastic than deterministic nature, as well as kinematic hardening and softening material properties.

Some further simplifying assumptions are made in practical calculations. A rail is considered as a prismatic beam on an elastic foundation [4]. The residual stresses are assumed to be the same in each cross section. The true wheel/rail rolling contact is replaced by a parabolic pressure distribution approximating the Hertzian elastic one.

Despite all these limitations, we may in this way quite well predict and evaluate the true residual stresses acquired in a railroad rail in service conditions, provided that the initial (manufacturing) residual stresses are known. A reliable numerical approach has been developed, based on the formulation mentioned above [1-3,5,6]. Three different methods (finite element, boundary element, and finite difference) may be used, yielding numerically verified independent solutions.

The numerical analysis has proved to be a valuable tool for practical applications,

J.J. Kalker et al. (eds.), Rail Quality and Maintenance for Modern Railway Operation, 273–285.

as demonstrated in this paper. The basic method is used to evaluate how the residual stress distribution is influenced by four factors: (1) yield strength; (2) load magnitude and sequence; (3) presence of initial residual stresses; and (4) lateral wheel wandering. The method[1] is also used to analyze and evaluate experimental residual stress data obtained from destructive examination of a rail [7].

Investigation program

A trial analysis of a rail has been conducted with a numerical model of 329 finite elements and 384 nodes [5]. The same model was used in the present investigation where only qualitative results were sought. Conversely, cases requiring numerically stable quantitative results were analyzed with a model of 1600 elements and 1719 nodes (Figure 1). Though all stress components $\sigma_{xx}, \sigma_{yy}, \sigma_{xy}, \sigma_{zz}$ were calculated, mainly the tensile part of the axial component σ_{zz} is presented here. A variety of particular cases (Table 1) were analyzed for different vertical load locations (Figure 2), magnitudes, and flow stress values.

Flow stress

Flow stresses five percent above and below a 70 ksi (483 MPa) nominal yield strength were assumed. An increase of the flow stress reduces the residual stress magnitude (Figure 3.)

Load magnitude and sequence

Solutions were obtained for four maximum wheel load values: V = 30, 33, 35, and 40 kips (133, 147, 156, and 178 kN). The corresponding magnitudes of axial residual stress increase as the maximum wheel load increases (Figure 4). The tensile zones are quite similar in size and shape, but the maximum residual tension occurs at a greater depth as the maximum wheel load increases (Figure 5).

Sequence is important from a theoretical point of view, since it may cause a non-proportional loading program. In the present analyses, however, sequence was found to have almost no effect on the final results.

The third and fourth cases of series 3 (Table 1) illustrate the effect of loads applied in sequence at adjacent locations (7 and 8, versus 8 and 7). For these cases the maximum magnitudes computed for residual axial compression were 27.53 and 27.46 ksi, respectively, while the maximum values for axial residual tension were 7.62 and 7.79 ksi.

The third line of series four represents another kind of sequence test. In this case, analyses were performed with loads of 30 and 33 kips magnitude in one location (7). The computed residual stresses were found to be exactly the same for both sequences (i.e., 33 kips first or second).

---

1 See the companion paper by Karmowski, Magiera, and Orkisz.

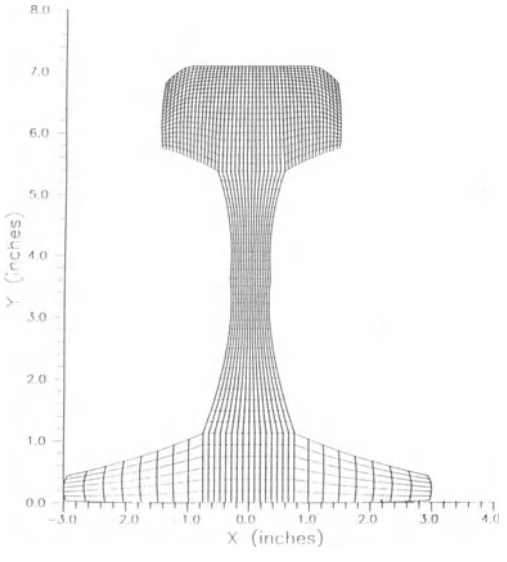


Figure 1. Refined finite element model of rail.

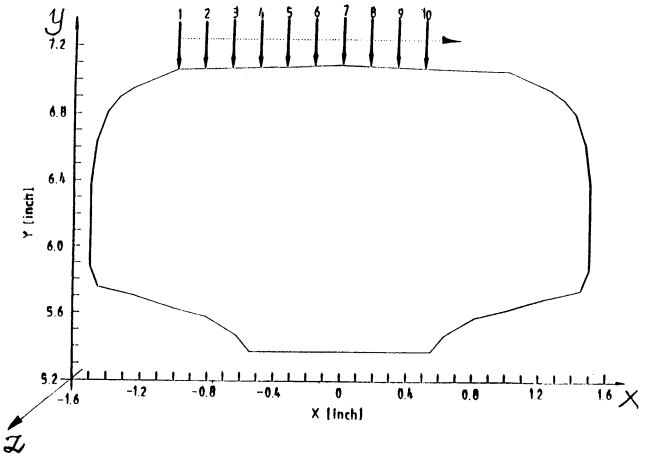


Figure 2. Locations assumed for center of wheel/rail contact.

Table 1. Summary of cases analyzed.

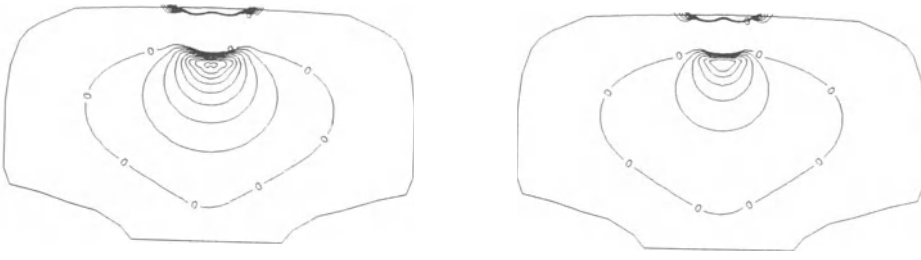
Test ser.	Fig.	Mesh[2]	$\frac{\sigma_o}{Y}$ [3]	Load (kips)	Last loc.[4]	Prior loc.[5]	Main purpose of the test Effect of
1	3a 3b	R	0.95 1.05	33	7	---	Yield strength
2	5a 5b 5c 5d	R	1	30 33 35 40	7	---	Load magnitude
3		R	1	33	7 8 7 8	--- --- 8 7	Load sequence in different locations
4		R	1	30 33 30,33	7	--- --- 7	Load sequence in the same location
5	6a	C	1	33	1 4 7 10	---	Load location
6	6b 7a	C	1	33	1 4 7 10	--- 1 1,4 1,4,7	Wandering load (coarse steps)
7	8a,9a 6c,8b,9b 8c,9c 7b,8d,9d 6d 10a	C	1	33	1 4 7 10 7 4 1	--- 1,2,3 1 - 6 1 - 9 1-10-8 1-10-5 1-10-2	Wandering load (fine steps)

2 C - coarse (329 elements); R - refined (1600 elements)

3 σ_o - flow stress; Y - yield strength, 70 ksi (483 MPa) assumed.

4 See Figure 2.

5 Load locations (see Figure 2) assumed for calculation of residual stress states preceding the calculation for the final location.



(a) $\sigma_o = 0.95Y$; $\sigma^{(r)}_{zz\,max}$ = *8.3 ksi* *(b)* $\sigma_o = 1.05Y$; $\sigma^{(r)}_{zz\,max}$ = *5.9 ksi*

Figure 3. Effect of yield strength on tensile axial residual stress. (Outer contour at zero stress; contour interval 1 ksi; 1 ksi ≅ 6.9 MPa.)

Presence of initial stresses

Initial stresses may influence the final residual stress state when, for example, the loading program consists of a sequence of loads of the same magnitude applied to different locations. Figure 6 illustrates the computed residual axial tension for several such cases. The contours in Figure 6(a) reflect loading at a single location (4, Figure 2) and thus serve as a reference. The other cases consist of loading at other locations to establish different initial stress states, followed by final loading at location 4. The prior locations are: 1 in Figure 6(b); 1-2-3 in Figure 6(c); and 1-2-3-4-...-10-9-...-5 in Figure 6(d). The change in the maximum stress magnitude apparent in these results is not observed when the initial-stress loading program remains within the envelope of the final program.

Wheel wandering

The term "wheel wandering" refers to the observed tendency of wheels with different worn profiles to contact the rail at different lateral locations. Strictly speaking, wheel wandering establishes a non-proportional loading program, violating one of the basic assumptions of the theoretical approach. However, we may obtain useful practical results by following a heuristic procedure based on moving the lateral location of wheel contact, the existing cyclic loading program, and an assumption that the residual stresses

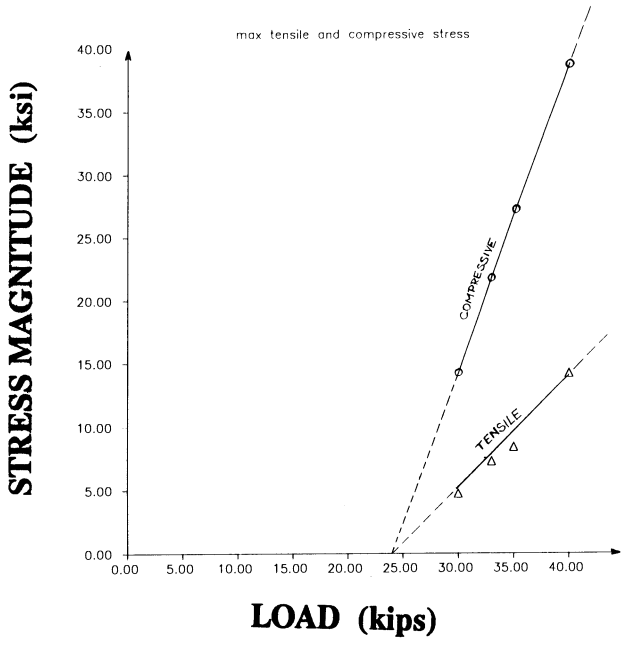


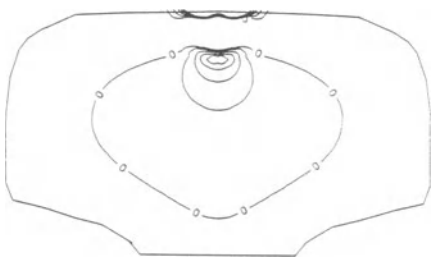
Figure 4. Effect of wheel load on residual stress. (1 kip ≅ 4.5 kN; 1 ksi ≅ 6.9 MPa.)

computed for prior load locations are the initial stresses for the current load.

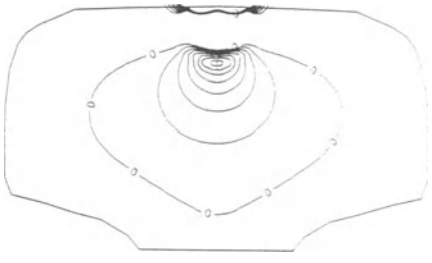
A shakedown state for the rail may then be estimated, for example, from programs in which the load is moved from left to right, vice versa, or in random order, to determine whether an essential change of residual stress is observed. If all these programs yield no significant differences, then the results may be interpreted as an estimate for the true shakedown stress state.

Two such cases are compared here: load at locations 1, 4, 7, and 10 (coarse steps); and load at locations 1, 2, 3, ... , 9, 10, 9, ... , 2, 1 (fine steps). Figure 7 illustrates the different compressive axial residual stress patterns computed from these two programs. The progressive development of the residual stress field calculated from the fine-step program is illustrated in Figures 8 (compression) and 9 (tension). It is important to note here that the return of the load from location 10 to 1 had no significant effect on the residual stresses established in the first half of the program. Almost the same results were obtained when the program started at location 10, moved to 1, and back to 10 again.

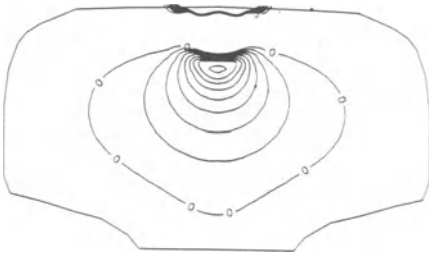
It is worth stressing here that the residual stress distribution and magnitude obtained from the fine-step wheel wandering simulation, Figure 10(a), displays much more



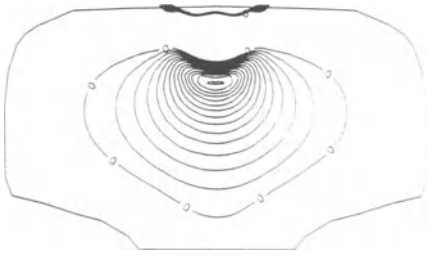
(a) V = 30 kips; $\sigma^{(r)}_{zz\,max}$ *= 4.7 ksi*



(b) V = 33 kips; $\sigma^{(r)}_{zz\,max}$ *= 7.3 ksi*

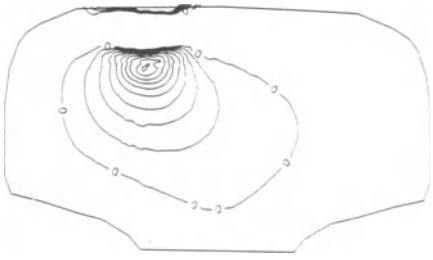


(c) V = 35 kips; $\sigma^{(r)}_{zz\,max}$ *= 8.4 ksi*

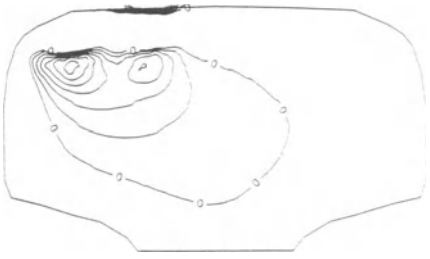


(d) V = 40 kips; $\sigma^{(r)}_{zz\,max}$ *= 14.2 ksi*

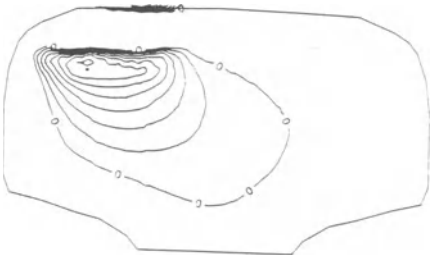
Figure 5. Contours of axial residual tension. (Outer contour at zero stress; contour interval 0.5 ksi for (a), 1 ksi for (b,c,d); 1 ksi ≅ 6.9 MPa.)



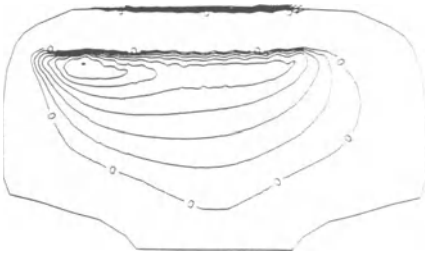
(a) sequence 4, $\sigma^{(r)}_{zz\,max} = 4.70$ *ksi*



(b) sequence 1-4, $\sigma^{(r)}_{zz\,max} = 6.55$ *ksi*

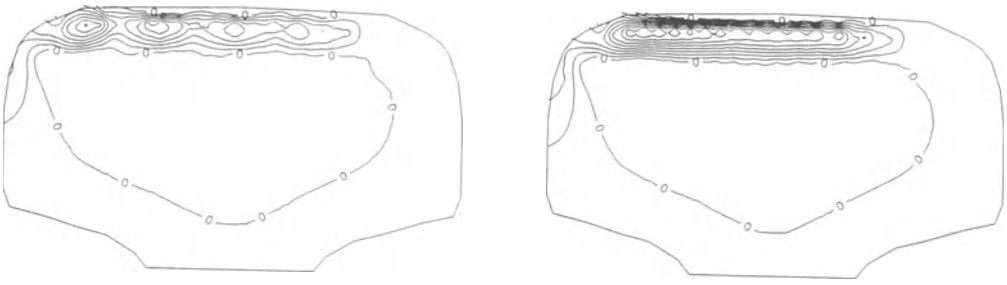


(c) sequence 1-2-3-4, $\sigma^{(r)}_{zz\,max} = 7.28$ *ksi*



(d) seq. 1-..10 ..-5-4, $\sigma^{(r)}_{zz\,max} = 7.09$ *ksi*

Figure 6. Effect of initial stresses.
(Outer contour at zero stress; contour interval 1 ksi; 1 ksi ≅ 6.9 MPa.)



(a) coarse steps *(b) fine steps*

Figure 7. Simulations of wheel wandering.
(Outer contour at zero stress; contour interval 3 ksi; 1 ksi ≅ 6.9 MPa.)

resemblance to experimental stress analyses of rails taken from service, Figure 10(c,d), than the results computed from a loading program with a fixed wheel location, Figure 10(b).

Conclusions

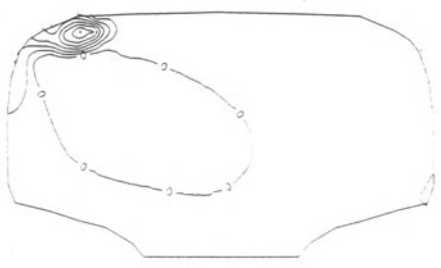
Several conclusions useful from an engineering viewpoint may be drawn from the results of the investigation.

The higher the rail yield strength, the smaller will be the magnitude of residual stresses caused by wheel contact. This is a natural result of the fact that wheel contact is a load-control situation. Conversely, one should expect the residual stress magnitude to increase, when yield strength is increased, for deflection-control situations such as rail straightening.

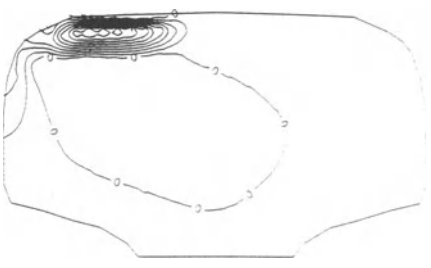
Increasing the load magnitude increases the residual stress magnitude. Loadings of the largest magnitude have the most influence on the final residual stress pattern.

In practice, the load sequence does not significantly affect the residual stress pattern or magnitude.

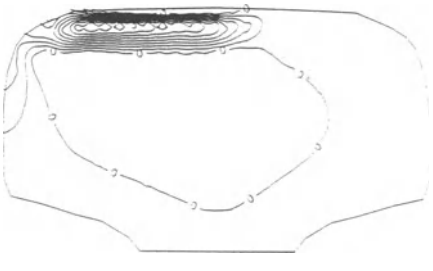
Initial residual stresses may essentially change the final residual stress pattern and magnitude, unless the initial stresses are the result of loads remaining within the same envelope as the current loading program. In particular, the initial stresses created in the fine-step simulation of wheel wandering caused the magnitude of axial residual



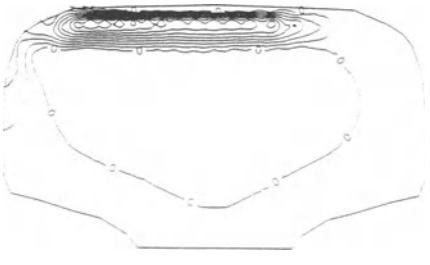
(a) load at location 1



(b) load at location 4

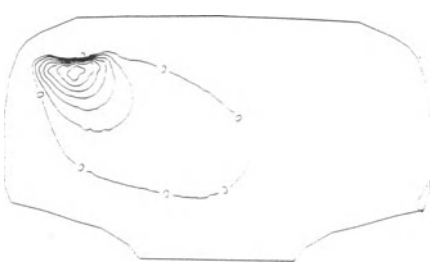


(c) load at location 7

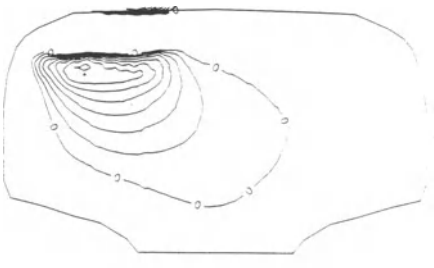


(d) load at location 10

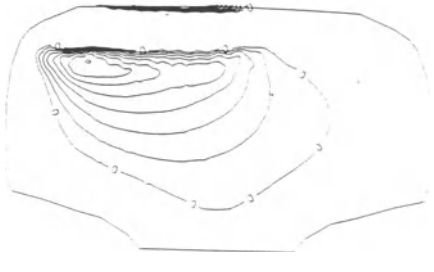
Figure 8. Development of compressive axial residual stresses. (Outer contour at zero stress; contour interval 3 ksi; 1 ksi ≅ 6.9 MPa.)



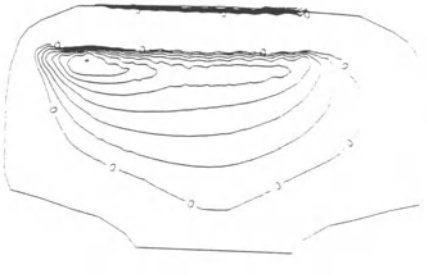
(a) load at location 1



(b) load at location 4

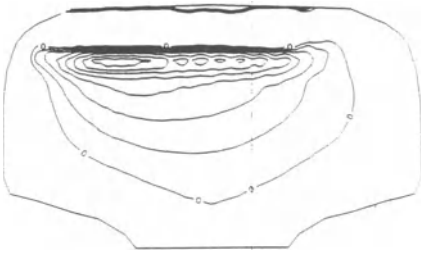


(c) load at location 7

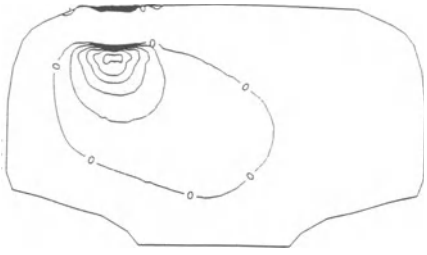


(d) load at location 10

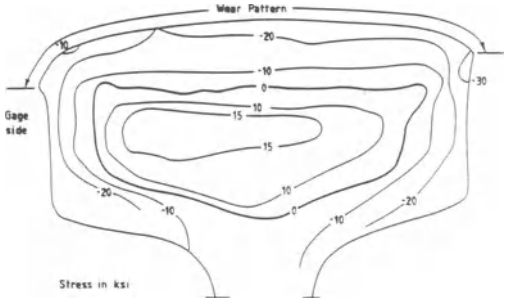
Figure 9. Development of tensile axial residual stresses.
(Outer contour at zero stress; contour interval 1 ksi; 1 ksi ≅ 6.9 MPa.)



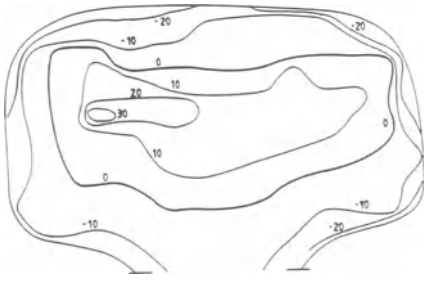
(a) wheel wandering simulation (axial tension, contour interval 1 ksi)



(b) load fixed at loc. 4 (axial tension, contour interval 1 ksi)



(c) 132 RE rail from tangent track on a North American mixed freight line [7]



(d) 136 RE rail from tangent track at the Transportation Test Center [8]

Figure 10. Comparison of theoretical and experimental analyses of rail residual stress. (Stresses in ksi; 1 ksi ≅ 6.9 MPa.)

tension to increase by about 50 percent.

When lateral wandering of wheels is taken into account, the predicted residual stresses are in good qualitative agreement with experimental analyses of residual stresses in rails taken from service. These results were obtained for a 33-kip load magnitude. Although the maximum computed residual tension was only 7 ksi, Figure 10(a), in comparison with measured values from 15 to 30 ksi, Figure 10(c,d), from Figure 4 one may expect a load of 42 kips to double or 51 kips to triple the computed stress. It is worth noting that rails in North American service are subjected to 33-kip static wheel loads, and that dynamic loads of 40 to 50 kips are experienced routinely.

References

[1] Orringer, O., Orkisz, J., and Świderski, Z. (eds.), *Residual Stress in Rails - Vol. I: Field Experience and Test Results*, Kluwer Academic Publishers, The Netherlands, 1992.

[2] Orkisz, J. et al., "Discrete analysis of actual residual stresses resulting from cyclic loadings", *Computers & Structures 35(4)*, 397-412 (1990).

[3] Orkisz, J., "Prediction of actual residual stresses by constrained minimization of energy", in: *Residual Stress in Rails - Vol. II: Theoretical and Numerical Analyses*, op.cit., 101-124.

[4] Perlman, A.B. and Gordon, J.E., "Application of the constrained minimization method to the prediction of residual stresses in actual rail sections", in: *Residual Stress in Rails - Vol. II*, op.cit., 151-177.

[5] Hołowiński, M. and Orkisz, J., "Hybrid finite element method for estimation of actual residual stresses", in: *Residual Stress in Rails - Vol. II*, op.cit., 125-149.

[6] Cecot, W. and Orkisz, J., "Estimation of actual residual stresses by the boundary element method", in: *Residual Stress in Rails - Vol. II*, op.cit., 179-190.

[7] Groom, J.J., "Determination of residual stresses in rails", Battelle Columbus Laboratories, Columbus, OH, report no. DOT/FRA/ORD-83/05, May 1983.

[8] Cundiff, C.H. and Rice, R.C., "Comparative evaluation of several alternative methods for measuring rail residual stress", in: *Residual Stress in Rails - Vol. I: Field Experience and Test Results*, op. cit., 121-142.

A NEW APPROACH TO ENHANCEMENT OF EXPERIMENTAL DATA

W. Karmowski, J. Magiera, and J. Orkisz
Cracow Institute of Technology
Cracow, Republic of Poland

Abstract

The so-called global-local method of experimental data enhancement is used as an approach to combining all the information available about a physical problem. It is applied here to enhance the approximation of residual stress distributions determined from experimental data measured by means of the strain gauge or moire interferometry techniques. In this case, the enhancement involves combination of enforcement of the equilibrium equations with optimal smoothing of the data. A benchmark problem is considered, in which experimental data is simulated by randomizing a known exact solution. The method is then applied to two problems with actual data from rail residual stress measurements. The enhancement method is shown to improve the quality of the residual stress estimates in comparison with the traditional approach, i.e., contour plotting by interpolation of experimental data.

Introduction

A new approach to enhancement of experimental data has been recently proposed [1,2] following some previous ideas [3]. It uses at once all the information available for the investigated problem, resulting from both the various experimental measurements as well as the theory. The approximation obtained in this way is physically based, since all basic physical relations relevant to the problem may be taken into account. Moreover, results are obtained within the error bounds determined by the appropriate physical statistics.

The problem is given in the following general way:

find the stationary point of the energy functional

$$\Phi = \lambda \Phi^E + (1-\lambda)\Phi^T, \quad \lambda \in [0,1] \tag{1}$$

satisfying the theoretical constraints (e.g., equilibrium equations for stresses)

$$A(\sigma) = 0 \tag{2}$$

and the experimental constraints (e.g., allowed deviation e_i from the measured values of σ_i)

$$B(\sigma_i) \leq e_i, \quad i = 1,2,\cdots,n \tag{3}$$

J.J. Kalker et al. (eds.), Rail Quality and Maintenance for Modern Railway Operation, 287–296.

Here $\Phi^E(\sigma)$ and $\Phi^T(\sigma)$ are the experimental and theoretical parts of the functional, σ is the required solution, and λ is a scalar weighting factor.

The idea of a combined experimental-theoretical approach may be realized in many ways. Some possible general formulations are proposed in detail in [2]. The objective of this paper is to present, test, and use the recent version of the so-called global-local method [2,4]. Though information from the entire domain is taken into account then, only a local approximation of experimental data is done. In this approach, stresses are calculated point by point where required. The method is a generalization of the weighted minimization technique [5-8].

It is first applied to a benchmark problem, for which the exact solution is known and "experimental data" is simulated. Subsequently we use this method to solve two problems with actual experimental data from rails. In the first one the measurements were made with strain gauges [9], and in the second with moire interferometry [10].

Global-local method for residual stress analysis

In this method we search for a local surface that, in a neighborhood of a considered point, best fits the unknown required solution. In order to formulate a method within the general approach outlined by equations (1)-(3), a functional Φ as well as constraints $A(\sigma)$ and $B(\sigma)$ need to be defined.

In the residual stress analysis, the equality constraints (2) are given as equilibrium equations and static boundary conditions. The inequality constraints (3) are defined by the requirement that deviation of a computed value of a function of stress $f_n(\sigma)$ from its measured value $f_n(s)$ at a point r_n

$$|f_n(\sigma) - f_n(s)| \le e_n, \quad n = 1, 2, \cdots, N \tag{4}$$

is limited by an admissible error e_n.

The functional (1) is formed as a weighted averaged error function that has to be minimized. At first, the sought stress field σ is expanded around the examined point into a Taylor series. Unknown coefficients of the series are chosen in such a way that strict satisfaction of the equilibrium equations is enforced. Then summed are the squares of the differences $f_n(\sigma) - f_n(s)$, between the approximating and experimental values, at all points where measurements have been made. Each term in this sum is weighted by means of a special function. As mentioned before, the solution is obtained at one point at a time, but in this way we may find the residual stresses in the whole domain.

Thus, the functional has the form:

$$\Phi(r, \sigma) = \frac{1}{N} \sum_{n=1}^{N} v(\rho_n) \frac{1}{I} \sum_{i=1}^{I} [f_n(\sigma_i) - f_n(s_i)]^2 \tag{5}$$

where $\rho_n = |r - r_n|$ is the distance between the examined point r and point r_n,

while I is the number of experimental data points available. The weighting function should reduce the influence of data measured at points distant from the point considered. It may be defined in many ways [2]. The following form of this function was assumed here:

$$v(\rho_n) = \left(\rho_n^2 + \frac{\rho_n^2}{\rho_n^2 + g^2} \right)^{-3} \tag{6}$$

where g is a free parameter of approximation, providing a variety of solutions ranging from a local interpolation ($g = 0$) to a global polynomial approximation ($g = \infty$). Somewhere in between we may find the best approximation (g_{OPT}).

Thus, g should be chosen as a finite value and established by an optimization procedure which takes into account the scattering of experimental data. It is proposed to find g by means of one of the following optimization problems:

(i) find

$$\max_g g \tag{7}$$

satisfying the inequality constraints (4), or

(ii) find

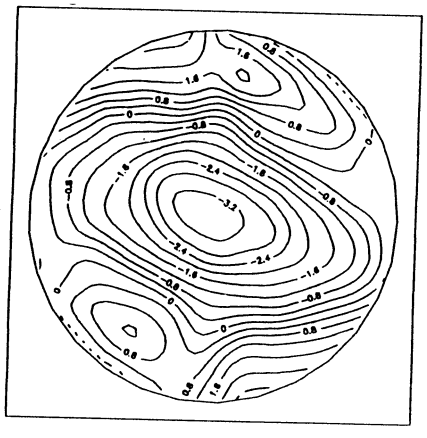
$$\min_g \Phi(r, \sigma) \tag{8}$$

where measured values at all experimental points are taken into account except the value at the point currently considered.

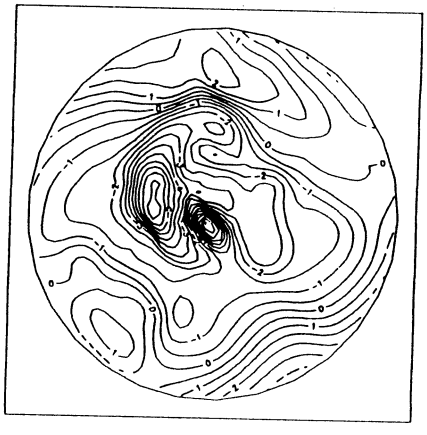
Having found g, we minimize $\Phi(r, \sigma)$ and find local values of the residual stresses σ. This procedure must be repeated at each point in the domain where residual stresses are required.

Test

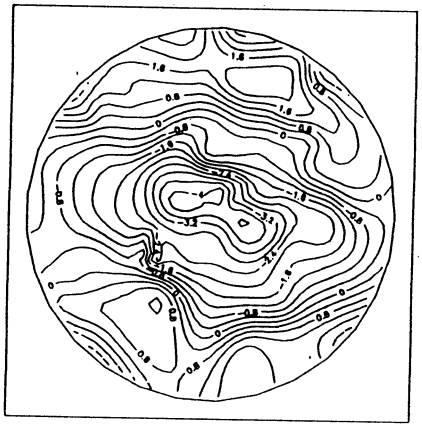
The approach outlined above has been tested in the case of a circular plate loaded in such a way that the exact solution is known, Figure 1(a). It was achieved by assuming an Airy stress function. Computer simulation of pseudo strain gauge experimental data was done by means of randomization of the exact values. These data were then approximated by the global-local method in three ways as follows: (1) interpolation with equilibrium equations not enforced, Figure 1(b); interpolation with equilibrium equations satisfied ($g = 0$), Figure 1(c); and the best approximation ($g = g_{OPT}$) with equilibrium equations satisfied, Figure 1(d). The figure illustrates only the σ_{xx}



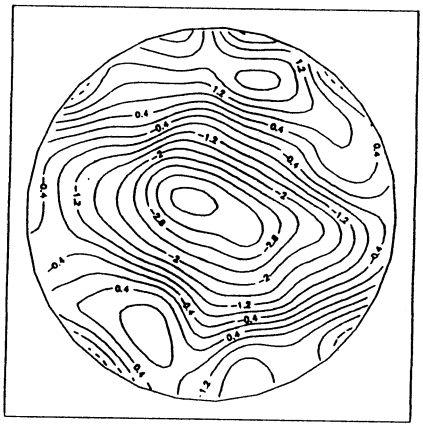
(a) exact solution



(b) interpolated data



(c) interpolated data ($g = 0$) with equilibrium satisfied



(d) best approximation ($g = g_{OPT}$) with equilibrium satisfied

Figure 1. Contours of stress σ_{xx} in loaded circular plate.

component, but similar quality results were obtained for the other components σ_{xy} and σ_{yy} as well.

The following observations can be made from the test results. Enforcement of the equilibrium equations may essentially change the results of the approximation. However, interpolation with or without satisfying the equilibrium equations is not adequate. The best global-local approximation ($g = g_{OPT}$) yields good quality results, as shown by comparison of Figure 1(a) with 1(d).

Enhancement of rail experiments

Two rail experiments were analyzed to provide examples of practical application of the global-local method. The first results were obtained from a service-worn rail by means of the combined Meier and Yasojima-Machii (Y-M) cutting techniques, with strain gauges used on the Y-M slice [9]. In the second experiment, a Y-M slice from a stress-relieved new rail was measured by means of a moire interferometry technique [10].

Figure 2 illustrates the first example with plots of the lateral stress component σ_{xx} surfaces. The rail in this case was a 136 RE section, removed from tangent track after 83 million gross tons of mixed North American freight traffic, and referred to as specimen 1 in the original experiment report [9]. Figure 2(a) illustrates the interpolation of σ_{xx} without satisfying equilibrium.[1] Figure 2(b) plots σ_{xx} as obtained by interpolation with equilibrium satisfied, while the best approximation for σ_{xx} is shown in Figure 2(c).

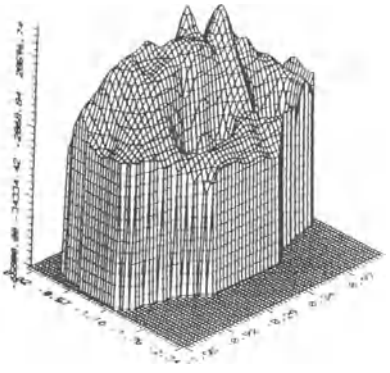
These surfaces have been re-plotted as contours in Figure 3(a,b,c), together with the corresponding results for the vertical stress component σ_{yy} in Figure 3(d,e,f). For both components but especially for σ_{yy} , comparison shows that enforcement of the equilibrium equations may introduce essential changes. It is also apparent that optimal smoothing introduces further significant changes.

In the second application, we used the lateral and vertical displacements, measured by a moire interferometry technique, on a slice of stress-relieved new rail of UIC 60 section, rolled by Huta Katowice [10]. This example was limited to a rectangular domain at the upper center of the rail head, Figure 4(a), the region where moire data was taken. (The boundary conditions for this region are not known.) Again, the same three approaches were applied, with the results for the lateral stress σ_{xx} shown in Figure 4(b,c,d).

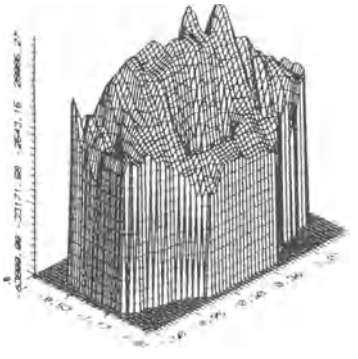
Comparison shows how much the final results obtained here differ from the residual

---

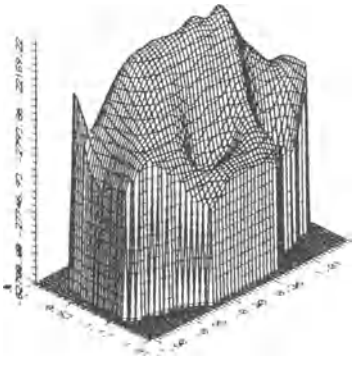
1 This is equivalent to the original experimental plots which, however, were made in terms of principal instead of cartesian stress components.



(a) interpolation of original data [9]

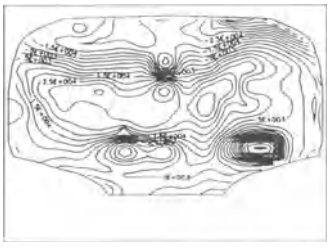


(b) interpolation (g = 0) with equilibrium satisfied



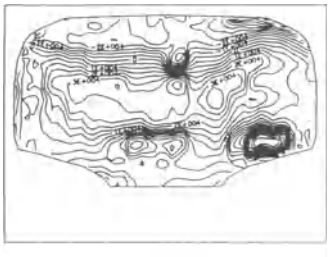
(c) best approximation ($g = g_{OPT}$) with equilibrium satisfied

Figure 2. Surfaces of lateral residual stress σ_{xx} in a 136 RE rail taken from service [9].

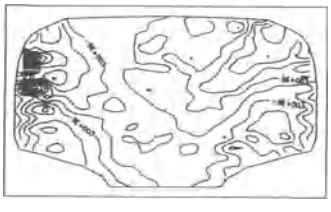


(a) σ_{xx} *interpolation of original data [9]*

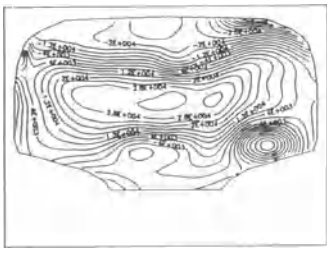
(d) σ_{yy} *interpolation of original data [9]*



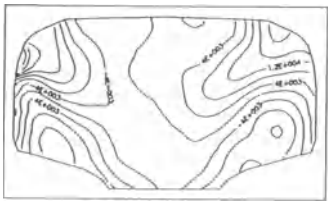
(b) σ_{xx} *interpolation (* $g = 0$ *) with equilibrium satisfied*



(e) σ_{yy} *interpolation (* $g = 0$ *) with equilibrium satisfied*

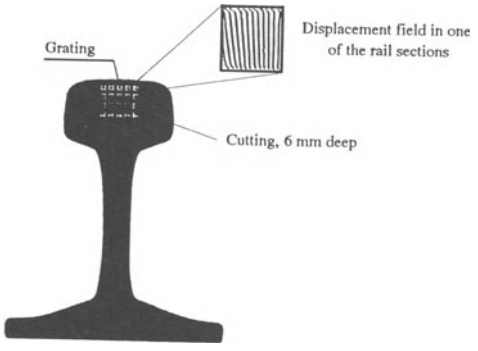


(c) σ_{xx} *best approximation (* $g = g_{OPT}$ *) with equilibrium satisfied*



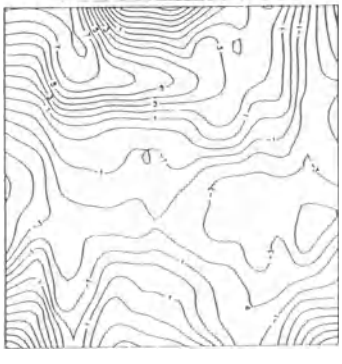
(f) σ_{yy} *best approximation (* $g = g_{OPT}$ *) with equilibrium satisfied*

Figure 3. Contours of lateral residual stress σ_{xx} *and vertical residual stress* σ_{yy} *in a 136 RE rail taken from service [9].*

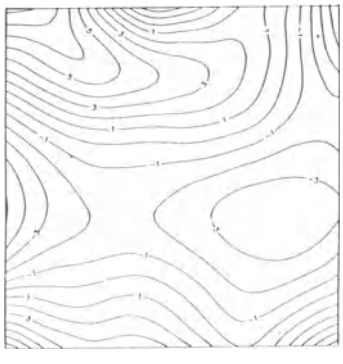


(a) moire grid location

(b) interpolation of original data [10]



(c) interpolation ($g = 0$) with equilibrium satisfied



(d) best approximation ($g = g_{OPT}$) with equilibrium satisfied

Figure 4. Surfaces of lateral residual stress σ_{xx} in a stress-relieved new UIC 60 rail [9].

stresses obtained in the traditional way. Following the benchmark test, we may expect that the level of measurement errors was significantly reduced by the combination of optimal smoothing and satisfying equilibrium.

Final remarks

The most recent version of a global-local method for enhancement of experimental results has been presented. This method has been formulated to work within the general concept of using all available information, i.e., both experimental and theoretical (or even heuristic), in order to obtain the optimum result. It is also worth stressing that the whole procedure is fully automatic.

Both the benchmark test and example applications show that the final optimal results differ essentially from those obtained by the traditional approach. Enforcement of the equilibrium equations and optimal smoothing together reduce the errors from measurement.

In the next stage of the research, the global-local method will be applied to experimental results from laboratory tests in which UIC 60 rails are subjected to controlled loading programs. The design for this series of experiments was presented earlier [11], and the results from the first stage of testing are expected this year. It is expected that the method will also work well when applied to enhancement of results obtained by experimental techniques other than the strain gauge or moire interferometry.

References

[1] Karmowski, W., Magiera, J., and Orkisz, J., "Enhancement of experimental results by constrained minimization", in: *Residual Stress in Rails - Vol. II: Theoretical and Numerical Analyses* (O. Orringer, J. Orkisz, and Z. Świderski, eds.), Kluwer Academic Publishers, The Netherlands, 207-217, 1992.

[2] Karmowski, W. and Orkisz, J., "A physically based method of enhancement of experimental data - concepts, formulation, and application to identification of residual stresses", *IUTAM Symposium on Inverse Problems in Engineering Mechanics*, Tokyo, May 1992.

[3] Karmowski, W. and Orkisz, J., "Fitting of curves and surfaces based on interaction of physical relations and experimental data", *Appl.Math. Modelling 7*, XX-XX (1983).

[4] Karmowski, W., "Determination of stress and strain fields by physically based interpretation of moire patterns", *20 Convegno Nazionale AIAS*, Palermo, September 1991.

[5] Liszka, T. and Orkisz, J., "The finite difference method at arbitrary irregular grids and its application in applied mechanics", *Computers & Structures 11*, 83-95 (1980).

[6] Liszka, T., "An interpolation method for an irregular net of nodes", *Int.J.Num.Meth.Eng. 20*, 1599-1612 (1984).

[7] Shepard, D., "A two dimensional interpolation function for irregularly spaced data", *Proc. 23rd National Conf. A.C.M.*, 517-523 (1965).

[8] Lancaster, P. and Salkauskas, K., *Curve and Surface Fitting*, Academic Press, Calgary, 1990.

[9] Groom, J.J., "Determination of residual stresses in rails", Battelle Columbus Laboratories, Columbus, OH, report no. DOT/FRA/ORD-83/05, May 1983.

[10] Czarnek, R., Lee, J., and Lin, S.-Y., "Moire interferometry and its potential for application to residual stress measurements in rails", in: *Residual Stress in Rails - Vol. I: Field Experience and Test Results*, op.cit., 153-167.

[11] Świderski, Z. and Wòjtowicz, A., "Plans and progress of controlled experiments on rail residual stress using the EMS-60 machine", in: *Residual Stress in Rails - Vol. I: Field Experience and Test Results*, op.cit., 57-66.

EFFECT OF GRINDING STRATEGY ON RESIDUAL STRESS IN THE RAIL HEAD

A.B. Perlman
Tufts University
Medford, MA 02155, U.S.A.

J.E. Gordon and O. Orringer
Volpe National Transportation Systems Center
Cambridge, MA 02142, U.S.A.

Abstract

Three scenarios representing an unground rail, a rail ground by long-train practice to approach conformal contact, and a rail ground for two-point contact were analyzed by means of the constrained energy minimization method to predict residual stress states in the rail head. Only normal loads were considered, an idealization which approximates lubricated conditions. The predicted residual stresses were found to be consistent with field observations of transverse defect propagation. Comparison of the results suggested that grinding to approach conformal contact could minimize the adverse effect of residual stress on defect propagation.

Introduction

Several of the railroads in North America have been conducting trials of scheduled rail grinding programs to control wear and shelling on heavy haul freight lines since circa 1985. Each program involves an unique combination of grinding frequency, depth of material removed, number and configuration of stones, etc. In some of these programs, the ground contour is intended to relieve the gauge corner from loading by splitting wheel/rail contact into two regions: a primary load-bearing zone toward the rail crown, and a flange contact zone on the upper part of the gauge face. On the whole, grinding programs appear to achieve their wear and fatigue objectives.

However, there have been recent reports of some ground rails apparently developing transverse fatigue cracks ("detail fractures") of an unusual character.[1] On one hand, the cracks seem to occur at about the same frequency as one would expect for detail fractures in rails not ground during service. On the other hand, the shells from which these cracks branch appear to originate at points much closer to the gauge corner than the origins of common detail fractures (Figure 1), and their rates of propagation appear to be much faster than the rates which have been measured for common detail fractures.

Earlier work has shown that typical detail fractures, such as the one in Figure 1(a), are driven to propagate in part by the effect of axial residual tension, which develops in the rail head as a consequence of rolling contact fatigue [1]. The apparently faster propagation rate for cracks in rails ground for two-point contact therefore suggests the possibility that there is a greater than usual concentration of tensile axial residual stress in the rail head.

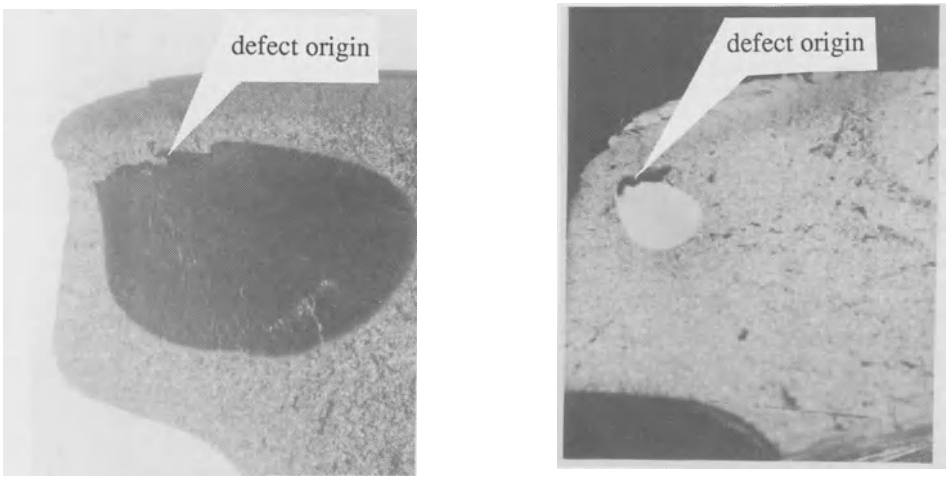
---

1 R.K. Steele, Chief Metallurgist, Chicago Technical Center, Association of American Railroads, private communication, 1991.

J.J. Kalker et al. (eds.), Rail Quality and Maintenance for Modern Railway Operation, 297–306.

The two-point contact grinding practice was originally developed to make the best use of the relatively short grinding trains which were first introduced into railway maintenance. There were believed to be no penalties on wear or shelling control, so long as the two contact points remained well separated. However, residual stresses greater than normal might be expected if the contact points were to approach each other as could happen, for example, in a practice where long intervals are allowed between re-grinding.

This paper presents a comparison of residual stresses predicted for three scenarios simulating the contact conditions associated with ground and unground rails. The chosen scenarios are intended to approximate: (a) an unground rail; (b) a rail ground by long-train practice to approach conformal contact; and (c) a rail ground for two-point contact. The theoretical method for estimating shakedown residual stress states, developed by the Cracow Institute of Technology [2,3], is applied to a finite element model of a rail. The model is subjected to simplified loading programs intended to approximate the three scenarios for heavy haul unit train conditions on well maintained track, and comparative results are presented.



(a) rail not ground *(b) rail ground for two-point contact*

Figure 1. Comparison of detail fractures in unground and ground rails.

Wheel/rail load simulation

For the purpose of the simulations, the track structure is assumed to consist of 132 RE rail of old standard composition with 70 ksi (483 MPa) yield strength on 3000 lb/in/in (21 MPa) foundation to represent typical conditions on North American heavy haul lines. Forces are assumed to be applied to the rail through wheels of 36-inch (0.91-m) diameter with 33-ton (30-tonne) static loads per axle.

Contact scenarios

The contact scenarios are created from normal pressure distributions based on the Hertz theory. However, theoretical calculations make use of the wheel profile radius (R'_2) as well as the rolling radius (R_2). Also, both wheel radii as well as the rail profile radius (R_1) are varied to represent different center-of-contact locations. This is different from the usual design procedure, in which only the wheel tapeline rolling radius and rail crown radius are considered.

Figure 2 summarizes the geometric conditions assumed for the three scenarios. In each case, the left diagram illustrates a typical contact situation, while the right diagram shows the range assumed for center-of-contact locations.

For the unground rail, Figure 2(a), the wheel profile is assumed to be worn to a constant radius of 13.25 inches (336.6 mm), i.e., slightly larger than the 12-inch (304.8-mm) 132 RE rail crown radius. The center of contact is allowed to be located anywhere from the crown to 3/4 of the distance to the gauge face. The latter limit is located at the junction of the nominal blend and corner radii of the rail profile. The rail profile radius actually used in the simulation is assigned a value between the crown and blend radii, as a function of the contact center location. At any location, the applied load is a force normal to the local profile with a magnitude such that the vertical component is 33 kips (147 kN).

For the conformally ground rail, Figure 2(b), the limits for center-of-contact travel and the rail profile radius function are the same as those for the preceding case. However, nearly conformal contact is approximated by taking $1.04\ R_1 \leq |R'_2| \leq R_1$, the larger values being adopted to keep the contact zone size realistic.

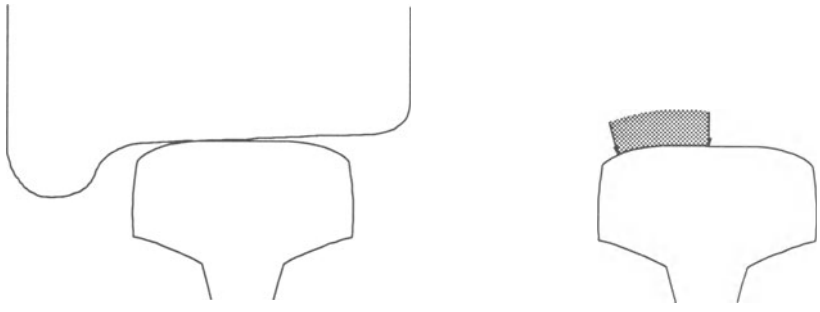
For the two-point contact scenario, Figure 2(c), center-of-contact travel is restricted to 58% of the distance to the gauge face. The wheel and rail profile radii are assigned the same values they would have for the same location in the unground scenario. The second contact zone is placed on the gauge face, just below the rail gauge corner, where the profile radii are assumed to be somewhat larger than the nominal gauge corner radius, and the wheel rolling radius is increased to represent contact on the flange. The load magnitude is the same as in the preceding scenarios, except for the extreme contact center locations.[2] In the latter two positions, the load magnitude is reduced to a value corresponding to a 20 kip (89 kN) vertical component, and a second normal force with a 20-kip lateral component is applied to the gauge face. These adjustments are motivated by empirically based curving load calculations, which suggest that the lateral/vertical load ratio is typically about 0.7 to 0.75, and that about 40 % of the vertical load is borne by friction as a tangent load on the gauge face.[3]

Table 1 summarizes the key contact parameters. These include the radii defined earlier and the angle ϕ between the normal force and the vertical. Values are shown for three primary contact positions, as well as the secondary positions used in the two-point contact scenario.

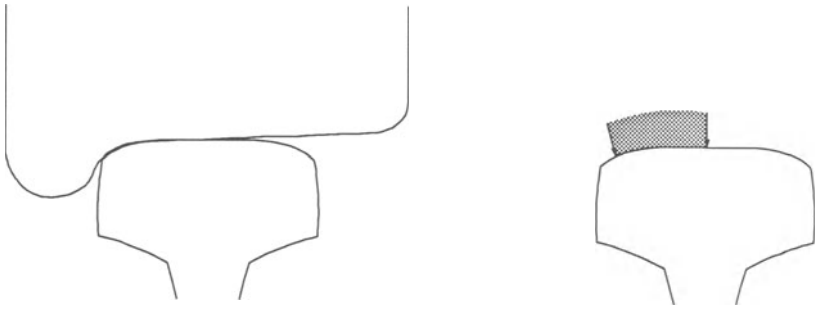
---

2 The center-of-contact position limits are based on visual observations of the surfaces of rails worn under typical North American heavy haul conditions.

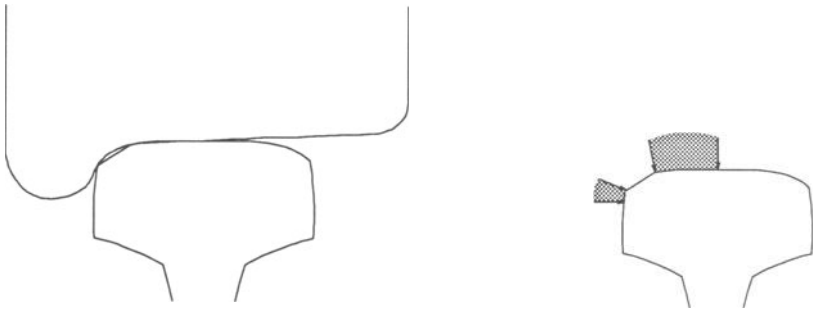
3 Tangent forces on contact zones have been neglected in all cases in these analyses (see the following section).



(a) unground rail



(b) rail ground for conformal contact



(c) rail ground for two-point contact

Figure 2. Contact profiles and center-of-contact bounds.

Table 1. Parameters for contact scenarios.

(a) unground rail

DISTANCE FROM CENTERLINE in (mm)	0 (0)	-0.7500 (-19.0)	-1.1250 (-28.6)
R_1 in (mm)	10 (254)	5 (127)	1.25 (31.8)
R_2 in (mm)	18 (457)	18 (457)	18 (457)
R_2' in (mm)	-13.25 (337)	-13.25 (337)	-13.25 (337)
ϕ	0	0	20.05
p_0/k	3.38	5.82	6.0

(b) rail ground for conformal contact

DISTANCE FROM CENTERLINE in (mm)	0 (0)	-0.7500 (-19.0)	-1.1250 (-28.6)
R_1 in (mm)	10 (254)	5 (127)	1.25 (31.8)
R_2 in (mm)	18 (457)	18 (457)	18 (457)
R_2' in (mm)	-12.5 (-318)	-5.72916 (-146)	-1.30208 (-33.1)
ϕ	0	0	20.05
p_0/k	3.20	3.42	3.74

(c) rail ground for two-point contact

DISTANCE FROM CENTERLINE in (mm)	0 (0)	-1.4373 (-36.5)	-0.7500 (-19.0)	-0.8750 (-22.2)	-1.4402 (-36.6)
R_1 in (mm)	10 (254)	0.5 (12.7)	5 (127)	1.25 (31.8)	0.5 (12.7)
R_2 in (mm)	18 (457)	18.5 (470)	18 (457)	18 (457)	18.5 (470)
R_2' in (mm)	-13.25 (-337)	-0.6875 (-17.5)	-13.25 (-337)	-13.25 (-337)	-0.6875 (-27.5)
ϕ degree	0	81.0	5.3	13.02	81.0
p_0/k	2.86	6.0	5.82	6.0	6.0

Numerical analysis and results

The distribution of residual stress within the rail head results not from a single extreme loading but from an progression of plastic deformations which occur during wheel passages [4]. To simulate the wandering behavior of wheel contact, the loads were applied in sequence at sets of load points that spanned the shaded regions shown in Figure 2. Figure 3 illustrates a typical sequence. Since the residual stresses correspond to estimated shakedown states, the results of these scenarios are independent of the order in which any specific load in a set is applied.

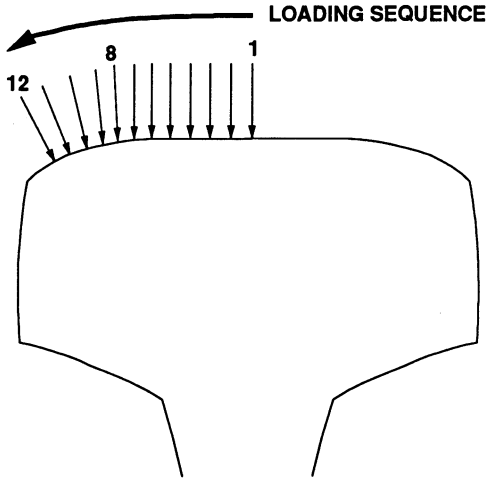


Figure 3. Sequence of load application for unground rail.

The elastic stresses for each loading, as well as the corresponding residual stresses calculated for the previous position, are required to estimate the residual stresses for each state in the sequence. The elastic stresses are computed by adding stresses from contact to the corresponding components caused by bending of a rail represented as a beam on a continuous elastic foundation [1]. While the analysis considers three dimensional states of stress, the model is assumed to be independent of the longitudinal location along the rail so that points in only a single cross section are used in the calculation. The contact stresses are computed numerically by summing the effects of the contact tractions, using the Boussinesq solution for a point load on a half space as a Green's function [5]. To simplify the calculations for this initial study, tangential tractions due to creep and friction have not been included. Only normal tractions, applied over a rectangular region equivalent to a Hertz contact ellipse, have been considered.

Each simulation was performed by calculating residual stress states for a sequence of eleven or twelve load positions from the crown to the left limit of the primary contact center (see Figure 2). The contact parameters were linearly interpolated for positions between those in Table 1. For the simulation of two-point contact, gauge face loading was included when the primary load was located either at the crown or at its leftmost position.

Figure 4 presents the results for the unground rail as a contour plot of axial residual stress after all the loads in the set have been applied. Only the tensile contours are displayed. Comparison of this plot with Figure 1(a) is convincing evidence of the role residual stresses play in the propagation of the detail fracture. The proximity of the largest stresses to the top of the rail head near the gauge corner is consistent with the pattern of flaw growth found in rails modeled by this scenario. Peak values of the estimated stresses are lower than measurements of axial residual stress in rails in this type of service and condition [6]. Higher force magnitudes associated with vehicle dynamic amplification of static loads and the augmented effect that tangential tractions would produce lend credibility to this estimate.

CONTOUR LEVELS

	(MPa)
A	75
B	60
C	45
D	30
E	15
F	0

Figure 4. Contours of axial residual stress for the unground rail

Corresponding results for the conformally ground rail are shown in Figure 5. While the extent of the region of tensile axial residual stress is similar to that of the preceding scenario, there is a conspicuous contrast in the level of these stresses (note the change of stress scales). As one would anticipate, spreading the same loads over larger areas reduces the likelihood that regions in the interior of the head will be plastic. One measure of the load intensity is the ratio of maximum pressure, p_0, to the shear strength of the material k. As Table 1(b) suggests, when p_0/k is relatively low (less than 3.35) no point in the elastic field reaches an octahedral stress[4] as high as the material yield strength. As the intensity increases, additional plasticity is predicted in the head and higher values of peak axial tension ensue.

---

[4] $S_{oct} = \frac{1}{2}\sqrt{(S_1 - S_2)^2 + (S_2 - S_3)^2 + (S_3 - S_1)^2}$

where S_1, S_2 and S_3 are the principal stresses.

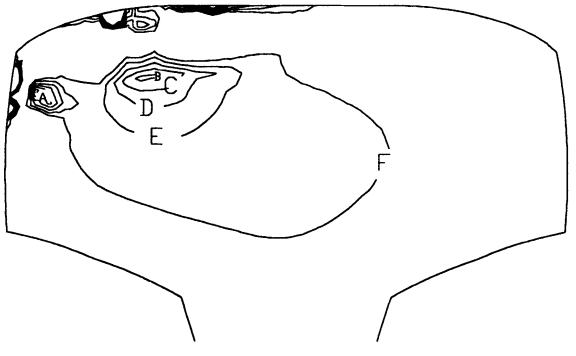
The values of p_0/k associated with these results are substantially higher than those of two dimensional analyses of contact stresses [7]. In the three dimensional case, elastic material is found to the sides as well as behind and ahead of the plastic zone.

CONTOUR LEVELS

	(MPa)
A	25
B	20
C	15
D	10
E	5
F	0

Figure 5. Contours of axial residual stress for the conformally ground rail

When a second contact point is created after grinding a rail head corner, the estimated distribution of axial tensile stress changes character. Comparison of the contours in Figure 6 with those of Figures 4 and 5 illustrates the difference. The presence of normal force on the side of the head creates a tensile pocket with a substantial peak level. The extent of the this region covers an area comparable in size and location to that of the flaw pictured in Figure 1(b).



CONTOUR LEVELS

	(MPa)
A	75
B	60
C	45
D	30
E	15
F	0

Figure 6. Contours of axial residual stress for the rail ground for two-point contact

Discussion and conclusions

Vertical loads used in this investigation have been limited to 33 kips (147 kN), the maximum static wheel load allowed under North American interchange rules. This representation deliberately underestimates extreme conditions, since vertical peak dynamic loads 50 to 100 percent higher than static levels are common on North American freight railway lines. In addition, all effects that transmit load tangentially to the surface of the rail have been ignored. Together, these assumptions imply that the predicted residual stresses should be lower than the actual levels found in rail heads. However, inferences drawn from comparison of the scenarios are valid. The effects of higher load levels would increase the stress magnitudes in all of the scenarios without changing the relations between them.

With this qualification in mind, particular attention should be focused on some details in the results. For example, there were no differences between the residual stresses estimated when load was applied to the gauge face, in conjunction with a vertical load on the crown, or when the gauge face load was associated with the nearest primary load. These calculations suggest that the loading conditions assumed in these simulations provide no analytical evidence of a strong influence of load proximity on the nature of residual stresses in rail heads, except as noted below.

Peak magnitudes of axial residual tension in the interior of the head are as large in the two-point contact scenario as they are for unground rail. There is, however, a distinction between the two cases in the location and shape of the tensile "pocket". A load of only 20 kips (89 kN) acting on the side of the head does as much damage to the rail as one of 33 kips (147 kN) that is applied on the top. The wheel and rail radii that intensify the gauge face load are not known with sufficient precision to fix a quantitative limit on this effect. However, it is obvious that the radii of curvature for both will be relatively small when contact occurs near the gauge corner.

The simulation of two-point contact represented the rail as ground, not a ground rail that has since worn and allowed the two loads to approach more closely. It is likely that such proximate loads will amplify the residual stress. One may expect, therefore, that larger stresses will be found when grinding is infrequent, and that rails will have a shorter detail fracture life under such conditions. While the results to date are preliminary, it does appear that conformal contact is a better goal than two-point contact, at least from the viewpoint of protection against rail failure.

Acknowledgement

This work was sponsored by the Office of Research and Development, Federal Railroad Administration, United States Department of Transportation.

References

[1] Orringer, O., Morris, J.M., and Jeong, D.Y., "Detail fracture growth in rails", *Theoretical and Applied Fracture Mechanics 5*, 63-95 (1986).

[2] Orkisz, J., "Prediction of actual residual stresses by constrained minimization of energy", in: *Residual Stress in Rails - Vol. II: Theoretical and Numerical Analyses* (O. Orringer, J. Orkisz, and Z. Swiderski, eds.), Kluwer Academic Publishers, The Netherlands, 101-124, 1992.

[3] Holowinski, M. and Orkisz, J., "Hybrid finite element method for estimation of actual residual stresses", in: *Residual Stress in Rails - Vol. II*, op.cit., 125-149.

[4] Swiderski, Z. and Wójtowicz, A., "Plans and progress of controlled experiments on rail residual stress using the EMS-60 machine", in: *Residual Stress in Rails - Vol. I*, op.cit., 57-66.

[5] Perlman, A.B. and Gordon, J.E., "Application of the constrained minimization method to the prediction of residual stresses in actual rail sections", in: *Residual Stress in Rails - Vol. II*, op.cit., 151-177.

[6] Groom, J.J., "Determination of residual stresses in rails", Battelle Columbus Laboratories, Columbus, OH, FRA/ORD-83/05, May 1983.

[7] Johnson, K.L., *Contact Mechanics*, Cambridge University Press, pp 288-291, 1985.

RESIDUAL STRESS MEASUREMENTS IN RAILS BY NEUTRON DIFFRACTION

P J Webster, G Mills, X Wang and W P Kang

Department of Civil Engineering, University of Salford, Salford, M5 4WT, UK

ABSTRACT

The relatively new technique of neutron strain scanning has several unique advantages over most traditional methods of measuring internal strains. It is non-destructive and, in principle, measurements can be made and repeated at any point in any direction within a sample. In practice its range of its application is limited in thick section samples, such as long lengths of rail, by neutron beam attenuation. In such cases it is often necessary to reach a compromise between what is ideally preferred by engineers for use in stress analysis calculations and what it is practical or economic to measure using the technique. Examples are given of how neutron strain scanning may be applied, in a cost effective manner, to the problem of the measurement of residual stresses in railway rails. Results are presented of representative longitudinal, transverse and vertical stresses measured down the centre-lines of rail sections and of stress contours measured in rail heads. The data are of sufficient quality and quantity to be used to validate theoretical calculations and reveal details of the residual stress distributions in rails not obtainable by other methods.

INTRODUCTION

Residual stresses, which are introduced into railway rails during the manufacturing process and are then subsequently modified in service, have an important influence on the rate and mechanisms of rail failure. The combination of the residual stresses with repeated applied loads from traffic in service can cause rolling contact fatigue which may result in cracking, flaking and possibly eventual rail fracture. An accurate knowledge of the residual stresses in rails, and their origins and evolution, is important for the safety and economic utilisation of rail transport. Stresses calculated using computer models and numerical techniques are being increasingly used for predictive purposes but require high quality experimental data for their validation. Traditional strain relaxation measuring techniques are often inadequate for this purpose, being destructive and restricted to a small number of surface measurements. Neutron strain scanning is a non-destructive technique capable of providing high quality three-dimensional internal strain data.

NEUTRON STRAIN SCANNING

The neutron technique uses the principles of diffraction to measure the spacings between atomic planes in materials. In crystalline materials the constituent atoms are regularly

J.J. Kalker et al. (eds.), Rail Quality and Maintenance for Modern Railway Operation, 307–314.

arranged on the crystal lattice which has dimensions that are very precisely defined at a any particular temperature. If, however, the material is strained the lattice will be distorted. In a region of tensile strain the atoms will be further apart, in the direction of the strain, than the equilibrium unstrained value, and in a compressed region they will be closer together. In neutron strain scanning the crystal lattice is used as an internal three-dimensional strain gauge which is "read" at selected points, in chosen directions, using neutrons [1-8].

When a beam of neutrons is incident upon a polycrystalline material, such as rail steel, the neutrons are "Bragg" scattered in certain specific directions defined by the Bragg equation:

$$2d\sin\theta = \lambda \qquad (1)$$

where **d** is the interplanar spacing, λ is the neutron wavelength and 2θ is the angle through which the neutrons are scattered. If the material is strained, so that the lattice dimension is changed by δd, the strain ε is given by:

$$\varepsilon = \delta d/d = -\delta\theta.\cot\theta \qquad (2)$$

In a strain scanner a "gauge volume" is defined by collimation of the incoming and outgoing neutron beams and the strains at any point within a component are measured by scanning the component through the gauge volume as indicated in figure 1.

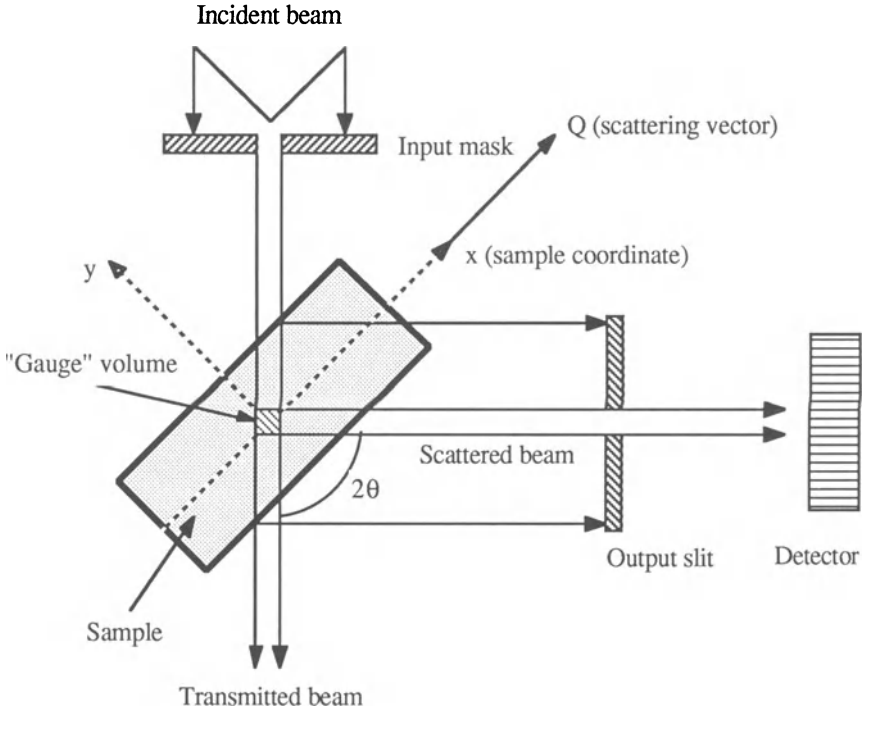


Figure 1. Principles and outline of the neutron strain scanner.

MEASUREMENTS

Measurements were made on samples of new and used head-hardened, roller-straightened rail designated 370 BHN, made by Thyssen Stahl for the ORE D 173 test programme. The used rail had been installed in a region of curved track at a test site in Poland and had been subjected to about 60 Mtonnes of traffic before removal [9].

Neutrons can penetrate several centimetres into most engineering materials but the beam attenuation is exponential and large thicknesses require excessively long counting times. For steel every 20 mm attenuates the beam by a factor $\times 0.1$. As the neutron path length through the thick region of a rail head would be about 100 mm the attenuation there would be $\times 0.00001$. Consequently an alternative approach was employed in this investigation. Transverse and longitudinal sections, nominally 10 mm thick as shown in figure 2, were cut from samples of new and used rail. The transverse slices were used to determine the transverse and vertical strains across the full section, and the plate for the longitudinal strains down the centre line. The measured strains are not exactly those appertaining to the original rail, but they are representative and do give a semi-quantitative measure and reveal substantial qualitative detail not otherwise observable [9-12].

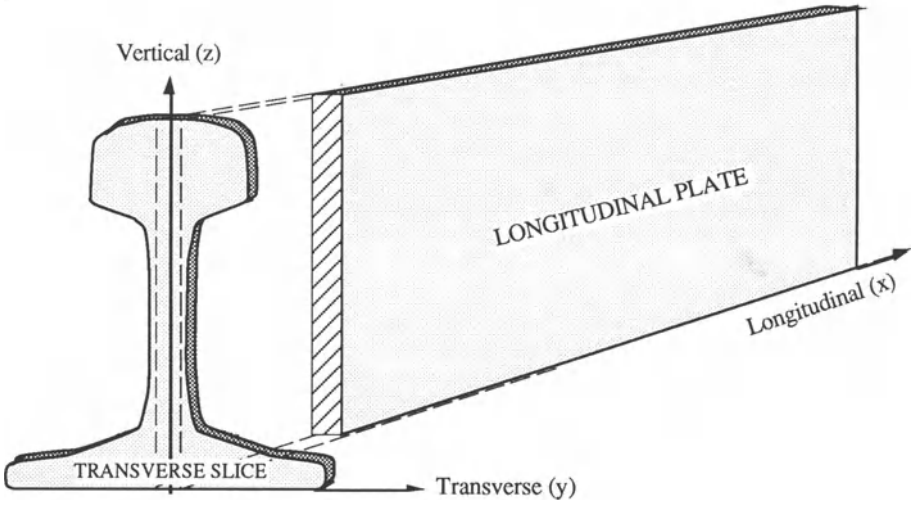


Figure 2. Transverse slice and longitudinal plate samples cut from a rail.

A gauge volume of $2\times2\times2$ mm^3 was used and series of scans were made at the positions in the rails shown in figure 3. Stresses were calculated from the strains, using a Young's Modulus of 209 GPa and a Poisson's Ratio 0.3, assuming that the symmetry directions x, y and z were coincident with the principal stress directions.

RESULTS

Longitudinal stresses from neutron data are compared with equivalent sectioning results in figure 4. Good agreement is observed at the surfaces but there are differences internally.

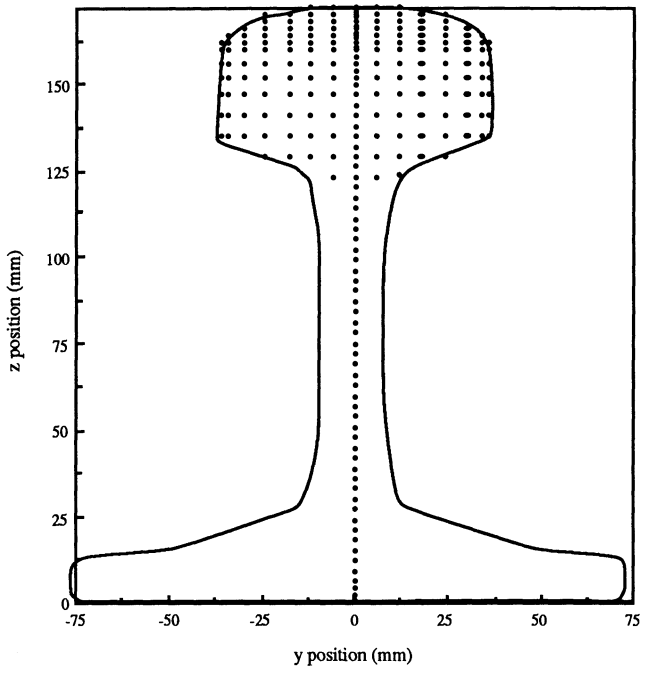


Figure 3. Strain scan locations on slice and plate samples.

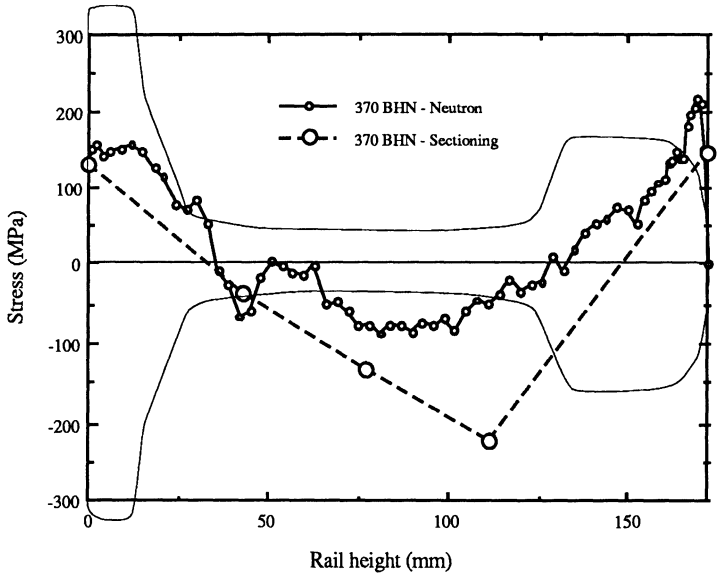


Figure 4. Longitudinal stresses in new 370 BHN rails measured by neutron and sectioning methods.

The most apparent differences between the neutron and strain gauge results are the relative number of data points, the possibility to choose the accuracy of the data and the capacity to repeat measurements. The fine detail shown by the multiplicity of points on the neutron scan is not revealed by the mechanical method.

Figure 5 shows the centre-line stresses measured in the 300 mm long plate sections cut from new and used 370 BHN head-hardened, roller-straightened rails. Both show similar tensile regions in the foot. The webs are mostly in compression with, at the lower end, a marked compressive peak but with differences in detail in the centre. In the heads there is low stress near the top surface but tension in the interior. In the new rail there is a region of high tensile stress just below the surface that begins to diminish at a depth below about 4 mm. In the used rail the significantly tensile region starts at a depth of about 4 mm but remains at a high value over most of the interior of the head.

Figures 6 and 7 show the corresponding centre-line stresses measured in the transverse and vertical directions in the transverse slices. Once again the corresponding new and used patterns are generally similar but there are differences in detail. In particular, in the used rails the tensile stresses in the heads are greater than in the new rails.

Figures 8 and 9 show the substantially symmetrical two-dimensional stress contours derived by interpolation from a series of parallel scans across the head of the new rail slice. The compressive layers at the top surface are probably largely due to the head-hardening process. The shapes of the internal tensile regions are broadly similar to those observed in other new rails and are probably more related to the original cooling regime than the subsequent head-hardening or roller straightening.

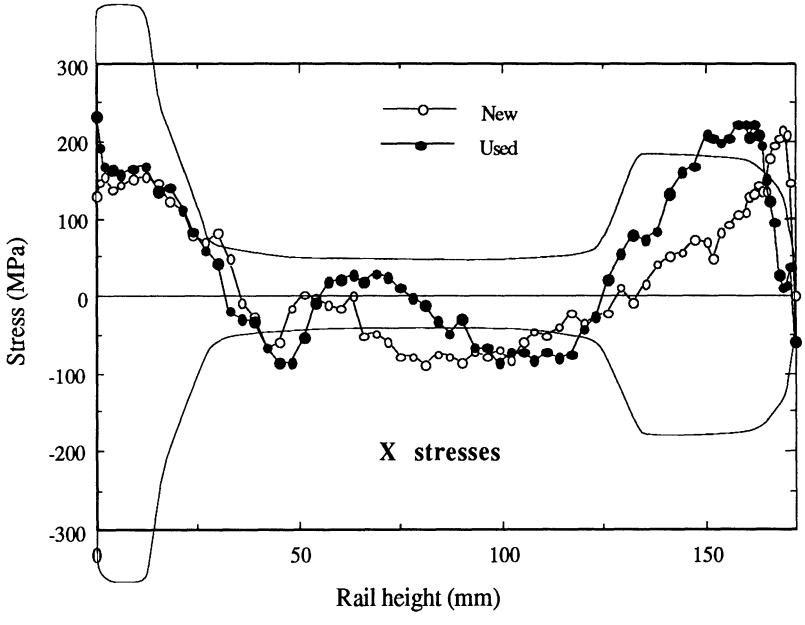


Figure 5. Longitudinal residual stresses in new and used 370 BHN rail plates.

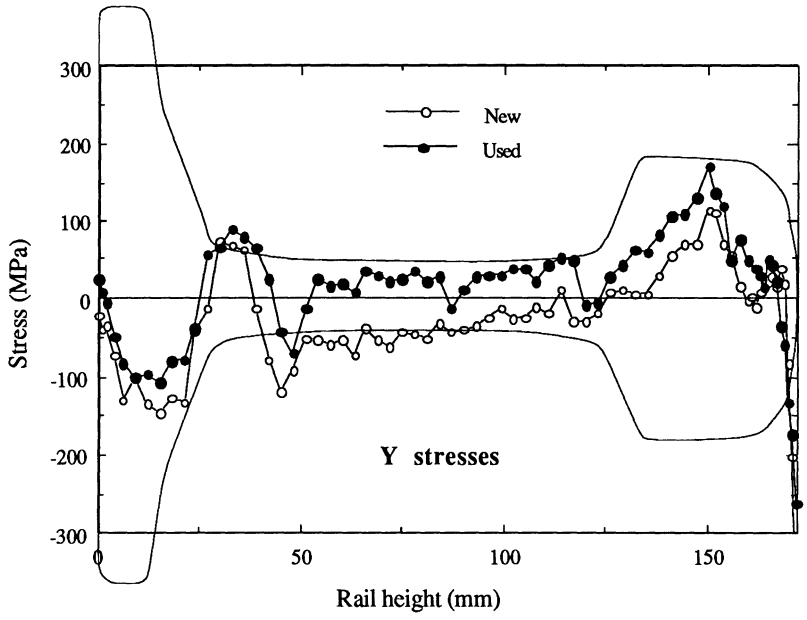


Figure 6. Transverse residual stresses in new and used 370 BHN rail slices

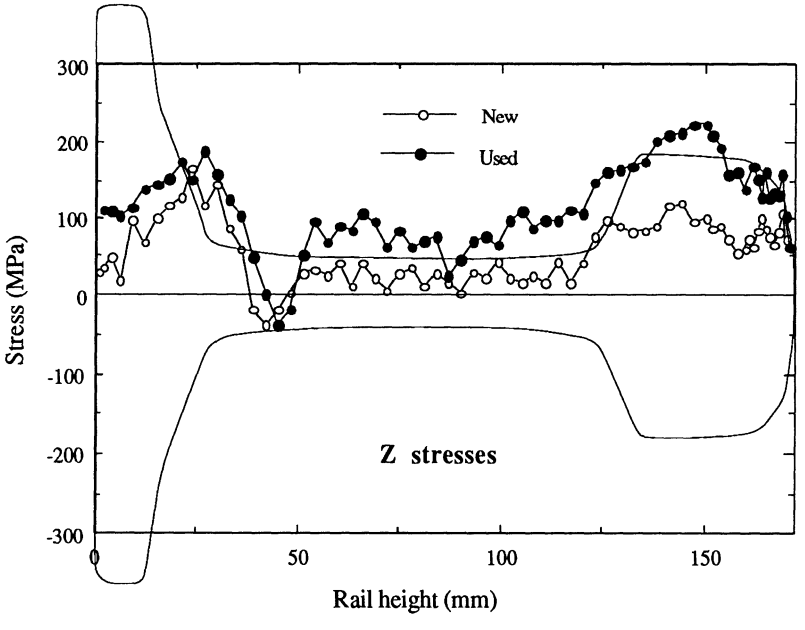


Figure 7. Vertical residual stresses in new and used 370 BHN rail slices

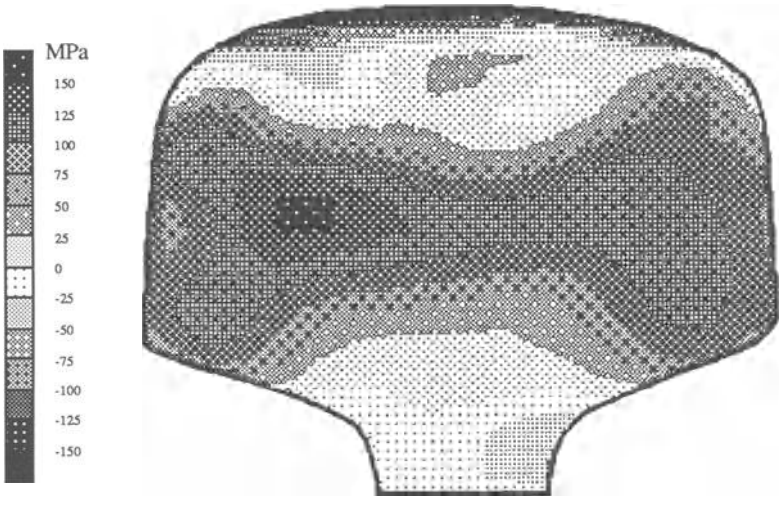


Figure 8. Transverse residual stresses in a new 370 BHN rail slice

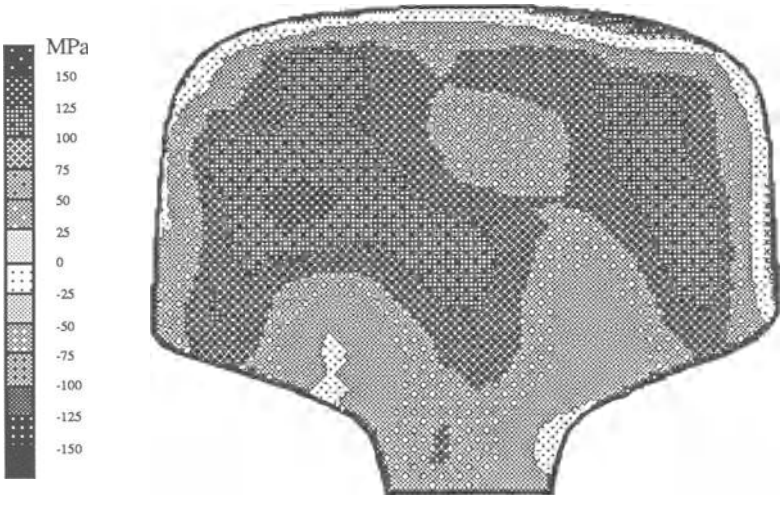


Figure 9. Vertical residual stresses in a new 370 BHN rail slice

DISCUSSION AND CONCLUSIONS

Examples have been shown of residual stresses in railway rail sections derived from neutron strain scanning measurements. Although the technique is still developing it is rapidly becoming established as a routine engineering procedure with unique capabilities, providing, non-destructively, detailed internal strain data which is suitable for validation of finite element codes. Although the technique is restricted when applied directly to thick samples selective sectioning enables significant semi-quantitative detail to be obtained. Selected confirmatory strain measurements may then be made economically in near surface regions of the head or through thinner regions such as the web or foot of a rail.

ACKNOWLEDGEMENTS

The work discussed here has utilised facilities at the ILL, Grenoble. The railway rails were supplied through British Rail on behalf of ORE as part of a Co-operative project with Imperial College. Financial support was provided by the UK Science and Engineering Research Council via grants GR/D63196, GR/E55556 and GR/F02427.

REFERENCES

1. Allen, A., Andreani, C., Hutchings, M. T. and Windsor, C. G. (1981) 'Measurement of internal stress within bulk materials using neutron diffraction' NDT International 14, 249-254.
2. Pintschovius, L., Jung, V., Macherauch E. and Vohringer, O. (1983) 'Residual stress measurements by means of neutron diffraction' Mater. Sci. Eng. 61, 43-50.
3. Allen, A., Hutchings, M. T., Windsor, C. G. and Andreani, C. (1985) 'Neutron diffraction methods for the study of residual stress fields' Adv. Phys. 34, 445-473.
4. Stacey, A., MacGillivray, H. J., Webster, G. A., Webster, P. J. and Ziebeck, K. R. A. (1985) 'Measurement of residual stresses by neutron diffraction' J. Strain Analysis 20, 93-100.
5. Krawitz, A. D. and Holden, T. M. (1990) 'The measurement of residual stresses using neutron diffraction' MRS Bull. XV, 57-64.
6. Lorentzen, C. (1988) 'Non-destructive evaluation of residual stresses by neutron diffraction' NDT International 21, 385-388.
7. Webster, P. J. (1990) 'The neutron strain scanner: a new analytical tool for engineers' Steel Times 218 No. 6, 321-323.
8. Webster, P. J. (1991) 'Neutron strain scanning' Neutron News 2, 19-22.
9. Webster, P. J., Wang, X., Mills, G., Kang, W. and Webster, G. A. (1991) 'Residual stress measurements on ORE Railway Rails by neutron diffraction' Final Report BRSUIC02/91, September 1991.
10. Webster, P. J., Low, K. S., Mills, G. and Webster, G. A. (1990) 'Neutron measurement of residual stresses in a used railway rail' Mat. Res. Soc. Symp. Proc. 166, 311-316.
11. Webster, P. J., Wang, X. and Mills, G. (1991) 'Problems with railway rails' NATO Advanced Research Workshop on Measurement of Residual Stress using Neutron Diffraction, Oxford, March 1991, NATO-ASI-E Series, Kluwer, Holland (In press).
12. Webster, P. J., Wang, X., Mills, G. and Webster, G. A. (1991) 'Residual stress changes in railway rails' International Conference on Neutron Scattering, ICNS'91, Oxford, August 1991, Physica B (In press).

Neutron Diffraction Measurements of Residual Stress in Rails

J.H. Root, T.M. Holden, R.J. Klassen
AECL Research, Chalk River, Ont. Canada, K0J 1J0

C. Smallman
Nova Corp., PO 2535, Calgary, Alt., Canada, T2P 2N6

B. Maxfield
Industrial Sensors and Actuators
400 Hester St., San Leandro, CA, USA 94577

N.R. Gore
Transportation Development Centre
200 René Lévesque Blvd. W, West Tower, Suite 601
Montreal, Que., Canada, H2Z 1X4

ABSTRACT

Neutron diffraction has been applied as a non-destructive probe of residual stresses in rails. As a first example, the preparation of rails for fracture-toughness testing by introducing a saw-cut through the rail web redistributes the residual stresses in the original rail. The resulting stress profile has been measured by neutron diffraction. The uniaxial longitudinal compression of -110 MPa at mid-height of an unmodified rail was altered by superimposing a biaxial stress state near the tip of the saw-cut as the tensions in the head and foot of the rail were relieved. The vertical component of stress reached a tensile value of 220 MPa, tending to open the saw-cut further, and the longitudinal component of stress reached a tensile value of 120 MPa. The decay of the stress concentration was consistent with the expected inverse square-root behaviour, falling to zero approximately 30 mm from the tip of the saw-cut in a 3 m section of rail. As a second example, ultrasonic shear-wave birefringence is being developed as a technique to assess residual stresses in rails. The difference in the principal stress components perpendicular to the direction of sound propagation has been measured in 38-mm-long sections of rail head. Neutron diffraction measurements, which determined the three principal stress components along two axes within the sections, verified the ultrasonic results.

J.J. Kalker et al. (eds.), Rail Quality and Maintenance for Modern Railway Operation, 315–324.

INTRODUCTION

Thermal neutrons are produced by nuclear research reactors. Diffraction from a monochromating crystal produces a neutron beam with a single wavelength, λ, suitable for lattice-spacing measurements on crystalline materials, analogous to measurements made by X-ray diffraction. Unlike X-rays, however, neutrons penetrate easily through most engineering materials, so it is possible to map out the lattice spacing as a function of position inside intact objects. Even after penetrating through 20 mm of iron, 10% of the neutron beam intensity remains to produce a measurable diffraction peak.

With masks that are made from neutron-absorbing cadmium, incident and diffracted neutron beams may be defined with small cross sections. The volume of space defined by the intersection of the incident and diffracted beams is denoted as the sampling volume. Setting the object at a number of positions, by a computer-controlled XYZ translation system, permits the scanning of the sampling volume throughout the volume of the object. The spatial variations of diffraction-peak parameters, such as the mean scattering angle, 2θ, and the integrated intensity, may then be determined. These parameters are obtained by fitting a model function, a gaussian peak with a linear background, to the raw data of neutron counts versus scattering angle.

Residual stresses arise from spatial variations of plastic deformation in an object with a particular thermo-mechanical fabrication history. The stresses cause slight compressions or expansions of crystal lattice spacings, d(hkl), which are determined from careful measurements of 2θ, through Bragg's law,

$$d(hkl) = \lambda / 2\sin(\theta) \qquad (1).$$

Comparing the lattice spacing at a point within the object to the value, $d^{o}(hkl)$, measured in a stress-free specimen of the same material, one defines the elastic strain, ε, as

$$\varepsilon = d(hkl)/d^{o}(hkl) - 1 \qquad (2).$$

The strain is determined in a direction that is parallel to the bisector of the incident and diffracted neutron beams. To evaluate the stress at a point in a rail, measurements of strain in at least three directions must be made. Figure 1 shows the orientations of a rail segment necessary to determine strains in three directions: longitudinal (L), which is parallel to the long dimension of the rail, normal (N), which is perpendicular to the surface of the rail web, and vertical (V). Having measured the three strain components,

ε_L, ε_N and ε_V, the stresses may be calculated through

$$\sigma_L = \frac{E}{(1+\nu)} \left[\varepsilon_L + \frac{\nu}{(1-2\nu)} (\varepsilon_L + \varepsilon_N + \varepsilon_V)\right] \quad (3),$$

with cyclic permutation of the subscript. This stress equation assumes that L, N and V are principal directions of stress and that the material may be treated as a homogeneous

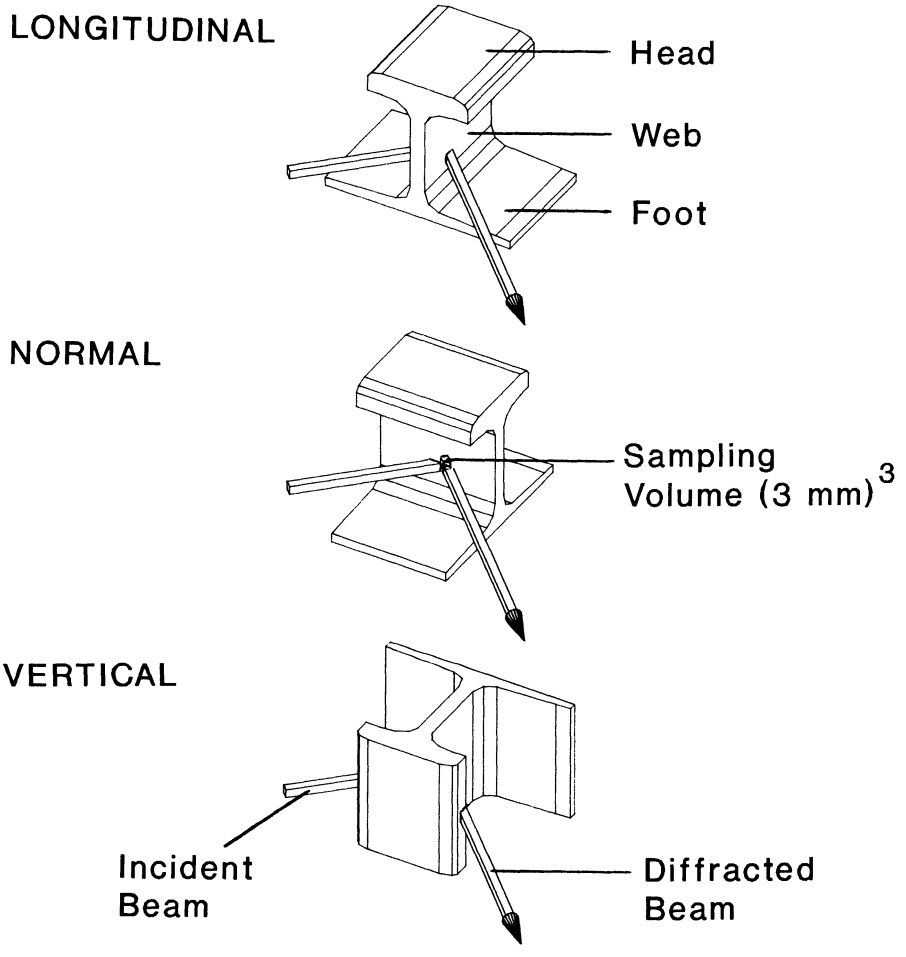


Fig. 1 Orientations of a rail with respect to the neutron beams to measure strains in the longitudinal, normal and vertical directions.

elastic continuum. The elastic constants, Young's modulus, E, and Poisson's ratio, ν, depend on the choice of diffraction line that is studied,[1] but in cubic materials, the (112) and (110) reflections exhibit elastic responses that are within 3% of bulk values.

STRESS NEAR A SAW-CUT

Residual stress distributions have been linked to longitudinal web fractures in high-strength roller-straightened rails.[2] Making a saw-cut through the rail web allows the longitudinal residual stresses, present in the unmodified rail, to relax and redistribute. In particular, a tensile vertical component of stress appears near the tip of the saw-cut and tends to promote longitudinal crack propagation.[3] The characterization of the residual stress distribution in a rail containing a saw-cut is readily achieved by the neutron diffraction method.

Experiment

A rail segment 3 m long, with a web thickness of 9 mm, was supplied by the Sydney Steel Corporation.* At one end of the rail, a 76-mm-long saw-cut was made at mid-height of the web, 86 mm from the rail foot. Strain measurements, with a $(3\ \text{mm})^3$ sampling volume, were made along a longitudinal locus at mid-thickness of the web, beginning at the tip of the saw-cut. Measurements were also made, 180 mm from the unmodified end of the rail, to assess the pre-existing residual stresses at mid-height of the web. According to finite-element predictions,[4] this position is sufficiently far from the end of the rail to avoid the end-effects that are expected in the residual stress distribution. All strains were evaluated from shifts in 2θ of the (112) diffraction peak. To determine 2θ and $d^{\circ}(hkl)$ of the stress-free material, measurements were made in a small cube, extracted from an offcut of the same rail. Cutting the small cube away from its surrounding material was assumed to remove its residual stresses.

The three principal components of residual stress are presented in Fig. 2. At 180 mm from the unmodified end of the rail, measurements were made at four positions through the thickness of the web, all at mid-height. No significant through-thickness variation in stress was found, so all the data were averaged to obtain the values that are shown on the figure as open symbols. The error bars on these data

*Sydney Steel Corporation, Sydney, Nova Scotia, Canada

are twice the standard deviations of the averages. A uniaxial longitudinal compressive stress of value -110 ± 15 MPa is observed. This level of compressive stress at mid-height of the web is in the middle of the range observed in roller-straightened rails[3] and is consistent with finite-element predictions.[4]

Approaching the tip of the saw-cut, at mid-thickness of the web, a marked stress concentration is shown in Fig. 2. There is a tensile vertical component of stress,

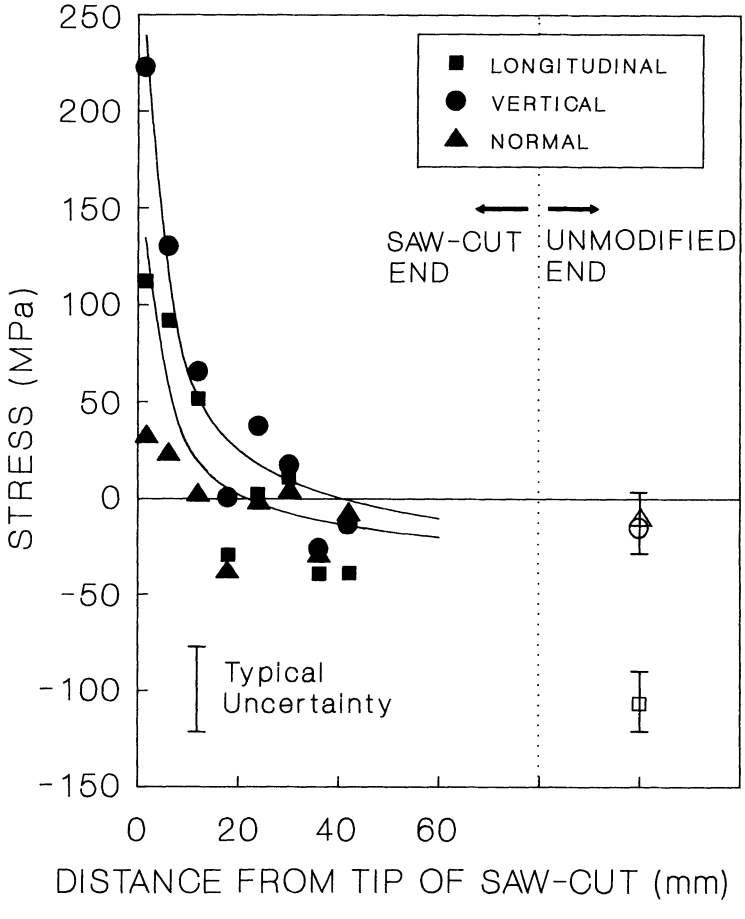


Fig. 2 Three components of residual stress at mid-height and mid-thickness of the web of a 3-m rail segment.

+220 ± 25 MPa. The vertical stress is expected, since the saw-cut permits relaxation of the tensile stresses that existed, in the unmodified rail, in the head and foot. The head and foot regions peel away from the saw-cut until a tensile vertical stress builds up in the web to oppose the change in shape. Vertical tension tends to open horizontal web cracks that originate at the tip of the saw-cut, and thus reduces the fracture toughness of the modified rail. The longitudinal component of stress also shifts from compression, in the unmodified rail, to a tension of +120 ± 25 MPa near the tip of the saw-cut. The normal stress is still zero, within the uncertainty of measurement. The typical uncertainty is shown in Fig. 2 as a bar of length twice the standard error that is traced, through Eqs. 2 and 3, to the fitting of a gaussian peak to the raw data. Continuous lines in the figure are best-fits of a $1/\sqrt{x}$ function to the longitudinal and vertical stresses at distances, x, from the tip of the saw-cut. The functions represent the respective data, within the experimental precision.

Discussion

The tasks of characterizing residual stress in the webs of roller-straightened rails and determining the stress concentration near a saw-cut through a rail web are amenable to neutron diffraction analysis. Since the material strongly scatters neutrons and is fine-grained, the neutron diffraction measurements are rapid, requiring between 15 and 30 minutes for each strain measurement. No special preparation of the surfaces nor labour-intensive application of strain gauges is required. Since neutron diffraction is non-destructive, it is possible to perform subsequent tests on a rail whose residual stress profile has been characterized. Neutron diffractometers are inherently large in scale, because of the shielding needed to reduce background radiation fields. With much space available for manipulation of large samples, rail segments of a length sufficient to avoid end effects can be examined in all of the orientations necessary to obtain a complete set of strain measurements. The algebraic calculation of stress is straightforward, because the principal components of strain can be measured directly.

Practical limitations of the neutron diffraction method arise from the low intensity of neutron beams. The ratio of signal to noise and the limited availability of neutron instruments limit the precision of stress determinations to a level of ± 15-20 MPa and the spatial resolution to volumes of at least $(2\ \text{mm})^3$. These limitations nevertheless permit the collection of much relevant information to benefit underlying research and process development.

VERIFYING ULTRASONIC STRESS DETERMINATIONS

Because specimens must be transported to a neutron scattering facility, neutron diffraction is not a practical tool for the characterization of residual stresses during fabrication or service. One non-destructive technique that is portable and provides information on residual stress as a function of position in rails is based on ultrasonic shear-wave birefringence.[5] Since neutron diffraction unambiguously measures lattice strains and hence residual stress, it can verify that the ultrasonic method reliably separates residual stress information from other material properties that influence ultrasonic waves, including texture and microstructure.

Experiment

The head of a roller-straightened rail was removed from the web and subdivided into 38-mm-long segments. One segment was further machined to produce cylindrical specimens 12 mm in diameter and 30 mm long, parallel to the longitudinal direction of the original rail. These cylinders were assumed to be stress-free for the determination of $d^{o}(hkl)$ of the material. In addition, one of the cylinders was examined in detail, to establish that the rail head contained no significant preferred crystallographic orientations of grains. Neutron diffraction strain measurements were made with a sampling volume defined by beams 3 mm wide and 5 mm high. Measurements were made on a horizontal locus at mid-length and mid-height of the rail head. All strains were evaluated from shifts, in 2θ, of the (110) diffraction peak.

The three principal components of residual stress are presented in Fig. 3a. The typical uncertainty in the stress values is shown as a bar that is twice as long as the standard error associated with fitting the gaussian model to the raw diffraction-peak data. A triaxial compression of -100 ± 25 MPa occurs near the middle of the rail head, but all stress components approach zero towards the edges. The residual stresses in the head of an intact rail are expected to differ from those measured in this specimen, which has been removed from many of its constraints.

The difference in velocities of orthogonally polarized ultrasonic shear-waves is proportional to the difference in the stress components that are perpendicular to the direction of wave propagation. Comparisons are made with a stress-relieved reference sample, to eliminate the effects of texture and microstructure on the wave velocities. For this experiment, an electromagnetic transducer created a

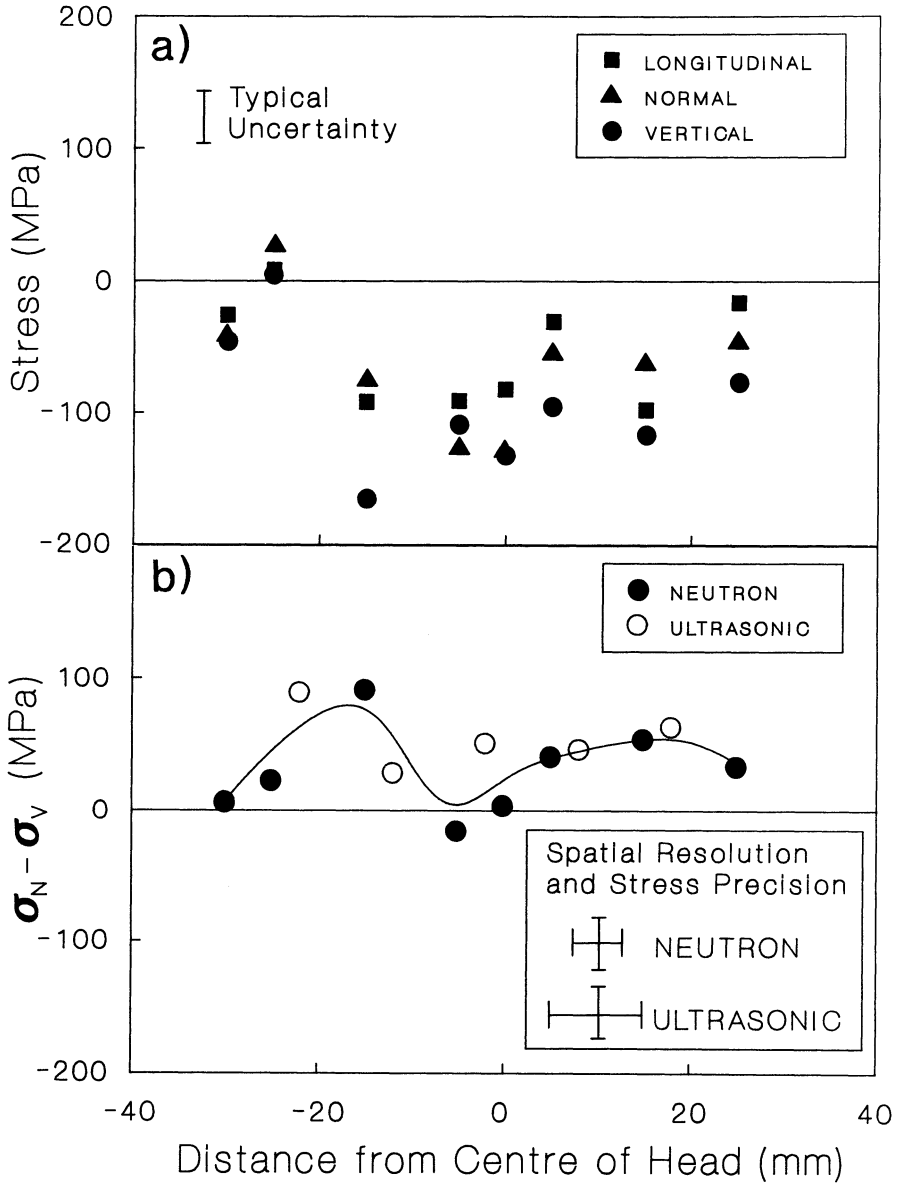


Fig. 3 (a) Measurements of stress in a rail head by neutron diffraction. (b) The difference between normal and vertical stress components by two methods. The continuous line is a guide to the eye.

shear wave that propagated in the longitudinal direction of the rail head segment. The difference between the normal and vertical stress components, $\sigma(N) - \sigma(V)$, was determined. A comparison with the difference, $\sigma(N) - \sigma(V)$, measured by neutron diffraction is shown in Fig. 3b. The results of the two methods agree within the experimental precisions. However, the neutron method gives absolute stresses, which directly influence the performance of the rail, while the ultrasonic method yields only a difference in orthogonal stress components, which could be misleading. For example, although there is a compressive stress near the mid-point of the rail head, $\sigma(N) - \sigma(V)$ is nearly zero.

Discussion

The stress-measurement precision of the two methods was comparable, but the spatial resolution of the ultrasonic technique was not as good as the neutron diffraction method. With the neutron method, longitudinal variations of the stress could be investigated, while the ultrasonic probe integrated over the longitudinal dimension. More development is needed before the ultrasonic method will serve as an in-factory or in-service, non-destructive probe of residual stress in rails. However, the agreement between the two techniques gives confidence that such a development would be worthwhile.

CONCLUSIONS

Neutron diffraction is a powerful tool for determining the volume distribution of residual stresses in rails. The stress-measurement precision and spatial resolution are well suited to the typical stress patterns that occur in rail sections. It is therefore appropriate to apply neutron diffraction data to check finite-element calculations of stress, or to determine the typical stress state arising from a particular fabrication route. Since the technique is non-destructive, the development of residual stresses at various stages of fabrication of a single specimen could also be monitored to gain an understanding of this subject. Neutron diffraction is suitable to validate other techniques that promise to measure residual stresses non-destructively in the field or at the point of fabrication. One can conclude, for instance, that the measurement of differences in the principal components of stress by ultrasonic shear-wave birefringence is correct, so that further development of an ultrasonic probe to measure residual stresses conveniently in rails is warranted.

REFERENCES

1. H. Behnken and V. Hauk, *Z. Metallkde.* 77 (1986) 620-625.

2. O. Orringer and P. Tong, "Investigation of catastrophic failure of a premium-alloy railroad rail", *Fracture Problems in the Transporation Industry*, Proc. ACSE Conference, Detroit, MI (1985).

3. S.J. Wineman and F.A. McClintock, *Theoretical and Applied Fracture Mechanics* 8 (1987) 87-99.

4. S.J. Wineman and F.A. McClintock, *Theoretical and Applied Fracture Mechanics* 13 (1990) 29-37.

5. B.W. Maxfield and A. Kuramoto, TP 10716E, "Residual Stress Measurement in Railroad Rails - Phase I", Industrial Sensors and Actuators, 400 Hester St., San Leandro, CA, USA 94577, March 1991.

ACKNOWLEDGEMENTS

The measurements of residual stress at the NRU Reactor of AECL Research benefit greatly from the work of L. McEwan, who designs and constructs customized fixtures to attach samples to the neutron diffractometer. The multiwire detector, which makes possible the rapid measurement of strain, was designed and constructed by D. Tennant, J. Bolduc, G. Sims and H. Spenceley.

RESIDUAL STRESSES AND CATASTROPHIC RAIL FAILURE

J.O. Igwemezie, Ph.D., P.Eng., President,
Applied Rail Research Technologies Inc.,
7030 Woodbine Ave., Markham, Ontario, Canada, L3R 6G2

S.L. Kennedy, BSc. Eng., Researcher,
Canadian Institute of Guided Ground Transport,
Queens University, Kingston, Ontario, Canada, K7L 3N6

N.R. Gore, C. Eng, Ing, Senior Development Officer,
Transport Development Centre, Transport Canada,
200 Rene Levesque Blvd., Montreal, Quebec, Canada, H2Z 1X4

ABSTRACT

Problems associated with internal stresses (otherwise referred to as residual stresses) introduced during the making of rails have been evident for the past decade. The fact that few derailments have been attributed to it stems more from the lack of understanding of the phenomenon than from its absence in the causality of derailments. Although a number of articles have been written on the subject of residual stress, the main area of concern, for obvious reasons, has been the heat-affected zones such as within the rail head and at weld locations.

Manufacturing and heat treatment processes to produce head-hardened rail with improved properties while maintaining geometrical tolerances often produce residual stress in excess of acceptable values.

Research aimed at obtaining a better knowledge of rail residual stress phenomena in relation to manufacturing processes, to develop residual stress acceptance standards and to contribute to the development of residual stress quality control test methods have been undertaken.

Results of this research have shown that rail manufacturing processes control the residual stress pattern in the rail while roller straightener settings modify these stress patterns. Also, roller straighteners can cause weakening of the rail web and fracture to the rail base. The damage to the web is such that the toughness difference that would otherwise arise due to differences in alloy content and manufacture are mitigated.

It is shown that rail with stress intensities due to residual stress that is greater than the fracture toughness will fracture very readily along the web, whereas those with residual stress intensities less than the fracture toughness will not readily allow catastrophic crack propagation.

J.J. Kalker et al. (eds.), Rail Quality and Maintenance for Modern Railway Operation, 325–336.

INTRODUCTION

In track, rail is the most important element in the transfer of load from the train wheels to the track structure. The demands on rail have changed continuously, with the aim of increasing tonnage carried while reducing maintenance costs and prolonging rail life. Premium or alloy rails have attempted to address these ever-changing demands but not without problems. By changing the chemistry of the rail steel to make it harder and more wear-resistant, ductility has been compromised. With irregular and rapid cooling, tremendous internal stresses are locked within the rail. These internal stresses (also known as residual stresses) cause the rail to distort. Consequently, the rail often requires roller straightening to achieve geometrical tolerance required by the users. Similar to what happens when a wire coat hanger is straightened, plastic deformation and release of energy as plastic strains allows the rail to be straightened.

The question then is, "where are the plastic deformations taking place?" The answer to this question will be addressed later. Since residual stresses will affect the fatigue and crack propagation properties of rail steel, it has been an object of study, especially in the Soviet Union [1].

Mier [2] stated that,

> *Residual stresses exist in a free, unloaded workpiece, when parts of the workpiece body are prevented, by their mutual connection, from assuming the size that they would take up if they were free from restraints, i.e., when released from the restraint. Residual stresses in a railway rail are distributed in a highly complicated manner, with respect to their magnitude and direction, within the rail body. It is possible to assume that the stresses that run parallel to the rail longitudinal axis are the maximum stresses.*

With the introduction of new rail manufacturing and head-hardening techniques and modifications of existing rail-making hardware to meet the new rail standards, the rail manufacturing industry is sometimes forced to take shortcuts to meet the ever-changing demands of the end rail user. These shortcuts can lead to compromises that result in rails having excessive internal stresses. To overcome the significant distortions of the rail associated with these stresses, roller-straightening forces of tremendous magnitudes are used. Sometimes the stresses induced by the rollers have exceeded the ultimate extreme fibre stresses in the head and foot of the rail. Considering that steel yields at 0.2 percent strain, it is obvious that to straighten rail, a depth of material in the head and foot of the rail will inevitably be plasticized. What is the extent of the plasticization?

Our experience has shown that in cases where residual stresses are significant, higher roller-straightener settings have been used in straightening the rail. This is usually accompanied by cracks in the rail base (broken base, see **Figure 1**). The stress intensity due to residual stresses at the tip of these base cracks will sometimes overcome the toughness of the rail material with ease in the presence of external loading such as a wheel load.

This can result in a catastrophic fracture that often starts vertically in the base and then change direction once it enters the web, and run along the web in the longitudinal direction of the rail. The same phenomenon can also take place when the rail fails from say, a transverse head defect or at a joint as in the case of the infamous Amtrak derailment near Marshall, Texas in 1983 (**see Figure 2**).

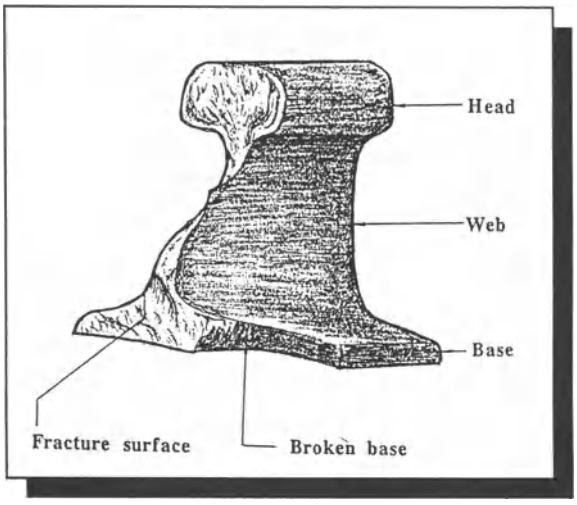


Figure 1: Rail with broken base from roller straightening

ROLLER STRAIGHTENER EFFECTS ON RAIL

Deroche et al. [3] reported that during roller-straightening, low-intensity internal stresses are converted into high-intensity stress patterns that invariably compromise the fatigue strength of the rail. Although this is generally true, high stresses are introduced in rails during the cooling associated with the rail making process. This is particularly worse in rails where rapid cooling is used to achieve the proper material transformations during head-hardening. Roller-straightening will only alter the internal stress by the formation of plasticized

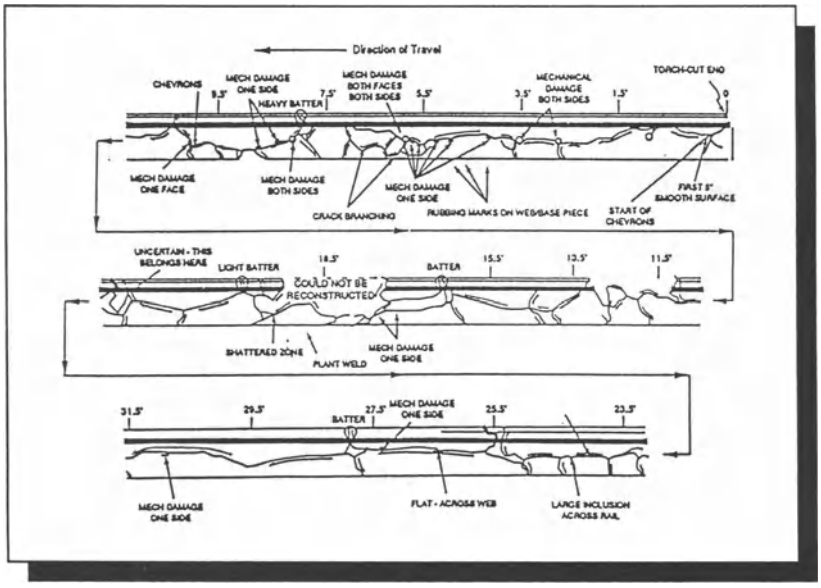


Figure 2: Shattering of Amtrak 21 rail [4]

zones at the top, bottom and along the web of the rail. Micro cracks formed along the web during roller-straightening have been documented by Deroche et al., and are shown in **Figure 3**. The micro cracks they reported are often internal and not visible from the outside. One might speculate that these macro cracks are formed by the brittle fracture of segregates in the rail web. These cracks are prevented from propagating through to the side of the rail web by the ductile (ligaments) material surrounding the segregates. Sometimes where ligaments holding these cracks are broken, rusting can take place on the crack surface. The difference in the coloration of the surface after it is broken open

is usually a good indication of whether the crack propagated though to the surface or not (see **Figure 4**). Kolmogorov et al. [5] demonstrated that the residual stress pattern developed during the rail cooling as a result of the irregular rail shape are redistributed and can often be intensified by roller-straightening in such areas as the tip of the rail base. Also the stress at the head and portions of the base are increased. They concluded that the pattern of residual stresses in rails have a peculiar configuration related to the shape of the cross section.

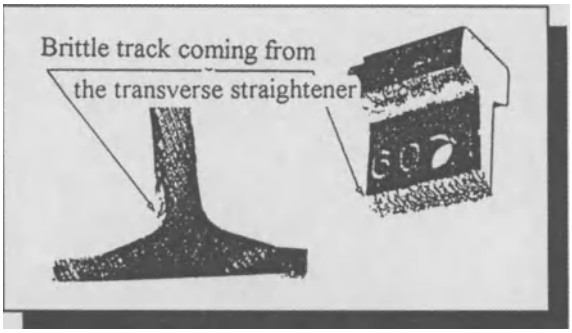


Figure 3: Crack generated by roller straightening

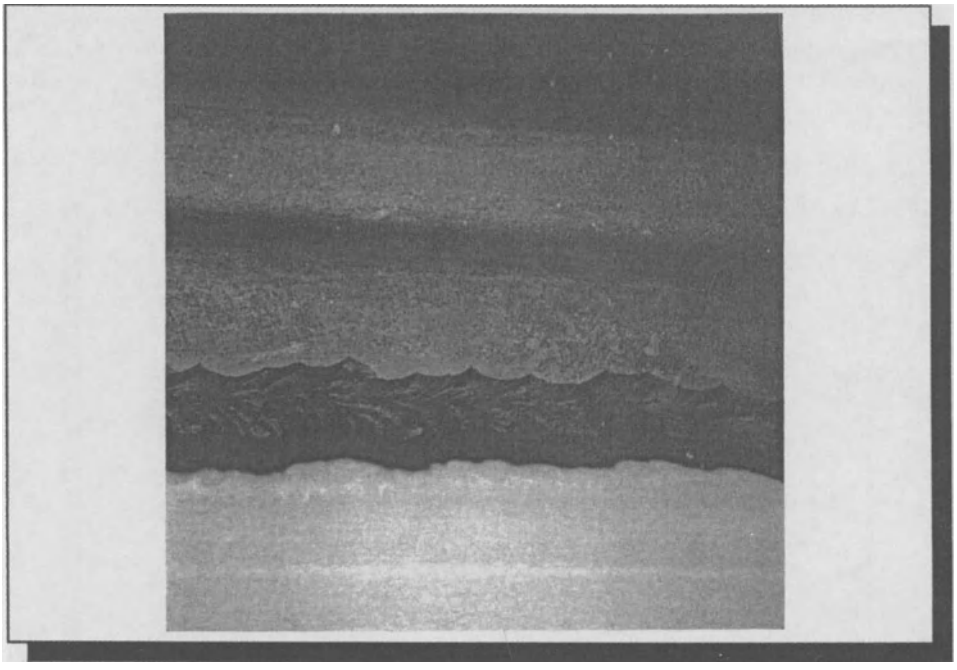


Figure 4: Rusting on surface of broken rail prior to failure

Where vertical rollers are predominantly used, the failure plane often tends to be more horizontal than vertical. Where both vertical and horizontal straightening is carried out, angular failure planes are predominant. The angle of the failure plane (**Figure 5**) will be controlled by the combinations of the magnitude of Modes I and III (see **Figure 6**) present in the rail.

Wineman and McClintock [6] investigated the propensity of residual stresses to drive web cracks in rails recovered after the Amtrak Eagle derailment, in 1983, and concluded that Mode II would have a tendency to change the direction in which the crack is propagating thereby retarding the crack growth. From the fractured rail pieces

recovered from the derailment site, the direction of the cracking clearly showed the varying intensities of Modes I and III that combined to cause the failure.

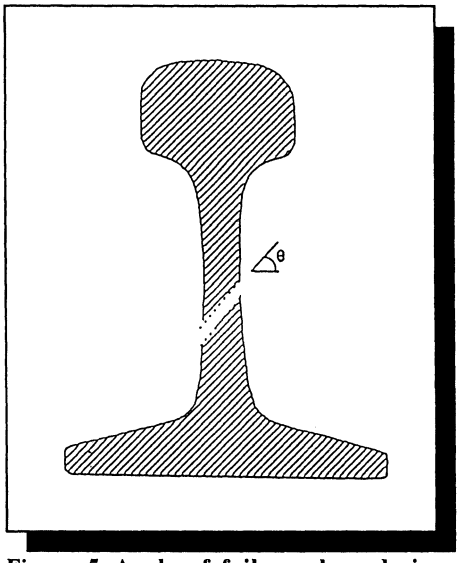


Figure 5: Angle of failure plane during catastrophic web failure

Similarly, our experience confirms the above claims. In analysing stresses due to roller straightening using roller displacements shown in **Figure 7**, the horizontal stress values which are presented show that the rail head and base can sustain fracture. Although these values were obtained from a linear elastic analysis, they nevertheless illustrate the point. We have observed such fractures in numerous occasions. While measuring residual stresses in rails, we have also observed significant lateral displacement and rotation of sawn halves of some of the rail specimens with respect to one another. The stresses unlocked in these rails, when combined with the axle load induced stresses, give principal stress values that sometimes exceeded the yield stress value. This can cause a high-stress-low-cycle fatigue failure of the rail.

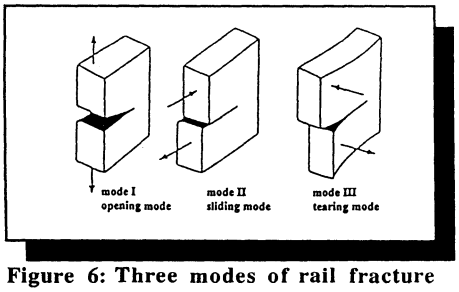


Figure 6: Three modes of rail fracture

RAIL FAILURE FROM RESIDUAL STRESS

The failure of structural components always follow the path of least resistance. This path often presents the least cross-section of material in the structure. In rails, failure in a vertical direction across the web section is usually the case. In residual stress related rail failures, that the authors have experienced, the failure has been along the rail web and can transverse the entire rail length. The broken half of the rails are then curled in the manner shown in **Figure 8**. The curvature of these halves indicate the level of stress holding the two sections together. It leads to the conclusion that the web of the rail must have been compromised during the rail making process.

Figure 4 also shows the failure surface of rail that failed from residual stresses. During roller straightening, the maximum shear stress in the rail section occurs in the web and is a function of the normal stresses. Therefore, if the normal stress exceeds the yield stress value by a large margin, the shear stress will also be expected to exceed the yield stress. This confirms the findings of Deroche et al. [3]. It also explains the Chevron patterns shown in Figure 4. These could be caused by a combination of bending stresses and difference in angular velocity between the upper and lower rollers (assuming that they both have the same diameter). The unidirectional characteristics of the Chevron also show that the damage was done while the rail was travelling in one direction.

EXISTING TECHNIQUES FOR EVALUATING RESIDUAL STRESSES

To date, several techniques have been put forward for use in the evaluation of rail residual stresses. These include the strain gauging, saw cutting, and hole drilling methods, which are destructive, and the neutron diffraction, Barkhausen noise and wave propagation techniques such as acoustic birefringence and shear wave propagation speed, which are non-destructive. The techniques are briefly described below.

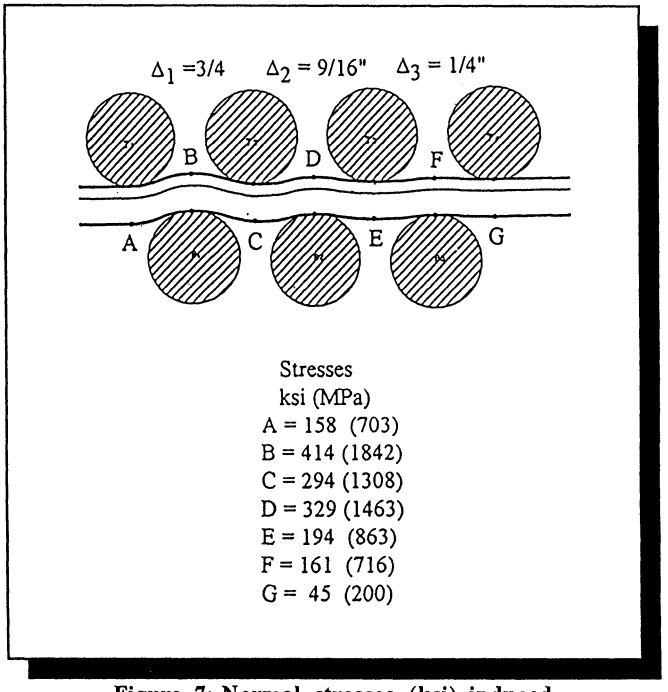


Figure 7: Normal stresses (ksi) induced by roller straightening

Strain Gauging Method

Developed by Battelle Columbus Laboratories, in-plane stresses are measured by strain gauge rosettes placed on quarter inch grids. The residual stresses are relieved by precision slicing of the specimen through these grids. The response of each slice is measured by the strain gauge attached to it and provides a measure of the residual stress that is relieved.

Saw Cutting Method

Similar to the strain gauging method, this technique was developed in the USSR and involves making saw cuts at strategic locations in the longitudinal, transverse and vertical planes of a rail sample. Sometimes the cut is made right through the specimen and the curvature of the specimen or distance between demec points on the specimen is measured before, during and after the cut. Strain gauges are also used to measure the strains. The strain values are then converted to stresses using bulk elastic properties of the material. Alternatively, analytical methods can be used to evaluate the stresses by using the measured deformations as input.

Hole Drilling

This is similar to the two methods above and involves placement of strain gauges in a circular arrangement, drilling a hole through the centre of the circle formed by the strain gauges and measuring the relaxation of the stresses near the hole. Again, bulk elastic properties of the material are used to convert these strain values to stresses.

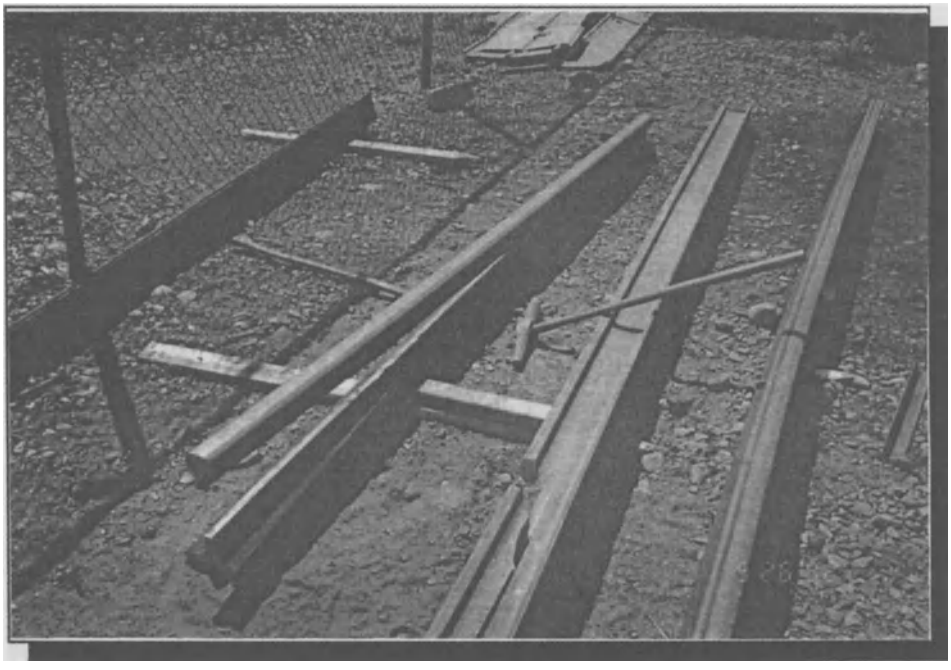


Figure 8: Photograph of longitudinal rail web failure

Acoustic Birefringence Technique

The acoustic birefringence technique is based on the time of travel of acoustic shear waves transmitted by EMAT coils on one side of the material and received by another on the opposite side. This method is sensitive to specimen size and the frequency of the current to the transmission coil. The transit time of the acoustic wave in materials is stress-dependent, decreasing with increased compression and vice versa. From the transit time data obtained from a destressed specimen, the calibration for the zero stress transit time is obtained. The difference between destressed and stressed transit times is used to obtain the stress state in the material.

Neutron Diffraction Method

It is based on the principle that the scattering angle of diffraction peaks of a neutron beam is related to the interplanar distance between atomic planes and the wavelength of the incident neutron beam through Bragg's Law. This inter-planar distance is then controlled by the state of stress in the material. By measuring the inter-planar distance, the strain in the lattice structure can be obtained. Conversion to stress uses Hooke's Law, three-dimensional elasticity theories and bulk elastic properties of the material.

Barkhausen Noises

This technique is based on the principle of magneto-elastic interaction between magnetostrictive and elastic lattice strains. This implies that if a piece of ferro-magnetic steel is magnetized, it will elongate in the direction of the applied magnet field and vice versa. By reversing the magnetic field in the material, a stress dependent, low strength noise (Barkhausen Noise) burst is emitted.

Shear Wave Propagation

Similar to the acoustic birefringence method, wave propagation speed is stress-dependent except that the wave input and pick-up are on the same side of the specimen. It also requires a completely destressed sample of steel for calibration and is sensitive to material properties as are all the other non-destructive techniques.

ARRT METHOD OF EVALUATING RESIDUAL STRESSES

The authors have been involved in residual stress related work that has examined rails from different manufactures world wide. The unlocking of these residual stresses has been achieved through the web saw-cutting method. This method is simple and when proper measurements are made, provide information on the rail quality.

The procedure for determining the residual stress in the rail and assessing whether the rail can fail catastrophically has involved:

(i) Performing saw cuts at the locations shown in **Figure 9**. Measurements are made with a special digital caliper of (a) the saw cut opening or closure every 150 mm (6 inches) for 450 mm (18 inches) and; (b) the lateral movement and rotation of one sawn half relative to the other.

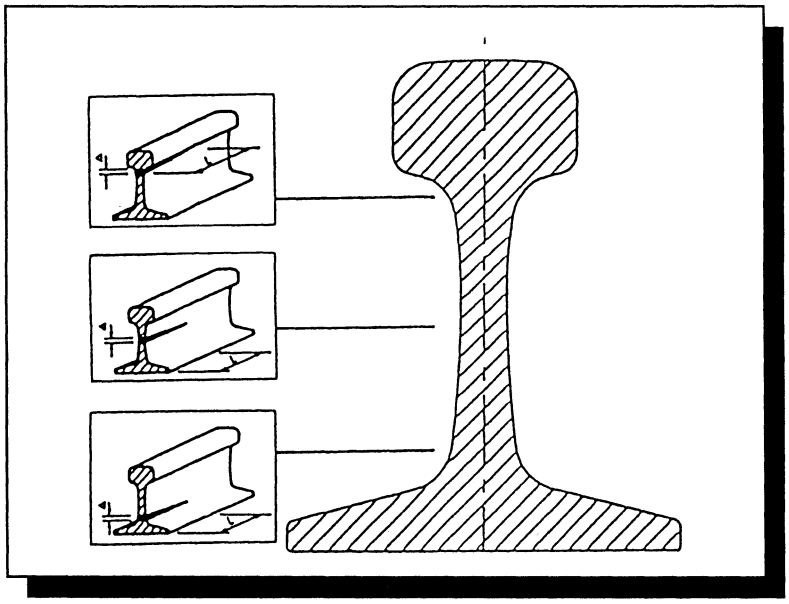


Figure 9: Saw-cut locations

(ii) Using linear and non-linear analytical and fracture mechanics techniques to convert the measured displacements into various stress components and stress intensities.

(iii) Performing stress versus strain tests in samples extracted from the locations shown in **Figure 10**.

(iv) Performing sharpy V. Notch and K_{IC} tests on samples extracted from the locations shown in Figure 10.

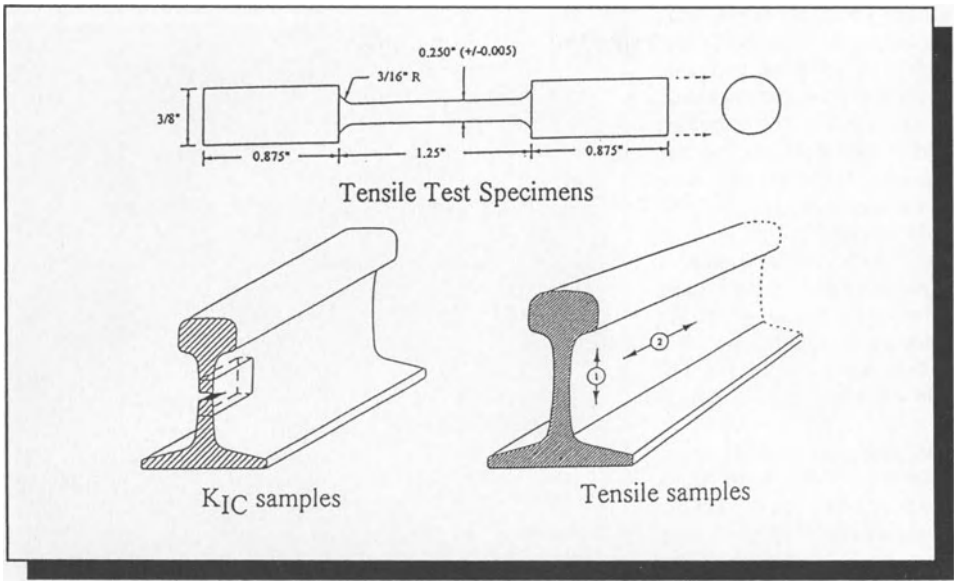


Figure 10: Locations from where Sharpy, K_{IC} and Tensile samples are to be taken

FAILURE CRITERIA

Having performed the above exercise, two failure criteria are used to verify the susceptibility of the rail to catastrophic failure. Firstly, the ARRT Stress Coefficients are compared to the yield stress values of the material at various locations. According to this failure criteria, the material would be considered as having failed if the ARRT Stress Coefficients exceed the yield stress value. Secondly, the calculated stress intensities are compared to the fracture toughness of the rail material at the various sections. If the material toughness is exceed at any of these sections, the rail is also considered to be prone to failure because it does not have the ability to resist rapid crack growth should one develop.

RESIDUAL STRESS IN RAILS OF VARIOUS MANUFACTURES

Over one hundred rails from several different manufactures have been examined by the authors. These rails include 115 lb, 132 lb and 136 lb RE rails. **Table 1** show typical results of the residual stress evaluations. More work is underway. It can be seen from Table 1 that in cases where the ARRT Stress Coefficient exceeds the yield strength, the fracture toughness is also exceeded. These rails have failed by rapid propagation of web cracks.

A simple test for such rails is to drive a wedge into the saw-cut. Rails with high residual stresses will propagate the crack along the longitudinal axis of the rail with ease. Rails whose ARRT Stress Coefficients are lower than their yield strength would propagate the crack (with difficulty) about 2 to 6 inches (50 - 150 mm) before the crack changes direction breaking to the head or foot of the rail.

It has been widely reported in the literature [7] that rail steel have toughness that varies between 25 and 50 ksi$\sqrt{in}$ (27 and 55 MPa$\sqrt{m}$) with the average for plain carbon of 35

ksi√in (38.5 MPa√m). Rail specimen D is a standard carbon head-hardened rail with a fracture toughness of 63 ksi√in (68 MPa√m) which is almost double the value listed in the literature. Some of the rails in specimens A to C are made of standard carbon and others are alloy rails yet their fracture toughness lie within a 38 - 42 ksi√in (41 - 46 MPa√m) band. **Why?**

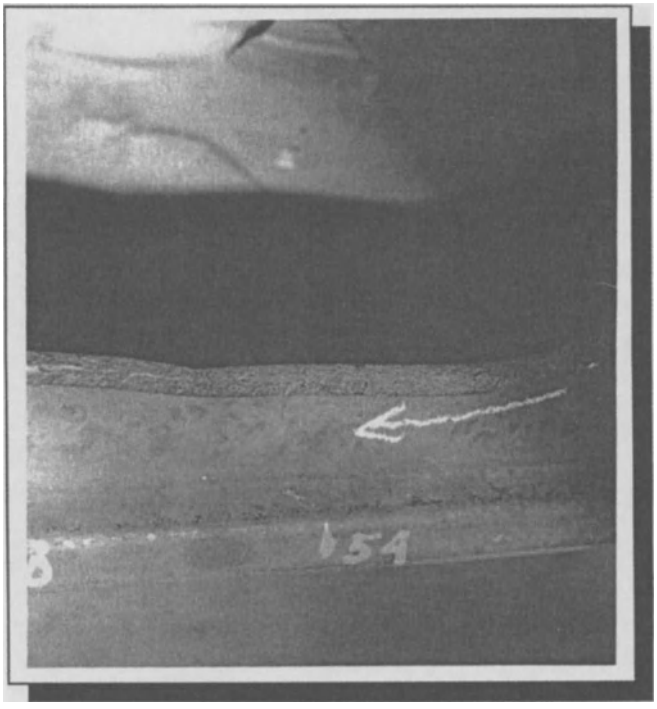


Figure 11: Surface of rail that failed the ARRT test

The common denominator between rails A to C is that they are roller straightened after they were head-hardened while specimen D was not. The inference can be made that roller straightening does enough internal damage to the rail to mitigate any differences in toughness that would normally arise as a result of differences in chemistry. As can be seen, specimens A, B and D all have residual stresses that is lower than their yield strength and stress intensities at the saw-cut tip that is lower than their fracture toughness and hence are not prone to catastrophic failure. Rail specimen C failed both tests and also propagated cracks very readily. **Figure 11** shows such a crack which in this case has a flat surface.

TABLE 1: SUMMARY OF RESIDUAL STRESS TEST RESULTS

Rail Make	Saw-Cut Location	Saw-cut opening at 18 in. (450 mm) length of cut. inches (mm)		Relative rotation of one half (degrees)	Max. normal stress ksi (MPa)	Max. vertical stress ksi (MPa)	ARRT Stress Coeff. ksi (MPa)	Yield strength of web ksi (MPa)	Avg. stress intensity ksi√in (MPa√m)	Avg. Fracture toughness ksi√in (MPa√m)	Rail type	Type of manufacture /Type
		Vertical	Lateral									
A	Below Head	0.15 (3.8)	0.04 (1.0)	-	57 (393)	14 (97)	36 (248)	78 (538)	25 (27)	42 (46)	136	STHH/ on-line
B	Above Foot	0.24 (6.1)	-	-	80 (552)	35 (241)	83 (572)	99 (683)	36 (39)	38 (41)	136	STHH/ off-line
	Above Foot	0.15 (3.8)	0.04 (1.0)	0.02	59 (407)	23 (159)	50 (345)	100 (690)	22 (24)	38 (41)	136	LAHH/ off-line
C	Neutral Axis	0.29 (7.4)	-	1.54	98 (676)	84 (579)	141 (972)	85 (586)	70 (76)	39 (42)	115	LAHH/ off-line
D*	Neutral Axis	0.06 (1.5)	-	-	21 (145)	18 (124)	28 (193)	92 (634)	20 (22)	63 (68)	136	STHH/ off-line

STHH = Standard Carbon Head-Hardened, LAHH = Low Alloy Head-Hardened, * = no roller straightening after Head Hardening

DISCUSSION

Given the technology presently available to us, are we skirting the envelope of rail-making and design? The rail-making industry is faced with having to achieve higher depths of hardness in the rail head while keeping the residual stresses low and the web and foot of the rail ductile. This may be very difficult but must be achieved if one wants to avoid a repetition of the Amtrak Eagle derailment and other more recent catastrophic rail failures that are residual stress related.

Finally, railway manuals or codes contain no established standards for dealing with residual stresses in rail. Apart from the Amtrak derailment in Texas in 1983, residual stresses have rarely been a source of grave concern to the railways. Our experience shows that to avoid catastrophic rail failures, the area of concern in residual stress evaluations should be the web and base of rail instead of the head. Until such provisions are made in the standards, each individual railway needs to set its own criteria, depending largely on several factors that are peculiar to that individual railroad.

A committee has been set up in Canada to investigate residual stresses in rail and to develop a standard for effective measurement of residual stress. This committee comprise representatives from the Canadian National Railways (**CN**), Canadian Pacific Railways (**CP**), Quebec Cartier Mining Railway (**QCM**), Transport Canada (Transportation Development Centre (**TDC**) and Railway Safety Directorate (**RSD**), Quebec Ministier du Transport (**MTQ**), Canadian Institute of Guided Ground Transport (**CIGGT**), and Applied Rail Research Technologies Incorporated (**ARRT Inc.**).

CONCLUSIONS

Given the information provided in this paper, the following conclusions can be drawn.

1) Residual stress in head-hardened rail is a problem that demands concern.

2) It is relatively simple and easy to check for excessive residual stresses. This should be done during manufacturing and not after the rail has gone in to the track.

3) A non-destructive technique for evaluating residual stresses on-line during rail manufacture needs to be developed. This will reduce the chances of rails of unsatisfactory quality being placed in service, the consequence of which can be catastrophic.

4) If a manufacturer changes his process either to improve the quality of the rail or to increase the quantity of his production, the customer should ask for residual stress evaluations from every batch of rail purchased. Measurements should include the saw-cut openings at the locations shown in Figure 9 up to 18 inches (450 mm) in length, tensile strength tests as well as sharpy and fracture toughness tests for the locations shown in Figure 10.

REFERENCES

1. Konyuklov, A. D., Reikhart, V. A. and Kaportsev, N.V., "**Mechanical Methods of Evaluating Residual Stressed State of Rails**", Towdy TsNII MPS, 1973, No. 491, pp 5-10.

2. Mier, H., "**Residual Stresses in Rails**", Orga F. d Fortschritte D. Eisenbahnwesens, 1936. vol. 91, No. 15, pp. 320-329.

3. Deroche et al. "**Stress Releasing and Straightening of Rails by Stretching**", *International Heavy Haul Railway Conference*, Colorado Springs, Colorado, U.S.A., September 1982, pp. 158-169.

4. Steel, R. K., Joerme, M. w., Utrata, D., and Carpenter, G.F. "**Catastrophic Web Cracking of Railroad Rail: A Discussion of the Unanswered Questions**", Association of American Railroads, May 29, 1990.

5. Komogorov, S.V., Makarov, Yu.D., and Mikhailov, O.N. "**Formation of Residual Stresses in Heat Treatment Hardened Rail**", *Steel in the USSR*. UDC 621.771.26: 625.143.2, June 1987, 17 (6), pp. 271-273

6. Wineman, S.J., and McClintock, F.A. "**Rail Web Fracture in the Presence of Residual Stresses**", *Theoretical and Applied Fracture Mechanics*, 8 (1987), pp. 87-89.

7. Orringer, O., Morris, J.M., and Steel, R.K. "**Applied Research on Rail Fatigue and Fracture in the United States**", *Theoretical and Applied Fracture Mechanics*, 1, 23-49, 1984

ACKNOWLEDGEMENTS

The authors acknowledge the following organizations for their contributions to the knowledge. The Quebec Cartier Mining Railway (QCM), The British Columbia Railway (BCR), Sidney Steel Corporation, Transport Canada (TDC), Queens University, Quebec North Shore and Labrador Railway (QNS & L), Canadian National Railways (CN), Canadian Pacific Railways (CP), the staff at Applied Rail Research Technologies Incorporated (ARRT Inc.), and the Canadian Institute of Guided Ground Transport.

BENEFICIAL RESIDUAL STRESSES AT BOLT HOLES BY COLD EXPANSION

Len Reid, Vice President, Engineering
Fatigue Technology, Inc.
Seattle, WA USA

ABSTRACT

Cracking at rail-end-bolt holes is a major safety issue and cause of premature rail replacement, imposition of rail speed restrictions and a significant factor in rail inspection and maintenance costs. The presence of residual compressive stresses around these bolt holes has been shown through exhaustive testing and field evaluation to minimize fatigue cracking and thereby extend inspection intervals and allow for higher axle loads. Split sleeve cold expansion is an economical and reliable method to pre-stress bolt holes during routine track maintenance or manufacture of new rail, joints and switches.

This paper reviews the evaluation of methods to overcome bolt hole fatigue failure. It describes the split sleeve cold expansion technique in detail and the mechanism by which these beneficial residual stresses effectively nullify fatigue causing cyclic tension loads and virtually eliminate the rail-end-bolt hole cracking problem.

Introduction

The problem of rail bolt hole cracking is not unique to any specific railroad, region or country but is recognized as a worldwide problem. Railroad tracks are continually subjected to high loads generated by the passage of rolling stock. Flexing and displacement of rails at bolted joints combine to induce high cyclic tensile and shear loads in the joint bar, or fish plate, which are transferred to the attaching bolt holes. Numerous studies and reports (see for instance [1] and [2]) detail research conducted to define the problem of rail-end-bolt hole cracking and to address practical solutions designed to minimize or prevent rail joint bolt hole failure. As shown in Figure 1, the most dangerous situation arises in jointed track when cracks occur at a bolt hole because fracture usually results in a piece of rail becoming detached.

In 1974, the U.S. National Transportation Safety Board (NTSB) identified broken rails as the largest single cause of train accidents [3]. Between 1982 and 1988, track-caused-accidents represented between 30 and 40 percent of the total number of the reported accidents [4] and total incurred costs of recovery. Specifically, Federal Railroad Administration statistics for 1988 [5] showed bolt-hole-caused derailments accounted for a high cost per accident and a significant 10 percent ($2,136,221) of the total cost of rail and joint bar defects.

J.J. Kalker et al. (eds.), Rail Quality and Maintenance for Modern Railway Operation, 337–347.

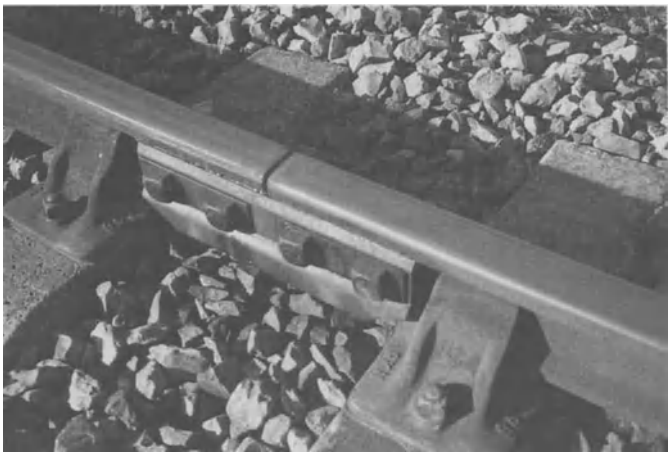


Figure 1 Typical Bolted Rail Joint

Similar statistics have been recorded in the U.K. In the early 1980's, over 3000 cracked and broken rails of different types were reported each year, with the highest number of incidents occurring on middle-speed range, heavily loaded track. Of these, about 25 percent were caused by cracks originating at rail-end-bolt holes. Between 60 and 70 percent were detected while still in the cracked stage before actual fracture.

More recently, tests were carried out as part of the Heavy Axle Load (HAL) Program for the U.S. rail industry [6] to investigate the effect of increased axle loads and speeds on existing track. It was found that an increase in axle load of only about 20 percent precipitated serious cracking in bolt holes in turnout frogs and switches.

Loading of the rail web and additional localized loads significantly increase the magnitude of the applied stress acting on a typical bolted joint when axle loads and speeds are increased. These loads are further magnified by dynamic effects in turnouts and curves or where the rail bed fails to adequately support the rail tie or sleepers.

Mechanism of Joint Failure

Cracks originating from rail joint bolt holes are the result of web shear stresses. In the U.S. Department of Transportation report [2] it was concluded that these stresses develop in the rail end when the joint bar transmits the bending moment across the joint through concentrated point contact at the rail and the end of the joint bar (Figure 2). This condition occurs whenever the joint bar becomes loose or, on installation, was not nested correctly between the head and flange of the rail. British Railways, in their studies on the effect of track and vehicle parameters on wheel/rail vertical dynamic forces [7], concluded the cause of bolt hole cracking is shear stress associated with dynamic wheel/rail forces generated by a dynamic dip at the joint.

The shear stress in the rail web is then magnified by the stress concentration effect of the bolt hole and other surface defects such as corrosion pits and material imperfections that may be present. Additional dynamic loads on the rail joint, including those due to loose bolts and

collapsed rail bed, add to the hoop surface stress of the hole. Two cracks generally initiate, one growing towards the rail head and the other one towards the base of the rail. They propagate along a 45 degree plane, as shown in Figure 3. If cracks are not detected before they become critical, usually 5-10 mm long, fracture can occur, often dislodging a triangular piece of rail which can lead to derailment.

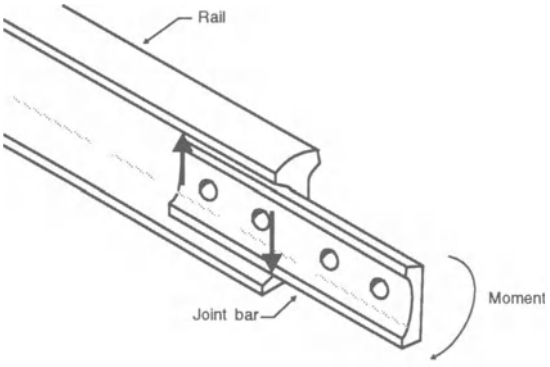


Figure 2 Loads Applied to a Typical Bolted Rail Joint [2]

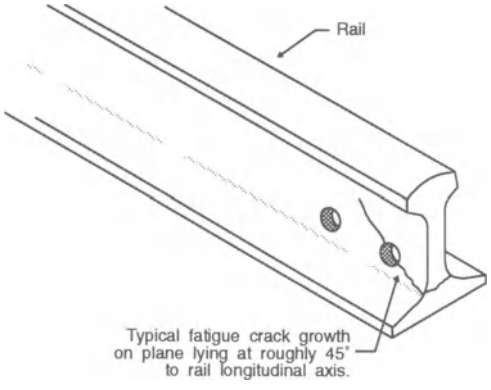


Figure 3 Typical Cracks Originating at Rail-End Bolt Hole

Fatigue Life Improvement of Rail Joints

Over the past three decades, a number of attempts have been made to mitigate the rail-end-bolt hole cracking problem. In the 1960's, British Rail increased rail-web thickness in the hope that it would proportionally reduce the high stresses and thus reduce the incidence of bolt hole cracking. About the same time, a hole "broaching" process was tried in laboratories to "work harden" the hole surface and locally increase its fatigue strength; however, the process was difficult to control and eventually proved to be ineffective.

Some of the options that are currently available to minimize the occurrence of rail-end-bolt hole failures include: (1) replacement with continuous welded rail, although some joints may still remain; (2) increased inspection frequency, which will not prevent fatigue crack initiation; (3) reducing the usage or loading on the track, which will minimize the stress cycles at the bolt hole and therefore the failure probability, but carries an economic penalty; and (4) modify the bolt hole locally by for example, introducing compressive residual stresses to inhibit crack propagation from it. The latter approach is highly cost effective and reliable under most circumstances.

The technique of mechanically pre-stressing holes to induce favorable compressive residual stresses is successfully used by other industries, notably aerospace. Methods such as shot peening and blasting, coining, ballizing, roller burnishing, and mandrelizing, have all been evaluated. Of these, mandrelizing, i.e. , the effect of forcing a tapered or multi-flanged expansion mandrel through a hole, initially appeared to show the most potential. This method results in the material around the hole being plastically
deformed. However, the amount of circumferential and radial expansion of the hole is very much dependent on hole quality. Careful lubrication of the hole is also necessary to avoid material being pushed through the hole rather than being radially expanded. In its application to bolted rails, one additional requirement is the capability to reliably and repeatedly carry out the process in the field under adverse environmental conditions. For these reasons, in practice, the mandrelizing process has not performed satisfactorily.

In 1975, the U.S. Department of Transportation (DOT) sponsored a study [2] to investigate the split sleeve cold expansion process and a pad-coining method as an alternate means of pre-stressing bolt holes. The rail web thickness and unevenness in the web surface reduced the effectiveness of the pad coining process. Also, the need for coining forces exceeding 100,000 pounds (445,000 N), precluded use of this method with conventional equipment. On the other hand, in laboratory studies and later, field trials, the split sleeve cold expansion process showed significant fatigue life improvement over non-cold expanded bolt holes. Results from the DOT study are shown in Figure 4.

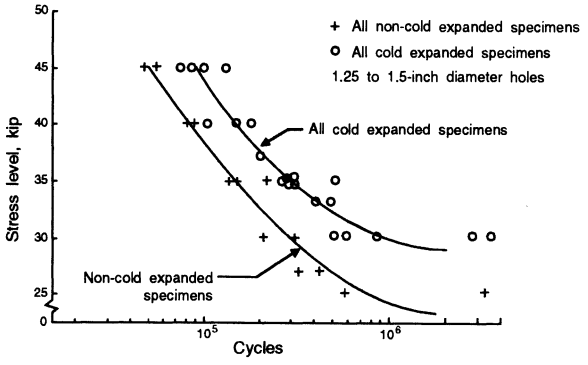


Figure 4 Increase In Fatigue Life For Cold Expanded Holes Vs. Non-Cold Expanded [2]

The results of additional extensive British Rail trial and evaluation of the split sleeve cold expansion process, including theoretical modeling and laboratory and in-service tests, is documented in [1]. Both the aforementioned U.S. Department of Transportation program and the British Rail Board trials have confirmed that split sleeve cold expanded rail-end bolt holes could increase the life of a bolted rail joint by a factor of 10 or more by reducing or eliminating bolt hole fatigue failure.

The Split Sleeve Cold Expansion Process

The split sleeve cold expansion process previously described is accomplished by pulling an oversize tapered mandrel, pre-fitted with a dry-film lubricated split sleeve, through the bolt hole using a specially designed hydraulic puller, as shown in Figure 5.

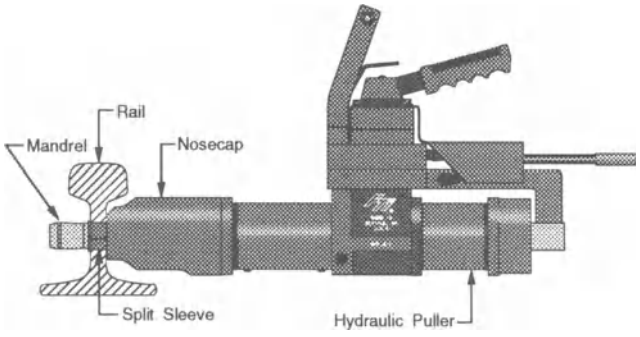


Figure 5 Schematic Of Split Sleeve Cold Expansion Process

The disposable lubricated sleeve is crucial to the process, because it reduces the pull force required and prevents local shearing of rail material as the mandrel is pulled through the hole. The combination of mandrel diameter and sleeve thickness creates very high radial pressure on the hole, and expands the hole to well beyond the yield strain of the rail steel. After the mandrel passes through the hole, the area surrounding the hole remains residually stressed in compression to a distance roughly equal to the hole radius (Figure 6). The magnitude of the peak compressive stress approaches the material compressive yield stress. This pre-stress effectively shields the hole from crack growth by counteracting the stresses imposed by the applied service loads.

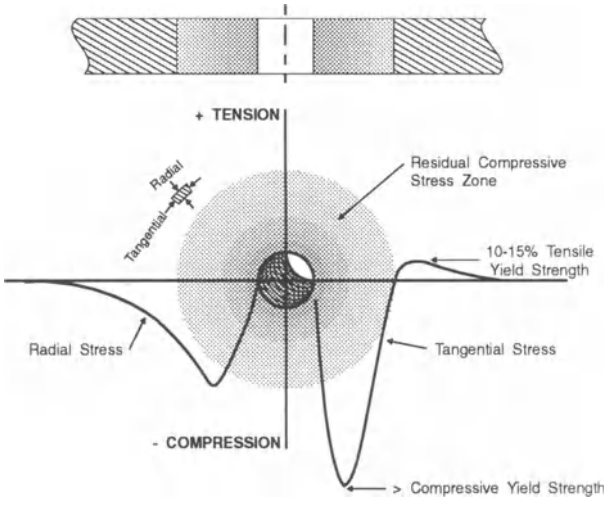


Figure 6 Distribution Of Residual Stress Around A Split Sleeve Cold Expanded Hole

Fatigue Technology, Inc. (FTI) of Seattle, Washington, USA has developed the RailTec™ Cold Expansion system of tooling to facilitate cold expansion of bolt holes in existing track or in new production rail. In a typical field application as used by British Rail, each fish-plated joint is dismantled, followed by cleaning of bolt holes and adjacent areas with scrapers and cloths. Holes are measured and cleaned up with a bridge reamer to a nominal size and then each hole is cold expanded using RailTec tooling as shown in Figure 7. Finally, the joint is re-assembled. All operations proceed simultaneously, and a production rate of about 10 joints per hour is readily achievable.

Figure 7 RailTec System In Use On Existing Track

The estimated total cost of servicing the joints during routine maintenance, including cold expansion of all holes, is about $3,000/km. This compares to the cost of converting jointed track to Continuous Welded Rail (CWR) at around $180,000/km [8], which can be prohibitively high for relatively low-revenue earning lines, and requires the track to be out of service for a much longer period of time.

Following its participation in the U.S. evaluation program, Union Pacific Railroad implemented the split sleeve cold expansion process on all new rail, switches and crossings.

British Rail also currently has a number of specialist field teams in all of its regions implementing this process. Surveys of the British Rail Western region since the RailTec system was introduced a few years ago, show a virtual elimination of occurrences of fatigue cracking from these holes. Results for the Plymouth to Penzance route (21 km), and Exeter to Sherbourne route (38 km) are shown in Figures 8 and 9 respectively [9]. Considering that some of these existing rail joint holes may have contained small or undetected cracks at the time of cold expansion, the outstanding benefits of this process are clearly evident.

Date	Star Cracks	Miles Cold Expanded	Total Miles Cold Expanded
1988	2	0	0
1989	22	1.58	1.58
1990	28	7.68	9.26
1991	7	3.83	13.09

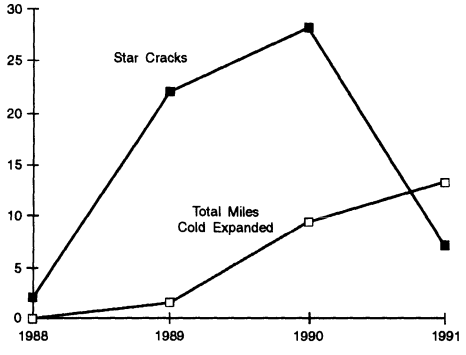


Figure 8 Survey Of Results After Cold Expansion / British Rail Plymouth To Penzance Route

Date	Star Cracks	Miles Cold Expanded	Total Miles Cold Expanded
1987	21	0	0
1988	24	8.51	8.51
1989	12	11.35	19.86
1990	6	3.64	23.5
1991	1		

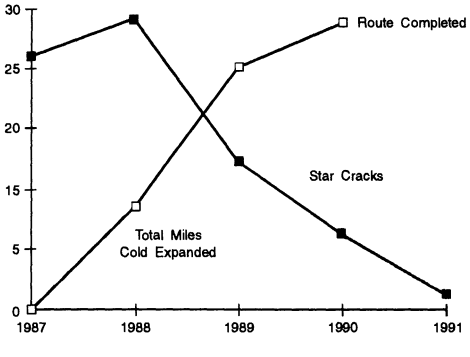


Figure 9 Survey Of Results After Cold Expansion / British Rail Exeter To Sherbourne Route

Mechanics of Cold Expansion

During the split sleeve cold expansion process, the mandrel and sleeve combination radially expands the hole beyond the yield point of the rail steel, by a prescribed amount. Plastic deformation progresses outwardly from the hole. Following travel of the mandrel through the hole, the bulk of the material, which lies beyond the plastic zone, is still in an elastic state and attempts to force the permanently deformed material in the periphery of the hole to return to its original position. The resulting effect is the creation of a band of material around the hole in a residual compressive circumferential stress state, as shown previously in Figure 6.

Residual stresses are self-equilibrating stresses existing in a material under uniform temperature conditions without external loading. In a cold expanded hole, this means that the resultant forces over a radial element produced by the cold expansion residual stresses, must be zero. This explains the small amount of residual tension away from the hole. Figure 10 shows the residual strain pattern created by cold expansion obtained by means of a photoelastic coating.

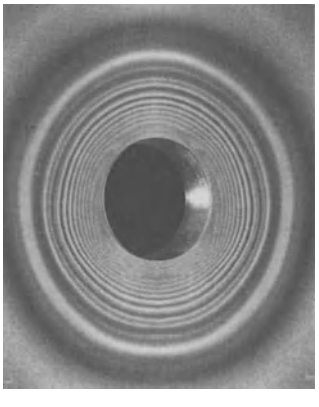


Figure 10 Residual Strain Pattern Around Cold Expanded Hole Seen As Photoelastic Fringes

For typical rail steels, the hole needs to be expanded at least 2.3 percent of its diameter to ensure adequate local yielding of the steel adjacent to the hole. When the rail-end-bolt holes incorporating these residual stresses are externally loaded by the passing of rolling stock, the effective stress state will be roughly the algebraic sum of the applied stresses and the residual stresses. With the magnitude of the residual compressive stress adjacent to the edge of the hole approximately equal to the compressive yield stress of the rail steel, the net effect of the combined stresses under in-service loads is to move the mean stress towards or into compression. This reduced net stress retards crack initiation and inhibits crack growth.

An important consideration is that cold expansion of the hole alone may not prevent (although it may retard) crack initiation. The possible presence of machining defects created during drilling or reaming, the existence of corrosion pits, and metallurgical characteristics of the material, all significantly influence crack initiation. The primary effect of the cold expansion residual stresses is to reduce crack growth rates by reducing the stress intensity factor range (ΔK) for existing cracks, as shown in Figure 11. Similar arrest of crack growth due to stress

intensity reduction and distribution of "K" was reported by British Rail [1] in their research. Additionally, the presence of residual stresses may change the critical crack length for unstable fracture. The lower crack growth rates and greater critical crack lengths can be used to extend non-destructive inspection intervals for rail joints, and/or simplify inspection procedures.

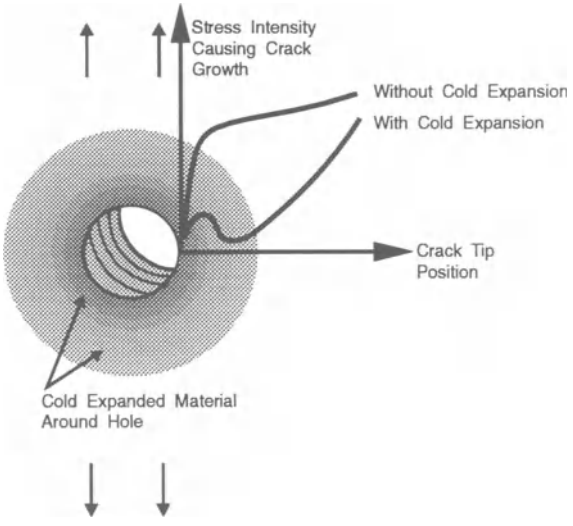


Figure 11 Reduction In Stress Intensity Factor Range (ΔK) Under Residual Compressive Stress From [10]

In a number of fatigue tests conducted by Fatigue Technology and other companies on typical aerospace materials, small pre-existing cracks were totally arrested after inducing compressive residual stresses by split sleeve cold expansion. Only a large increase in external load caused further crack growth.

Analytical modeling methods are available to predict the effect of cold expanded holes on fatigue crack growth [10, 11]. These models, and results from field implementation of the split sleeve cold expansion system confirm that the introduction of beneficial residual compressive stresses around rail-end bolt holes can delay the onset of crack initiation and slow down or totally arrest crack growth.

Conclusions:

The problem of rail-end-bolt hole cracking has been overcome by introduction of residual compressive stresses around the hole. These residual stresses lower the magnitude of the applied stresses acting near the hole, which in turn inhibits crack growth. The most cost effective method for inducing these residual compressive stresses on existing track in the field, and in production of new or replacement bolted track, switches and crossings, is split sleeve cold expansion.

The results of British Rail field surveys and observations from other users of the process confirm split sleeve cold expansion of rail-end bolt holes can be a viable economic alternative to replacement of existing bolted track. The overall result is a greatly extended fatigue life of bolted track, safer and more economical rail operation, reduced routine or special joint maintenance costs, and extended joint inspection intervals.

References:

[1] Cannon, D.F., Sinclair, J., and Sharp, K.A., "Improving the Fatigue Performance of Bolt Holes in Railway Rails by Cold Expansion." Presented at: Fatigue Corrosion Cracking, Fracture Mechanics and Failure Analysis, International Conference and Exposition, December 1985, Salt Lake City, Utah USA.

[2] Lindh, D.V., Taylor, R.D., and Rose, D.M., "Sleeve Expansion of Bolt Holes in Railroad Rail," U.S. Department of Transportation, Final Report, December 1977.

[3] Broken Rails: A Major Cause of Train Accidents, U.S. Department of Transportation NTSB RSS-74-1, January 1974.

[4] Zarembski, A.M., "Track Caused Derailments," RT&S, October 1990.

[5] "Accident/Incident Bulletin, Calendar Year 1988," U.S. Department of Transportation, Federal Railroad Administration, Office of Safety, Bulletin No. 157, June 1989.

[6] Read, D.M., "Fast/HAL Turnout Performance Experiment," American Railway Engineering Association; Bulletin No. 728 Vol. 91 (1990)

[7] Jenkins, H.H., et al, "The Effect of Track and Vehicle Parameters on Wheel/Rail Vertical Dynamic Forces", Railway Engineering Journal (IMechE), January 1974.

[8] Tolley, K.H., "Extending Track Life with Improved Safety at Reduced Maintenance Costs," International Railway Congress Association Intersessional Seminar; Budapest, May 23-24, 1991.

[9] Survey of Results from British Rail Western Region provided by British Rail Research, Derby

[10] Clark, G., "Modeling Residual Stresses and Fatigue Crack Growth at Cold Expanded Fastener Holes" Fatigue Fracture of Engineering Materials Structures. Vol. 14, No. 5.

[11] Rufin, Antonio C., "Extending the Fatigue Life of Aircraft Engine Components by Hole Cold Expansion Technology", ASME Paper 92-GT-77, presented at the 1992 Gas Turbine and Aeroengine Congress and Exhibition, Cologne, Germany

Residual Stress Analysis in Rolling Contact

Huseyin Sehitoglu[+], Y. Roger Jiang
Department of Mechanical Engineering, University of Illinois, Urbana, IL61801
[+]temporarily, Director, Mechanics and Materials Program, National Science Foundation, Washington, DC 20550

ABSTRACT: Based on a stress invariant hypothesis and a stress/strain relaxation procedure, an analytical approach is developed for approximate determination of residual stresses and strains in elastic-plastic rolling contact. It is shown that for the line contact problems, the proposed method provides residual stress results comparable to published results predicted by the finite element method (FEM). As compared with the residual stress results obtained by the Merwin and Johnson method and the McDowell and Moyar method, the current approach yields closer results to the finite element predictions. The developed analytical approach, utilizing a two surface plasticity model with Mroz type hardening rule, is applied to the calculation of combined rolling and sliding cases using 1070 steel properties typical of railroad wheels. It is shown that the component experiencing $Q/P<0$ (driven), where Q and P represent the total tangential force and normal force respectively, experiences higher plastic shear strains compared to the driving ($Q/P>0$) component, which implies, for example, that the driven wheel is inferior to the driving wheel in fatigue life. This prediction agrees with the experimental observations. Moreover, residual stresses obtained by the proposed method are compared with the experimental results from the literature, and very close agreements are demonstrated.

Introduction

Upon rolling contact of ductile materials, if the load is under the shakedown limit, residual stresses will result but no cumulative plastic deformation will accrue upon passages of load. However, if the contact load exceeds the shakedown limit, incremental displacements, reversed plasticity and residual stresses evolve with time. Contact fatigue is a dominant cause of failure in bodies undergoing repeated rolling contact; therefore, quantitative information elucidating residual stress/strain and stress/strain histories is fundamental[1].

The coordinate system adopted in this study is given in Fig.1. The rolling direction for the material is shown. The maximum normal pressure p_0 , and the positive stress directions are indicated. The positive Q/P case shown is representative of a driving wheel.

J.J. Kalker et al. (eds.), Rail Quality and Maintenance for Modern Railway Operation, 349–358.

Due to the complexity of the stress/strain state involved in the rolling contact problem, accurate elastic-plastic stress analysis of contact is difficult. Merwin and Johnson [2] developed an approximate method for the determination of plastic deformation in rolling contact by assuming that the strain cycle remains identical with the elastic strain cycle. With this method, Johnson and co-workers[2-3] obtained residual stress and strain results which were later confirmed with the finite element method (FEM) [4]. Upon noting the lower surface displacement rate predicted in Ref.[2] compared to the FEM, Hearle and Johnson [5] forwarded a model using orthogonal shear strain component as the only plastic strain occurring throughout the cycle. Bower and Johnson [6] used this condition, and assumed that within the thin layer just beneath the contact surface, the stresses corresponded to the elastic solution. With this assumption, the stress histories on the thin layer of contact can be readily determined. However, the demarcation of "surface," "subsurface," and "near surface", and the lack of continuity in the stress and deformation fields, may pose a dilemma.

McDowell and Moyar's work [7] generated further discussion and interest in this field. McDowell and Moyar proposed that under elastic-plastic contact loading, the two stress components σ_z, the stress in radial direction, and τ_{xz}, the orthogonal shear stress, are identical to the elastic solutions. In addition, the deformation rates in both the rolling (x) and axial (y) directions are assumed zero. Their model captures the essential features of residual stress profiles, with the exception of those at the contact surface. Their work emphasized the pressing need for constitutive models suited to handle the material behavior under contact type loadings.

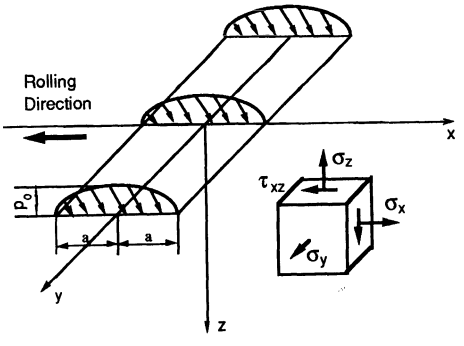


Fig.1 Contact Loads, Coordinate System, and Orthogonal Stresses

Bhargava and Ham et al.[4,8-10] have conducted elastic-plastic stress analyses of rolling line contact using the finite element method (FEM) using elastic-perfectly plastic [4] material properties. We note that, to obtain accurate results, the mesh size requirement with FEM is rather stringent. References [8-10] employed a bilinear constitutive relation in FEM. The residual stress and strain magnitudes obtained using a strain hardening material model were much smaller than those obtained from previous analyses where elastic-perfectly plastic material properties were assumed. These results indicate that the details of constitutive behavior has a profound

influence on the residual stress magnitudes and cumulative shear strain of rolling contact. Therefore, more emphasis should be placed on constitutive model development while advancing simpler numerical algorithms for contact analysis. In this paper, we present a new analytical approach for treating elastic-plastic stress analysis of rolling contact which overcomes the limitations of the previous analytical approaches and provides exceptional agreement with the finite element results.

Analytical Approach and the Plasticity Model

In the new approach, it is assumed that the stress cycle during elastic-plastic rolling contact is equal to that for the purely elastic case. Expressing this hypothesis mathematically,

$$\sigma_{ij}=\sigma_{ij}^{el} \tag{1}$$

where σ_{ij} is the stress tensor and *el* denotes the elastic solution. The relaxation procedure is as follows: The stresses and strains at a point are relaxed proportionally to meet the residual stress and strain conditions upon passage of load.

When applying the proposed method to the rolling line contact problem where plane strain is assumed, Eq.(1) is expressed as,

$$\sigma_x = \sigma_x^{el};\ \sigma_z = \sigma_z^{el};\ \tau_{xz} = \tau_{xz}^{el} \tag{2}$$

where σ_x^{el}, σ_z^{el}, and τ_{xz}^{el} are obtained by using the formula derived by Smith and Liu [11]. Referring to the coordinate system shown in Fig.1, the plane strain condition requires, $\dot{\varepsilon}_y = 0$, where a dot above a symbol means derivative with respect to time. In the elastic region the above equation will result in,

$$\dot{\sigma}_y = \mu(\dot{\sigma}_x + \dot{\sigma}_z) \tag{3}$$

where m is the Poisson's ratio. When the deformation is elastic-plastic, the stress increment $\dot{\sigma}_y$ can be expressed as,

$$\dot{\sigma}_y = \frac{(\mu - \frac{E}{h} n_x n_y)\,\dot{\sigma}_x + (\mu - \frac{E}{h} n_y n_z)\,\dot{\sigma}_z - \frac{2E}{h} n_{xz} n_y)\,\dot{\tau}_{xz}}{1 + \frac{E}{h} n_y^2} \tag{4}$$

where E is the elasticity modulus, h is the plastic modulus function, and n_{ij} represents the unit normal in the plastic strain rate direction on the yield surface. In Eq.(4) we used the plastic strain rates obeying the normality flow rule based on a kinematic hardening von Mises yield surface. The normality flow rule is expressed as,

$$\dot{\varepsilon}_{ij}^{p} = \frac{1}{h}\langle \dot{S}_{ij} n_{kl} \rangle n_{ij} \tag{5}$$

where v=<u> denotes v=u if u≥0 (plastic loading) and v=0 if u<0 (elastic unloading). $\dot{S}_{ij}$ in Eq.(5) is the deviatoric stress tensor rate and can be expressed in terms of the orthogonal components as,

$$\dot{S}_{ij}=\dot{\sigma}_{ij}-\frac{1}{3}\delta_{ij}\dot{\sigma}_{kk} \tag{6}$$

The unit normal, n_{ij} , in Eq.(4) and Eq.(5) can be expressed as,

$$n_{ij}=\frac{S_{ij}-\alpha_{ij}}{\sqrt{2}k} \tag{7}$$

where α_{ij} is the deviatoric back stress tensor and k is the yield stress in shear.
The yield surface is defined as,

$$f=\frac{1}{2}(S_{ij}-\alpha_{ij})(S_{ij}-\alpha_{ij})-k^2=0 \tag{8}$$

The method of determining the increment of the deviatoric back stress depends on the plasticity hardening rule used. In the finite element analyses[8-10] from which the results are to be used as the references for comparisons, the Prager-Ziegler hardening rule had been used. Prager-Ziegler hardening rule can be expressed as:

$$\dot{\alpha}_{ij}=<\dot{S}_{kl}n_{kl}>n_{ij} \tag{9}$$

We now proceed with an alternate plasticity formulation better suited to handle non-proportional loading histories.

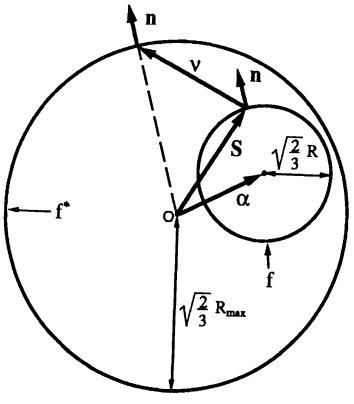


Fig.2 Yield and Bounding Surfaces in Deviatoric Stress Space

A two surface plasticity model is shown in Fig.2. The terms **n**,**S**,α in Fig. 2 represent second order tensors. The increment of deviatoric back stress in the two surface model is expressed according to Mroz hardening rule [12], i.e.

$$\dot{\alpha}_{ij}=\frac{\dot{S}_{kl}n_{kl}}{\nu_{mn}n_{mn}}\nu_{ij} \tag{10}$$

where

$$\nu_{ij}=S^{*}_{ij}-S_{ij} \tag{11}$$

is the direction of translation of the inner surface;

$$S^{*}_{ij}=\frac{R_{max}}{R}(S_{ij}-\alpha_{ij}) \tag{12}$$

where S^{*}_{ij} is a point on the bounding surface f* . This surface is defined as

$$f^{*}=\frac{3}{2}S^{*}_{ij}S^{*}_{ij}-R^{2}_{max}=0 \tag{13}$$

where R_{max} is the bounding surface radius given by,

$$R_{max}=\begin{cases}R^{*} \text{ if } \max[\sqrt{\frac{3}{2}S_{ij}S_{ij}}]\le R^{*}\\ \max[\sqrt{\frac{3}{2}S_{ij}S_{ij}}] \text{ if } \max[\sqrt{\frac{3}{2}S_{ij}S_{ij}}]>R^{*}\end{cases} \tag{14}$$

For the sake of simplicity, the plastic modulus function suggested by Drucker and Palgen[13] is used,

$$h=\frac{2}{3}H=\frac{W}{(\sqrt{J_2}/k)^{b}} \tag{15}$$

where W and b are constants and J_2 is the second stress invariant, h is the plastic modulus function, and H is the plastic modulus under uniaxial tension-compression loading.

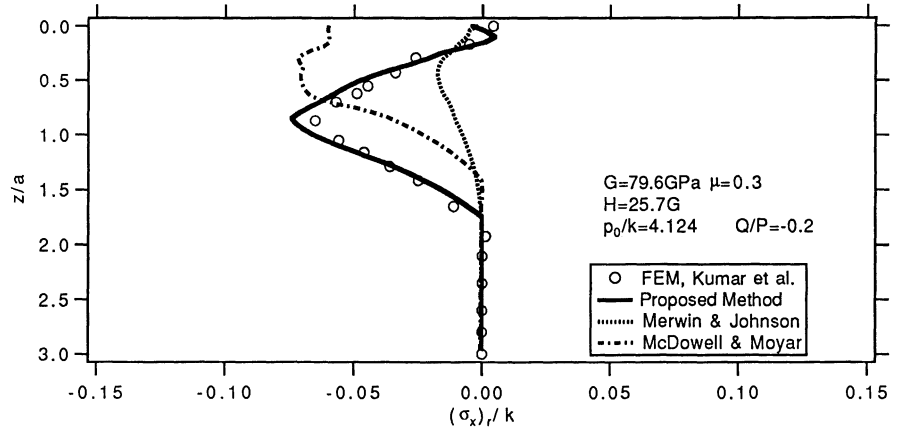
For 1070 steel, typically used in railroad wheels, we use the cyclic stress-strain parameters are as following:

G=80GPa; μ=0.3; k=139MPa; $\frac{H}{G}=\frac{4.11}{(\sqrt{J_2}/k)^{2.1}}$; $R^{*}=4R$

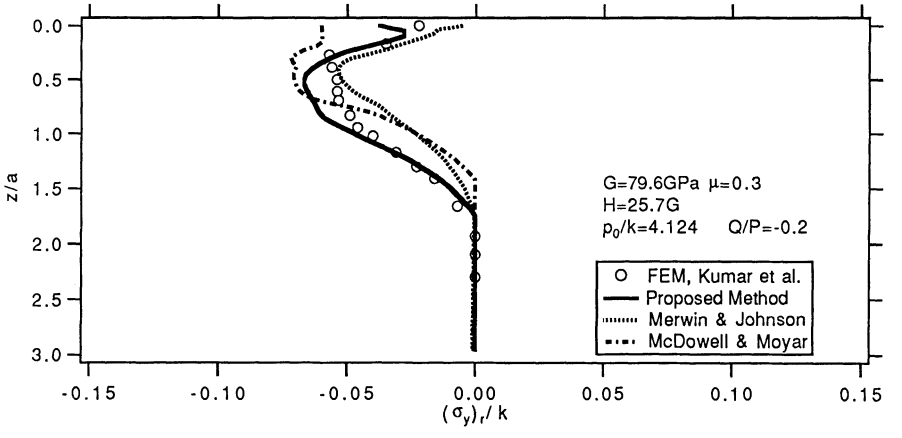
Results and Discussion

In Fig. 3, results from Kumar et al.'s [10] FEM analyses are reproduced and compared with those obtained using thethree approaches. Both residual stresses in the rolling direction, $(\sigma_x)_r$, and the axial direction, $(\sigma_y)_r$, are shown. The material properties of a bearing steel are used, which has a yield strength in pure shear, k, of 606MPa and a plastic modulus, H, of 25.7G, where G is the elasticity modulus in shear. The yield stress is approximately the proportional limit according to Kumar et al.[10]. The tangential force is negative according the coordinate system shown in Fig. 1. The tangential force is assumed to be proportional to the normal load. The same stress-strain relation and hardening rules are utilized in the calculations using the three

approcahes. The residual stresses after the 11th passage of the load are taken to be the steady state for both Merwin and Johnson, the proposed method, and the FEM analyses. The residual stress state after the 21st passage of the load is taken to be the steady state when the McDowell and Moyar method is used. More passages of the load are usually necessary to reach a steady state when the McDowell and Moyar method is employed. Normalized quantities are employed. Stresses and Hertzian pressure are normalized with respect to the yield stress in pure shear, k. Plastic modulus is normalized with respect to the elasticity modulus in shear, G. All lengths are normalized with respect to a, the half width of the contact area.



(a) Residual Stress in the Rolling Direction



(b) Residual Stress in the Axial Direction

Fig.3 Variations of Steady State Residual Stresses $(\sigma_x)_r$ and $(\sigma_y)_r$ Predicted by Different Methods

In Fig.3 the Merwin and Johnson method predicted virtually the same residual stress zone size as the FEM solution but underestimated greatly the residual stress in the rolling (x) direction. The locations of maximum residual stress in the rolling (x) direction is about 0.5a beneath the contact surface according to the Merwin and Johnson method while the FEM and the proposed method predicts the value to be about 0.9a. Based on the McDowell and Moyar method, the residual stress in the rolling (x) direction is the same as that in the axial (y) direction, and the maximum residual stresses in both the rolling (x), and axial (y) directions appear on the contact surface. On the other hand, the FEM and the other two methods predict the maxima at the subsurface.

The residual stress results predicted by the proposed method are in very close agreement with the FEM predictions. Furthermore, the locations of the maximum residual stress and the size of the residual stress zone predicted by the proposed method conforms to the FEM results. More comparisons, available in Ref.[14], lead to the same conclusion.

The FEM results [8-10] indicate that the residual stress magnitudes are highly dependent on the plastic modulus; with the same load ratio p_0/k, the larger the plastic modulus, the smaller the residual stresses.

The theoretical validity of Eq.(1) and the relaxation procedure, on which the proposed approach is based, will not be discussed in this paper. Because it is an approximate approach, more concerns are placed on the agreement between the results obtained from the approach and those obtained from finite element analyses. However, several characteristics of the proposed method should be pointed out. The fact that both Merwin and Johnson and the proposed method can predict the residual stress zone sizes virtually identical to those predicted by the FEM indicates that beyond the plastic zone, the stresses can be accurately represented by Smith and Liu formulas[11]. Because the plastic zone is small and the stresses among the elastic and plastic regions should satisfy continuity, it is appropriate to characterize the stresses in the plastic zone with elastic solutions. Then, relaxation is imposed to meet the boundary conditions. The exceptional agreements of the results of the proposed method and those from the FEM as shown in Fig. 3 provide strong positive support for the assertion. The proposed method, due to the elastic stress assumption, should only be applied to the cases where the material strain hardens. This restriction is not severe as elastic-perfectly plastic behavior could be nearly simulated with a low hardening rate. Moreover, it has been noted that the proposed theory is limited to small plastic deformation problem, where the constraint within the plastic zone is amenable to elastic approximations.

Presented in Fig.4 are the analytical predictions of residual stresses in circumferential and axial directions with the experimental results from Pomeroy and Johnson [15]. Experimental measurements of the circumferential and axial components of residual stress were taken in an aluminum alloy disc under sliding rolling contact. Because the material constants used in the elastic-plastic stress analysis were not available in the Ref.[15], the stress-strain relation of the material is assumed to be bilinear. The plastic modulus of the aluminum alloy is assumed to be 2.0G, where G is the elasticity modulus in simple shear. A two surface plasticity model

described previously is used in the analysis. Q represents the total tangential force and P represents the total normal load. The distribution of the tangential force is assumed to be proportional to the normal load. The model leads to predicted residual stress results that agree closely with the experimental data both in magnitude and distribution pattern.

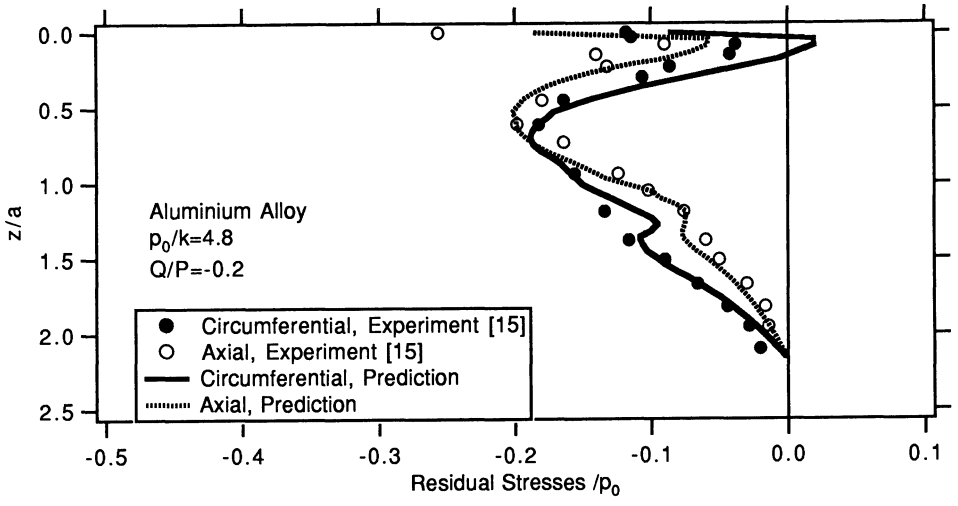


Fig.4 Comparison of Experimental Results with Theoretical Predictions with Aluminum Alloy

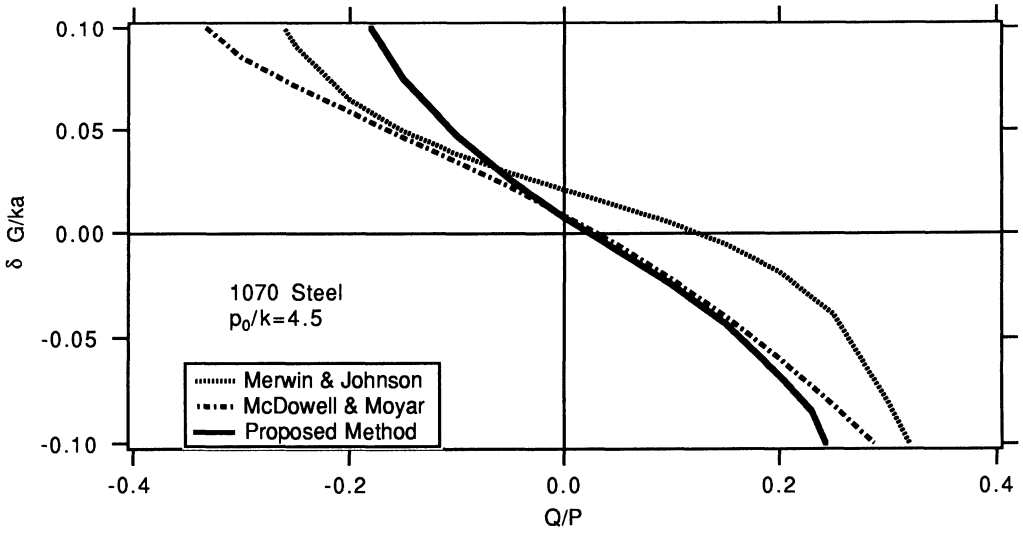


Fig.5 Comparison of the Relative Surface Displacement per Passage of Load Predicted by Different Methods

The surface displacement rate is related to the residual shear strain increment, $\Delta(\gamma_{xz})_r$, in the following manner, $\delta=\int_0^\infty \Delta(\gamma_{xz})_r dz$, where δ is the surface displacement rate. A positive δ value is in accord with a forward flow according to the coordinate system in Fig.1. The results of the surface displacement rate predicted by the three methods are compared in Fig.5. The model results show that the direction of the displacement is predicted to be the same by both the Merwin and Johnson method and the proposed method; however, the proposed method predicts a larger surface displacement rate. Note that the strain hardening material properties of 1070 steel have been used for all the three methods. According to the coordinate system in Fig.1, a negative Q/P ratio corresponds to the load on a driven wheel and a positive Q/P corresponds to the contact of a driving wheel. At the same Q/P values (i.e. positive versus negative) both Merwin and Johnson and the proposed methods predict that a driven wheel experiences greater ratchetting rates compared to the driving wheel which implies earlier failure for the driven wheel. This prediction is in agreement with the experimental observations [16]. Surface displacement obtained from the McDowell and Moyar method is approximately the same as that predicted by the proposed method when the Q/P ratio is small. However, the McDowell and Moyar method predicts that a driving wheel experiences greater surface displacement rate compared to the driven wheel.

Conclusions

Based on the hypothesis of elastic stresses and the employment of a stress and strain relaxation procedure, an analytical approach has been developed for the approximate determination of residual stresses and strains in rolling contact problems. For line contact cases, the proposed method has been shown to provide residual stress results comparable to the finite element method predictions. The proposed method predicts that a driven wheel experiences greater plastic deformation than the driving wheel. This calculation concurs with the experimental observations. The analytical predictions of the residual stresses by the proposed method are substantiated with the experimental findings. The proposed approach is especially appropriate for cases where plastic deformations are small. Further research on the role of constitutive models on the stress results, surface dispacement rates, and fatigue life prediction is in progress.

Acknowledgments This research was supported by Association of American Railroads (AAR), Technical Center, Chicago, Illinois. The corporation of Dr. Dan Stone, Executive Director and Mr. Michael Fec, Senior Engineer, of AAR, is appreciated. Collobarations with Dr. Gerald Moyar, Moyar Technical Services, and Dr. Roger Steele, of AAR, were valuable.

References

1. Stone, Daniel H. and Moyar G.J., "Wheel Shelling and Spalling--An Interpretive Review", 1989 ASME W.A.M

2. Merwin, J.E. and Johnson, K.L., "An Analysis of Plastic Deformation in Rolling Contact", Proceedings, Institution of Mechanical Engineers, Vol.177, No.25, 963, pp.676-685
3. Johnson, K.L. and Jefferis, M.A., "Plastic Flow and Residual Stresses in Rolling and Sliding Contact", Proceedings of the Symposium on Fatigue in Rolling Contact, March 1963, Paper 5, pp.54-65
4. Bhagava, V., Hahn, G.T., and Rubin, C.A., "An Elastic-Plastic Finite Element Model of Rolling Contact. Part I: Single Contacts, Part II: Repeated Contacts", Journal of Applied Mechanics, Trans. ASME, Vol.52, 1983, pp.66-74, pp.75-82
5. Hearle, A.D. and Johnson, K.L., "Cumulative Plastic Flow in Rolling and Sliding Line Contact", Journal of Applied Mechanics, Trans. ASME, Vol.54, March 1987, pp.1-7
6. Bower, A.F. and Johnson, K.L., "The Influence of Strain Hardening on Cumulative Plastic Deformation in Rolling and Sliding Line Contact", Journal of Mechanics and Physics of Solids, Vol.37, No.4, 1989, pp.471-493
7. McDowell, D.L. and Moyar, G.J., "Parametric Study of Cyclic Plastic Deformation in Rolling and Sliding Line Contact with Realistic Nonlinear Material Behavior", Wear, Vol.144, 1991, pp.19-37
8. Hahn, G.T., Bhargava, V., Rubin, C.A., Chen, Q., and Kim, K., "Analysis of the Rolling Contact Residual Stresses and Cyclic Plastic Deformation of SAE52100 Steel Ball Bearings", Journal of Tribology, Trans. ASME, Vol.109, Oct., 1987, pp.618-626
9. Ham, G.L., Hahn, G.T., Rubin, C.A., and Bhargava, V., "Finite Element Analysis of the Influence of Kinematic Hardening in Two-Dimensional, Repeated, Rolling-Sliding Contact", Tribology Transactions, Vol.32, No.3,; 1989, pp.311-316
10. Kumar, A.M., Hahn, G.T., Bhargava, V., and Rubin, C., "Elasto-Plastic Finite Element Analyses of Two-Dimensional Rolling and Sliding Contact Deformation of Bearing Steel", Journal of Tribology, Trans. ASME, Vol.111, April 1989, pp.309-314
11. Smith, J.O. and Liu, C.K., " Stresses due to Tangential and Normal Loads on an Elastic Solid with Application to Some Contact Stress Problems", Journal Applied Mechanics, Trans ASME, Vol.20, June 1953, pp.157-166
12. Mroz, Z., "On the Description of Anisotropic Workhardening", Journal of Mechanics and Physics of Solids, Vol.15, No.3, May 1967, pp.163-175
13. Drucker, D.C. and Palgen, L., "On Stress-Strain Relations Suitable for Cyclic and Other Loading", Journal of Applied Mechanics, Trans. ASME, Vol.48, Sept.,1981, pp.479-485
14. Jiang, Y. Roger and Sehitoglu, H., "Fatigue and Stress Analyses of Rolling Contact", Materials Engineering-Mechanical Behavior, Report No.161, Fracture Control Program, College of Engineering, University of Illinios at Urbana-Champaign, Feb. 1992
15. Pomeroy, R.J. and Johnson, K.L., "Residual Stresses in Rolling Contact", Journal of Strain Analysis, Vol.4, No.3, 1969, pp.208-218
16. Shima, M., and Okada, K., " Measurements of Subsurface Plastic Flow in Rolling Contact", Journal of JSLE, Int. Ed., No.2, April 1981, pp.75-80

CRACK PROPAGATION

Fatigue Crack Growth in Rail Steels

Anthony W. Thompson, Diane E. Albert* and G.T. Gray III†
Dept. of Materials Science and Engineering
Carnegie Mellon University
Pittsburgh, PA 15213, USA

*Now at: Institut National Polytechnique de Grenoble
†Now at: Los Alamos National Laboratory, Los Alamos, NM

Abstract

Considerable understanding of fatigue crack initiation and crack growth has been achieved in recent years, with a synthesis of concepts now available which includes crack initiation phenomena and also the growth behavior of short cracks and long cracks. This understanding has been shown to apply to rail steels in both the theoretical and experimental senses. Experimental results illustrating such extension are presented and discussed for both plain carbon and alloy rail compositions, with microstructural dependence of properties featured as a characteristic amenable to exploitation for improvement of rail performance.

Introduction

The failure of railroad rails in service can cause catastrophic derailments of trains, resulting in extensive material damage and human injury. There has accordingly been considerable work in the railway engineering, mechanical engineering, and metallurgy communities to study this problem, to improve understanding of rail behavior and, in particular, to attempt some predictability of rail failures. In the United States, such research has since 1975 been guided and in large part sponsored by the Association of American Railroads (AAR), the American Railway Engineering Association (AREA), and the U.S. Dept. of Transportation (DOT). Similar work in Europe is described in Esveld's book [1]. The research has been aimed at inspection frequency and methodologies, as well as at mechanics approaches to understanding the mechanical environment in which cracks in rail grow and eventually cause failure of the rail, so that particular inspection results can be used to suggest rail replacement schedules.

Work of this kind has a fairly long history, with reasonably sophisticated work dating back to Timoshenko and Langer's work in 1932 [2]. The ASTM conference book, Special Technical Publication 644, entitled *Rail Steels – Developments, Processing and Use* [3], was important in bringing together a wide range of technical viewpoints on the problem. Included were an effort to apply fracture mechanics to rail performance [4], as well as analyses of rail metallurgy issues [5,6] and data on fatigue behavior [7-9]. At about the same time, Hyzak and Bernstein [10] conclusively determined for the first time the virtually unique independence of the microstructural bases for yield strength and toughness in rail

J.J. Kalker et al. (eds.), Rail Quality and Maintenance for Modern Railway Operation, 361–372.

steels, the former property depending only on the interlamellar spacing of the pearlite microstructure, the latter depending only on prior austenite grain size. (These microstructural aspects are discussed further below). This result had important implications for rail steel processing [5,11] as well as for understanding of the relation between properties and metallurgical microstructure in rail steels.

The major concern about rail failure rests with cracking transverse to the long axis of the rail. Such a crack, typically initiated in the rail head from shelling cracks near or in the running surface of the rail, is called a "detail fracture" in North America, and that terminology is used here. An example of a detail fracture, taken from a rail in service, is shown in Fig. 1. Several reviews have addressed the construction of an integrated analysis of this type of fracture [4,12-16]. Very helpful data for such analysis came from tests at the U.S. Transportation Test Center in Pueblo, Colo., conducted in the track of the Facility for Accelerated Service Testing (FAST) [15,17].

The analytical approach which has emerged combines the following factors. First, the rail stress environment is quantified, a complex problem involving not only the cyclic, longitudinal stresses arising from live loads due to passage of trains [4,12,14-18], but also the residual stresses, initially from processing and subsequently developing in service from wheel/rail contact [1,5,12,15,17], and the thermal stresses [12,14], diurnal and seasonal, particularly in continuous welded rail (CWR). Modeling and measurement can provide the spatial variation of each of these stresses within the rail section, particularly the rail head where detail fractures develop. The result will typically be a model of the following form,

$$\text{Net (cyclic) stress} = \Sigma\,(\text{Live, residual, thermal stresses}) \qquad \{1\}$$

where the summation indicated is not simply an algebraic summation but is the appropriate combination of the relevant stresses. Once these stresses are determined, one can apply fracture mechanics to determine crack growth driving forces, and then utilize laboratory measurements of fatigue cracking in rail steels to obtain crack growth rates.

The latter measurements have gradually become more detailed and sophisticated as

Fig. 1. An example of an incipient detail fracture of a rail in service. The rail is a Colorado Fuel & Iron 133 RE (AREA 133 lb./yd.) section, from Union Pacific Railroad track near Laramie, Wyoming, in service from 1957 to 1975. After detection of the internal crack, a section of rail containing the defect was removed from the track and broken at the AAR Chicago Test Center to reveal the fatigue crack submerged in the rail.

knowledge about fatigue behavior of structural materials has deepened. Following early work [7-9,19], summarized and averaged by Orringer, *et al.* [12], much more detailed work has been carried out [20-27], taking advantage of the microstructural knowledge obtained by Hyzak and Bernstein [10]. Some of these measurements have concentrated on the behavior of small cracks, a potentially important subject in rail fatigue life, as discussed further below.

There is a long-standing problem in understanding the interrelations among rail wear, initiation and growth of shell cracks parallel to the rail surface, and the change of shell cracking to detail cracking [5,12,14-16,28,29]. Work on this broader area, like that on detail fractures alone, spans the range from metallurgical fundamentals to applied mechanics [28-36]. The analytical models now appear to obtain results which are in reasonable agreement with service observations [29,37,38], although material deformation properties are typically used only in a superficial manner. The role of spectrum loading of rail during train passage has also been addressed in a series of laboratory studies [39-43], and reasonable agreement found with data from FAST [14,16,42]. Interpretations of laboratory results have included such concepts as crack closure and crack retardation by crack tip plastic zones, concepts which are addressed further below.

Each of the foregoing issues has evident consequences for rail service, from installation conditions and service loadings, through inspection intervals and procedures, to rail replacement or failure [1,44,45]. Some can likewise be extended to address related microstructures in railroad materials, such as wheel steels [46]. Most of these are beyond the scope of the present paper, although all have been treated to varying degrees in the references cited in this paper.

Fatigue Crack Propagation

Macroscopic fatigue cracks are now known to grow in response to the range of applied stress intensity, ΔK, with crack length a increasing with cycles N in a linear dependence on a power of ΔK, over a wide range of growth rates. This is called the Paris law [47],

$$da / dN = C (\Delta K)^m; \quad \text{alternatively,} \quad da / dN = C (\Delta K)^m / (1 - R), \qquad \{2\}$$

where $R \equiv \sigma_{min} / \sigma_{max} = K_{min} / K_{max}$ and C and m are constants (see Fig. 2). This applies to a wide variety of materials, including rail steels [5,7-9,12,14,22]. Typical values of m are about 4, but there is a very large variation in fitted values of C [14], evidently the result of limited ranges of data. Clearly, data extending over the whole accessible range of ΔK in the linear or Paris law region of the da/dN vs. ΔK curve are essential to accurate fitting.

Thus once a fatigue crack has initiated in a material and undergone the earliest stages of growth, it becomes subject to fracture mechanics in its further growth. From the laboratory perspective, then, determination of total fatigue life requires measurement of both the fatigue crack growth behavior of long cracks, which are under ΔK control, and that of shorter cracks, near to initiation conditions. The latter situation does not clearly fall into fracture mechanics, and it is possible that stress control may apply for initiation and earliest growth of fatigue cracks [22,48]. A "short crack" is one which is short relative to normal fracture mechanics size requirements [48]. For example, it may be short with respect to the relative size of the plastic zone; many short cracks have computed plastic zone sizes of a tenth or more

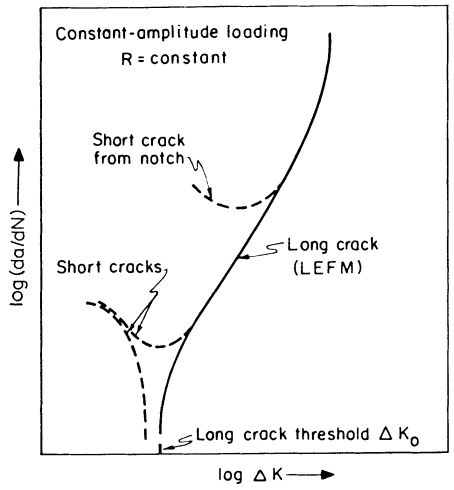


Fig. 2. Schematic illustration of the dependence of crack growth rate, da/dN (a = crack length, N = number of stress cycles) on applied stress intensity, ΔK, for both long and short cracks. The linear region of the long-crack or LEFM (linear elastic fracture mechanics) line is called the "Paris law region" after its discoverer, P.C. Paris [47,48]. From ref. 48.

of the crack length. It may also be considered short if it is comparable in size to the microstructural dimensions which control mechanical behavior of the material. In such conditions, Fig. 2 shows that short cracks can grow significantly faster than long ones.

Experimental Results

The microstructure of North American, and some European rail steels is composed entirely of a constituent called pearlite. This comprises fine lamellae of alternating ferrite (low-carbon iron) and iron carbide (Fe_3C) or cementite. Figure 3 illustrates an actual microstructure and, schematically, shows the three size scales on which this microstructure may be described: the prior austenite grain size, the size of pearlite colonies, and the spacing of the pearlite lamellae or plates. These three dimensions are determined in a commercial material by processing conditions, and in the laboratory can be manipulated over certain ranges by appropriate heat treatment [10,22,49-51]. Details of deformation behavior in this microstructure have been studied by transmission electron microscopy [34-36].

It might be expected that microstructural variations could readily affect fatigue crack propagation, particularly for short cracks. This is certainly true for rail steels, in which microstructural control of yield strength and toughness [1,5,10,49,50] has effects on fatigue

Table I
Chemical Composition of Rail Steels, wt. pct.

Material*	C	Mn	Si	S	P	Cr	Mo	Al	N	Fe
AISI 1080	0.80	0.84	0.17	0.013	0.018	-	-	-	-	bal.
Cr-Mo	0.75	0.58	0.39	0.023	0.021	0.61	0.20	0.015	0.008	bal.
1CR	0.73	1.3	0.3	0.026	0.023	1.25	-	-	-	bal.
CTL	0.76	1.24	0.84	0.009	0.023	-	-	-	-	bal.

*Alloy rail descriptions in text

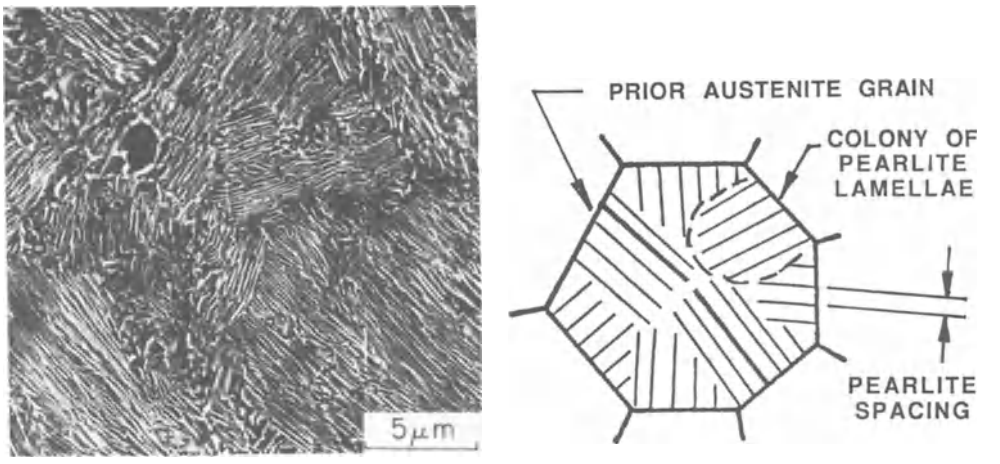


Fig. 3. Microstructure of pearlite, typical of rail steels. (a) Scanning electron micrograph showing lamellae of ferrite and cementite. (b) Schematic microstructure; dark lines are cementite.

behavior [20-22]. Experiments to test this possibility were conducted on several microstructures of the same AISI 1080 rail steel, Table I, with the mechanical properties [51] shown in Table II. Long cracks were found to propagate at rates which are sensitive to the prior austenite grain size, as shown in Fig. 4, and to be insensitive to pearlite spacing. This result arises from crack closure [23], not in the original Elber sense [52] of what is now called "plasticity-induced closure" but due to roughness of crack surfaces at low crack growth rates [48]. This conclusion is supported by tests at high R ratio, in which the grain size dependence disappears, and by tests in helium to rule out closure due to fretting oxidation [22,48], as shown in Fig. 5. Analytical models have also concluded that plasticity-induced closure does not play a major role in fatigue of rail steels [53].

Table II
Example Properties of Rail Steels

Microstructure*	Hardness, Rc	Charpy Impact Energy, joules	Yield Strength, MPa	Endurance Limit, MPa
1080: FG / FC / FS	27.4	2.71	344	593
1080: FG / FC / CS	17.3	2.03	293	406
1080: CG / CC / FS	30.0	1.35	379	643
1080: CG / CC / CS	21.5	1.35	293	457
1080: CG / FC / FS	26.5	1.35	344	617
1080: CG / FC / CS	18.2	1.35	282	411
1080: AR	24.1	1.35	310	510
Cr-Mo: AR	36.0	5.42	448	772

*FG, CG = fine, coarse prior austenite grain size; FC, CC = fine, coarse pearlite colony size; FS, CS = fine, coarse pearlite spacing; AR = as-received.

There seem to be unresolved questions about crack growth rates under spectrum loading, in particular as to whether plastic zones created during pre-cracking could affect subsequent results [42,43]. There are procedures to minimize such effects, such as measuring da/dN during load shedding (reducing ΔK) in steps of 10 pct of σ_{max} or less, then allowing crack growth for a distance of at least three times the cyclic plastic zone size generated at the previous (higher) load level [22]. Such procedures, combined with closure measurements, should suffice to determine if plastic zone effects are significant under spectrum loading.

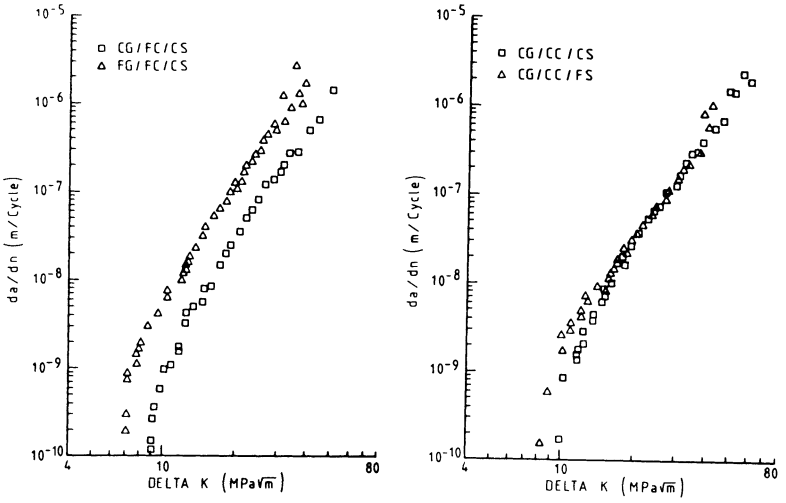
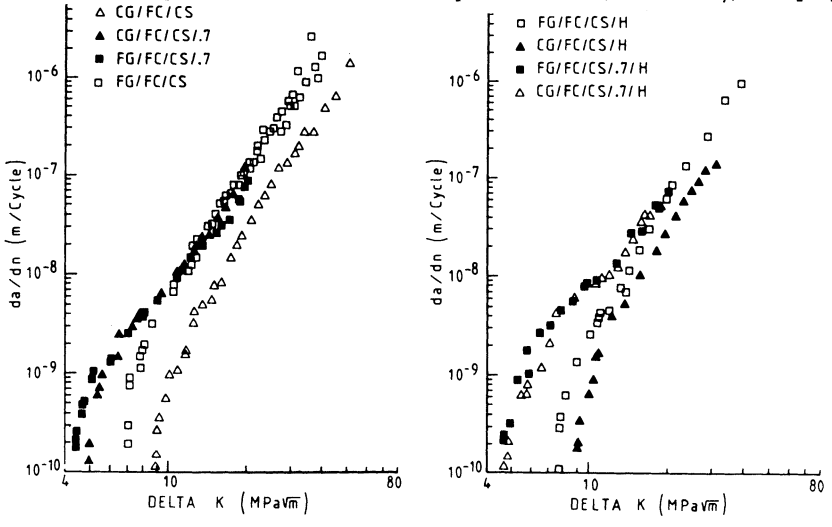


Fig. 4. (Left) Dependence of crack growth on grain size. **(Right)** Dependence of crack growth on pearlite spacing. Microstructure coding shown in Table II. From Gray, *et al.* [22].

Fig.5. (Left) Dependence of crack growth on R ratio, for values of 0.05 and 0.7. **(Right)** Dependence of crack growth on environment (H = purified helium). From Gray, *et al.* [22].



In light of results such as Figs. 4 and 5, it may be an oversimplification to use an "average" fit to crack growth data [12,14], instead of recognizing the effects of metallurgical microstructure. This point is not a new one (cf. the discussion by Fletcher [54] to Barsom and Imhof's paper [8]), but with modern experimental precision, should be revisited. Barsom and Imhof's assertions that microstructure played no role in their results is patently unsupported. In their reply to Fletcher, they cited *interlaboratory* experimental variations in an effort to dismiss Fletcher's statistical analysis [54] of the results from their laboratory alone. Fletcher's analysis appears sound, and like the results presented here, shows that careful analysis of metallurgical factors (composition, processing, and microstructure) is needed. Not all metallurgical factors necessarily result in fatigue behavior differences, but each case must be examined separately.

Initiation of cracks in rail steels has often been studied by smooth-bar tests, such as rotating bending or push-pull loading. Results [8,26] typically show a microstructural dependence; Fig. 6 is an example. Also included here is a result for Cr-Mo rail steel, which is discussed below. Endurance limit data, S_e, in Table II were fitted to a regression equation on the yield strength, σ_y (regressions on ultimate tensile strength were also succesful),

$$S_e = 0.42\ \sigma_y + 103.5, \qquad \{3\}$$

where the stresses are in MPa [51]. The correlation coefficient for this fit was r = 0.9707. Smooth-bar and notched-bar specimens have been compared in their initiation behavior, and the classical effects of stress concentration [55] were found relevant to interpretation of the results [23]. This is important for rails, since surface flaws, wear, and local hardening quickly provide notch-like surface defects which can act as crack initiators.

Alloy rail steels were also studied [51,56], with compositions as shown in Table I. Here the Cr-Mo is a weldable Colorado Fuel and Iron product, with nominally 0.6 Cr and 0.2 Mo; the 1CR steel is a British Steel Corp. development with nominally 1% Cr; and CTL is a high-silicon, weldable rail from Nippon Steel [57]. The 1CR steel is of particular interest, since the usual rail steel used in Britain conforms to B.S. 11, a 0.5% C composition which is not fully pearlitic [1,30,58]. The S-N performance of the Cr-Mo steel is shown in Fig. 6; the

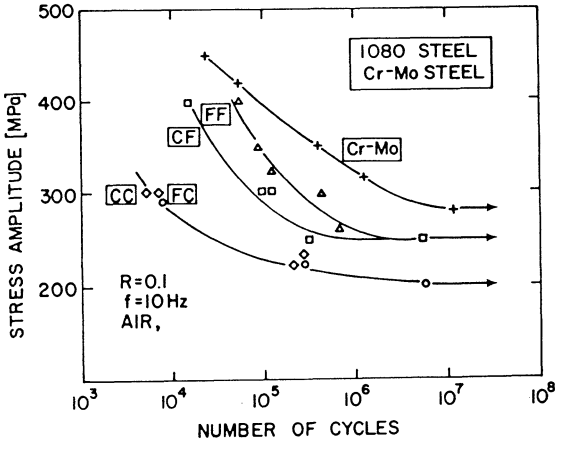


Fig. 6. S-N curves for smooth-bar, constant stress amplitude tests of rail steels. Each point plotted averages several tests. Microstructures as listed in Table II, except coding is simplified; here FC = fine grain, coarse spacing. From Daeubler, *et al.* [26].

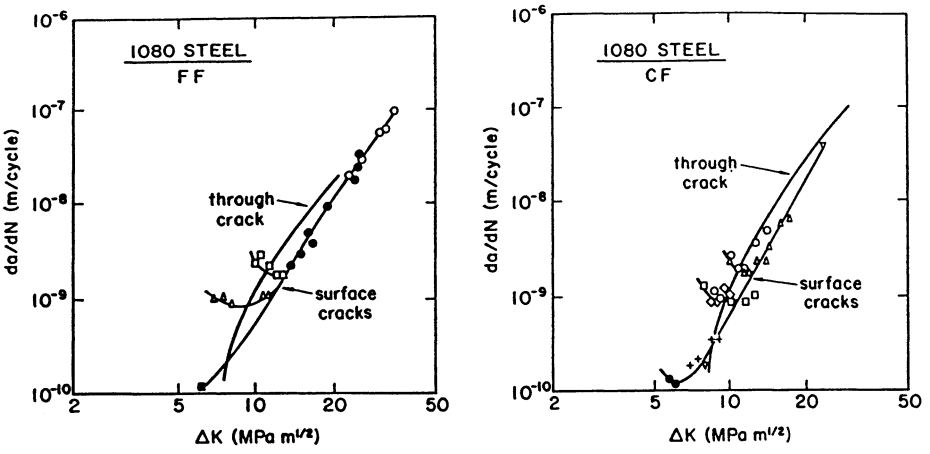
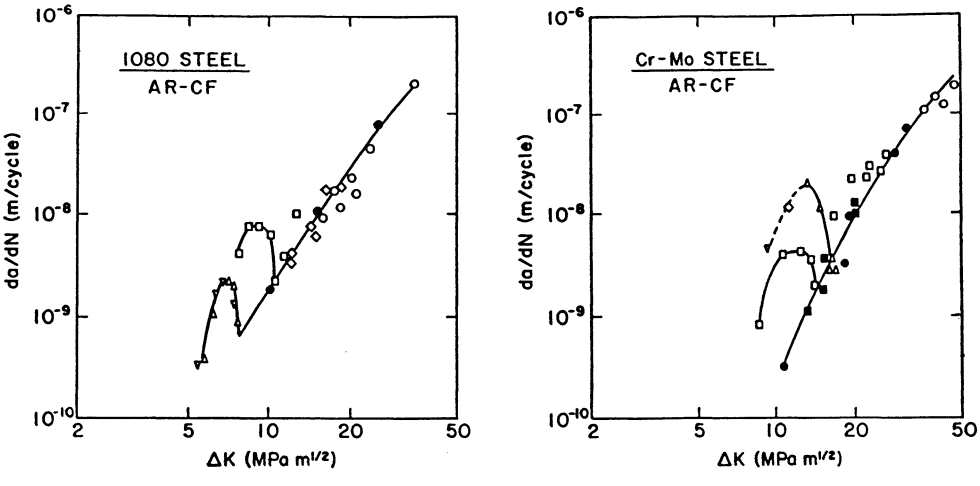


Fig. 7. Examples of behavior of short cracks for two microstructures, coded as in Fig. 6, compared to trend lines for long cracks as measured in ref. 22. Filled data points represent measurements by "load shedding," i.e. with long, not short, cracks. From Daeubler, *et al.* [27].

short-crack results are shown in Fig. 7, with additional results on as-received (coarse grain, fine spacing) materials in Fig. 8. It will be noted that the short-crack data follow the pattern expected [48] for such cracks in fatigue, Fig. 2, growing significantly faster until they join the trend for long cracks. Finally, the 1CR and CTL steels [56] are presented in Fig. 9, showing again that short cracks tend to propagate faster than the long-crack trend at low values of ΔK. These latter data are sparse, but suggest that additional research, particularly as a function of microstructure in alloy steels, would be of value.

Fig.8. Short-crack data for plain carbon rail (left) and for Cr-Mo alloy rail (right), in the as-received condition (coarse grain size, fine spacing). Coding as in Fig. 6. From Daeubler, *et al.* [27].



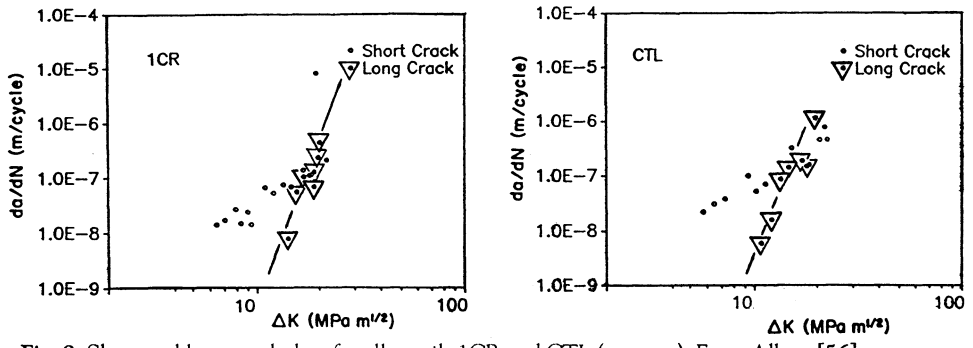


Fig. 9. Short and long crack data for alloy rails 1CR and CTL (see text). From Albert [56].

Discussion

The performance of rail steels is now well known to depend on the metallurgical microstructure. This is because mechanical properties depend on microstructure; microstructure in turn depends on composition and processing history. Thus it is inappropriate to "average" data for different steels unless careful analysis demonstrates that microstructure in fact plays no role. At least some effort should be made to determine whether any differences in performance of different steels should be interpreted in terms of their metallurgy (and, as necessary, in terms of differences in testing technique, which can vary from laboratory to laboratory, despite current recommendations for test methods, such as ASTM E647). Combining all available data, from a large number of individual samples, into a single plot naturally tends to obscure real differences among materials and create a "band" of data points [8,12], creating the false impression that an average trend adequately represents behavior.

One microstructural feature which can cause closure effects is grain size, e.g. Fig. 4. Effects of R ratio have been seen before, particularly in moist air [12], for which closure concepts [48] would predict retardation, not the observed acceleration [12] of growth (raising the possibility of direct effects of moisture or hydrogen). It is essential to account for all variables which could produce crack closure, and microstructure is clearly one of these.

Experimental data are still needed on the role of residual stresses, either from processing or more importantly, from service, on the propagation of fatigue cracks. Acquisition of such results would be complex in that the most convincing data would be measured on steels with varying amounts and patterns of residual stress, no small matter experimentally. The same can be said of thermal stresses. Since extant models of rail behavior appear to depend significantly on residual stress, it would seem particularly important to obtain relevant measurements with which to compare the model results. Similar remarks could be made about rail welds. Microstructures developed in flash or thermite welds [1] need exploration. Properties are known to be degraded from parent rail, and fatigue data seem essential.

Conclusions

The following conclusions can be drawn from the results shown and discussed here, and

also supported by the numerous other studies cited in the appended references.

(1) Properties of rail steels, such as yield strength, endurance limit, toughness (whether plane strain fracture toughness (K_{Ic}) or Charpy V-notch energy), and various parameters for fatigue crack growth, depend on metallurgical microstructure. Microstructure in turn depends on rail composition and processing history. Thus variations in properties among various published studies should not necessarily be attributed to "scatter" but should be assessed for microstructural effects.

(2) Crack closure does occur in rail steel, although for reasons of fracture surface roughness, not plasticity-induced closure. It acts to decrease da/dN at low R ratio, below what would be inferred from the Paris-law behavior, equation {2}, and thus is conservative in nature, compared to use of equation {2} alone.

(3) There now exist extensive data on microstructure effects in plain-carbon rail steel, but data are still sparse on alloy rail compositions. Moreover, there is an evident need to acquire such data for in-service rail, rail welds, and for residual stress effects, particularly from straightening and from changes experienced by rail in service.

Acknowledgements

We appreciate helpful discussions with I.M. Bernstein, D.H. Stone, and R.K. Steele, and bibliographic help from O. Orringer. Experimental work on rail steels at Carnegie Mellon University has been sponsored by the AAR Chicago Technical Center, through an AAR Affiliated Laboratory, with D.H. Stone as technical liaison. Preparation of this manuscript was supported by the AAR and by the David Jackson Fund for Scholarly Endeavors.

References

1. C. Esveld, *Modern Railway Track*, Thyssen Stahl AG, Duisburg (1989).
2. S. Timoshenko and B.F. Langer, *Trans. ASME*, 54 (1932) 277-93.
3. D.H. Stone and G.G. Knupp, eds., *Rail Steels – Developments, Processing and Use*, STP 644, ASTM, Philadelphia, 1978.
4. P.M. Besuner, in *Rail Steels – Developments, Processing and Use*, STP 644, D.H. Stone and G.G. Knupp, eds., ASTM, Philadelphia (1978), 303-329.
5. D.H. Stone and R.K. Steele, *ibid.*, 21-62.
6. G.K. Bouse, I.M. Bernstein and D.H. Stone, *ibid.*, 145-61.
7. G.J. Fowler and A.S. Tetelman, *ibid.*, 363-86.
8. J.M. Barsom and E.J. Imhof, *ibid.*, 387-413.
9. C.E. Fedderson and D. Broek, *ibid.*, 414-29.
10. J.M. Hyzak and I.M. Bernstein, *Metall. Trans. A*, 7A (1976) 1217-24.
11. W. Heller, R. Schweitzer and L. Weber, *Canadian Metallurgical Quart.*, 21 (1982) 3-15.
12. O. Orringer, J.M. Morris and R.K. Steele, *Theor. Appl. Fracture Mechanics*, 1 (1984) 23-49.
13. A.S. Blicblau and C.G. Chipperfield, *Theor. Appl. Fracture Mechanics*, 1 (1984) 61-71.
14. O. Orringer, J.M. Morris and D.Y. Jeong, *Theor. Appl. Fracture Mechanics*, 5 (1986) 63-95.

15. O. Orringer and R.K. Steele, in *Fracture Mechanics: Nineteenth Symposium*, STP 969, T.A. Cruse, ed., ASTM, Philadelphia (1988) 260-78.
16. O. Orringer, Y.H. Tang, J.E. Gordon, D.Y. Jeong, J.M. Morris and A.B. Perlman, "Crack Propagation Life of Detail Fractures in Rails," Report DOT/FRA/ORD-88-13, Transportation Systems Center, U.S. DOT, Cambridge, MA, Oct. 1988.
17. P. Clayton and Y.H. Tang, in *Residual Stresses in Rails*, Vol. I, O. Orringer, J. Orkisz and Z. Swiderski, eds., Kluwer, Dordrecht (1992) 37-56.
18. G.C. Sih and D.-Y. Tzou, *Theor. Appl. Fracture Mechanics*, 1 (1984) 103-15.
19. G.J. Fowler, "Fatigue Crack Initiation and Propagation in Pearlitic Rail Steels," Ph.D. thesis, University of California at Los Angeles, 1976.
20. G.T. Gray, A.W. Thompson, J.C. Williams and D.H. Stone, *Canadian Metallurgical Quart.*, 21 (1982) 73-78.
21. G.T. Gray, A.W. Thompson, J.C. Williams and D.H. Stone, in *Fatigue Thresholds*, J. Backlund, A. Blom and C. J. Beevers, eds., Vol. 1, EMAS Ltd., Warley, U.K. (1982) 345-361.
22. G.T. Gray, J.C. Williams and A.W. Thompson, *Metall. Trans. A*, 14A (1983) 421-433.
23. G.T. Gray, A.W. Thompson and J.C. Williams, *Metall. Trans. A*, 16A (1985) 753-760.
24. G.T. Gray, A.W. Thompson and J.C. Williams, in *Fatigue Crack Growth Threshold Concepts*, D.L. Davidson and S. Suresh, eds., TMS-AIME, Warrendale, PA (1984) 131-143.
25. M.A. Daeubler and A.W. Thompson, in *Small Fatigue Cracks* (Engineering Foundation Conference Proceedings), R.O. Ritchie and J. Lankford, eds., TMS-AIME, Warrendale, PA (1986) 157-164.
26. M.A. Daeubler, A.W. Thompson and I.M. Bernstein, *Contact Mechanics and Wear of Rail/Wheel Systems II*, G.M.L. Gladwell, H. Ghonem and J. Kalousek, eds., Univ. of Waterloo Press, Waterloo, Ont. (1987) 451-461.
27. M.A. Daeubler, A.W. Thompson and I.M. Bernstein, *Metall. Trans. A*, 21A (1990) 925-933.
28. T.N. Ferris, L.M. Keer and R.K. Steele, *J. Mech. Phys. Solids*, 35 (1987) 677-700.
29. T.N. Ferris, Y. Xu and L.M. Keer, in *Residual Stresses in Rails*, Vol. II, O. Orringer, J. Orkisz and Z. Swiderski, eds., Kluwer, Dordrecht (1992) 23-43.
30. D.H. Stone, S. Marich and C.M. Rimnac, Transportation Research Record 744, *Railroad Track and Facilities*, Washington (1980), 16-21.
31. P. Clayton and M.B.P. Allery, *Canadian Metallurgical Quart.*, 21 (1982) 31-46.
32. S. Kumar, V. Aronov, B.R. Rajkumar and R. Margasahayam, *Canadian Metallurgical Quart.*, 21 (1982) 59-66.
33. M. Dollar, A.W. Thompson, I.M. Bernstein and D.H. Stone, *Contact Mechanics and Wear of Rail/Wheel Systems II*, G.M.L. Gladwell, H. Ghonem and J. Kalousek, eds., Univ. of Waterloo Press, Waterloo, Ont. (1987) 315-323.
34. M. Dollar, I.M. Bernstein and A.W. Thompson, *Acta Metall.*, 36 (1988) 311-320.
35. M. Dollar, I.M. Bernstein, M. Daeubler and A.W. Thompson, *Metall. Trans. A*, 20A (1989) 447-451.

36. M. Dollar, A.W. Thompson and I.M. Bernstein, in *Strength of Metals and Alloys* (Proc. ICSMA 8), P.O. Kettunen, T.K. Lepistö and M.E. Lehtonen, eds., Pergamon, Oxford (1988) Vol. 1, 647-652.
37. A.F. Bower, *J. Tribology (Trans. ASME)*, 110 (1988) 704-11.
38. M. Olzak, J. Stupnicki and R. Wojcik, in *Residual Stresses in Rails*, Vol. II, O. Orringer, J. Orkisz and Z. Swiderski, eds., Kluwer, Dordrecht (1992) 45-62.
39. B.G. Journet and R.M. Pelloux, *Theor. Appl. Fract. Mechanics*, 7 (1987) 19-22.
40. B.G. Journet and R.M. Pelloux, *Theor. Appl. Fract. Mechanics*, 8 (1987) 117-23.
41. D. Jablonski, Y.H. Tang and R.M. Pelloux, *Theor. Appl. Fract. Mechanics*, 14 (1990) 27-36.
42. Y.H. Tang, A.B. Perlman, O. Orringer and D.A. Jablonski, *Theor. Appl. Fract. Mechanics*, 15 (1991) 1-9.
43. D.A. Jablonski and R.M. Pelloux, in *Residual Stresses in Rails*, Vol. I, O. Orringer, J. Orkisz and Z. Swiderski, eds., Kluwer, Dordrecht (1992) 81-98.
44. D.D. Davis, M.W. Joerms, O. Orringer and R.K. Steele, *Proc. Third Internat. Heavy Haul Railway Conf.*, Vancouver, Oct. 1986. (Also available as AAR Report R-656)
45. O. Orringer, "Control of Rail Integrity by Self-adaptive Scheduling of Rail Tests," Report DOT/FRA/ORD-90-05, Transportation Systems Center, U.S. DOT, Cambridge, MA, June, 1990.
46. A. Cabral, A.W. Thompson, I.M. Bernstein and D.H. Stone, *Materials Science and Engineering*, 93 (1987) 73-82.
47. P.C. Paris and F. Erdogan, *Trans. ASME (J. Basic Eng.)*, 85 (1963) 528-34.
48. S. Suresh and R.O. Ritchie, *Internat. Metals Reviews*, 29 (1984) 445-76.
49. J.J. Lewandowski and A.W. Thompson, *Metall. Trans. A*, 17A (1986) 1769-1786.
50. D.J. Alexander and I.M. Bernstein, *Metall. Trans. A*, 20A (1989) 2321-35.
51. G.T. Gray, "Influence of Microstructure on Fatigue Initiation and Propagation Properties of Pearlitic Steels," Ph.D. thesis, Carnegie Mellon University, Pittsburgh, 1981.
52. W. Elber, *Eng. Fract. Mech.*, 2 (1970) 37-45.
53. G.C. Sih and D.Y. Jeong, in *Residual Stresses in Rails*, Vol. II, O. Orringer, J. Orkisz and Z. Swiderski, eds., Kluwer, Dordrecht (1992) 63-85.
54. F.B. Fletcher, in *Rail Steels – Developments, Processing and Use*, STP 644, D.H. Stone and G.G. Knupp, eds., ASTM, Philadelphia (1978), 410-11.
55. N.E. Frost, K.J. Marsh and L.P. Pook, *Metal Fatigue*, Clarendon Press, Oxford (1974) 149-56.
56. D. E. Albert, "Growth of Short Fatigue Cracks in Pearlitic Alloy Rail Steels," M.S. research report, Carnegie Mellon University, Pittsburgh, 1988.
57. H. Masumoto, K. Sugino, H. Kageyama, S. Nishida, C. Urashima, M. Hattori, and T. Terada, "Basic Properties of High Strength Rail Steels and New Rail Development," Technical Report, Yawata Works, Nippon Steel Corp., Oct. 1982.
58. G.W.J. Waldron and S. Wise, *The Metallurgist*, January, 1973, 15-21.

CRACK BEHAVIOUR UNDER ROLLING CONTACT FATIGUE

M.C Dubourg*, J.J. Kalker**

* I.N.S.A., Laboratoire de Mécanique des Contacts, U.R.A CNRS 856, 20 Av. Einstein, 69621 Villeurbanne Cédex, France.
**Faculty of Technical Mathematics and Informatics, Delft University of Technology, Delft, The Netherlands.

Abstract

A theoretical approach on surface crack influence on rolling contact stress field is presented. A steady rolling contact model and a fatigue crack model are combined. Connection between the two problems is introduced through surface geometry modification caused by displacements generated by cracks. Split up of contact area, maximum pressure increases are highlighted. No significant stress intensity factor variations are obtained.

Appendix

P	line load
$h(y)$	geometry of contacting surfaces
V	rolling speed
ξ	creep ratio
fs	surface coefficient of friction
$2a$	contact area
$p(y)$, $q(y)$	normal and tangential tractions
b_j	length of crack j
θ_j	inclination of crack j
Γ_j	line of crack j
μ	shear modulus
ν	Poisson's ratio
k	$3 - 4\nu$
f	interfacial crack coefficient of friction
K_{lm}	kernels expressed in axis (l,m)
σ^T, u^T, v^T	resulting stress and displacement fields
σ^{MC}, u, v	continuum stress and displacement fields
σ^C, u^C, v^C	crack stress and displacement fields
δun , δut	displacement discontinuities between crack faces
b_x, b_y	burger's vectors
K_I, K_{II}	stress intensity factors (SIFs) in mode I and II respectively
V^C_{surf}	half-plane surface deformation due to crack

J.J. Kalker et al. (eds.), Rail Quality and Maintenance for Modern Railway Operation, 373–384.

Introduction

The principal types of rolling contact fatigue defects which require rail replacements are subsurface - shells - and surface cracks - headcheacks (cf. fig 1) situated at or near the rail head gauge corner in curved, moderately loaded tracks, and squats (cf. fig. 2) situated in the running bands of rails in high-speed lines -.

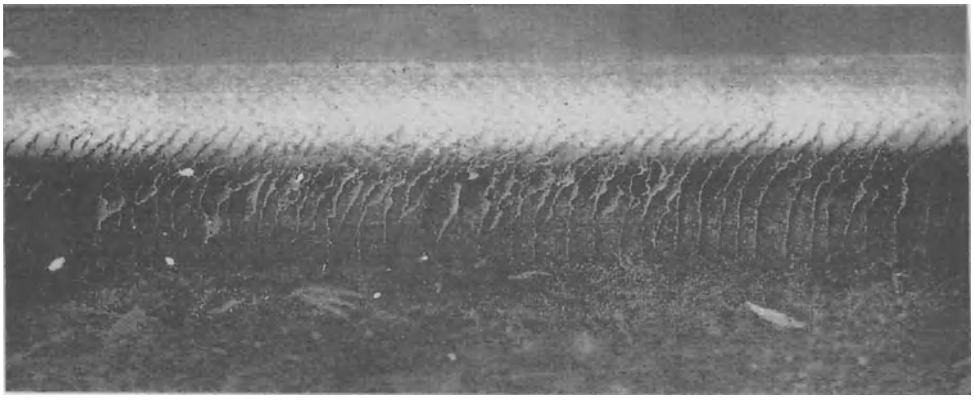


Figure 1. Headcheck. Traffic from right to left [1]

Figure 2. Vertical/longitudinal section through a squat type rolling contact fatigue defect *0.75. Traffic from right to left [1]

Linear fracture mechanics can be employed to study the propagation of such cracks [2-8]. For loading conditions characteristic of those experienced in rolling contact conditions, Bold et.al. [9] have demonstrated that crack growth is possible under shear mode loading without liquid. Bower [10] analysed arbitrarily inclined cracks, including crack face friction and lubricant effects. Multiple

interactive surface breaking or subsurface cracks, arbitrarily oriented, straight or bent were analysed in 2D by Dubourg et.al. [11-14]. Significant overloading or unloading may result from interaction between cracks. Olzak et. al. discussed the problem of the influence of a single crack on the wheel/rail rolling contact in the 2D and 3D (semi-elliptical surface breaking crack) configurations [15,16] . Their model is based on the finite element approach that implies very large computer time. Division of the contact area into two parts is observed but no comments are formulated on pressure and stress intensity factor modifications.
In this paper the analysis focuses on the mutual influence of multiple surface breaking cracks on the wheel/rail rolling contact, i.e possible split up of contact area, modifications of normal and tangential traction distributions, maximum pressure increases on one hand and on the other hand modifications of the stress intensity factors at crack tips. The wheel/rail rolling contact model and the crack model, presented successively, are based on half analytical and numerical approachs.
Steady state tractive rolling of a cylinder over an elastic half-space is considered. The tractive force applied is smaller than the limiting friction. The normal and tangential tractions, the contact and slip zones are determined numerically through unilateral contact analysis.
The half-space considered contains one or more frictional cracks. Open, slip and stick zone distribution, stresses and displacement discontinuities along crack faces and stress intensity factors (SIFs) in mode I and II at crack tips are determined. This crack model rests on the dislocation theory for modelling displacement jumps between crack faces and on the unilateral contact analysis for determining the open, stick and slip zone distribution along the crack line.
Displacements along crack faces modify the half-space surface geometry which in turn modifies the contact zone width, the normal and tangential tractions, and finally the SIFs at crack tips. This process goes on until convergence is attained, i.e when normal and tangential tractions exerted on the half-plane surface are stabilised.

I. Line contact of a cylinder over an elastic half-plane

The steady tractive rolling contact of a cylinder (wheel) on a strip of infinite length (rail) is considered. The thickness of the rail is such that plane strain assumptions are acceptable. Bending and shear stresses due to vertical and lateral wheel loads and residual stresses introduced during manufacture and service are not taken into account. No spin conditions are considered but unsaturated creep is taken into account. Perfectly elastic conditions are retained for simplicity. These simplifying assumptions are a first step towards gaining insight into the influence of cracks on the conditions at the interface between the wheel and the rail.
The steady rolling contact problem between the rail and the wheel is solved as a unilateral contact problem following the method developed by Carneiro Esteves et.al. [17]. As semi-infinite bodies are

considered, Boussinesq and Cerruti potentials are used. Relations between displacements and stresses are obtained. The potential area of contact is discretized into segments on which stresses are assumed constant. Results obtained for a given load step are the real area of contact (2a) and the slip zone, normal (p(y)) and tangential (q(y)) tractions and displacements.
The method used here for the unilateral contact solution between two rolling bodies can be applied to the interfacial crack problem solution. The former formulation is based on velocity differences etween the two bodies. The latter one is based on displacement differencies between crack faces. Nevertheless equalities and inequalities are solved in the same way.

II. Frictional fatigue crack model

The model used here is presented elsewhere [11-14]. It rests on the distribution dislocation theory, pionneered by Keer and Bryant [5,6] and on the unilateral contact analysis developed by Kalker [18,19]. Multiple interactive cracks, straight or kinked, embedded or surface breaking, situated in an isotropic elastic half-plane can be modelled. Friction is taken into account between crack faces. It leads to a drop of shear stresses and thus of KII. Load cycles are described with an incremental description which takes into account the load history, as hysteresis is introduced by the frictional contact between crack faces. The load is then applied in small steps.
Stress and displacement fields in the cracked half-plane are the sum of the uncracked half-plane (σ^{MC}, u,v) and the crack response (σ^{C}, u^{C},v^{C}) to the normal and tangential tractions p(y) and q(y) exerted on the half-plane surface.

II.1 Crack modelling

Stress and displacement fields corresponding to the uncracked half-plane are derived from Airy stress functions associated with concentrated normal and tangential loads.
Analytical expressions for stress and displacement fields corresponding to the crack response exist in the literature [5,6]. Displacement discontinuities (slip and opening) along crack faces are modelled with continuous distributions of dislocations bx and by. This technique, already used in previous models [2-8], has been modified to be adapted to the contact problem solution [14]. A single stress and displacement formulation for the whole crack, independent from the final contact zone distribution, has been developed. Dislocations distributed along all cracks contribute to the stress field at M situated on crack p, but only dislocations distributed along the crack itself cause relative displacements along its faces. Thus, stress and displacement expressions are (cf. figure 3):

$$(1)\quad \sigma^{C}_{ij}(x,y) = \frac{2\mu}{\pi(k+1)} \sum_{q=1}^{m} \int_{\Gamma q} [b_{qx}(\xi)\, K^{x}_{ij}(x,y_{\eta},\xi) + b_{qy}(\xi) K^{y}_{ij}(x,y_{\eta},\xi)]\, d\xi$$

with m : the crack number

i ,j = n,t: reference axis of crack p, $1 < p < m$

$$(2)\quad \delta u_{n}(x,y) = \delta v \cos\beta - \delta u \sin\beta$$

$$(3)\quad \delta u_{t}(x,y) = \delta u \cos\beta - \delta v \sin\beta$$

with,

$$(4)\quad \delta u\ (x,y) = \int_{\Gamma p} b_{x}(\xi)\, d\xi$$

$$(5)\quad \delta v\ (x,y) = \int_{\Gamma p} b_{y}(\xi)\, d\xi$$

$$(6)\quad u^{C}(x,y) = \frac{1}{2\pi(k+1)} \int_{\Gamma p} [b_{x}(\xi)\, D^{x}_{x}(x,y_{\eta},\theta,\xi) + b_{y}(\xi)\, D^{y}_{x}(x,y_{\eta},\theta,\xi)]\, d\xi$$

$$(7)\quad v^{C}(x,y) = \frac{1}{2\pi(k+1)} \int_{\Gamma p} [b_{x}(\xi)\, D^{x}_{y}(x,y_{\eta},\theta,\xi) + b_{y}(\xi)\, D^{y}_{y}(x,y_{\eta},\theta,\xi)]\, d\xi$$

The stress kernels K^{x}_{ij}, K^{y}_{ij} and the displacements kernels D^{j}_{i} are given in [13].

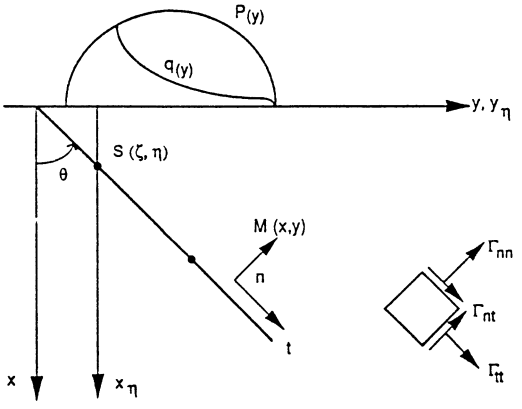


Figure 3: Notations employed

Numerical solutions employed to solve these integrals are those developed by Erdogan et al [20]. Discretized forms of equations (1), (4) to (7) are obtained. The next step is the determination of the distributions of dislocations b_{qx} and b_{qy} via the contact problem solution.

II.2 Contact problem solution

The distributions of normal σnn^T and tangential σnt^T tractions that satisfy the boundary conditions (listed below) at crack interfaces are solved using the unilateral contact technique. Continuum stress field σ^{MC} at load step j is the resulting stress field σ^T calculated at the previous load step (j-1). Note that a compressive traction is negative.

Contact zone :	(8) $\delta un = 0.$	(9) $\sigma nn^T < 0.$
Open zone :	(10) $\sigma nn^T = 0$	(11) $\delta un > 0$
Backward slip zone :	(12) $\sigma nt^T = f * \sigma nn^T$	(13) $\delta ut * \sigma nt^T > 0$
Forward slip zone :	(14) $\sigma nt^T = -f * \sigma nn^T$	(15) $\delta ut * \sigma nt^T > 0$
Adhesion zone	(16) $\delta ut = 0$	(17) $\lvert\sigma nt^T\rvert < f * \lvert\sigma nn^T\rvert$

The unilateral contact problem with friction was solved previously using the Panagioutopoulos process of alternatively calculating the normal problem (contact zone and normal traction determination) and the tangential problem (slip and stick zone, tangential traction determination) until convergence occurs [11-14] following [18,19].
It is solved now by using the Kombi process developed by Kalker [20] which is slower but more reliable. This process is briefly described below.

1- At the beginning of each load step, cracks are assumed to be closed and adherent. The traction bound g is estimated to be the value of the continuum stress field at the load step: $g = \sigma nn^{MC}$.

2-Corresponding equations are solved on the basis of the estimated traction bound. Equations (12) and (14) become: $\sigma nt^T = +/-f*g$. The algorithm determines the normal and tangential stresses σnn^T and σnt^T and the displacement discontinuities δun and δut.

3-The solutions are then tested to see if they satisfy the boundary conditions in terms of the inequalities:

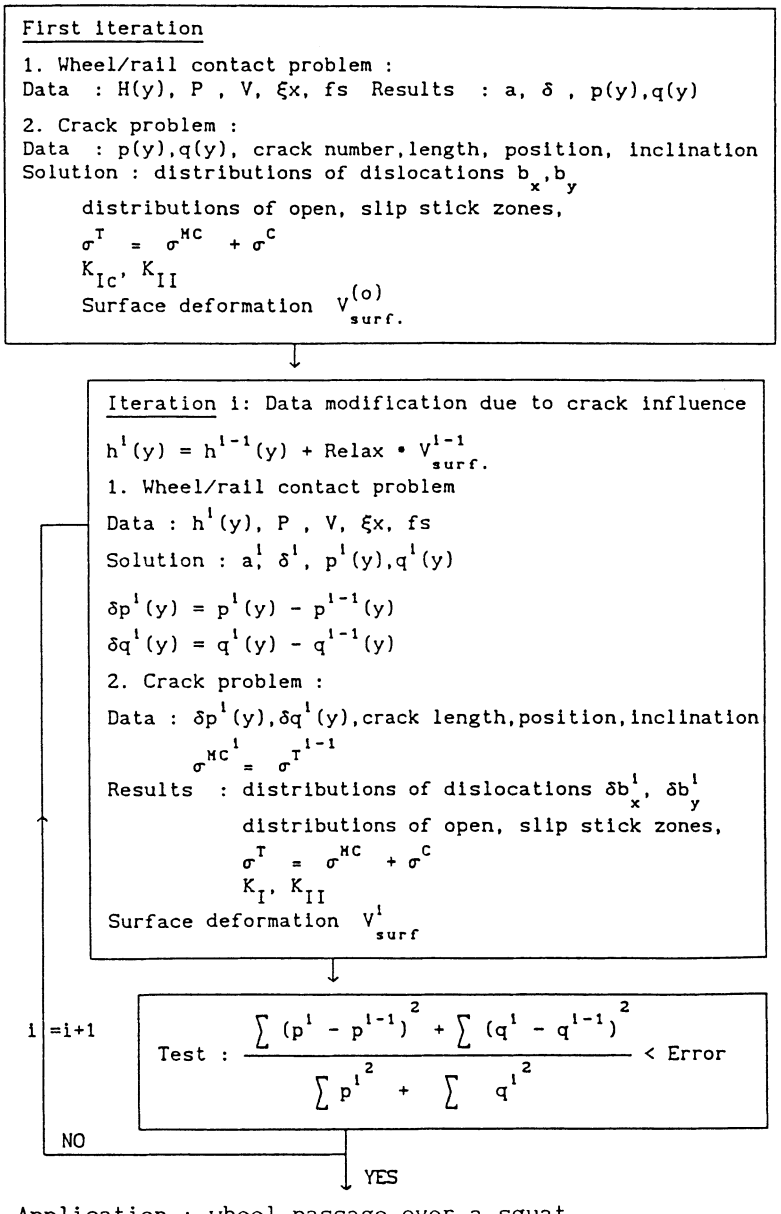
*For Qi belonging to a contact zone :
if $\sigma nn^T(Qi) > 0$, Qi is set to the open zone,

*For Qi belonging a stick zone:
if $\lvert\sigma nt^T(Qi)\rvert > f * \lvert\sigma nn^T(Qi)\rvert$, Qi is set
in a backward slip zone if $\sigma nt^T(Qi) < f* \sigma nn^T(Qi)$ or
in a forward slip zone if $\sigma nt^T(Qi) > - f* \sigma nn^T(Qi)$

Equations are solved again if any modification has taken place; if not,

*For Qi belonging to a slip zone:

tractions exerted on the half-plane surface are stabilised.
For the next load step the surface geometry is reinitiated to H(y).



IV. Application : wheel passage over a squat

The influence of cracks on the conditions at wheel/rail interface is illustrated in the case of the passage of a wheel over a squat. This defect occurs in the running band of some high-speed passenger lines

and appears to be related to high driving traction and creepage of the locomotive wheels.

The wheel/rail configuration chosen comes from [22]. A nominal load of 173 kN was applied , with Young's modulus and Poisson's ratio for steel taken as 297 GPa and 0.3 respectively. The wheel diameter is chosen to be equal to 1.25 m, the interfacial friction to be equal to 0.2. It ranges usually from 0.15 to 0.45 depending on the presence of contaminants on the rail head [21]. The rolling speed is taken as 50 m/s, the creep ratio is -0.001. 200 discretized points are distributed along the 22 mm potential contact area width. The maximum Hertzian pressure Po is equal to 1294 MPa and the half contact width a to 9.91 mm. The slip zone is situated at the trailing edge of the contact area (cf. figure 4). The tangential to normal load ratio is equal to 0.106. The squat configuration is represented with 2 cracks whose characteristics are:
crack 1 : b1 = 4 mm $\theta 1 = -50°$ f1 = 0.1
crack 2 : b2 = 5 mm $\theta 2 = 70°$ f2 = 0.1
distance between cracks d = 500 μm
The position of the wheel is defined with respect to the trailing edge of the loading zone y_T. The passage of the wheel is studied from y_T= -35 mm to y_T= 15 mm. The load advance step is 150 μm. 333 steps are thus observed.

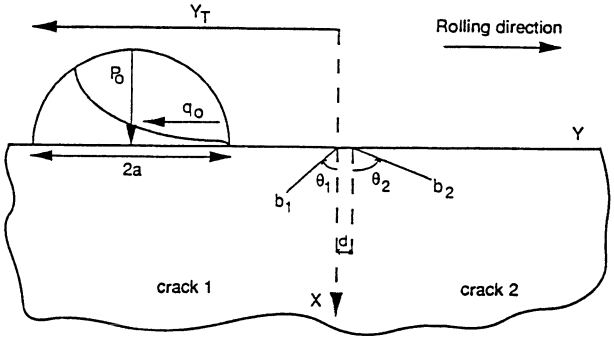


Figure 4: Wheel passage over a squat

The influence of cracks modifies the hertzian stress field as long as the contact zone is situated over the cracks, i.e from y_T =-25mm to y_T=0.5mm. This influence leads to :

-local overpressures of the same order of magnitude as Po without any modification of the contact area width (cf. fig.5). This behaviour

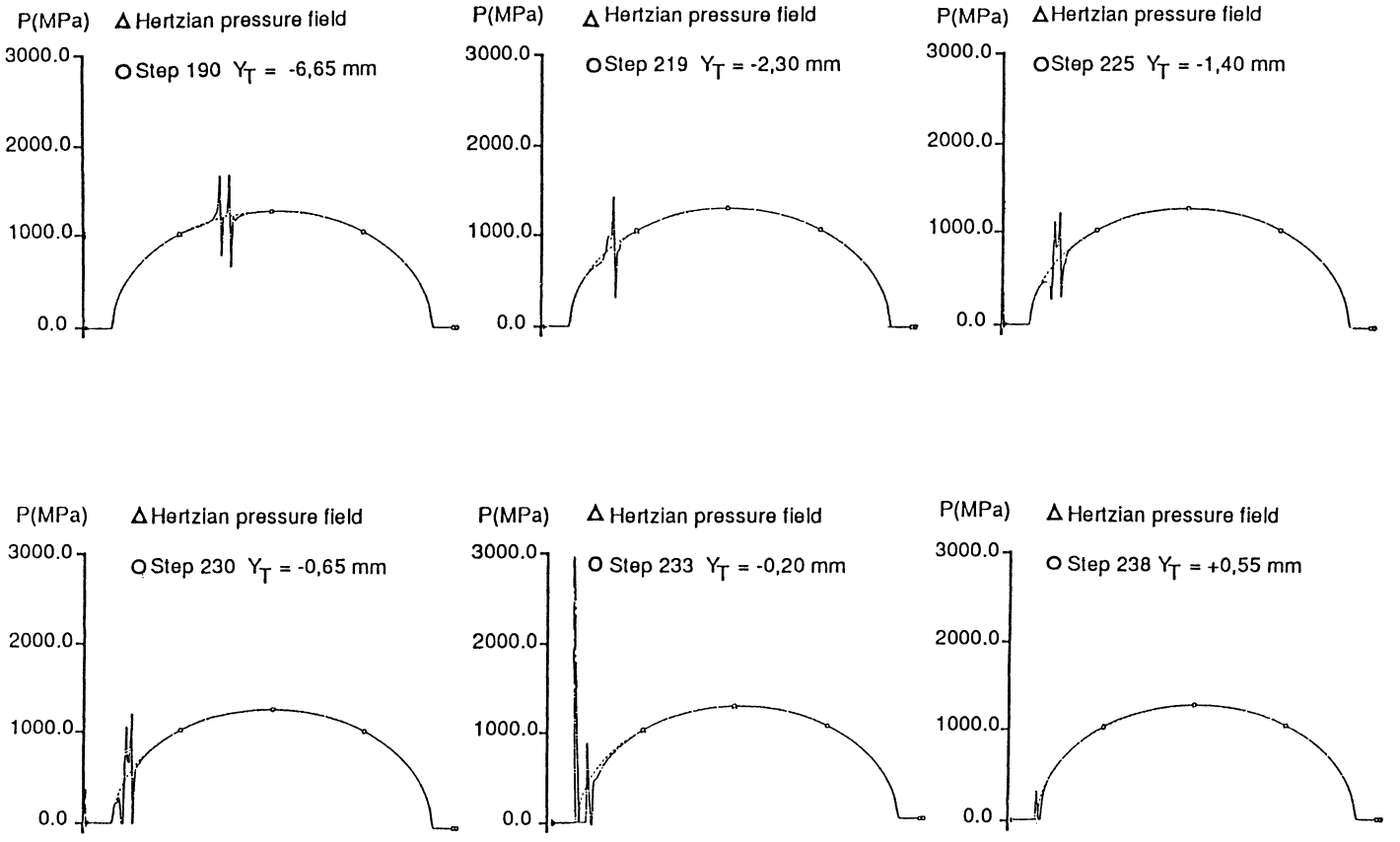


Figure 5: Evolution of the normal traction as the trailing edge of the contact zone moves over crack mouths.

is observed when edges of the contact zone are not situated in the vicinity of crack mouths. Note that maximum shear stresses corresponding to these local maxima are situated just below the surface and are of the same order of magnitude as the maximum shear stress corresponding to the macrogeometry.

-split up of the contact zone in 2 or 3 regions with important local overpressures. Figure 5 show the evolution of the normal traction as the trailing edge of the contact zone moves over both crack mouths. The maximum pressure Po' is equal to 3020 MPa instead of 1294 MPa. Danger zones induced by cracks are concentrated in a subsurface strip of about 100 μm in thickness.
In fact surface discontinuities caused by cracks have the same effect as micro-asperities.

-concerning cracks 1 and 2, the modifications of the contact width and of the normal and tangential tractions have no direct influence on the stress intensity factors as seen on figure 6. $K_I(1), K_{II}(1)$ and $K_I(2), K_{II}(2)$ are the stress intensity factor variations at the tips of cracks 1 and 2 respectivelly versus the load position when the influence of cracks is not taken into account. $K_{IC}(1)$, $K_{IIC}(1)$ and $K_{IC}(2)$, $K_{IIC}(2)$ are the SIFs variations when the influence of cracks are taken into account. Nevertheless new shear stress maxima localised just below the surface can influence crack path growth and cause crack initiation.

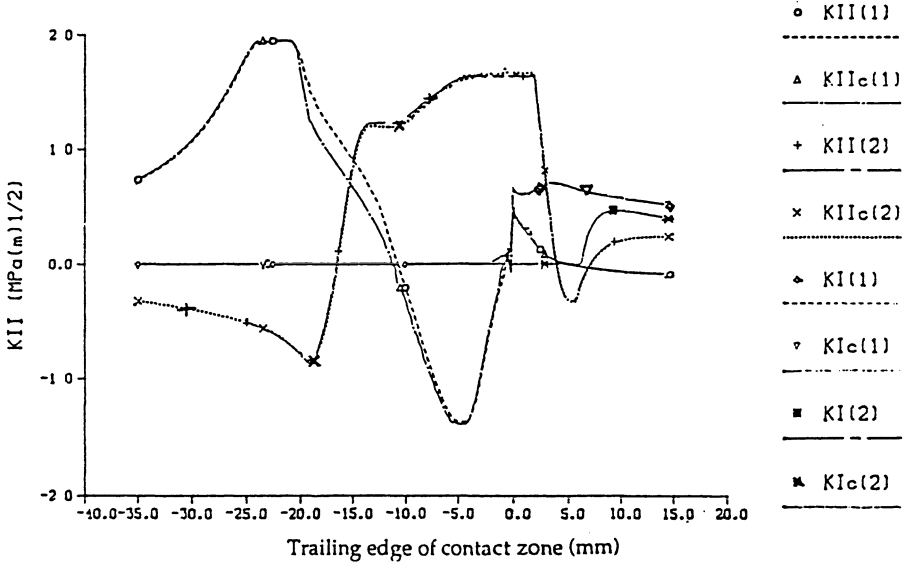


Figure 6: SIFs variations versus load position

Conclusion

A steady two dimensional elastic model based on half analytical and half numerical methods for the analysis of crack influence on rolling contact stress field has been developed. Results show that the presence of surface breaking cracks disturb the hertzian stress field. Split up of the contact area in two or three regions is observed. Significant local overpressures lead to maximum shear stress localised just below the surface in a strip of 100μm in thickness.
The current model will be combined with fatigue crack growth path criterion and rate laws.
It will be then refined to include transient effects.

REFERENCES

[1] Office for research and experiments of the international Union of ways. Question: Rolling contact fatigue, Report N°1, "Review of rolling contact fatigue in rails", 75 p., 1990.

[2] Comninou M. "The Interface Crack". ASME, Journal of Applied Mechanics, vol. 44, p. 631-636, 1977.

[3] Comninou M. "The Interface Crack in a Shear Field". ASME, Journal of Applied Mechanics, vol. 45, p. 287-290, 1978.

[4] Comninou M., and Schmueser D. "Frictional Slip Between a Layer and a substrate Caused by a Normal Load". Int. J. of Engineering Science, vol. 18, p. 131-137, 1980.

[5] Keer L.M., Bryant M.D. and Hiratos G.K. "Subsurface and Surface Cracking due to Hertzian Contact". ASME, Journal of Lubrication Technology, vol. 104, p. 347-351, 1982.

[6] Keer L.M. and Bryant M.D. "A Pitting Model for Rolling Contact Fatigue". ASME Journal of Lubrication Technology, vol. 105, p. 198-205, 1983.

[7] Hills D.A., Comninou M., "A Normally Loaded Half Plane with an Edge Crack". Int. J. of Solids Structures. Vol. 21, n° 4, p. 399-410, 1985.

[8] Hills D.A., Comninou M. "An Analysis of Fretting Fatigue Cracks During Loading phase" In Int. J. of Solids Structures, vol. 12, n°7, pp. 721-730,1985.

[9] Bold P.E., Brown M.W., Allen R.J. "Shear Mode Crack Growth and Rolling Contact Fatigue" Wear, vol. 144, pp. 307-317, 1991.

[10] Bower A.F. "The Influence of Crack Face Friction and Trapped Fluid on surface Initiated Rolling Contact Fatigue Cracks". ASME Journal of Tribology, vol. 110, pp. 704-711, 1988.

[11] Dubourg M.-C." Le contact unilateral avec frottement le long defissures de fatigue dans les liaisons mécaniques". Thèse de doctorat, Institut National des Sciences Appliquées de Lyon, 253 p. 1989.

[12] Dubourg M.-C., Villechaise B. "Analysis of Multiple Fatigue Cracks- Part I : Theory". ASME, Journal of Tribology, vol. 114,

pp. 455-461, 1992.
[13] Dubourg M.-C., Villechaise B., Godet M. "Analysis of Multiple Fatigue Cracks- Part II : Results". ASME, Journal of Tribology, vol. 114, pp. 462-468, 1992.
[14] Dubourg M.-C., Villechaise B. "Stress Intensity Factors in a Bent Crack : a model". Eur. J. Mech. A/Solids, vol. 11, n° 2, p.169-179, 1992
[15] Olzak M., Stupnicki J., Wojcik R. "Investigation of crack propagation during contact by a finite element method", Wear, Vol. 146, pp. 229-240, 1991.
[16] Olzak M., Stupnicki J., Wojcik R. "Numerical analysis of 3D cracks propagating in rail-wheel contact zone". To be published.
[17] Carneiro Esteves A., Seabra J., Berthe D. "Roughness frequency analysis and particle depth", In Interface Dynamics, proceedings of the 14ème Leeds-Lyon Symposium, Lyon, 8-11 Sept. Ed. by D. Downson, C.M. Taylor, M. Godet, D. Berthe. Amsterdam: Elsevier, p 209-213, 1988.
[18] Kalker J.J. "Two Algorythms for the Contact Problem in Elastostatics". Delft, Université de Technologie, Report of the Department of Mathematics and Informatics, n° 82-26, 8 p., 1982.
[19] Kalker J.J., "The Contact Between Wheel and Rail", Delft Université de Technology, Report of the Department of Mathematics and Informatics, n° 82-27, 36 p., 1982.
[20] Kalker J.J. "Three-Dimensional Elastic Bodies in Rolling Contact". Kluwer Academic Publishers, 1990, 314p.
[21] Erdogan F., Gupta G.D., Cook T.S.,Numerical solution of singular integral equations. Methods of analysis and solutions of crack problems. Edited by G.C. Sih, Leyden: Noordhoff International Publishing, 1973, p 368-425.
[22] Maitournam M.H., "Formulation et résolution numérique des problèmes thermoviscoplastiques en régime permanent." Thèse de Doctorat, Ecole Nationale des Ponts et Chaussées, 1989, 176p.

NUMERICAL ANALYSIS OF 3D CRACKS PROPAGATING IN THE RAIL–WHEEL CONTACT ZONE

Mirosław Olzak, Jacek Stupnicki, Ryszard Wójcik

Abstract

The paper presents the results of investigation of contact between the wheel and rail with semi-elliptical crack existing in contact zone under normal load. Shapes of the contact surface and the distribution of normal and tangential forces on the contact surfaces as well as interaction between the crack faces were determined. Finally, Stress Intensity Factors K_1, K_2 and K_3 were calculated along the front of the crack and plots of the maximum values of SIFs versus the position of the load in relation to the crack are presented.

Introduction

A significant number of failures of rails is caused by a fracture originated from near-surface initial cracks. Some theories, which try to explain the phenomenon of crack growth in contact have appeared. They are mainly based on numerical analysis of the stress state near the crack front. Keer et al (1982, 1983) [1,2] calculated the values of the Stress Intensity Factors (SIFs) at the front of the two-dimensional crack in elastic half space. They concluded that the crack propagates according to the so called second mode of opening. Bower (1987, 1988) [3,4] has also studied 2D cases where the effects of frictional interaction between the crack faces were taken into account.

Keneto and Murakami (1985, 1986, 1991) [5,6,7] discussed the problem of 2D and 3D surface breaking cracks in an elastic half space under elliptical or ellipsoidal hertzian pressure distribution. They also concluded that crack propagated according to the second mode of opening.

In the authors opinion the fundamental assumption accepted in most of the publications that the pressure distribution on the contact surface is ellipsoidal and is not disturbed by the cracks is incorrect. When the crack is situated in contact zone it locally modifies the stiffness of the surface layer and causes the changes to the distribution of normal and tangential interaction.

The purpose of the paper is to present results of numerical approach for the analysis of contact of wheel and rail with semi-elliptical surface breaking crack perpendicular to the contact surface. The solution was obtain for normal load and quasi-static rolling of the wheel over the cracked rail.

The shapes of the contact area between the rail and wheel as well as contact area between the crack faces were determined. The distributions of normal and tangential loads on the contact areas in consecutive stages of wheel position were

J.J. Kalker et al. (eds.), Rail Quality and Maintenance for Modern Railway Operation, 385–395.

obtained. The final step of the analysis was the calculation of the Stress Intensity Factors K_1, K_2, K_3 along the crack front.

System investigated

The system investigated comprised a wheel and section of rail containing a semi-elliptical crack as shown in the Fig. 1. The material is pure elastic and isotropic, the frictional forces on the contact surfaces and between the faces of the crack have been taken into account. The effect of rolling of the wheel was considered. The distribution of the contact forces and the state of stresses were determined by use of a numerical procedure based upon the FEM.

The mesh of finite elements in the vicinity of the crack is shown in the Fig. 2. Since the elements used for problems of fracture mechanics must be of a special type the standard finite element procedure was supplemented by special elements which allow the real singularity of the type $(r)^{-1/2}$ ahead of the crack tip to be represented.

Numerical procedure

For solving the contact problem with friction it is necessary to analyse the whole process of loading starting from the initial state and carrying on the calculation for rolling along the rail until the steady-state distribution of normal and tangential (longitudinal and lateral) contact forces are obtained. When the wheel approaches the crack the new distribution of normal and tangential interactions on the contact area as well as interaction of the faces of the crack must be determined.

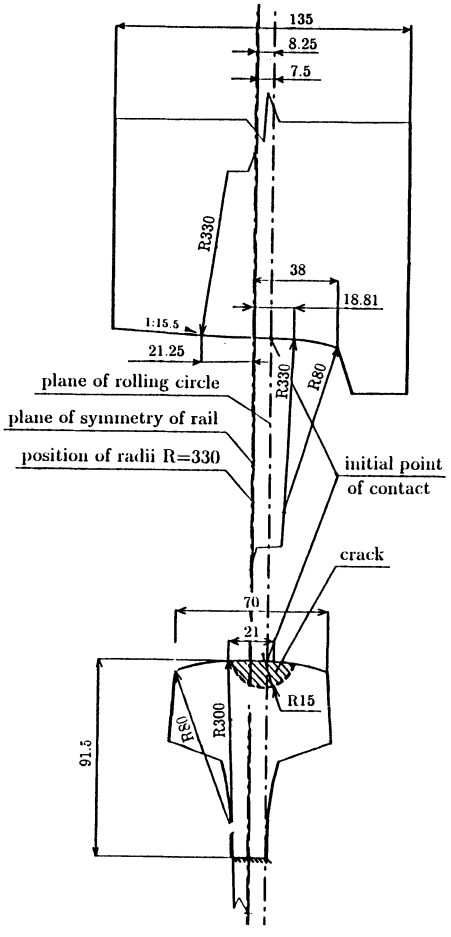


Fig. 1. The object of investigation

In the investigation the calculation was started from applying the normal load at the distance from the crack equal $x = 10b$ (where b is the half width of the contact area in the longitudinal direction) (Fig. 3). Then the simulation of rolling was performed. The steady-state distribution of tangential forces were obtained after rolling the distance $x = 8b$ with the step of calculation $\Delta x = b/50$. The significant influence of the existing crack on the contact was observed for the distance $-b < x < +b$.

The algorithm for solving the 3D contact problems with friction is given in the Appendix.

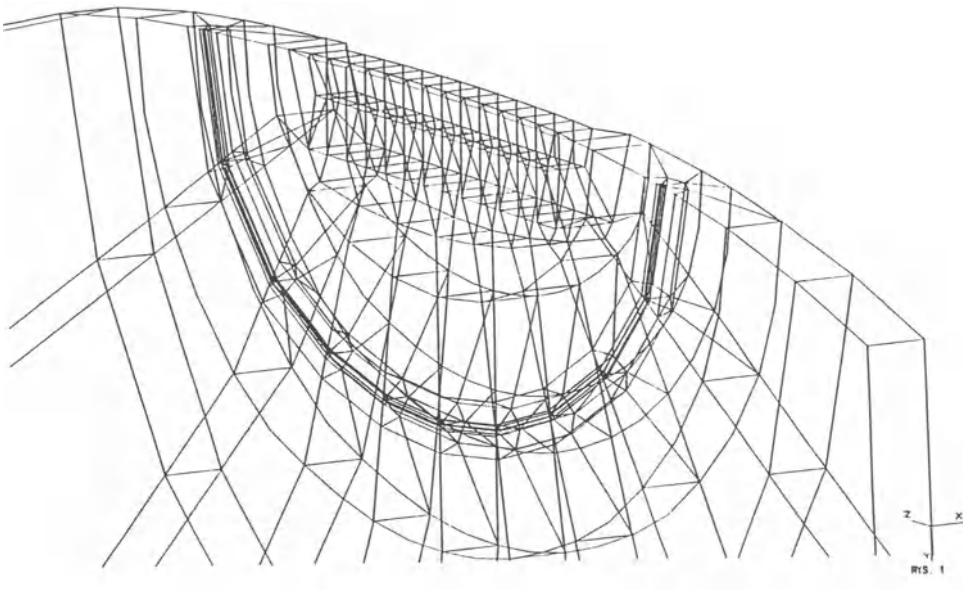


Fig. 2. The mesh of finite elements in the vicinity of the crack

Results

The solution was obtained for total normal load $R = 70$ kN, resultant tangential load in the longitudinal direction $K_x = 0$. The tangential load in lateral direction K_z was applied to provide the straight rolling of the wheel along the longitudinal axis of the rail $K_z = 0.01R$. The friction coefficient on the contact surface and between the forces of the crack was assumed to be $\mu = 0.4$.

Fig. 4a presents the distribution of normal forces acting on the contact area when the wheel is loaded by the normal force at a distance $x = 10b$ from the crack. The pressure distribution, which is a result of acting nodal forces, differs significantly from hertzian ellipsoidal pressure distribution. This is a result of the finite dimensions of cross-section of the rail and the rapid changes of the radii of the curvature of the surfaces of the wheel and the rail. The distribution of longitudinal and lateral forces on the contact surfaces is presented in Fig. 4b, 4c, respectively.

After the series of consecutive stages of quasi-static rolling with the maximum step of $\Delta x = b/50$ over the distance of $x = 8b$ the steady-state distribution of tangential forces was obtained. Fig. 5a

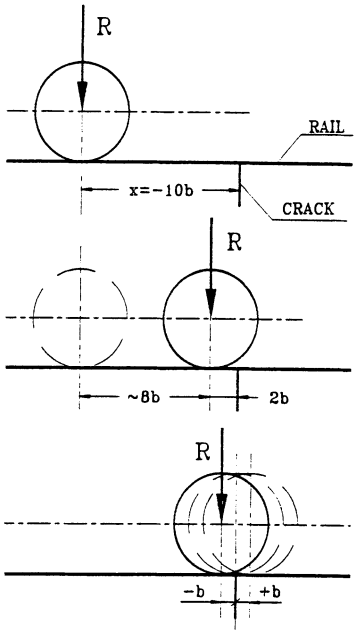


Fig. 3. Schematic representation of the loading process of the contact

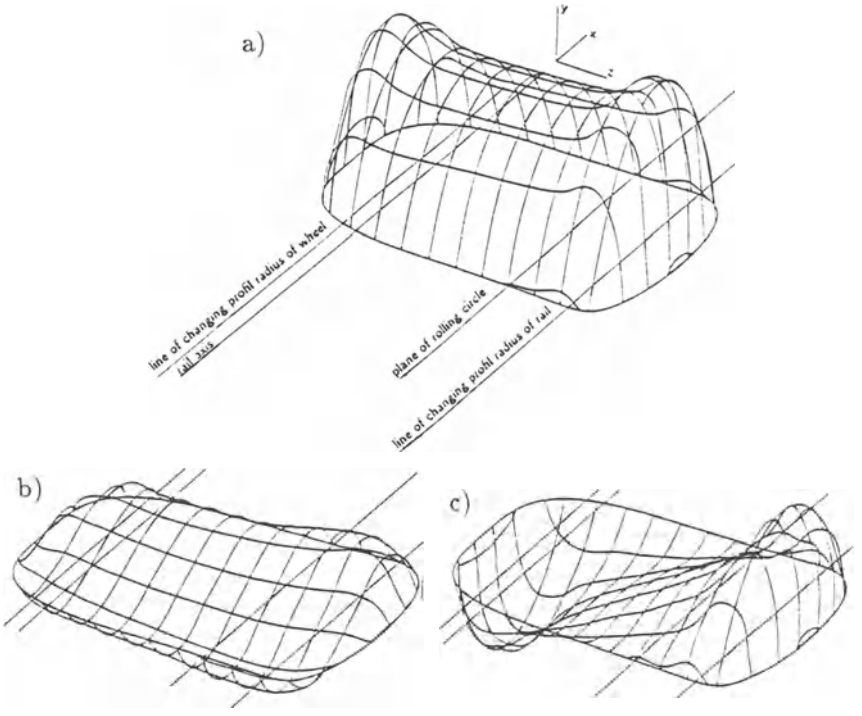


Fig. 4. Distribution of forces in contact at $x = -10b$. a) normal forces, b), c) longitudinal and lateral forces on the the contact surface before rolling,. Scale of the forces enlarged 50 times in relation to Fig. 4a.

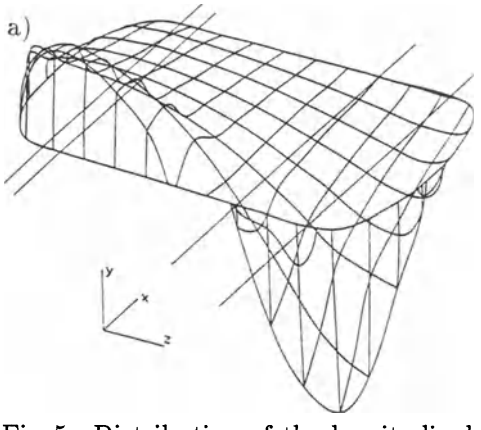


Fig. 5. Distribution of the longitudinal (a) and lateral (b) forces on the contact surface after rolling over the distance of 8b.

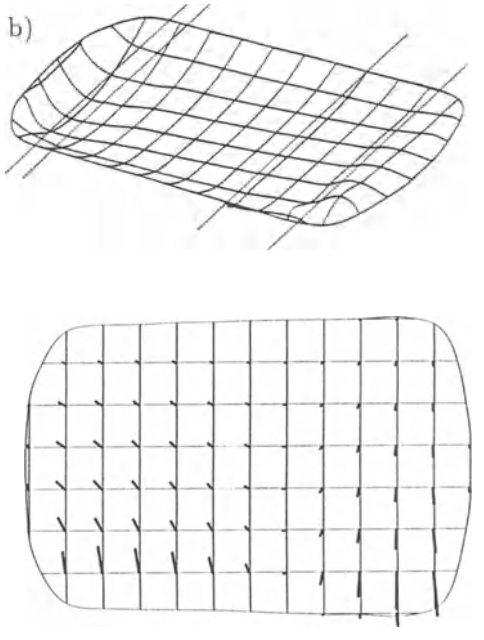


Fig. 6. The vectors of resultant frictional forces on the contact area after rolling.

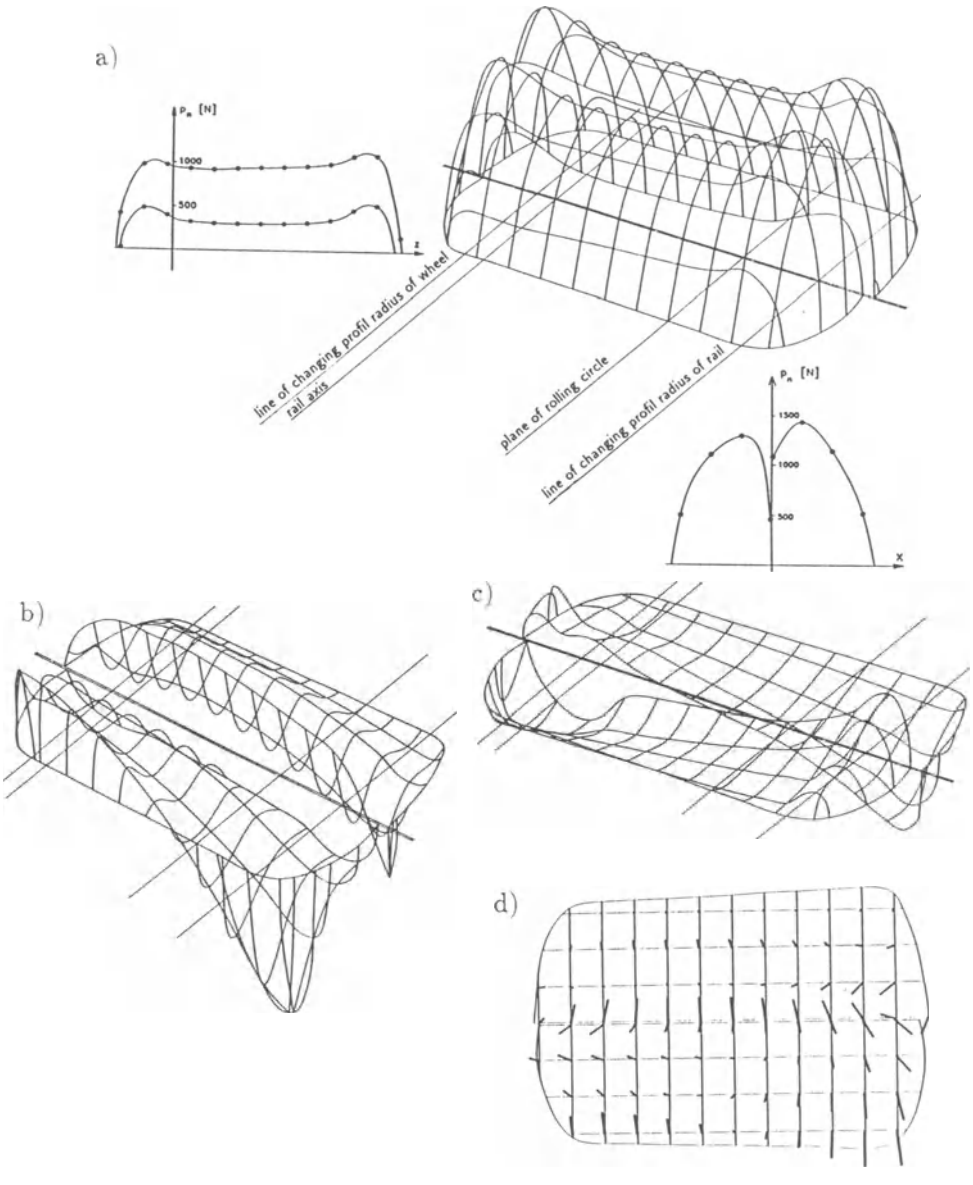


Fig. 7. Distribution of forces in contact for the wheel situated symmetrically relative to the crack: a) normal forces, b) longitudinal forces, c) lateral forces. The scale of the forces in Fig. 7b and 7c enlarged 10 times in relation to Fig. 7a. d) The vectors of the resultant friction forces on the contact surface for the wheel situated symmetrically relative to the crack.

and 5b show the distribution of the longitudinal and lateral forces on the contact surface, respectively. The distribution of normal forces and shapes of the contact area are practically not affected by the rolling of the wheel.

Fig. 6 shows the vectors of the resultant frictional forces on the contact area. The directions of vectors are the result of the micro-slip created by the conical shape of the wheel and the geometry of the rail.

The results for the symmetrical position of the wheel in relation to the crack are presented in the Fig. 7. They show respectively the shape of the area of the wheel/rail contact, the distribution of normal forces and the distributions of tangential, longitudinal and lateral forces. All of the information presented in Fig. 7 is of interest, however, special attention should be given to the distribution of the normal forces swonw in Fig. 7a. The existence of the crack in contact distinctly affects the stiffness of the surface layer of the rail in this region and causes the difference in the distribution of the contact forces in respect to that shown in the Fig. 4. A local minimum of the forces at the mouth of the crack is visible. In the cross-section perpendicular to the crack the pressure distribution seems to be rather double-ellipsoidal on both sides of the crack. Loads on the front side of the crack edge are higher than those on the rear one.

The transverse longitudinal and lateral forces on the contact area change from negative to positive values in both cases (Fig. 7b, 7c). Fig. 7d shows the vectors of the resultant frictional forces on the contact area for the symmetrical position of the wheel in relation to the crack.

The determination of the Stress Intensity Factors

The Stress Intensity Factors K_1, K_2 and K_3 were determined at nodes situated along the crack front for the various stages of rolling of the wheel. The method of determination of the SIF was discussed in [8]. Generally the values of K_1, K_2 and K_3 were evaluated from the Williams equation by substituting the displacements around the crack front obtained from the solution of the contact problem by use of FEM. The values of SIFs given on the plots are obtained as their average values on the region surrounding the considered node on the crack front. Fig. 8 presents the distribution of K_1, K_2 and K_3 at the moment when the wheel is located at the distance $x = -2b$ from the crack, Fig. 9 corresponds to the distance $x = -1b$ and Fig. 10 concerns the symmetrical position of the wheel relative to the crack ($x = 0$). In the Fig. 11 the Stress Intensity Factors are shown for the situation when the wheel is located behind the crack at the distance $x = +1b$. The plots presented show how the Stress Intensity Factors change along the crack front at selected stages of rolling.

In all the cases (with exception of the symmetrical position of the wheel in relation to the crack) the maximum value of K_2 corresponds to the nodes located at the center of the crack front. The maximum values of K_1 and K_2 (for the central node on the crack front) versus the position of the wheel in relation to the crack

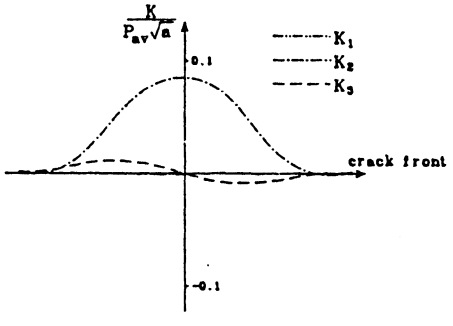


Fig. 8. The plot of Stress Intensity Factors K_1, K_2, K_3 for nodes situated along the crack front. The wheel in distance $x/b = -2$ from the crack.

is shown in the Fig. 12. It results from the plot that the Stress Intensity Factor K_2 exceeds the K_1 along the long distance in front and behind the crack. Only in the symmetrical position of the wheel to the crack, the K_1 reaches the maximum values and exceeds the K_2, which is close to zero at this moment.

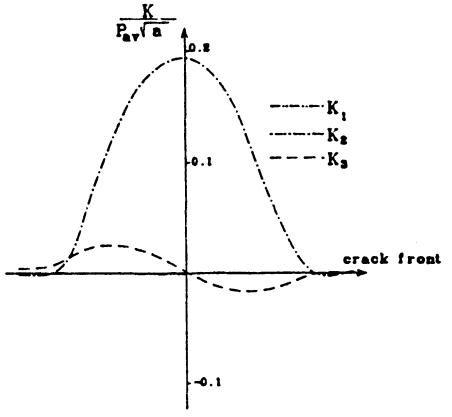


Fig. 9. The plot of Stress Intensity Factors K_1, K_2, K_3 for nodes situated along the crack front. The wheel in distance $x/b = -1$ from the crack.

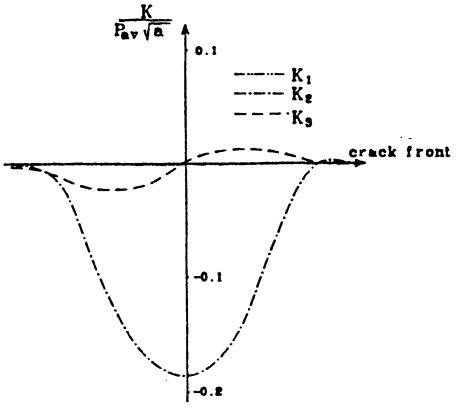


Fig. 11. The plot of Stress Intensity Factors K_1, K_2, K_3 for nodes situated along the crack front. The wheel in distance $x/b = +1$ from the crack.

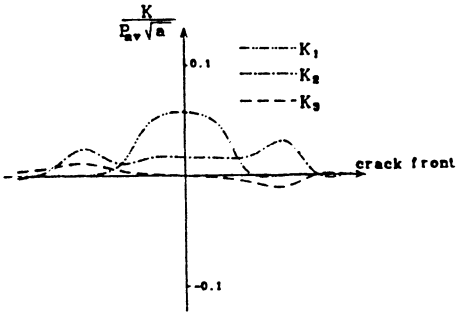


Fig. 10. The plot of Stress Intensity Facotrs K_1, K_2, K_3 for nodes situated along the crack front. The wheel situated symmetrically relative to the crack ($x/b = 0$).

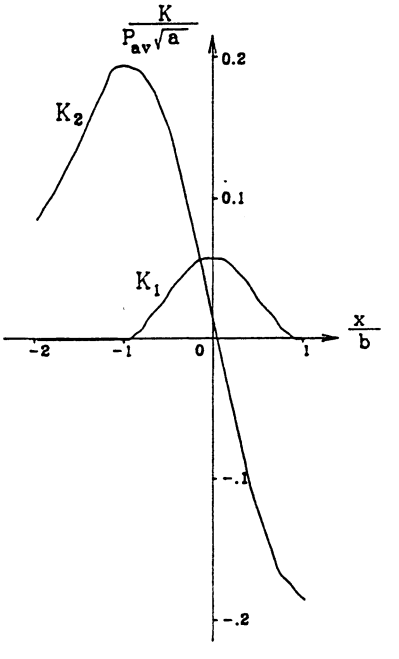


Fig. 12. Plot of Stress Intensity Factors K_1 and K_2 versus the position of the wheel relative to the crack.

Conclusions

The following conclusions can be formulated:

1. For the objects of real geometry, i.e. wheel and rail, the distribution of the pressure on the contact area differs from hertzian ones. The differences are the most significant in the region close to the crack ($-b < x < b$), where the pressure distributions have a discontinuity in the central region.
2. The shape of the contact area between the wheel and rail differs from elliptical in whole range of rolling. This shape is rather close to rectangular.
3. K_2 plays the dominant role in the propagation of cracks in the contact region. This is due to existence of significant shear stresses in the area of the crack front during the rolling of the wheel.
4. For the symmetrical position of the wheel in relation to the crack the value of K_2 is close to zero, however the factor K_1 caused by the tensile stresses near the crack front reaches a maximum.
5. For consideration the crack front propagation it is necessary to consider a combination of stress intensity factors.

Acknowledgement

The study was finacially supported by ORE D173 Committee and by Program of Fundamental Research

References

1. Keer L.M., Bryant M.D., Haritos G.K.: Subsurface and Surface Cracking Due to Hertzian Contact. ASME Journal of Lubrication Technology, Vol.104 (1982) pp.347–351.
2. Keer L.M., Bryant M.D.: A Pitting Model for Rolling Contact Fatigue. ASME Journal of Lubrication Technology, Vol.105 (1983) pp.198–205.
3. Bower A.F.: The Effects of Crack Face Friction and Trapped Fluid on Surface Initiated Rolling Contact Fatigue Cracks. University of Cambridge, CUED/C-Mech/TR.40 1987.
4. Bower A.F.: The Influence of Crack Face Friction and Trapped Fluid on Surface Initiated Rolling Contact Fatigue Cracks. ASME Journal of Tribology, Vol.110 pp.704–711.
5. Kaneta M., Murakami Y., Okazaki T.: Mechanism of Opening/Closure of a Subsurface Crack Due to a Moving Hertzian Loading. Proc. of the 12th Leeds-Lyon Symp. on Tribology. (1985) pp.29–38.
6. Kaneta M., Suetugu M., Murakami Y.: Mechanism of Surface Crack Growth in Lubricated Rolling/Sliding Spherical Contact. ASME Journal of Applied Mechanics, Vol.53 (1986) pp.354–360.
7. Kaneta M., Murakami Y.: Propagation of Semi-Elliptical Surface Cracks in Lubricated Rolling/ Sliding Elliptical Contacts, ASME Journal of Tribology, Vol.113 (1991) pp.270–275.
8. Olzak M., Stupnicki J., Wójcik R.: Investigation of Crack Propagation During Contact by a Finite Element Method, WEAR 146 (1991) pp.229–240.

Appendix

For solving the contact problem with friction is it necessary to know the whole process of loading starting from the initial state (usually zero load) to the state which is being solved.
For the most advanced 3D case with friction the set of equations (1.1) may be written as follows:

$$\begin{bmatrix} [A_{11}] & [A_{12}] & \{1\} & [A_{13}] & [A_{14}] & \{0\} & [A_{15}] & [A_{16}] & \{0\} \\ [A_{21}] & [A_{22}] & \{0\} & [A_{23}] & [A_{24}] & \{0\} & [A_{25}] & [A_{26}] & \{0\} \\ \lfloor 1 \rfloor & \lfloor 0 \rfloor & 0 & \lfloor 0 \rfloor & \lfloor 0 \rfloor & 0 & \lfloor 0 \rfloor & \lfloor 0 \rfloor & 0 \\ [A_{31}] & [A_{32}] & \{0\} & [A_{33}] & [A_{34}] & \{1\} & [A_{35}] & [A_{36}] & \{0\} \\ [A_{41}] & [A_{42}] & \{1\} & [A_{43}] & [A_{44}] & \{0\} & [A_{45}] & [A_{46}] & \{0\} \\ \lfloor 0 \rfloor & \lfloor 0 \rfloor & 0 & \lfloor 1 \rfloor & \lfloor 0 \rfloor & 0 & \lfloor 0 \rfloor & \lfloor 0 \rfloor & 0 \\ [A_{51}] & [A_{52}] & \{1\} & [A_{53}] & [A_{54}] & \{0\} & [A_{55}] & [A_{56}] & \{1\} \\ [A_{61}] & [A_{62}] & \{1\} & [A_{63}] & [A_{64}] & \{0\} & [A_{65}] & [A_{66}] & \{0\} \\ \lfloor 0 \rfloor & \lfloor 0 \rfloor & 0 & \lfloor 0 \rfloor & \lfloor 0 \rfloor & 0 & \lfloor 1 \rfloor & \lfloor 0 \rfloor & 0 \end{bmatrix} * \begin{Bmatrix} \{P\} \\ \{S\} \\ \Delta y \\ \{T_x\} \\ \{S_x\} \\ \Delta x \\ \{T_z\} \\ \{S_z\} \\ \Delta z \end{Bmatrix} = \begin{Bmatrix} \{y^0\} \\ \{\delta^0\} \\ R \\ \{x^0\} \\ \{\zeta^0\} \\ K_x \\ \{z^0\} \\ \{\eta^0\} \\ K_z \end{Bmatrix} \tag{1.1}$$

where
$[A_{ij}]$ are matrices of influence of i forces to the j dispalecements:
i, j = 1 normal to the rail surface
i, j = 2 normal to the crack surface
i, j = 3 tangential in longitudinal direction to the rail surface
i, j = 4 tangential in longitudinal direction to the crack surface
i, j = 5 tangential in lateral direction to the rail surface
i, j = 6 tangential in lateral direction to the crack surface,
for example $[A_{41}]$ is the matrix of influence of longitudinal forces at the crack surface on the normal displacement at the rail surface.

$\{P\}$, $\{S\}$ – vector of normal forces in contact and crack, respectively,

$\{T_x\}$, $\{T_z\}$ – vector of longitudinal and lateral forces in contact, respectively,

$\{S_x\}$, $\{S_z\}$ – vector of longitudinal and lateral forces in crack, respectively,

Δy, Δx, Δz – components of the vector of relative displacement of the wheel in normal, longitudinal and lateral directions, respectively,

$\{y^0\}$, $\{\delta^0\}$ – vector of initial distances between nodes of the wheel and rail and the nodes of the crack faces, respectively,

$\{x^0\}$, $\{z^0\}$ – vector of slip in longitudinal and lateral directions of nodes in contact surface, respectively,

$\{\zeta^0\}$, $\{\eta^0\}$ – vector of slip in longitudinal and lateral directions of nodes in the crack faces, respectively,

R, K_x, K_z – total normal, longitudinal and lateral loads, respectively.

The procedure presented below concerns the progress from the solution for the state determined by loads R', K'_x, K'_z to the solution for the state corresponding

to loads R'', K_x'', K_z'', assuming that the differences between those two states are small (R, K_x, K_z represent the normal, longitudinal and lateral loads of contact, respectively).
Let us assume that the solution for a certain case corresponding to the given loads (R', K_x', K_z') is known. This means that the following force and slip vectors are also known:

$$\{P'\} = \begin{Bmatrix} p'(1) \\ p'(2) \\ \vdots \\ p'(n) \end{Bmatrix} \qquad \{S'\} = \begin{Bmatrix} s'(1) \\ s'(2) \\ \vdots \\ s'(m) \end{Bmatrix} \qquad \{T_x'\} = \begin{Bmatrix} t_x'(1) \\ t_x'(2) \\ \vdots \\ t_x'(n) \end{Bmatrix}$$

$$\{S_x'\} = \begin{Bmatrix} s_x'(1) \\ s_x'(2) \\ \vdots \\ s_x'(m) \end{Bmatrix} \qquad \{T_z'\} = \begin{Bmatrix} t_z'(1) \\ t_z'(2) \\ \vdots \\ t_z'(n) \end{Bmatrix} \qquad \{S_z'\} = \begin{Bmatrix} s_z'(1) \\ s_z'(2) \\ \vdots \\ s_z'(m) \end{Bmatrix}$$

$$\{x^{0\prime}\} = \begin{Bmatrix} x^{0\prime}(1) \\ x^{0\prime}(2) \\ \vdots \\ x^{0\prime}(n) \end{Bmatrix} \quad \{\zeta^{0\prime}\} = \begin{Bmatrix} \zeta^{0\prime}(1) \\ \zeta^{0\prime}(2) \\ \vdots \\ \zeta^{0\prime}(m) \end{Bmatrix} \quad \{z^{0\prime}\} = \begin{Bmatrix} z^{0\prime}(1) \\ z^{0\prime}(2) \\ \vdots \\ z^{0\prime}(n) \end{Bmatrix} \quad \{\eta^{0\prime}\} = \begin{Bmatrix} \eta^{0\prime}(1) \\ \eta^{0\prime}(2) \\ \vdots \\ \eta^{0\prime}(m) \end{Bmatrix}$$

where:
n – number of rail-wheel pairs of nodes being in contact,
m – number of pairs of nodes being in contact on the faces of the crack.
Force vectors are the results of solving of the set of equation (1.1) in which slip vectors have been determined by the iteration procedure described below.

The first step of iteration procedure is the repeated solution of the set of equations (1.1) for the new loads: R'', T_x'', T_z''. The following vectors of the new forces acting in the pairs of nodes being in contact are the results of the solution:

$$\{P''\},\ \{S''\},\ \{T_x''\},\ \{S_x''\},\ \{T_z''\},\ \{S_z''\}.$$

The next step is to check if the following conditions are fulfilled for all pairs of nodes being in contact:

$$\begin{cases} p''(i) * \mu \geq \sqrt{\left(t_x''(i)\right)^2 + \left(t_z''(i)\right)^2}, & \text{for pairs located on the rail surface,} \\ s''(i) * \mu \geq \sqrt{\left(s_x''(i)\right)^2 + \left(s_z''(i)\right)^2}, & \text{for pairs located on the crack faces,} \end{cases} \tag{1.2}$$

where μ is coefficient of friction.

For the pairs of nodes located on the rail surface which do not fulfil the condition (1.2) the new values of slip components $x^{0\prime\prime}(i)$, $z^{0\prime\prime}(i)$ are introduced to the appropriate equation.
The new values are defined as follows:

$$\begin{aligned} x^{0\prime\prime}(i) &= x^{0\prime}(i) + \omega \cdot t''_x(i), \\ z^{0\prime\prime}(i) &= z^{0\prime}(i) + \omega \cdot t''_z(i), \end{aligned} \tag{1.3}$$

where ω is the number chosen in such a way so that the value of the expression

$$\sqrt{\left(t'''_x(i)\right)^2 + \left(t'''_z(i)\right)^2} \tag{1.4}$$

is reduced.
The above expression contains the results of the next solution of the modified set of equations (1.1).
The modification here consists in substituting new vectors $\{x^{0\prime\prime}\}$ and $\{z^{0\prime\prime}\}$ for vectors $\{x^{0\prime}\}$ and $\{z^{0\prime}\}$.
The slip components $\zeta^0(j)$, $\eta^0(j)$ for the pairs of nodes on the crack faces are modified similarly, if necessary.
The above procedure of altering the values of $x^0(i)$, $z^0(i)$ or/and $\zeta^0(j)$, $\eta^0(j)$ and checking the condition (1.2) is being continued until the condition (1.2) is fulfilled for all pairs of nodes being in contact. Additionally, the changes of values of $x^0(i)$, $z^0(i)$ and $\zeta^0(j)$, $\eta^0(j)$ must be chosen in such a way that after finishing the process, the another condition is fulfilled for all pairs being modified:

$$\begin{cases} p(i) * \mu - \sqrt{\left(t_x(i)\right)^2 + \left(t_z(i)\right)^2} \leq \varepsilon, & \text{for pairs located on the rail surface,} \\ s(i) * \mu - \sqrt{\left(s_x(i)\right)^2 + \left(s_z(i)\right)^2} \leq \varepsilon, & \text{for pairs located on the crack faces,} \end{cases}$$

where ε is the assumed accuracy of calculation of the tangential forces.
If during the iteration process the new pair of nodes (of the number k) comes to contact, such an intermediate values of load R^*, K_x^*, K_z^* should be determined for which the y component of the distance between the nodes of this pair is close to zero.
Then at the moment of the iteration procedure the new equation for this pair of nodes is added to the set of equations (1.1). The x and z components $\alpha(k)$, $\beta(k)$ of the distance between the nodes of the new pair are substituted for slips $x^0(k)$ and $z^0(k)$ in this equation (or $\zeta^0(k)$ and $\eta^0(k)$ if the pair is located in the crack faces). Next, the modified set of equations (1.1) is being solved for the new loads; R'', T''_x, T''_z and whole iteration process is continued from the beginning.
The procedure is similar for the case when the pair of nodes no. l being previously in contact lose their contact. However, in this case, such the loads (R^*, K_x^*, K_z^*) are being sought for which the normal forces in contact $p(l)$ (or $s(l)$ if the pair is situated on the crack faces) are close to zero, and the appropriate equation is removed from the set of equations (1.1).

MODELLING OF 3D-RAIL HEAD ROLLING CONTACT FATIGUE CRACKS.

P. E. BOLD
Department of Mechanical and Process Engineering,
University of Sheffield, P.O.Box 600, Mappin Street, Sheffield, S1 4DU, U.K.

Abstract

The methods of predicting the growth rates of rail head cracks are reviewed, with reference to the tache ovale and squat defects. Results from the finite element modelling of the stress in the rail head, a new stress intensity calibration for elliptical cracks in non-uniform shear stress fields, and fatigue testing under combined tensile and shear loads are discussed.

INTRODUCTION

Squat, shell, tache ovale and detail defects are an international railway problem, costing millions of pounds annually. If their growth rate could be predicted from the wide variety of loading parameters, then the economic costs of varying those parameters could be estimated, and the most economic policy of track loading, maintenance, and renewal could be evaluated. In particular it would be possible to predict the costs or savings associated with these defects due to changing the quality of steel used, the rail section, the stiffness of the track supports, the stress free temperature, residual stresses, static axle loads, dynamic load characteristics, track curvature, or wheel/rail conformity.

The costs are associated with the inspection of rails, the repair of cracks and the replacement of rails. A predictive model will therefore consider how many defects need to be repaired, how frequently the track needs to be inspected and how many rails need to be replaced per year. This will be done by considering the initial defects that are present in a rail, or the maximum size of defect that might have avoided detection. The model would then determine whether that defect will start to grow, if so how fast, and at what point it would fracture the rail.

Addressing these issues involves two basic stages. Firstly the crack tip conditions due to the track loading need to be evaluated, usually in terms of stress intensity factors. Secondly the consequences of these crack tip conditions, in terms of whether a crack starts growing, how quickly it grows and whether it causes fracture need to be determined. This is achieved by either reproducing the conditions in a laboratory where the growth can be observed, or by making predictions using growth rules developed from laboratory tests using similar conditions.

Various attempts at preforming this type of modelling have been made, and the work of Orringer *et al* is acknowledged as a good example of the general approach

J.J. Kalker et al. (eds.), Rail Quality and Maintenance for Modern Railway Operation, 397–406.

that should be used [1]. In this paper the various methods of predicting the stress intensities will be discussed, the limitations of traditional laboratory testing techniques will be pointed out, and recent work that has substantially increased the potential accuracy of these modelling techniques will be described. The paper concentrates on the tache ovale defect, as this is the simplest to model. Squats will also be mentioned, and the particular problems associated with predicting their growth will be discussed. No work has been done at Sheffield specifically directed at the detail and shell defects, but the general approach discussed here could also be applied to these cracks.

THE PREDICTION OF STRESS INTENSITY CYCLES FOR TACHES OVALES.

Taches ovales are internal defects, initiating 10-15mm below the running surface. When they were first discovered, they were initiating from hydrogen shatter cracks, which could give starter cracks of up to 5 mm diameter. Recent improvements in rail manufacture have prevented these shatter cracks occurring, but taches ovales still appear. The cause of this is as yet unknown.

For internal defects the best way of calculating the stress intensity cycle is a three stage process. Firstly the stress field for the uncracked rail head is calculated. The stress intensity for a crack in that stress field is then found, using established stress intensity calibrations. Finally corrections are made for crack closure, crack face friction, and for the finite width of the rail, if necessary.

THE CALCULATION OF STRESS IN A RAIL HEAD.

There are three different methods of predicting the stress field in a rail head under rolling contact loading. Each method has its own advantages and disadvantages. The easiest of these approaches is to treat the rail as a simple beam with the sleepers providing an elastic foundation, and the head as a beam with the web providing an elastic foundation. The equations for this method are well established [1,2], and the calculations required are quick. However it does not model the three dimensional nature of the problem or the contact stress field.

The second approach is to model the contact load as a Hertzian stress on an elastic or elastic-plastic half space [3]. This can model the contact zone very accurately, but the mathematics involved is hard, and the three dimensional rail geometry cannot be examined.

The third approach is to use the finite element method. This is the only one that can model the geometry of the rail and the contact patch. The weaknesses of the method are that highly inaccurate answers can be generated by its incorrect usage, and that large amounts of computer time will be required to predict the stress fields for all the possible loading configurations. However this method does appear to provide the best route for the prediction of tache ovale behaviour. A round

robin exercise to compare the rail stress predictions of a variety of finite element programmes and meshes would provide a check on their accuracy.

Some results found using the method are described here, which show the type of stresses that occur in the rail head, and therefore the type of stress intensity calibration that is required and the type of laboratory fatigue testing. The finite element programme used in this work was one written by British Rail specifically for calculating stresses in beams on elastic foundations. Unlike normal three dimensional finite element programmes, which use a three dimensional mesh structure, this programme uses a two dimensional mesh of the rail section, and then creates what are effectively very long thin elements in the longitudinal direction. These elements are given a displacement function in the longitudianal direction made up of any number of harmonics.

Fig. 1a. shows how the longitudinal stress 10mm below the rail surface varies across the width of the rail. Fig. 1b. shows how the longitudinal stress in the centre of the rail varies with height. For defects up to 4 mm in diameter, the stress could be modelled as a uniform stress, or a linearly varying stress. For defects up to 20 mm diameter, the stress could be described by a quadratic. If the stress intensities are to be accurately calculated they therefore need to be able to model stresses described by a quadratic.

In addition to the stresses due to the wheel load, there are thermal and residual stresses in the rail head. These need to be added to the wheel stresses, in order to predict the whole stress cycle. Fig. 2. shows the stresses on a plane 20° to the vertical on which the tache ovale grows. The tensile stresses perpendicular to the plane and the shear stress on the plane, including residual and thermal stresses, are plotted against distance from the contact load. This figure shows that the stress cycle experienced by a tache ovale contains significant tensile and shear stresses. Stress intensity calibrations and laboratory fatigue crack growth measurements are therefore required for both tensile and shear stresses.

In fracture mechanics a crack growing perpendicular to a tensile stress is said to be growing in mode I, Fig.3a, if it is growing parallel to a shear stress, it is growing in mode II, Fig.3b, and if it is growing perpendicular to a shear stress it is growing in mode III, Fig.3c. A tache ovale will experience all three types of loading, with the magnitude of the shear stress intensities varying with position around the crack circumference.

THE CALCULATION OF STRESS INTENSITIES.

The solution for the tensile stress intensities around an elliptical crack under uniform tension, or a linearly varying tension are well known, [4]. However the solution for any stress field that can be described by a third order polynomial in x and y has been produced by Shah *et al* [5], so the tensile stress intensities can be calculated with reasonable accuracy for any stress field in the rail head. However, the only solution for the shear stress intensities that could be found in the literature

was for a uniform stress field. This is presumably because under the majority of loading conditions, cracks loaded in shear branch so that they grow perpendicular to the maximum tensile stress, and so shear stress intensities are not of great interest. As taches ovales do experience a significant shear stress field without branching, a suitable calibration is required.

Such a calibration has now been produced at Sheffield, using the three dimensional stress functions used by Shah *et al* and Kassir and Sih [5,6]. This calibration gives the stress intensities due to any shear stress field that can be described by a third order polynomial in x and y, to compliment the tensile stress intensity calibration of Shah *et al.*

It should be noted that both of these stress intensity calibrations require a computer to solve the equations, and so the simpler calibrations for uniform stresses, or for circular cracks rather than for ellipses should be used where only an approximate answer is required.

CORRECTIONS FOR CRACK FACE CONTACT AND THE FINITE WIDTH OF THE RAIL.

The stress intensities for any stress field calculated by the finite element method can therefore be found. However, these assume that the crack faces do not touch and that the crack is in an infinite stress field. As the crack faces will touch when the thermal stresses are low, or when the crack grows out of the tensile residual stress region, compressive stresses and shear stresses will occur on the crack faces, making the real stress intensity ranges smaller than the theoretical ones. Also as a rail has a finite width the real stress intensities will be higher than those predicted for an infinite stress field, particularly for large cracks.

Corrections to the stress intensities for crack face contact could be calculated using a development of the iterative approach used by Bower to correct for crack face contact in his two dimensional model of a surface breaking rolling contact fatigue crack [7]. This involves using the stress functions to determine if the crack faces are overlapping, and correcting for this iterativley by adding crack face stresses.

Corrections for the finite width could be accurately calculated by developing the approach used by Shah and Kobayashi in calculating the stress intensities around an ellptical crack approaching the surface of a semi-infinite solid [8]. However for taches ovales the free surface will only become significant when the tache ovale height reaches about 10 mm. The current British Rail action criteria require that 10 mm high defects are removed from service. This is in accordance with approximate calculations that showed that a defect about 10 mm in height could fracture under the worst possible conditions. Such defects should therefore be removed before the effects of the free surface become significant. The above process is therefore unnecessary, unless more accurate calculations show a significantly different picture, or a tougher rail steel is used, or the loading conditions change.

THE GROWTH RATE OF CRACKS UNDER COMBINED TENSILE AND SHEAR STRESSES.

Fig 2. showed that the stress cycle experienced by a tache ovale had both tensile and shear components. Converting these into stress intensities gives a mixed mode I and mode II stress intensity cycle at the top and bottom of the crack, a mixed mode I and mode III stress intensity cycle at the two sides of the crack, and a mixed mode I, II and III around the rest of the crack.

No crack growth tests are recorded in the literature that have reproduced the type of stress cycles shown in Fig. 2. This is not surprising as it is not possible to do so with conventional fatigue testing machines. The nearest are those that have been performed at Sheffield [9], in which mode I and mode II cycles have been applied sequentially, as shown in Fig. 4. These tests were performed using a cruciform or cross shaped specimen. If tension is applied to the specimen horizontally and vertically, a crack at 45° will experience a mode I stress intensity. If tension is applied on one axis, and compression on the other, then the 45° crack will experience a mode II stress intensity.

The potential significance of the combination of the tensile and shear stress intensities is shown in Fig. 5, which shows the growth rates measured in two tests for the case where the mode II stress intensity was twice the mode I stress intensity. For comparison the growth rates that would be predicted for the mode I part of the cycle alone is also shown. It can be seen that the combination of mode I and II increases the growth rate by a factor of about 10.

Work is continuing at Sheffield University into this type of loading, and in particular into the effects of overlapping the cycles. In the future it is hoped that thresholds and mean streses will also be examined, and that similar tests under mixed mode I and III will be performed.

THE MODELLING OF OTHER DEFECTS.

So far taches ovales have been discussed. In principal the same type of method should be used to predict the growth rate of any rail head defect. However other defects do have particular problems that should be mentioned.

Squats have been studied in some detail at Sheffield, and it was the desire to predict their growth rate that initiated the mixed mode I and II loading studies that have already been mentioned [9]. However the calculation of the stress intensity cycle for a squat is a much harder process than for a tache ovale. In the first place they initiate and start to grow within the contact stress region which is difficult to model with finite elements, and where compressive stresses across the crack faces are high. Secondly it was shown by Way in 1935 that surface breaking cracks growing under a rolling contact load require a liquid if they are to propagate [10]. If the cracks are to be accurately modelled the effect of the liquid on the stress intensity cycle needs to be evaluated.

Significant progress was made by Bower in this area, by using the method of distributed dislocations developed by Keer [7,11]. Fig.6. shows one of his predictions for a stress intensity cycle. Similar models have been produced by Olzak *et al* [12] using finite elements. They also predict a mixed mode I and II cycle, but the shape of their cycle is somewhat different. Hopefully both of these approaches will be developed, and differences between them discussed in order to ascertian which one is the more appropriate tool for predicting the stress intensity.

Both Bower's and Olzak's models were two dimensional, and it is still uncertain as to whether a three dimensional predction method can be developed. Initial studies in this area have been made by Stupnicki *et al* [13] which suggested that the three dimensional results could be predicted from the two dimensional model, but further studies are hampered by the enormous amount of computer time required for the three dimensional finite element model.

Studies are continuing at Sheffield into the laboratory crack growth measurement tests, and in the near future these will be combined with Bower's prediction to make an initial squat growth prediction model.

CONCLUSIONS

Concerning taches ovales:

1. The use of finite elements is recommended for the calculation of the stresses in a rail head, for a three dimensional geometry.

2. The use of stress intensity calibrations based on three dimensional stress functions is recommended for the prediction of tensile and shear stress intensity factors for elliptical cracks in the stress fields calculated in 1.

3. The use of mixed mode I and II, and mixed mode I and III non-proportional loading tests is recommended for the prediction of the growth rates due to the stress intensity cycles predicted in 2.

4. If these recommendations are followed it should be possible to predict the economic costs associated with taches ovales.

Concerning squats:

1. Both the finite element method and the distributed dislocation method should be used to calculate two dimensional stress intensity cycles for squats.

2. These cycles should be combined with mixed mode I and II laboratory growth rate observations in order produce a growth rate and branching prediction model for squats.

3. If possible the model should be extened in the future to model three dimensional aspects of squats, as and when the stress intensities can be calculated.

ACKNOWLEDGEMENTS

Dr. Bold was sponsored by British Rail during the time this work was completed. The advice and assistance of the staff of British Rail Research in Derby is gratefully acknowledged.

References

[1] O.Orringer, Y.H.Tang, J.E.Gordon, D.Y.Jeong, J.M.Morris, A.B.Perlman (1988). Crack Propagation Life of Detail Fractures in Rails. U.S.Department of Transportation, Research and Special programs Administration, Transport Systems Centre, Cambridge, Massachusetts, 02142. Final Report.

[2] R.J.Roark (1989). Formulas for Stress and Strain. McGraw-Hill Publishing Company.

[3] M.T.Hanson, L.M.Keer (1991). Analysis of Edge Effects on Wheel Rail Contact. Presented at the Third International Conference on Contact Mechanics and Wear of Rail/Wheel Systems. Cambridge, 22-26 July 1990. Published in *Wear*, 144, pp39-55.

[4] *Stress Intensity Factor Handbook.* Ed. in Chief Y. Murakami. JSMS - Pergamom Press.

[5] R.C.Shah, A.S.Kobayashi (1971). Stress Intensity Factor for an Elliptical Crack under Arbitrary Normal Loading. *Engineering Fracture Mechanics.* Vol.3, pp71-96.

[6] M.M.Kassir, G.C.Sih, (1974). Embedded Elliptical Cracks. *Three Dimensional Crack Problems.* Noordhoff International Library. pp 74-117.

[7] A.F.Bower (1988). The influence of crack face friction and trapped fluid on surface initiated rolling contact fatigue cracks. *Trans. ASME J. of Tribology.* Vol. 110, pp704-711.

[8] R.C.Shah, A.S.Kobayashi (1973). Stress Intensity Factors for an Ellptical Crack Approaching the Surface of a Semi-Infinite Solid. *International Journal of Fracture* Vol. 9, No.2, pp133-146.

[9] P.E.Bold, M.W.Brown, R.J.Allen (1991). Shear Mode Crack Growth and Rolling Contact Fatigue. Presented at the Third International Conference on Contact Mechanics and Wear of Rail/Wheel Systems. Cambridge, 22-26 July 1990. Published in *Wear*, 144, pp307-317.

[10] S.Way (1935). Pitting due to Rolling Contact. ASME *Journal of Applied Mechanics*, Vol. 2, pp A49-A58.

[11] L.M.Keer, M.D.Bryant (1983). ASME *Journal of Lubrication Technology*, Vol.105, pp198-205.

[12] M.Olzak, J.Stupnicki, R.Wojcik (1990). Investigation of Crack Propagation in Rail due to the Contact Load of Wheel. Warsaw University of Technology. Report for ORE Committee D173.

[13] J.Stupnicki, M.Olzak, R.Wojcik (1991). The Investigation of Crack Propagation in Rails due to the Transient Load of Wheel. Warsaw University of Technology. Report for ORE Committee D173.

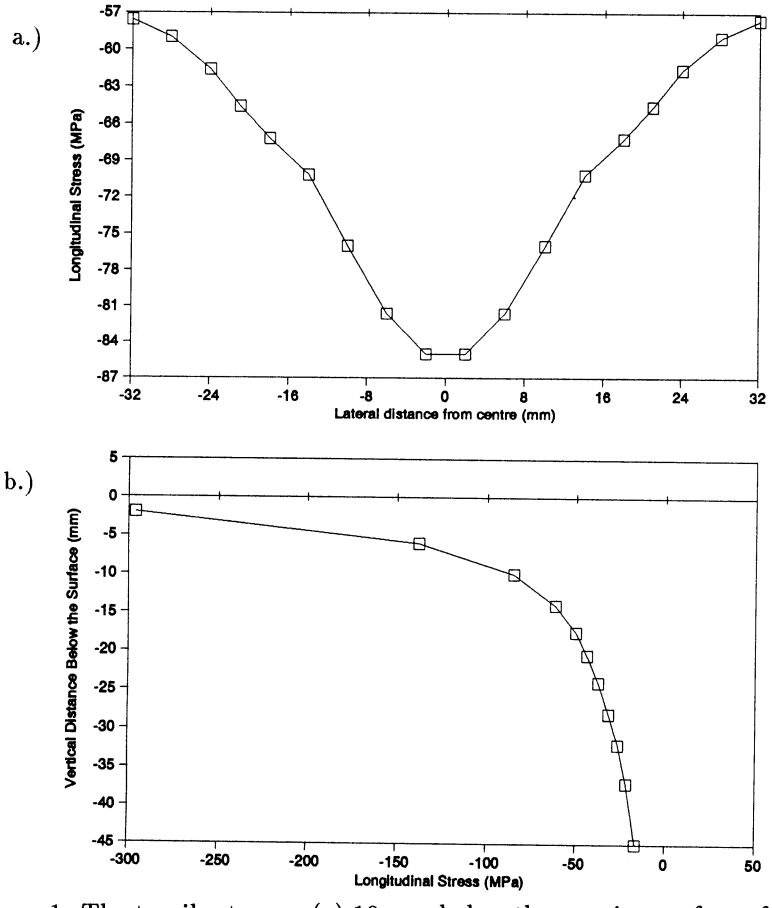


Figure 1: The tensile stresses (a) 10 mm below the running surface of a rail, and (b) on the centre line of the rail; due to a 100kN load, applied on the axis of symmetry, on BS113A section rails, with a track stiffness of 50 MN/m per sleeper end.

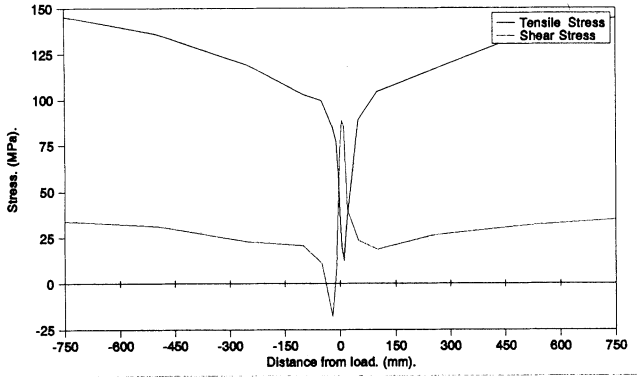


Figure 2: The tensile and shear stresses 10 mm below the running surface of a rail, due to a 100kN load, applied on the axis of symmetry, on BS113A section rails, with a track stiffness of 50 MN/m per sleeper end. A vertical residual stress of 50 MPa, a longitudinal residual stress of 100 MPa, and a longitudinal thermal stress of 50 MPa have been included.

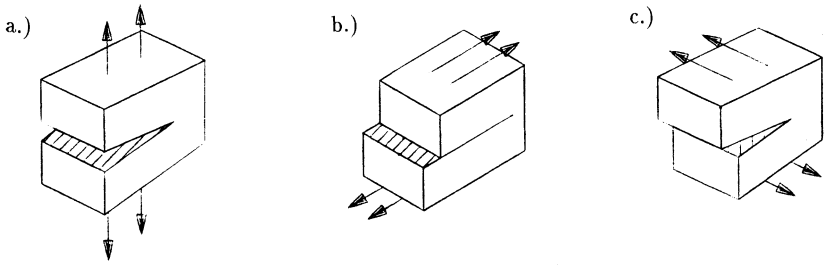


Figure 3: Modes of Loading. (a) Mode I, (b) Mode II, (c) Mode III

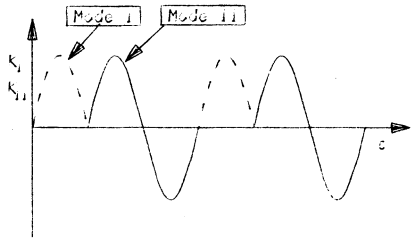


Figure 4: The Stress Intensity Cycles Used by Bold.

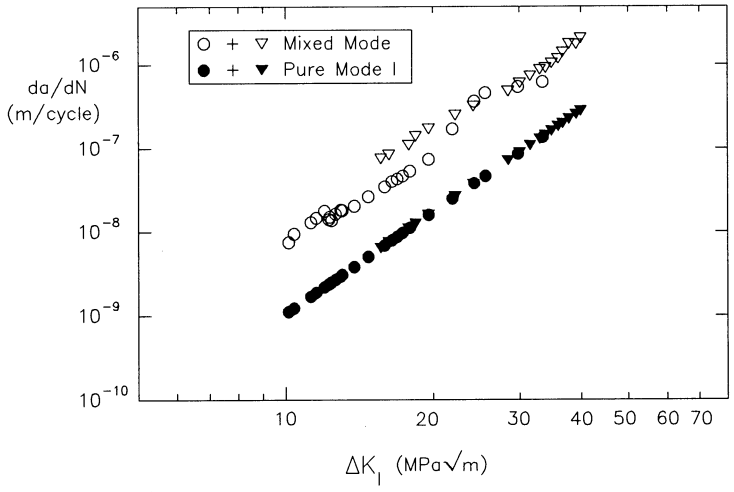


Figure 5: The Growth Rates Measured by Bold. The Open symbols are the observed growth rates, and the closed symbols are the growth rates that would be predicted for the mode I load alone.

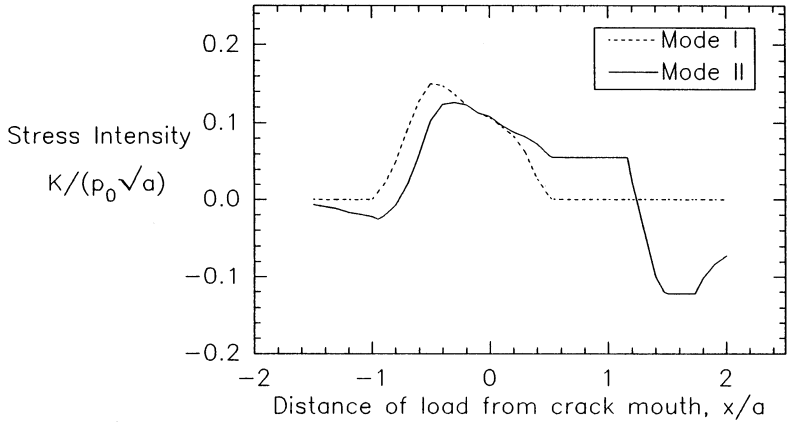


Figure 6: Stress Intensity Cycle Predicted by Bower, for Rolling Contact Fatigue. p_0 is the maximum Hertzian contact pressure, a is the contact patch radius, and x is the distance from the crack mouth. The crack makes an angle of 25° to the running direction, and the crack length is half the contact patch width.

Tribological Investigations of Crack Development on the Rail on an energetical basis

Fleischer, G., Technische Universität Magdeburg
Grohmann, H.-D., Deutsche Reichsbahn, Zentralstelle Umweltschutz und Materialprüfung Brandenburg-Kirchmöser
Weinhauer, D., Technische Universität Magdeburg

1. Subject of investigations

Narrow curved tracks cause high wear on rails necessitating frequent replacement of rails. As a countermeasure, a lubricant was fed into the tribological system wheel/rail leading to reduced wear, however, causing cracks on the gauge corner as shown in Figure 1. The problem with this damage, called "Head Checks", is that the degree of damage can only be determined by great effort. Consequently, it is also possible that the gauge corner will break out, considerably impairing both travelling comfort and safety of the train. These Head Checks occur with increased travelling speeds and axle loads - resulting in higher forces between wheel and rail.
Consequently, the question arises how the above mentioned damage can be reduced or even completely eliminated. To this end, it must be clear what leads to their development.
A possible approach to solve the problem is to apply Fleischer's energy theory of friction and wear which will be described below.

2. Energetic analysis

All investigations start from an analysis of the given tribological system. A tribological system (see Figure 2) is an integral part of technical or biological (which are not considered here) functional groups and presents a frictional contact. This means that in the case of an interaction of the system components (frictional elements 1 and

J.J. Kalker et al. (eds.), Rail Quality and Maintenance for Modern Railway Operation, 407–424.

2, intermediate element, surroundings) friction might occur and, consequently, also wear [1]. It is clear that friction and wear have a system character and, consequently, it is not enough to consider one system element alone. Moreover, friction and wear can be characterized as a temporal process, whereas the damage of the frictional elements is understood as being caused by partial energy accumulation in the frictional element material becoming apparent when a certain threshold value will be exceeded (e. g. the fracture energy density e_B). Thus it will be possible to comprehend crack beginning and crack development as a phenomenon of wear. Therefore, material separation in the micro-area is the common basis for the occurence of wear and cracks.
In an energetic analysis, the following states of wear are distinguished: abrasive wear, material building-up wear, deformation wear and material changing wear. These states of wear are similary classified in the DIN Standard. Thus, for example, abrasion can be compared with abrasive wear level 4 (see Table 1). The individual levels of abrasive wear may be related to frictional states and correspondig ranges of values of tribological parameters (e_R^* - apparent frictional energy density, $\tilde{\tau}_R$ - friction shearing stress, I_h - linear wear intensity, n_k - critical contacting number). A corresponding survey is given in Table 1 and Figure 3 [2]. The vitual frictional energy density e_R^* can be used as an important parameter characterizing the tribological system and as a comparing quantity between real and modelled systems.

On condition that direct proportionality exists between frictional work and friction volume we obtain:

$$W_R = \frac{W_R}{V_R} * V_R = e_R * V_R \qquad (1)$$

W_R — Frictional work
e_R — Frictional energy density
V_R — Friction volume

Analogously, the following equation may be derived:

$$W_R = \frac{W_R}{V_V} * V_V = e_R^* * V_V \qquad (2)$$

W_R — Frictional work
e_R^* — apparent frictional energy density
V_V — Volume of wear

Other than the volume of friction, the volume of wear can be directly determined by experiments, consequently it is better to use equation

(2). Equations (1) and (2) are linked by the definition of wear index ν_V.

$$\nu_V = \frac{V_V}{V_R} \tag{3}$$

Consequently, the coupling of frictional energy density is given as follows:

$$e_R = e_R^* * \nu_V \tag{4}$$

Defining the frictional work as given below

$$W_R = \int_{(s_R)} F_R * ds_R \tag{5}$$

F_R — Frictional force
s_R — Friction distance

and using the equations

$$\frac{W_R}{A_R} = \frac{F_R * s_R}{A_R} = e_R^* * \frac{V_V}{A_R} \tag{6a}$$

$$\frac{F_R}{A_R} = e_R^* * \frac{V_V}{A_R * s_R} = e_R^* * \frac{h_V}{s_R} \tag{6b}$$

h_V — Height of wear

the so-called basic equation of wear (after FLEISCHER) can be derived:

$$\tau_R = e_R^* * I_h \tag{7}$$

τ_R — Frictional shearing stress
e_R^* — Apparent frictional energy density
I_h — Linear wear intensity

The coefficients given in equation (7) describe a certain state of friction and wear of the system. If a particular system element is to be investigated in detail, the friction energie shall be accordingly distributed to the frictional elements.

$$W_R = \sum_{(i)} W_{Ri} \qquad (8)$$

(i) ——— Number of frictional elements

The following equation holds good for two frictional elements:

$$W_{R1} = \alpha_1 * W_R \qquad (9a)$$

$$W_{R2} = \alpha_2 * W_R \qquad (9b)$$

$$W_R = W_{R1} + W_{R2} \qquad (10)$$

where $$\alpha_1 + \alpha_2 = 1 \qquad (11)$$

For classification of the tribological system investigated, see Figure 3 or Table 1. The limits of the ranges are based on the evalution of many results obtained by experiment and are predominantly ensured for metallic materials. By classifying the system investigated, it is possible to compare it with other systems and to allpy the knowledge available in order to improve both frictional and wear bbehaviiour of the subject under investigation. When the system elemnts have been determined, it is necessary to determine those properties of the elements which can be characterized by shape and material parameters as well as other effects such as forces, moments, temperatures, velocities, ambient conditions etc.

It is the objective of this analysis to assess the conditions prevailing in the contact area in order to establish corresponding conditions for testing in the laboratory.
Considering the tribological system from a thermodynamic point of view means to deepen the system analysis. In the case that wear occurs in the frictional contact, it is necessary to make up the energy balance for the open thermodynamic system (compare Figure 2)[3].

$$\Delta U = U_2 - U_1 = -(H_2 - H_1) - Q_{1-2} + A_{t1-2} \qquad (12)$$

where: U - Internal energy
H - Enthalpy
Q - Heat
A_t - Technical work
1,2 - State 1, State 2

If we equate the frictional work W_R with the technical work, we get:

$$W_R = \Delta H + Q_{1-2} + \Delta U \qquad (13)$$

Under the condition that the kinetic and the potential energies of the system remain constant and the heat dissipation by wear particles is negligible, the following equation will hold:

$$\Delta H = \Delta H_V = W_{RV} \qquad (14)$$

ΔH_V — Wear

$$W_{RV} = h * m = h * \rho * V_V$$

V_V — Wear volume
ρ — Density
m — Mass
h — Specific enthalpy
W_{RV} — Frictional work required to produce wear

By inserting into and reordering the equations (13) and (14), we obtain:

$$e_R^* = \frac{W_R}{V_V} = \frac{h * \rho}{1 - \frac{Q_{1-2}}{W_R} - \frac{\Delta U}{W_R}} = \frac{e_R}{\nu_V} \qquad (15)$$

where the non-dimensional denominator corresponds to the wear index ν_V mentioned above (equations (1) to (4)).

Herewith it is proved that applying e_R^* as a tribological parameter is also justified from a thermodynamic point of view.

If we assume the frictional contact to be a source of energy of a determined power, by deviding both values of the quotient by $A * t_R$, where A - friction surface; t_R - friction time, e_R^* can be brought into the following form:

$$e_R^* = \frac{W_R}{V_V} = \frac{p * \mu * v_R}{\dot{h}_V} \qquad (16)$$

v_R — Frictional velocity
$\dot{h}_V$ — Wear velocity

This description of the apparent frictional energy density illustrates the interrelation between frictional power per unit area and velocity of wear. In the case of developing fatigue cracks, the time of crack beginning and wearing speed can be associated with each other as described below.

Consequently, it is possible to determine a characteristic value for e_R^* in the case of the "crack development". The contact-number of the frictional elements up to the moment the damage becomes obvious, can be determined by the apparent frictional energy density.

3. Modelling of the tribological system of wheel flange/rail gauge corner

It is the aim of this modelling (see Figure 4) to determine the trial parameters for a testing machine which was specially designed for these investigations. This machine serves to simulate the damaging process at defined energetic conditions. To this end, the following assumptions are made:

- From the point of view of gear technology, the system is rigid.
- The angular velocity of the wheel is constant.
- The axis of rotation of the wheel is perpendicular to the rail track.
- There is freedom from slipage between wheel and rail.
- There is line contact in zone A_1.
- Two level faces contact each other within zone A_2.
- F_a = axial force, F_r = radial force, F_z = tractive force with regard to one wheel.

Now it is easy to subdivide the contact zone A_2 into many individual points K. Figure 5 allows to illustrate how the relative velocity is directed at any contact point and how it may be calculated. From the point of view of gear technology, the system wheel/rail has been prepared here as a wheel rotating de-spin around D and a rail which is moving transversally.
Consequently, the following relations may be derived:

$$\overrightarrow{v_{K21}} = \overrightarrow{v_{K31}} + \overrightarrow{v_{K23}} \qquad (17)$$

(Rotational velocity of point K around the axis of rotation)

$$\overrightarrow{v_{K31}} = \omega_{21} * \overline{DM} \qquad (18)$$

(Travelling speed)

$$\overrightarrow{v_{K32}} = \omega_{21} * \overline{MK} \qquad (19)$$

(Relativ velocity between gauge corner and wheel flange)

In addition to the relative velocity in the contact point, the size of the contact area is required for determining the specific frictional power. The point K as such does not yet present an area, however, its relative velocity v_{K23} is assumed as being representative of the process of motion.

Assuming F_a = 15 kN, μ = 0.2, v_{K31} = 100 km/hour, MK = 100 mm, A_R = 125 mm², we obtain at v_{K23} = 6.2 m/sec a value of 149 W/mm² for the specific frictional power. Using data for the Hertzian stress in the rail gauge corner/wheel flange contact area from literature [4], as a rule values of approximately 1500 MPa are given. However, since it is very difficult to determine the correct value of the normal force at the contact and the pertinent contact area, this standard value will be used as a basis for further considerations. Moreover, for the friction index μ a range of 0.1 to 0.3 is used. For P_R/A_H this results is approx. 1 to 3 kW/mm². When comparing this value with specific powers which are introduced at laser treatment by the laser beam, it becomes event that the damage is not only due to mechanical, but also to thermal effects. In compliance with Figure 6 [5], for contact periods of 10^{-3} sec the phenomenon of transformation hardening can be expected which might be the cause of thermal crack development. However, crack growth is only possible if the abrasive wear velocity h_V is smaller than the velocity of crack growth. To this end, the stress level causing crack beginning in the frictional contact must have been reached. Proceeding from energy and damage accumulation, a certain period is necessary to creat the preconditions for crack development. In addition, it should be observed that the direction of abrasion and of crack development are normally different. As found in literature [6, 7], the velocity of crack growth at the **initial phase** of crack development is 10^{-6} to 10^{-7} mm/vibrating cycle and the directional course is approx. 45° relative to the main direction of stressing, i. e. approx. 45° relative to direction of wear progress. Consequently, it becomes obvious that the velocity of wear, relative to one vibrating cycle, must not become greater than approximately 4 * 10^{-7} mm/vibrating cycle. Under test rig conditions this means that at a setting of some 1...3 vibrating cycles/sec, a velocity of wear lower than approx. 5 * 10^{-7} mm/sec must be implemented.

Theoretical assessments of e_R^* give a range of 3 * 10^8 to 6 * 10^9 J/mm³. It is clear from Table 1 that relatively favourable tribological conditions are necessary for crack simulation.

However, since crack development in the initial phase (micro-range) cannot be easily measured and can be registered macroscopically only, more rapid velocities of crack development are assumed for the time of measurement. Consequently, for this phase of damage an apparent frictional energy density e_R^* which is smaller by several orders of magnitude is to be expected. It is only this level of e_R^* which can be determined by test evaluations.

Current experimental tests have to bear out these statements. Furthermore, it is necessary to determine the characteristic values for various combinations of materials to allow comparisons and to draw concluions for practical usage.

A testing machine was developed for investigation of the above described damage. A cylinder of dia. 10 mm X 40 mm and a cuboid of 10 mm X 10 mm X 150 mm are used as frictional elements (see Figure 7). This combination of forms corresponds to the roller/plane contact according to Hertz allowing to obtain a mean Hertzian stress of approximately 1500 MPa. In the following text, the frictional elements are designated as round and flat specimens. The flat specimen is mounted on a force measuring unit (3-Component-Dynamometer) which allows to record normal and frictional force at the same time. Furthermore it is possible to feed a lubricant into the tribological system. By setting a normal force and a number of revolutions of the carrier of the round specimen, a specific frictional power ranging from 37 to 2000 W/mm^2 (at a frictional index of 0.1 to 0.4) can be obtained. In addition, the body temperature of the flat specimen and the ambient temperature are recorded during the test.
A computer system serves to record all values measured at the test rig and to carry out control functions, thus allowing intermediate evaluations after adjustable cycle frequencies (10^5).
In order to be able to correctly allocate the measured values, it is necessary to mark the contact zones of the round and the flat specimens.

In Figures 8 and 9, K shall be the momentary contact point. A and A' are the initial zones and E and E' are the end zones of the contact area on the respective specimens. C and C' are the characteristic zones with regard to the course of motion between the round and flat specimens which shall not be discussed here in detail. Thus it becomes possible to determine the tribological data which are characteristic of the obtained zones of damage. The measured data available in digitalized form can be evaluated with a suitable software. Examples are given in Figures 10 and 11.

4. Evaluation of laboratory tests in compliance with the energy theory developed by FLEISCHER

Tests were made to investigate the rolling contact behaviour of various rail stells. These tests were carried out using

- small specimens on an Amsler testing machine (BR-BRR Derby) and
- original components on a testing machine according to the rolling load principle (by PKP - Polish Institute for Railway Research, Warsaw).

Steels of normal grade 900 A according to UIC 860 and the same quality, but hardened to 340 HB and 370 HB, were subjected to testing. Various lubricants including water were used on the Amsler testing machine. Applying the principles described in this paper, the apparent frictional energy density e_R^* was determined for the time of the macroscopic occurrence of rolling contact fatigue damage for two different laboratory testing methods. To this end, the factors causing

the damage such as contact pressing, coefficient of friction and wearing rate were determined in corresponding the different conditions of testing and used for calculating e_R^*.
The results obtained on the Amsler testing machine in accordance with the basic equation of wear (7) are shown in Figure 12. However, since a defined time of defect beginning was not recorded during the tests on the rolling load test rig, e_R^* will be described here depending on the load cycles (Figure 13).
The values calculated for the apparent frictional energy density e_R^* are mainly in the order of magnitude of 10^5 J/mm^3; with the values for steel of grade 900 A being below those for head-hardened steels of 340 HB and 370 HB. As expected, the values for e_R^* obtained during the tests on the Amsler machine for "lubricated" steel 900 A are higher than the ones obtained with "water", and as mentioned above, they were below the values for steel of 340 HB and 370 HB. The values for e_R^* increase with increasing hardness. Furthermore, the effect of lubrication is obvious for all three materials; On average I_h for (TEXACO) grease is smaller than for (ASEOL) oil. However, the statements on the effect of the lubricant are not guaranteed due to the relatively small number of tests.

The magnitude of 10^5 J/mm^3 for e_R^* was determined for both the tests on the Amsler test machine and the rolling load test machine. It shows that the conditions of the damaging process were similar in both testing methods. Since this amount corresponds with the values for e_R^* determined earlier under service conditions also with regard to its magnitude, the frictional and wear conditions prevailing during these testing rig trials were similar to field conditions, i. e. the results obtained can be compared with those occurring under practical conditions.

5. Conclusions

Concluding from the energetic consideration, it can be stated that
- a link has been established between laboratory and service conditions;
- individual elements of the tribological system can be comparatively investigated under various conditions; and
- applying this method to rolling contact fatique allows to determine the "rolling contact service life" in advance.

To this end, the tests mentioned in this paper are carried out under defined energetic conditions.

R e f e r e n c e s:

[1] Fleischer,G., Gröger, H., Thum, H.:
Verschleiß und Zuverlässigkeit.
VEB Verlag Technik, Berlin 1980
[2] Fleischer, G.: Zur Energetik der Reibung.
Scientific Journal of Technische Universität "Otto von Guericke"
Magdeburg, 34 (1990) Heft 8
[3] Fleischer, G., Sadowski, J.: Betrachtungen zur Reibungs- und
Verschleißenergetik.
Schmierungstechnik, Berlin 18 (1987) 4
[4] Kalousek, J.: Wear and Contact Fatigue Model for Railway Rail.
Canada 1986
[5] Bonas, C.M.: High Power Laser Welding.
Optical Engineering 17 (1978) 3
[6] Schatt, W.: Einführung in die Werkstoffwissenschaft.
VEB Deutscher Verlag für Grundstoffindustrie, Leipzig 1983
[7] Schott, G.: Werkstoffermüdung.
VEB Deutscher Verlag für Grundstoffindustrie, Leipzig 1976

Figure 1

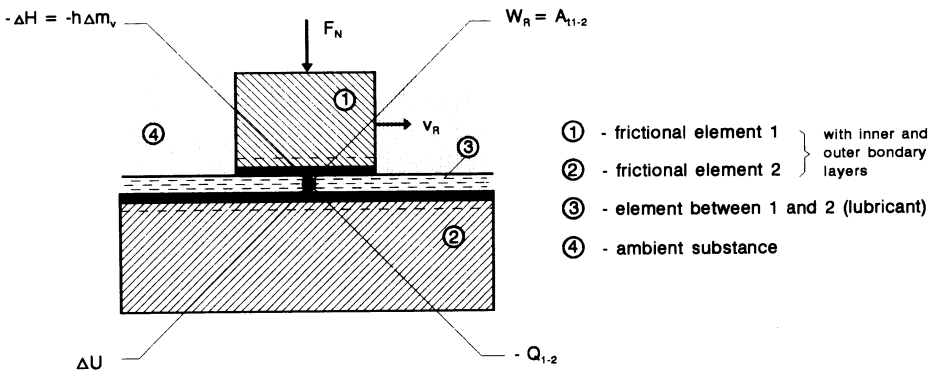
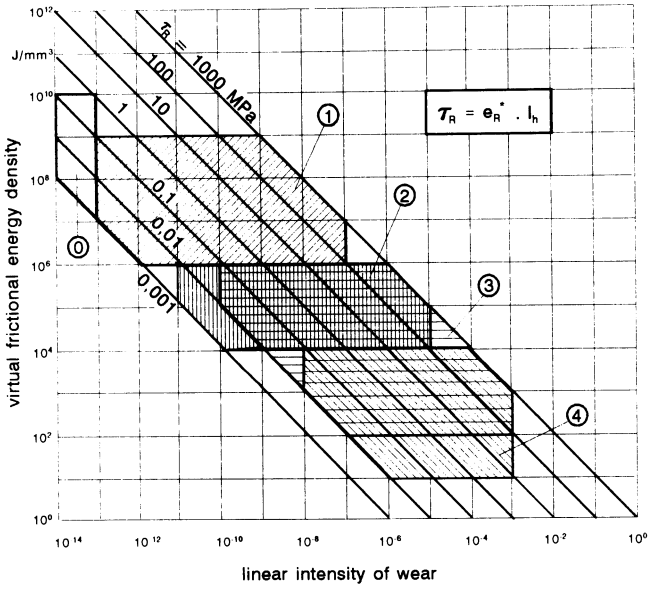


Figure 2



⓪ fluid friction (rheological deformation)

① apparent-fluid friction (elastic deformation)

② mixed friction (elastic-plastic deformation)

③ dry friction (elastic or plastic deformation and cut)

④ dry friction (shaven deformation)

Figure 3

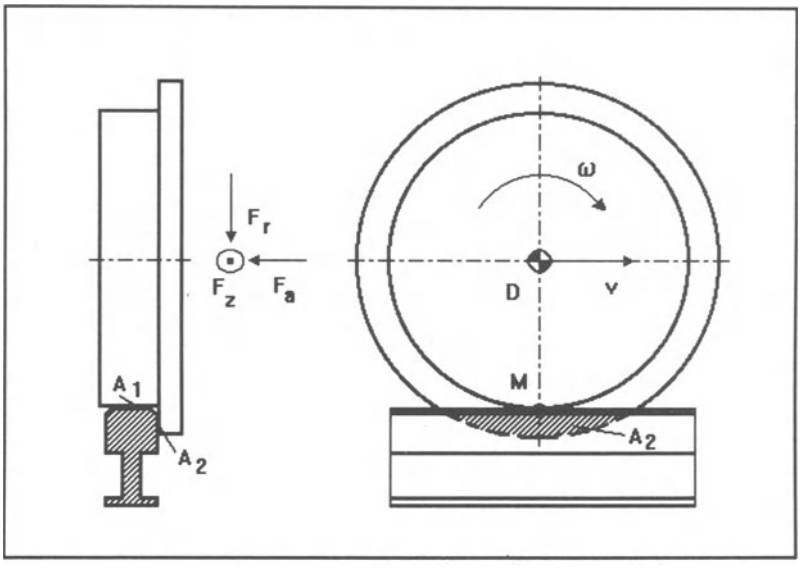


Figure 4

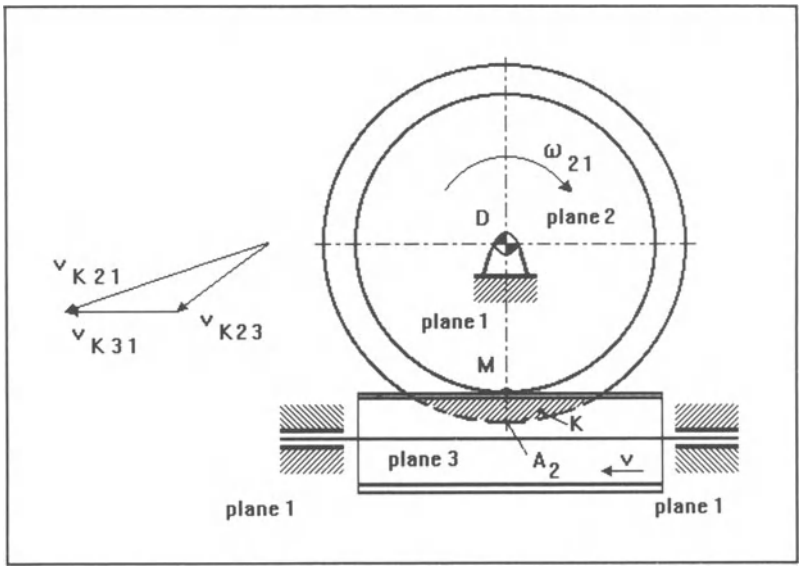


Figure 5

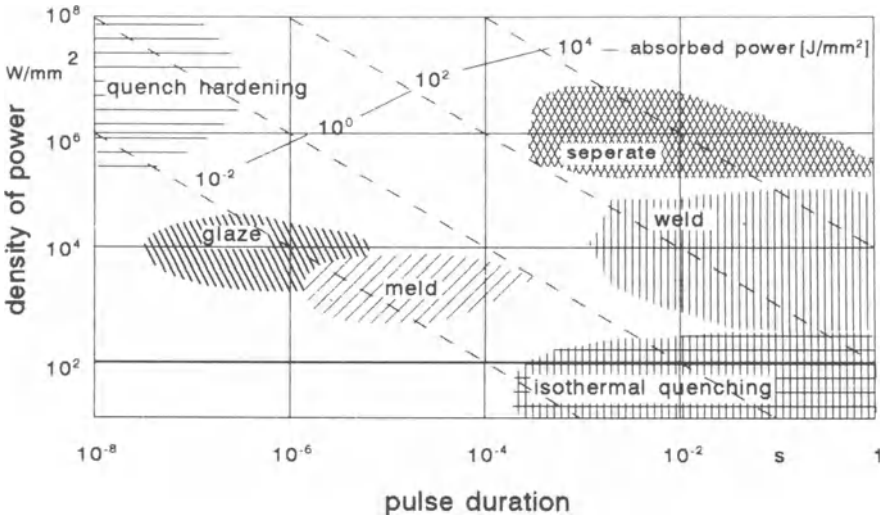


Figure 6

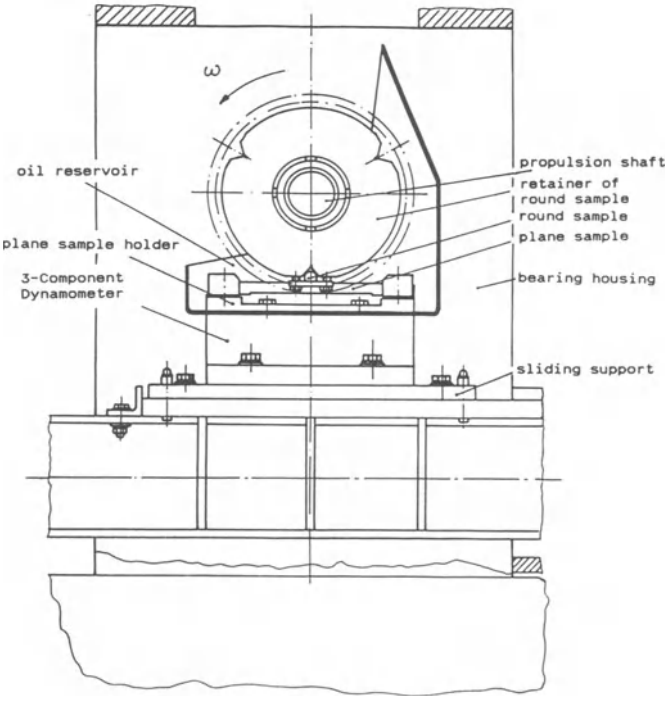


Figure 7

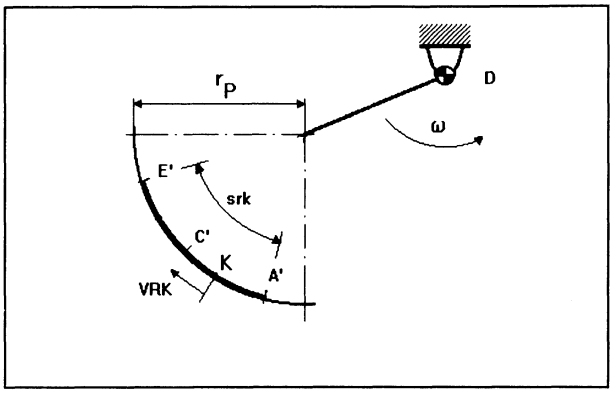


Figure 8

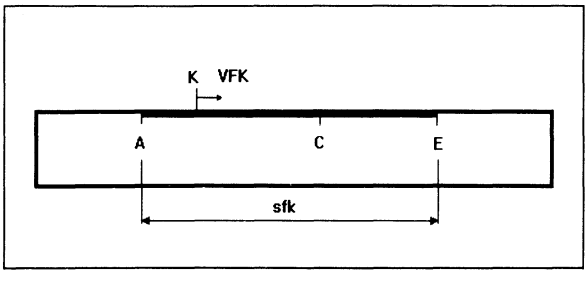
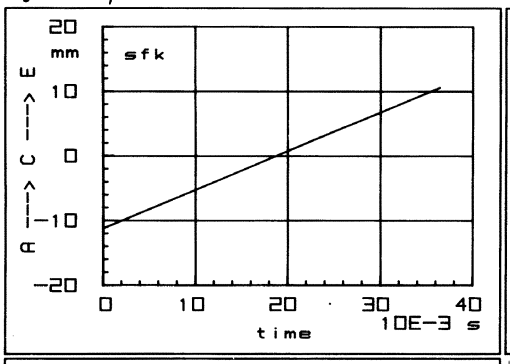
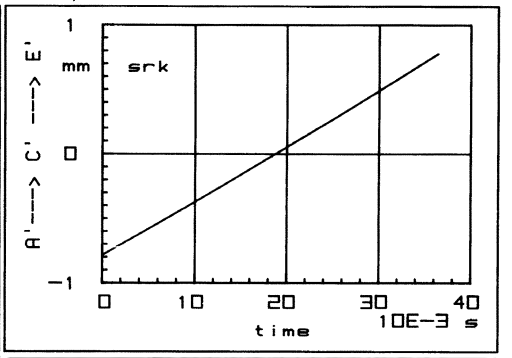


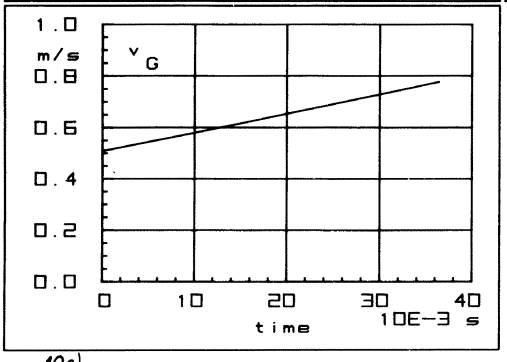
Figure 9

Figures:10a)

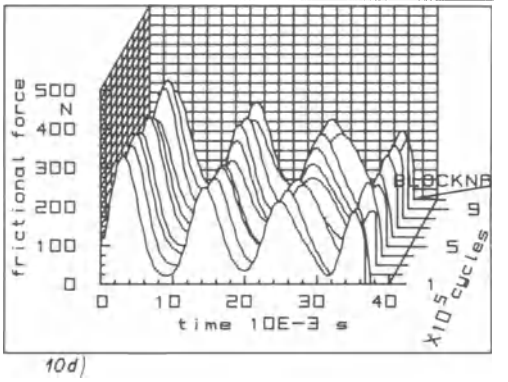


10b)



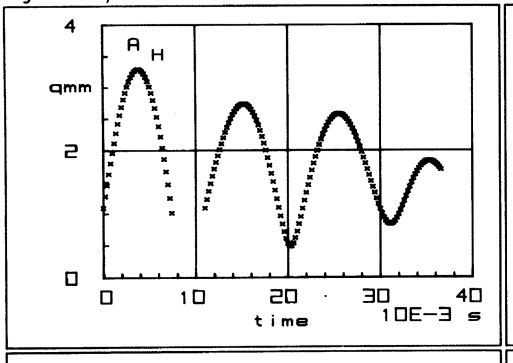


10c)

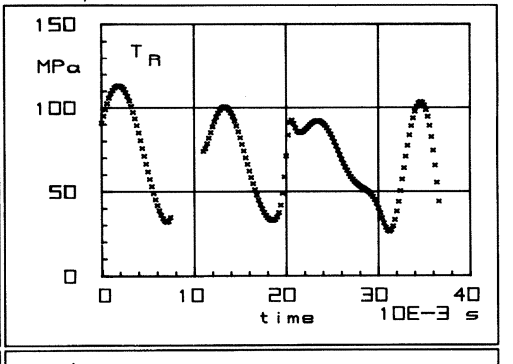


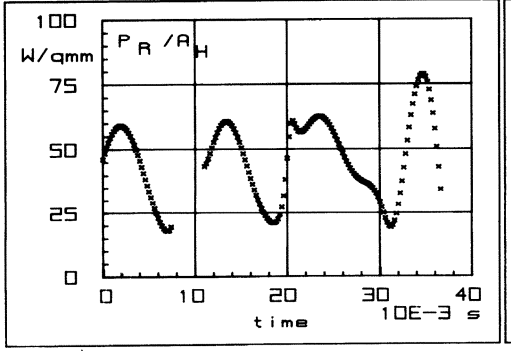
10d)

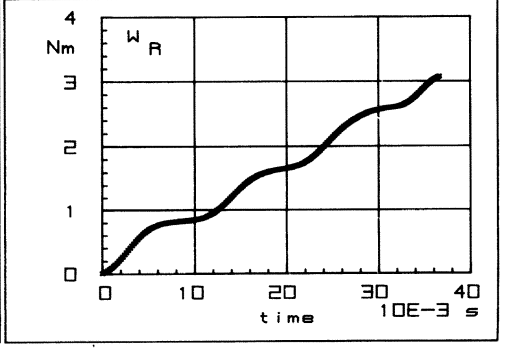
Figures: 11a)



11b)







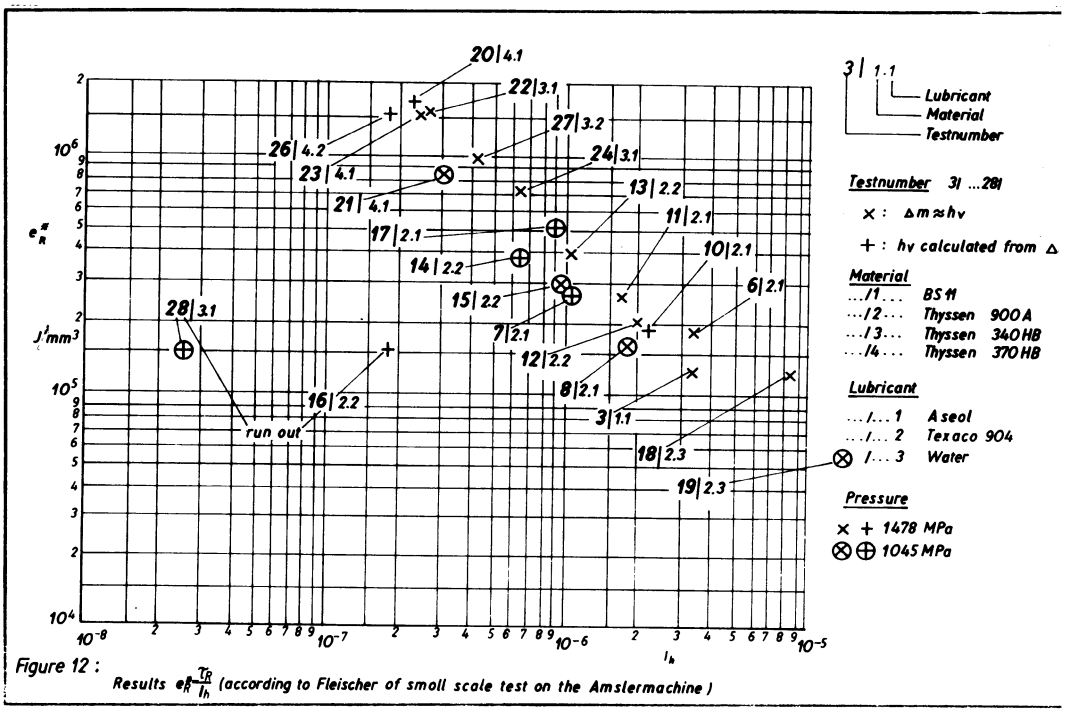


Figure 12 : Results $e_R^* \frac{T_R}{l_h}$ (according to Fleischer of small scale test on the Amslermachine)

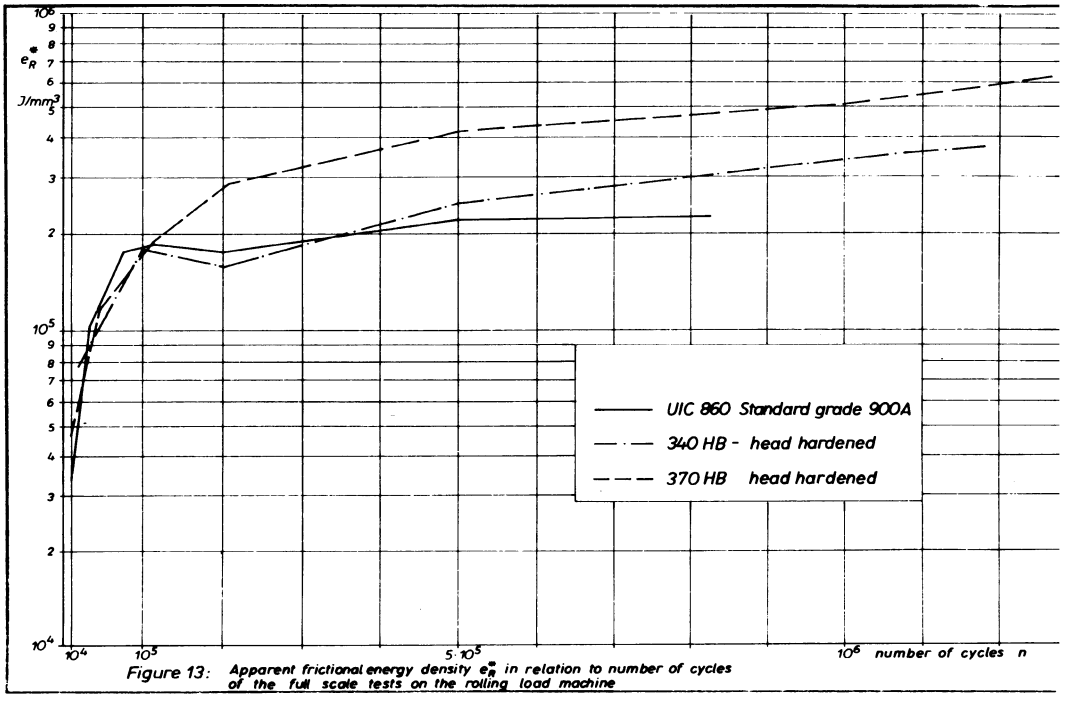


Figure 13: Apparent frictional energy density e_R^* in relation to number of cycles of the full scale tests on the rolling load machine

The Effect of Head Checking on the Bending Fatigue Strength of Railway Rails

Eisenmann J., Leykauf G.
Munich Technical University, Germany

INTRODUCTION

In sharply curved track the service life of rails is shortened by wear and possibly sub-surface initiated defects such as "shelling". However, in slightly or moderately unlubricated or lubricated curved track and sharply curved lubricated track the presence of "head checks" may cause premature removal from track. Although its appearance has been reported for about 15 years [1], the causes are still not understood: however, the initiation of these defects does not appear to be associated with metallurgical defects. The fine head check cracks occur on the gauge corner running surface of the high rail with a spacing of 2 to 10 mm and their angle against traffic direction is about 45° (Fig. 1) . These cracks may propagate and coalesce and spalling may occur (Fig. 2): in some cases a transverse fracture may occur but this appears to be relatively rare. A significant criterion for judging the remaining service life of head checked rails is the bending fatigue strength of the rail head: this has been investigated in laboratory tests forming part of the European Railway Research Institute D173 research programme. The financial support of ERRI for the work reported in this paper is gratefully acknowledged.

RAILS TESTED

A total of 30, 2 m long samples of UIC grade 900 A rail steel in the UIC 60 profile were tested in three groups:

- 10 rail sections(16 years old) with typical head check defects (Fig. 1), removed from a revenue track with a radius of r = 592 m after a total tonnage of about 290 . 10^6 gross tons (maximum speed v = 110 km/h).

- 10 rail sections (20 years old) with severe head check defects (Fig. 2), removed from a revenue track with a radius of r = 1600 m after a total tonnage of about 500 . 10^6 gross tons (maximum speed v = 135 km/h).

- 10 new rail sections

J.J. Kalker et al. (eds.), Rail Quality and Maintenance for Modern Railway Operation, 425–433.

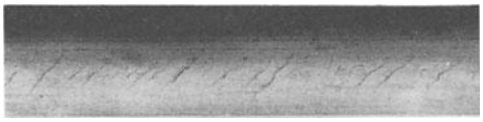


Fig.1: Typical head checks. Rail section UIC 60/900 removed from a track radius 592 m after 290 million gross tons. Traffic direction from right corner to left.

Fig.2: Severe head checks. Rail section UIC 60/900 removed from a track radius 1600 m after about 500 million gross tons. Traffic direction from right corner to left.

BENDING FATIGUE TESTS

Test Set-Up

The loading device for bending tests with fixed and loose rollers as shown in Fig. 3 ensures that all forces applied to the rail were vertical only and were free of excentricity. With the chosen outer span spacing of 1,9 m the rail head bending moment distribution in the test is similar to that created in track in front and behind of single axles as well as between two axles of a bogie with a distance of about 2,5 m to 3,5 m. The bending stresses were controlled by strain gages applied at rail head. The lower stress σ_u was kept constant in the test series (σ_u = 50 N/mm^2), the upper stress σ_o was varied to determine the Wöhler curve as per DIN 50 100. The experiment was terminated when 2.10^6 loading cycles had been reached without fracture of the rail. The loading frequency was 3 Hz.

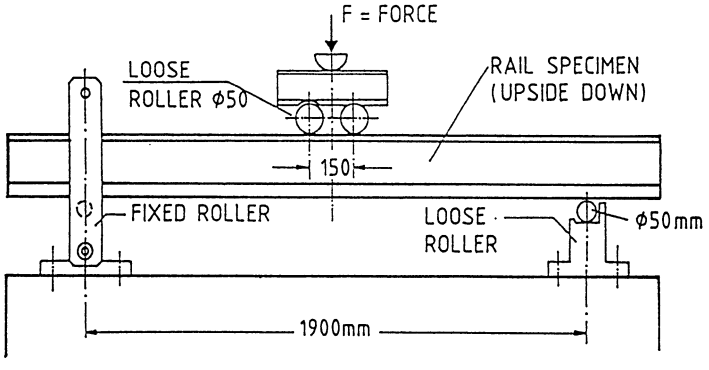


Fig. 3: Test set up for the repeated loading test; bending tensile stresses at rail head.

Test Results

In Fig. 4 the relationship between oscillating bending stress ($2.\sigma_A = \sigma_o - \sigma_u$) applied at the rail head and cycles to failure N is plotted. For every test series a regression analysis was performed to determine that regression line which fits the results best and hence shows the average dependence.

- New Rails
  $2 \cdot \sigma_A = 775 - 70 \cdot \log N$; $r^2 = 0,29$
- Rails with Typical Head Checks
  $2 \cdot \sigma_A = 821 - 111 \cdot \log N$; $r^2 = 0,96$
- Rails with Severe Head Checks
  $2 \cdot \sigma_A = 1005 - 138 \cdot \log N$; $r^2 = 0,94$

It is evident that the coefficient of correlation for the new rails (r = 0,54) is very poor thus demonstrating a great scattering of the test results, while for the rails with head checks a very good correlation is available. The reason for this fact may be on the one hand the polishing effect of wheel passages which leads to a reduction of the usual rail head surface roughness and on the other hand leads to a reduction of residual longitudinal tensile stresses in the rail head of new, roller straightened rail as a result of rolling contact and a reduction in the scatter of these stresses [2].

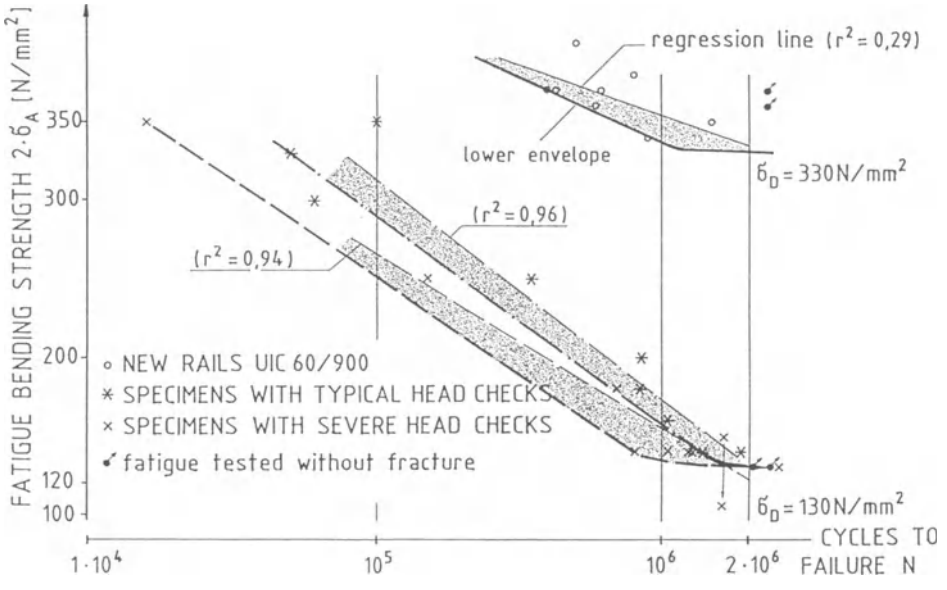


Fig. 4: Rail head bending fatigue results of the 3 test series. All rail samples are UIC 60 profile and grade 900 A.

In accordance with DIN 50 100 a lower envelope of the experimental points is taken as the Wöhler curve. The lower envelopes, which are also shown in Fig. 4, provide a greater statistical confidence than the regression lines. Taking the lower bound data the fatigue strength at 2.10^6 cycles for the new rails is $\sigma_D = 330$ N/mm² (which is the same value as determined for the foot of new rails, see [1]) while for the two series with head checks a much lower but identical value of $\sigma_D = 130$ N/mm² is determined.

All fractures happened where the maximum bending moment is activated, except one rail (marked with † in Fig. 4) which fractured in a distance of about 22 cm outside from a roller. Due to the reduced bending moment in that area the resulting fatigue strength would be only σ_D = 105 N/mm². Obviously this rail was affected by an extraordinary deep and severe head check crack at that spot, which is not typical for the test series and should be judged separately.

Mean Stress Effects

The results of the fatigue tests, together with the tensile strength and yield stress of the material (see below), can be used to construct a Smith's diagram (Fig. 5). This diagram provides a relationship between stress range (which corresponds to traffic load) and mean stress (which in this case corresponds to thermally induced stresses in continuously welded rail). It will be seen from the Smith diagram that as the mean stress increases (which would occur as the temperature of continuously welded rail falls) the cyclic, fatigue stress and consequently the permissable traffic load stress decreases. The diagram is limited by the yield stress $\sigma_{0,2\%}$ and the ultimate strength. These values were determined on tensile test specimens taken from the rail head in accordance with the UIC 860-0 specification.

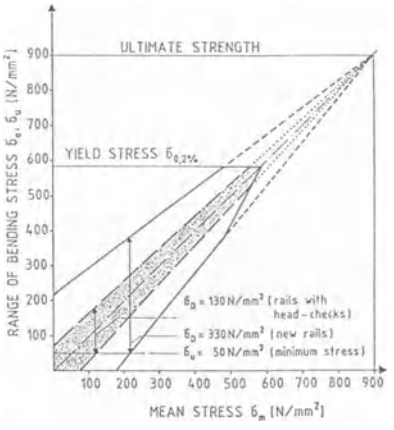


Fig. 5: Smith's diagram for fatigue strength at rail head

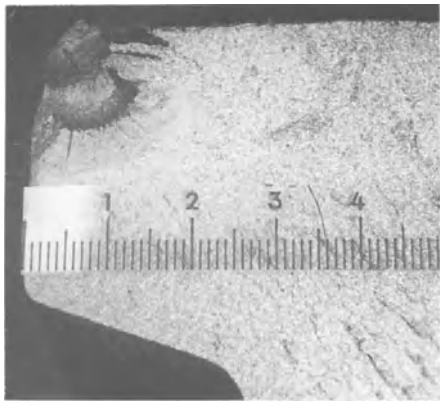


Fig 6: Transverse fracture surface of a rail with severe head checks, failed in the bending fatigue test

The following minimum values were obtained:

tensile strength : R_m = 927 N/mm²
yield stress : $\sigma_{0,2\%}$ = 580 N/mm²
elongation : A_5 = 11,6 %

These properties meet the requirements of UIC 860-0 for grade 900 A rails. No significant differences were found between the rails of the 3 series.

Crack Growth Characteristics

Typical examples of transverse fracture surfaces of the fatigue tested rails are shown in Fig. 6 and 7. There is a clear contour line between the fatigue fracture surface (fine) and the final brittle fracture surface. The fatigue propagation process starts from the root of a head check crack.

To determine the depth and course of these cracks more accurately longitudinal cuts parallel to rail axis in distances of about 3 mm were performed on a 30 cm long rail section with severe head checks, abutting a transverse fracture. It is clearly shown, that the head check cracks, which can be seen on the gauge corner running surface at an angle of about 45° against traffic direction, propagate inside the rail head at a shallow angle of about 10° to 15° in the direction of traffic. At a depth of 3 to 5 mm below the rail head surface the cracks with a length of about 35 mm partly turn over to a course parallel to rail head surface. Due to the spacing of head check cracks on the rail head surface of about 2 to 10 mm (Fig. 1) the cracks overlap below running surface (Fig. 8) and, as a result of high wheel contact pressures and increasing accumulation of service loads, steel particles finally spall from the surface (Fig. 2).

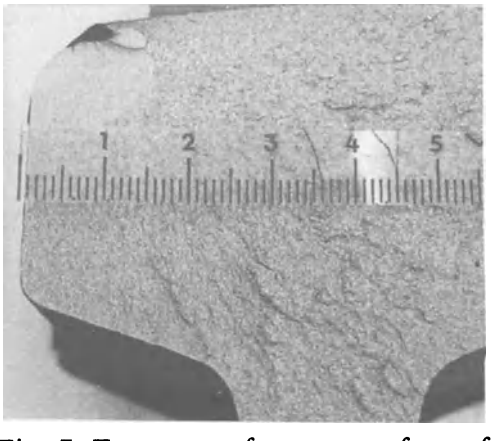


Fig. 7: Transverse fracture surface of a rail with typical head checks, failed in the bending fatigue test

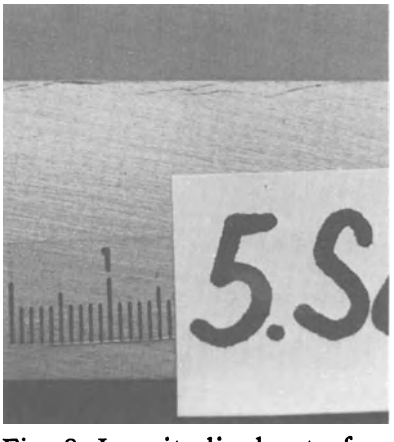
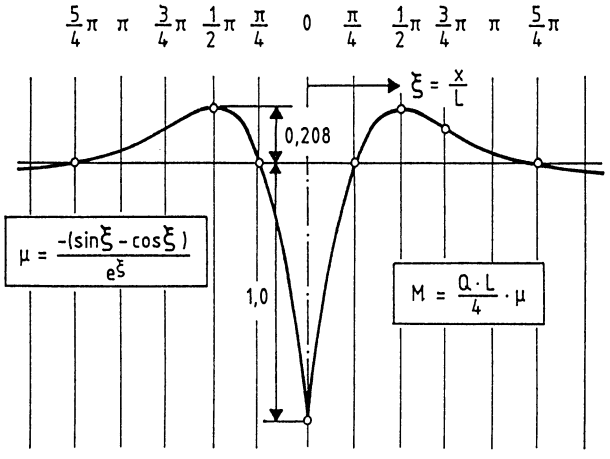


Fig. 8: Longitudinal cut of a rail section with severe head checks: course of check cracks below running surface

Further rail sections were taken from the outer support locations, where bending moments were negligible during the fatigue test, and these showed similar crack formation and crack depths in the rail head interior. Up to four overlapping cracks were found. Therefore it may be concluded that the existing cracks do not propagate to greater depth during the pulsating loading tests. When assessing the remaining service life of head checked rails it is necessary to consider both the considerable reduced rail head bending fatigue strength and the irregularities in the running surface due to spalling effects.

CALCULATION OF RAIL HEAD BENDING STRESSES

To assess the fatigue behaviour of rails in track it is necessary to compare the rail head bending fatigue strength with the tensile stresses created by traffic tracks. The calculation of the bending tensile stresses due to traffic load was performed according to theory of Zimmermann [2] (infinite longitudinal beam on Winkler foundation) as shown in Fig. 9. The negative bending moment M comes up to 21 % of the positive bending moment in a distance of $1/2 \cdot \pi \cdot L$ before and behind a single axle, where L is the fundamental value according to Zimmermann. The respective parameters for the calculation of L are outlined in Fig. 9, too. It should be mentioned that for the moment of inertia and the moment of resistance (related to the rail head) the values were reduced by 10 % compared with a new rail section UIC 60 to approximately correspond to the worn profile of rails with head checks. The maximum negative moment doubles if two axles with a spacing of $\pi \cdot L$ are acting!



Fundamental Value $L = \sqrt[4]{\frac{4EI}{b \cdot C}}$

NEW RAIL UIC 60

Modulus of Elasticity $E = 2{,}1 \cdot 10^5$ [N/mm²]

Moment of Inertia $I = 3055 \cdot 10^4$ [mm⁴]

Moment of Resistance $W_H = 335{,}5 \cdot 10^3$ [mm³] (rail head)

Modulus of Reaction C = 0,05 to 0,30 [N/mm³]

C [N/mm³]	L [mm]
0,05	1030
0,10	870
0,20	709
0,30	640

Effective Width $b = \frac{F}{2 \cdot a} = \frac{5500 \cdot 10^2}{2 \cdot 600}$

where F = tie contact area
a = tie spacing

Fig. 9: Influence curve for the bending moment M; track calculation according to Zimmermann

The axle load was assumed to be 200 kN (wheel load Q = 100 kN) which is typical for freight cars of DB where in general single axles are installed. To account for the centrifugal and dynamic effects vertical loads were increased by factors of 1,2 and 1,5, respectively. The dynamic factor of 1,5 is representative for a good maintained track with only small misalignments. Higher axle loads of up to 225 kN can be neglected as they mostly occur only in bogies with a favourable axle spacing which leads to a reduction of the bending moment.

Hence at the rail head, directly below the axle bending compression stresses of

$$\sigma_{AXLE} = -1{,}2 \cdot \frac{Q \cdot L}{4 \cdot W_H}$$

and at a distance of $1/2 \cdot \pi \cdot L$ from the axle, in the uplift wave, bending tensile stresses of

$$\sigma_{UPLIFT} = 0{,}208 \cdot \sigma_{AXLE}$$

are created.

With an axle spacing of $\pi \cdot L$ (which is, depending on the modulus of reaction, 2,0 m up to 3,1 m) the bending tensile stresses between the two axles would be doubled. However, this axle spacing only occurs on the DB network with locomotives and some special wagons, e.g. ore wagons (axle spacing between two bogies = 2,60 m and between two wagons = 3,65 m). Therefore, for the DB network this loading case should be considered mainly for special tracks where ore trains are in service.

For the consideration of dynamic effects the following is important: the probability that the dynamic peak value occurs at some locations such as to enhance both the peak compressive and tensile stresses in the cycle is very low. Hence for the determination of the maximum bending stress amplitude (Fig. 10) the dynamic factor of 1,5 is used only for the calculation of σ_{AXLE}, but not for σ_{UPLIFT}. The result is listed in Table 1.

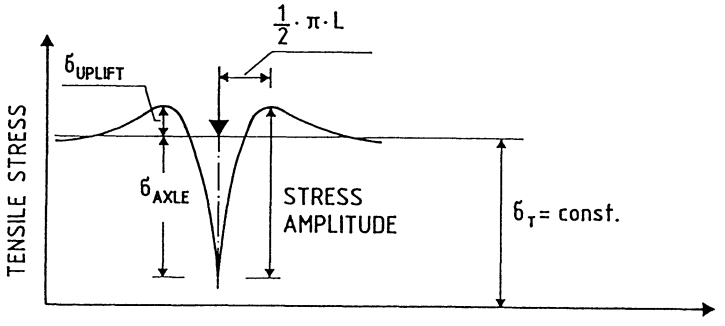


Fig. 10: Course of tensile stresses at a point of rail head under a passing axle, when constant tensile stresses σ_T due to temperature decrease are available

		EXISTING TRACKS		NEW HIGH-SPEED TRACKS	
	MODULUS OF REACTION C [N/mm³]	0,05	0,10	0,20	0,30
QUASI-STATIC BENDING STRESS	σ_{AXLE} [N/mm²]	-99,5	-83,7	-70,4	-63,6
	σ_{UPLIFT} [N/mm²]	20,7	17,4	14,7	13,2
MAX. BENDING STRESS RANGE	SINGLE AXLE $(-1{,}5\cdot\sigma_{AXLE} + \sigma_{UPLIFT})$ [N/mm²]	170	143	120	109
	AXLE SPACING = $\pi\cdot L$ $(-1{,}5\cdot\sigma_{AXLE} + 2\cdot\sigma_{UPLIFT})$ [N/mm²]	191	160	135	122

Table 1: Calculated maximum bending stresses at the rail head as a function of modulus of reaction (wheel load = 1,2 . 100 kN; dynamic factor = 1,5)

ASSESSMENT OF THE IN-SERVICE FATIGUE BEHAVIOUR

The assessment of fatigue behaviour of head checked rails involves comparing the calculated maximum bending stress amplitude with the permissible fatigue stress amplitude for a constant minimum or mean stress, resulting from the thermal loading of continuously welded rail. The winter tensile stresses in UIC 60 section rail do not exceed 100 N/mm² (stress free temperature = 18°C; minimum rail temperature = -30°C, i.e. ΔT = 48 K). This tensile stress can, as a first approximation be set as the mean stress in the Smith diagram (Fig. 5). Taking account of the unsymmetrical stress amplitude the permissible stress range turns out to be 130 - 140 N/mm².

A comparison with the calculated values of Table 1 shows that for tracks with a modulus of reaction C ≥ 0,15 N/mm³ (as expected with newly constructed tracks with modern design of track formation) no fatigue fractures due to head checks must be expected. Considering the calculation assumptions which are far on the sure side this statement also is correct for axle loads up to 225 kN (locomotives, ore wagons etc.) and for a lower modulus of reaction down to 0,10 N/mm³ (existing tracks).

It should be pointed out that the fatigue bending stress was determined for 2 . 10^6 loading cycles which is equivalent to a minimum of about 30 - 35 million gross tons during the cold season, i.e. representative of a minimum of 200 to 250 million gross tons. From Table 1 it is evident that for a modulus of reaction lower than 0,10 N/mm³ fatigue fractures must be expected. Therefore, under these conditions, and including a safety factor, an early renewal of rails with head check defects after

about 100 million gross tons should be scheduled. The urgency of withdrawal will also be indicated by the severity of running surface defects due to spalled off particels. As was evident from the fatigue tests (Fig. 4) one rail exhibited an extremely deep head check crack, resulting in a considerably reduced bending fatigue stress. Therefore isolated transverse fractures in continuously welded track cannot be excluded even before this tonnage; however, these isolated single fractures are not dangerous as they are not safety-relevant and are usually detected promptly by its effect on track circuits.

SUMMARY AND CONCLUSIONS

This investigation has shown that, provided rails are removed before extreme surface irregularities caused by head check spalling occurs, UIC 60, grade 900 A rails with head check cracks will not increase the number of rail failures in the DB network.

Special observation and - if necessary - speed restrictions should be imposed at critical track sections with head checks where increased negative bending moments may occur e.g. at bridge transitions, track overpasses or track sections with low bearing capacities where extremely poorly supported sleepers are to be expected.

For rails of less weight than that of the UIC 60 profile the bending stress range rises, e.g. by a factor 1,23 for the profile S 54. Therefore an increase of fractures due to head checks must be expected during winter if a track modulus of less than 0,10 N/mm^3 is available.

Periodic grinding of the rail head may retard or stop the initiation and/or development of head checks.

REFERENCES

[1] Masumoto, H., Sugino, K., Nishida, S., Kukihara, R. and Matsuyama, S.,
Some features and metallurgical considerations of surface defects in rail due to contact fatigue
ASTM STP 644, p. 233, 1978.

[2] Eisenmann, J.
The Rail as Support and Roadway - Theoretical Principles and Practical Examples.
in: Railroad Track, Frederick Ungar Publishing Co, New York, 1981

ROLLING CONTACT FATIGUE CRACKS ON S.N.C.F CONVENTIONNAL TRACKS

R.Y DEROCHE - J.P BETTEMBOURG (*UNIMETAL*)
B. PRASIL (*S.N.C.F*)
J.P BERTRAND - C. JUCKUM (*IRSID*)

ABSTRACT

Contact fatigue cracks, often called "SQUATS" or "BLACK SPOTS" appeared in the early eighties. For some years they have been, more or less, associated with wheel burns. Due to the nature of the fatigue cracks, SNCF became very concerned especially as they appear to be more frequent on high speed conventionnal lines (excluding T.G.V)
The results of the investigations are given. Initiation occurs close to the running surface. X Ray pole figures indicate a strongly work-hardened narrow strip. Inclusions do not seem to be involved in the fatigue process.
Proeutectoïd ferrite is involved in the first step .

1. INTRODUCTION

The rolling contact fatigue defect, the "Squat", is a growing concern on the SNCF network. It first appeared in the early eighties. For some time it was identified as wheel burns or impacts, until its common features with the English squat became obvious.

The distribution map of squats on the French Network shows that, as of now, they affect high speed conventionnal lines, (e.g. PARIS - BORDEAUX, PARIS - STRASBOURG), on flat tangent track or very large curves (in which the squat co exists with head-checking on the high leg) ; squats are also found on slopes used by low speed freight trains.
As for any Railway Company with passengers responsibility, the squat is a very serious problem since, if not detected on time, it degenerates into transverse fracture. SNCF has had to remove a quantity of rails affected by this defect, as a preventive action.
The economical consequence can be measured by the loss of rail lifetime which has been estimated at about 150 Million tons, i.e about 25 %. The most heavily affected grade is the ferrite pearlite UIC 700 MPa steel. The UIC 900 MPa grade is slightly affected.

J.J. Kalker et al. (eds.), Rail Quality and Maintenance for Modern Railway Operation , 435–448.

2. ON-TRACK AND LABORATORY INVESTIGATION OF SUBSURFACE DEFECTS.

An analysis on fully developped squats is impossible since the useful fractographic features are damaged by the hammering effect of the wheels. It was decided to isolate new defects in the initiation stage, either on or under the surface.
The first in-track inspection showed no initiation at the surface, however ultra sonic testing later revealed, cracks at a depth of about 3,0 mm under the running surface.

After opening , the internal crack has a "Coffee Grain" appearance, similar to shelling in the early stages, with a stringer of elongated silico-aluminate inclusions in the middle. (Figure 1)

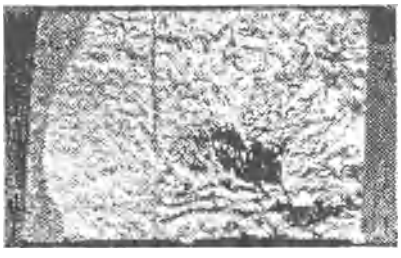


Fig. 1 "Coffee Grain" internal crack (Magnification 10)

This observation was repeated many times and the temptation was strong to consider the poor inclusions count of the Thomas Steel as the origin of the "coffee grain" crack (which is true) and the "coffee grain" itself as the origin of the squat (which is wrong). All the investigated "coffee grain" type cracks appeared to be totally confined and, none of them exceeded a maximum length of about 10 mm.

3. SURFACE INITIATED CRACKS

Further in track inspection near localised squats finally revealed cracks initiated at the surface. (Figure 2.)

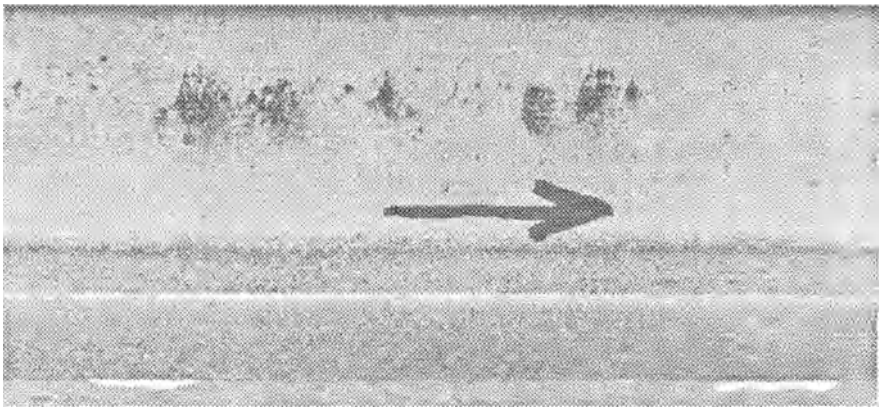


Fig. 2. Typical surface cracks (early stage squats) on the PARIS - BORDEAUX line

3.1 MORPHOLOGY OF THE SURFACE CRACK OF THE EARLY STAGE SQUAT

A typical surface crack goes down through a highly cold-worked surface layer (about 100 to 150 µm). The crack goes down, tilted in the rolling direction at an angle of about 20 to 25 degrees (Figure 3.).
In a ferrite-pearlite steel, such as the UIC 700 MPa grade, the crack follows the proeutectoïd layers of ferrite inclined by the cold working of the surface. These ferrite inclined layers are, in all the investigated samples, the preferential path of early propagation (Figure 3.)
It should be underlined that inclusions are not involved in the initiation process of crack formation.

Fig. 3. Initiation of a squat in a ferrite-pearlite steel grade UIC 700

In a pearlitic steel (Figure 4.) (one sample only) the crack goes down, tilted forwards, through a highly cold-worked surface layer 50 um only in depth, at an angle of 30 degrees ; the origin of initiation has not been identified.

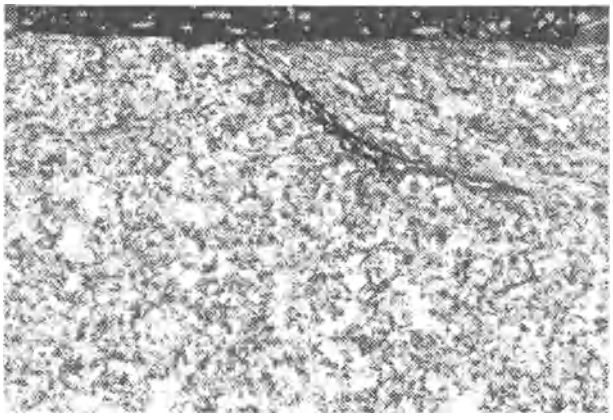


Fig 4. Initiation of a squat in a pearlitic steel grade UIC 900 B.

3.2 - WORK-HARDENING OF THE SURFACE

3.2.1. - MICROSCOPIC INVESTIGATIONS

We have focussed on three cases taken out of three typical main lines portions with squats. In every case, the running strip is narrow, 12 mm wide, and the new squats appear as inclined small cracks - 5 mm long - as shown in Figure 5.
No wear can be seen in the running strip. In some cases there is a white phase layer (about 150 µm in depth and 760 Vickers Hardness) ; in these cases, squat initiation is located at the interface, between the white phase and the steel substrate. There is no connection between the vertical cracks observed in the white phase, and the inclined cracks in the ferrite of the substrate steel. Squats exist independently of the presence of a white phase.

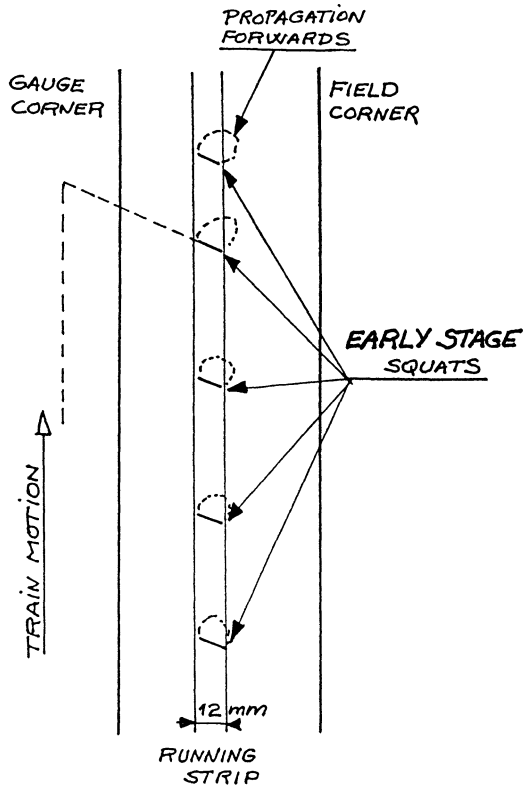


Fig. 5. Early stage squat inclination in the rolling direction ; the arrow shows the traffic direction.

In every case, the surface is strongly cold-hardened in the running strip. As shown in Figure 3, the grain boundaries and proeutectoïd ferrite are crushed and pushed backwards. The surface hardness of the running strip is about 300 to 330 Vickers, i.e 1000 MPa Tensile Strength (a new rail has a TS of about 750 MPa).

The grains at the surface, which were initially equiaxed, become elongated with an aspect ratio of 1.5 to 2 (Figure 6.). The direction of the long axis of the elongated grains is normal to the general direction of the surface crack. In Figure 6. it is possible to see some black spots at the trailing edges of the elongated grains ; SEM views, Figure 7a. show how the ferrite is drawn out of the boundaries, leaving voids, and how, Figure 7b., voids connect to each other to start an alignment which is the very beginning of a surface crack.

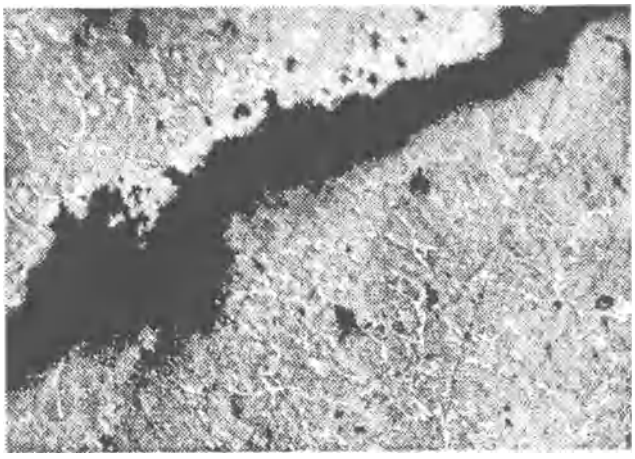


Fig 6. Elongation of the surface grain, the ferrite phase is drawn out.

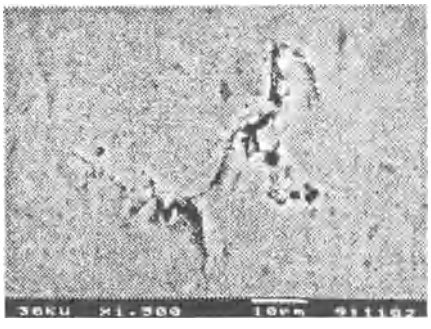


Fig 7a. The ferrite phase is drawn out.

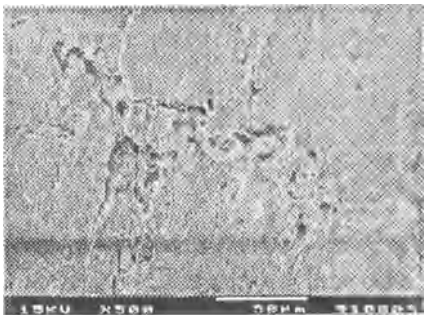


Fig. 7b Formation of an alignment

3.2.2 X- RAY DIFFRACTION INVESTIGATION.

X- Ray diffraction has been used for :

1 - X-Ray peak - broadening, to determine the width of the running strip at the surface of the rail.

2 - measurement of the principal and shear stresses on the surface, by the displacement of the diffraction peak

3 - texture analysis and pole figures in the running strip, and in the base metal, for comparison.

The peak - broadening results, on the rail head surface, show a significant increase over a strip 10 mm wide (Figure 8.). The same method applied 50 μm below the surface displays a quite different profile, especially in the center of the rail .

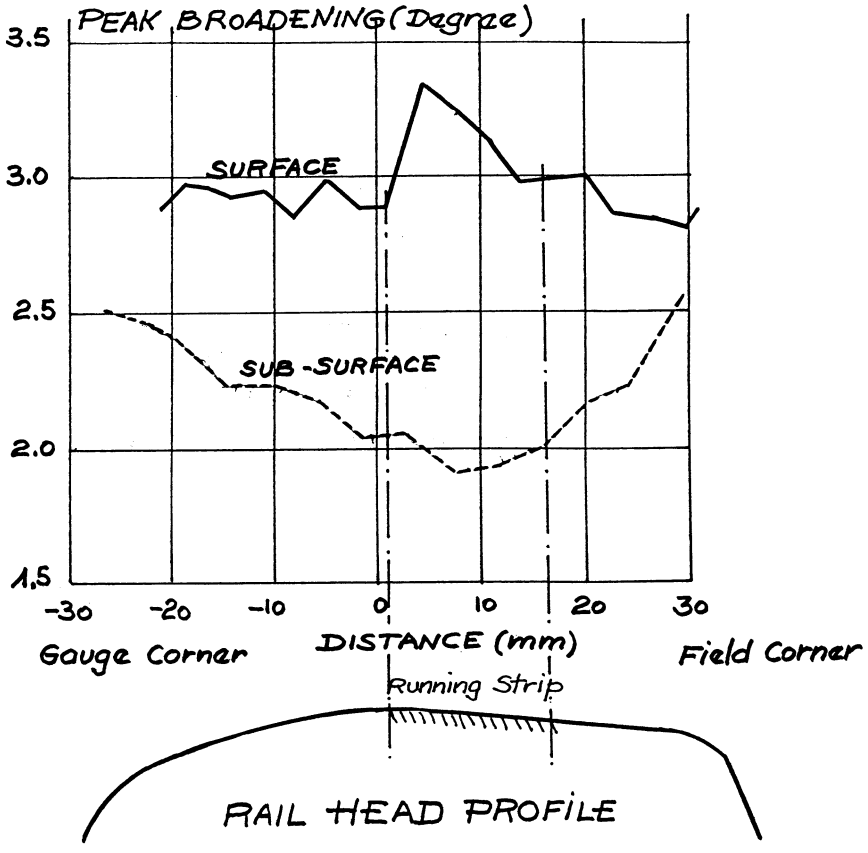


Fig. 8. Peak-Broadening (X-Ray diffraction)

Stresses at the surface of the rail head have been measured by X- Ray diffraction (Figure 9.).
High longitudinal and transverse compressive residual stresses reaching nearly 400 MPa in the running strip corresponding to the area of maximum peak broadening. Surface residual shear stresses, up to + 80 MPa on the gauge corner side and - 25 MPa on the field side have been also measured ; despite their low level, this may indicate that plastic flow occurs in one direction on the gauge corner side and in the opposite direction on the field corner side.

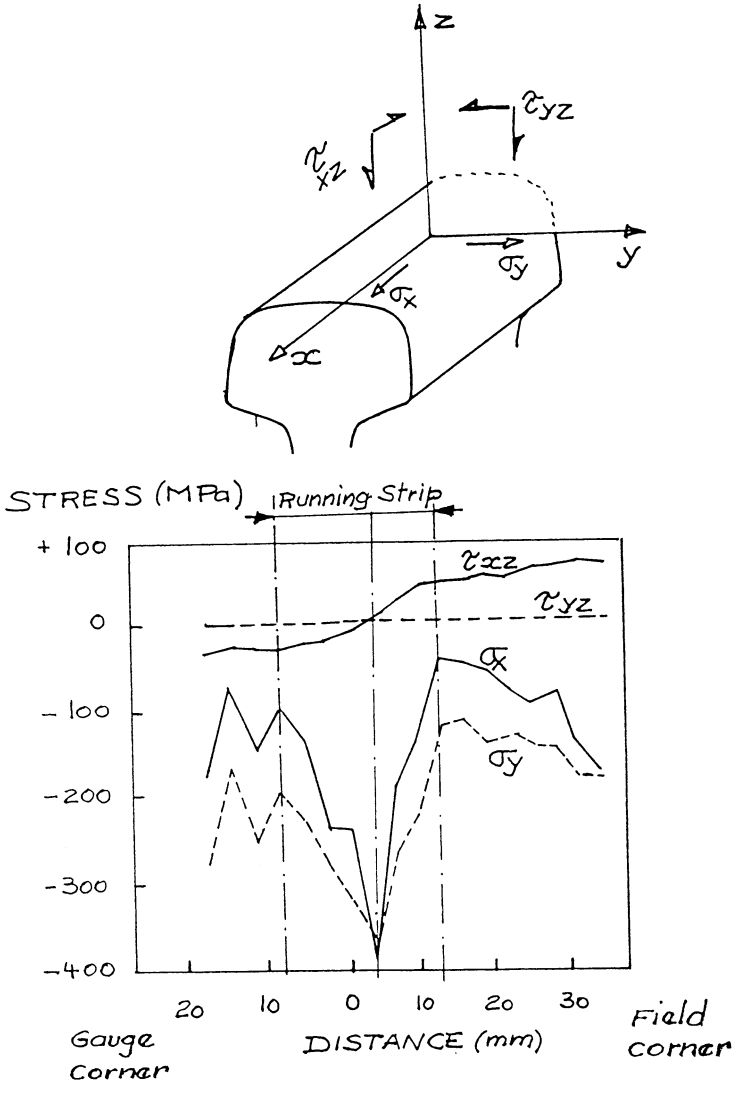


Fig. 9. Surface residual stresses induced by rolling.

X-Ray texture analysis has been carried out on samples taken from the running surface and other locations (surface, - 5 mm, - 15 mm) as indicated in Figure 10. Only one sample, taken from the running strip, displays a very pronounced texture ; no texture effect can be seen on any of the other samples.

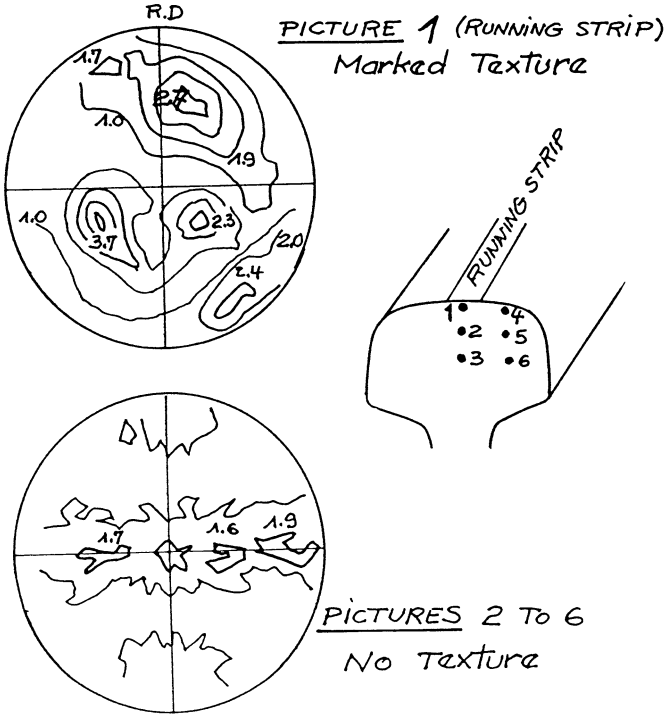


Fig. 10 . X-Ray texture analysis

The results so far indicate that :
- there is a very good agreement between the qualitative microscope observations (optical or SEM) and the quantitative measurements by X-Ray diffraction techniques.
- the surface is work-hardened to the maximum of the material's capacity (hardness 300 to 350 Vickers, and compressive residual stresses around 380 MPa), which means that the shake-down of the material has totally absorbed the ductility of the surface material and that there is no more ductility available for further strains.

4. SQUAT FORMATION PREDICTION.

To programme rail grinding operations when squats are in the early stage of formation (less than 100 µm), SNCF requires some kind of a prediction mechanism ; grinding when the depth of the crack exceeds 0.1 mm requires much more time.

It can be seen in Figure 11, that, near to a developping squat (0.2 to 0.3 mm long), six or seven small surface cracks (coming from ferrite voîds at the trailing edge of the grains) are in alignement parallel to the squat and are almost connected to each other : this is the very beginning of the squat process. Confirmation is given in the SEM image shown in Figure 12. : the ferrite phase as been removed, and the voids are connected in a micro-crack 320 µ m in length. Polishing with emery paper removes these superficial cracks.

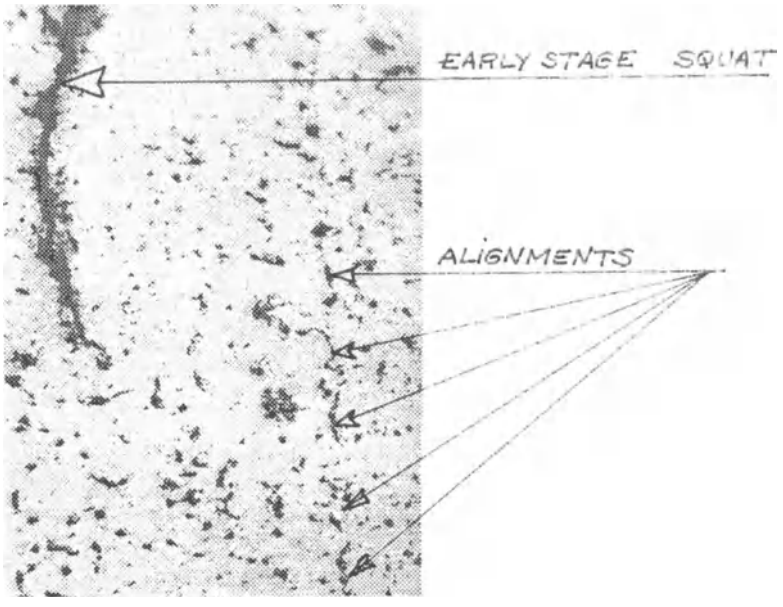


Fig. 11. Alignement of surface micro-cracks Optical microscope.

IRSID (the Central Research Laboratory of the French Steel Industry) is testing a replica technique in track in order to pick up these characteristic images at the early stage of squat initiation. The method is very simple, and the first results are extremely encouraging.

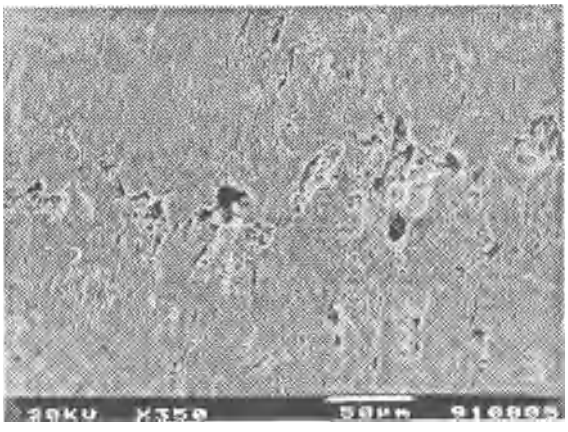


Fig. 12. Alignement of surface ferrite voids – SEM

5. QUALITATIVE DESCRIPTION OF PROPAGATION.

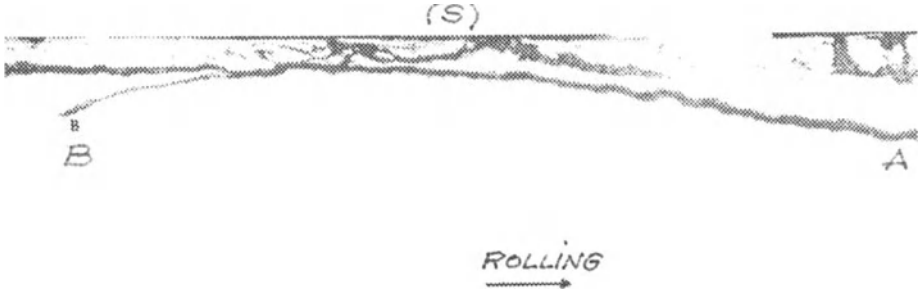


Fig. 13. Typical squat propagation

The surface (S) may be with or without a white phase layer.
Below the white phase there is a strongly work–hardened layer. The squat propagates along a crack network made of the main crack A propagating at 15 degrees into the rail and in the traffic direction, and a secondary crack B propagating in the opposite direction.

The crack propagates in the proeutectoid ferrite layers close to the surface, and continues either in the proeutectoid ferrite, in the pearlite proeutectoid ferrite interface or in the pearlite. In the pearlite (Figure 14.) the crack may propagate normally to the pearlite lamellae by breaking the cementite lamellae, or propagate parallel to the cementite lamellae, shearing the ferrite of the pearlite (Figure 15.).

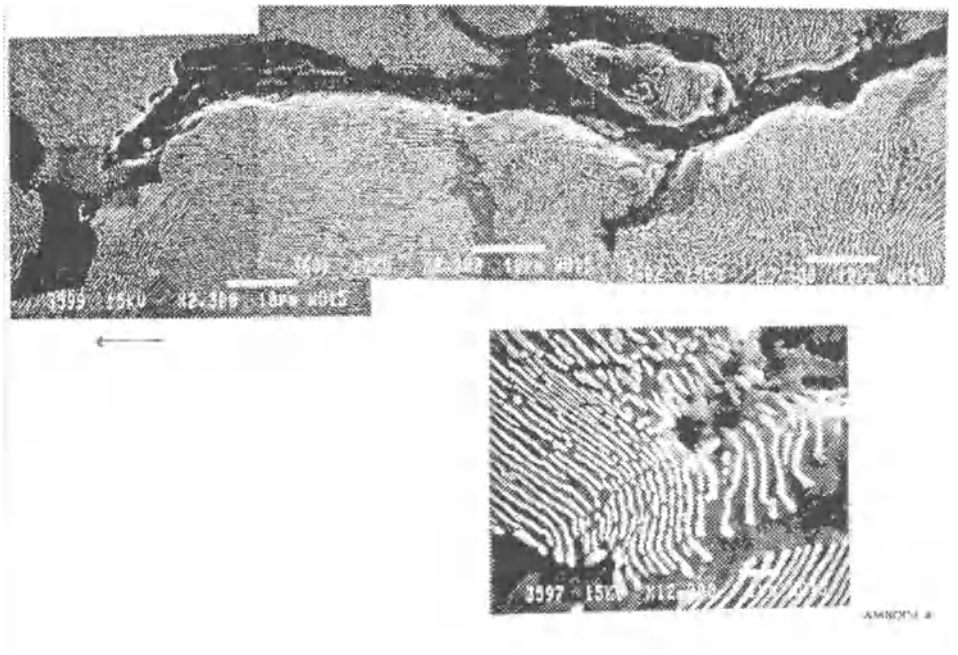


Fig.14. Crack trip propagating normally to pearlite lamellae

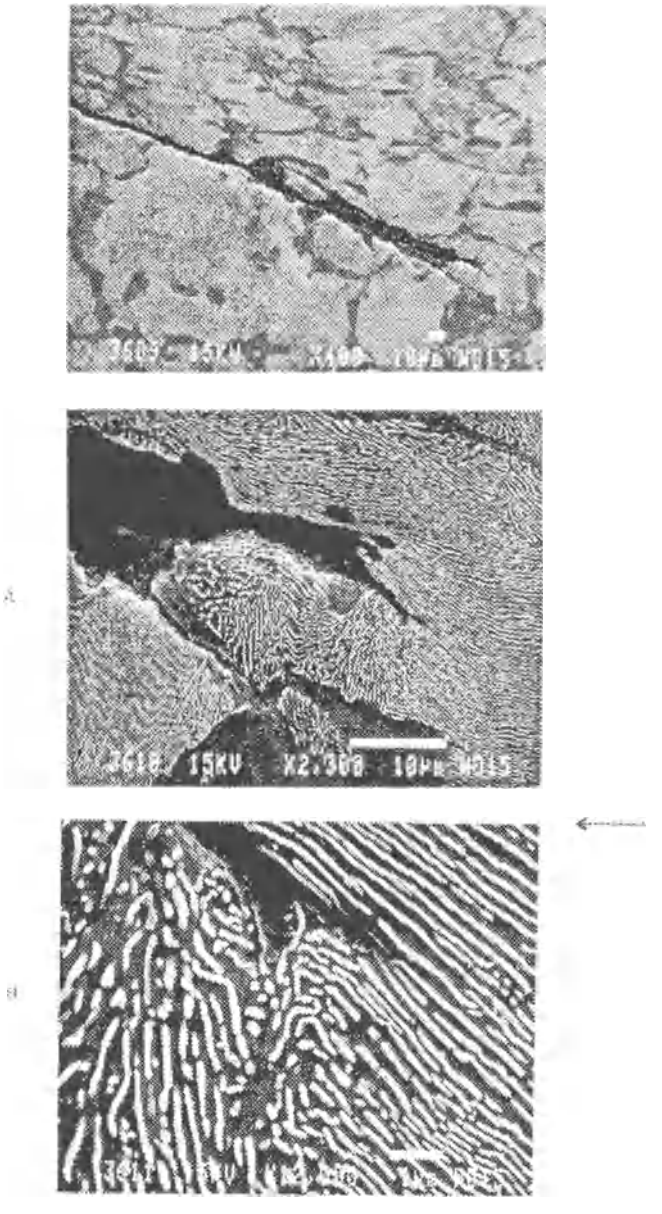


Fig.15. Crack tip propagating parallel to pearlite lamellae

6. CONCLUSIONS :

6.1 Early stage squats, observed on the French network, are aligned in a narrow rolling strip (12 to 15 mm).
Progagation occurs into the rail, and in the rolling direction (ahead of the wheels).

6.2 In ferrite–pearlite steels the cracks initiated at the surface occur in the heavily deformed proeutectoïd ferrite layers. In pearlite steels, no single source of initiation has, as yet, been identified. If a white phase layer is present, cracks in this layer may be the cause of squat initiation, but this is not always so.

6.3 The top layer (50 to 150 µm) of the surface on the running strip is work hardened up to the maximum capacity of the metal. The grains are heavily deformed and pushed backwards with respect to the rolling direction, as a result of high vertical and tangential stresses.
Squats may be found on high speed conventionnal lines where the aerodynamic resistance requires a high level of traction in the contact area (function of the square of speed), but they are also found in tangent track, where slopes require high traction levels to pull heavy trains.

6.4 The early stage of squat initiation is associated with grain boundary void alignment. A replica technique may be feasable to detect initial surface crack formation, and could be of use industrially as part of a preventive approach to removing squats by superficial grinding.

6.5 The squat formation mechanism seems to be based on a cumulative strain process, where shear stresses (Hertzian contact pressure) combine with high tangential stresses (traction) to locally exceed the ultimate shear strength of the steel.
If a steel only slightly work hardens, to a resistance level below the applied stress, squat formation will occur when plastic deformation is complete ; this is the case for the UIC 700 grade. If a steel presents strong work hardening, such as the UIC 900 grade, the shear strength of the steel increases up to a level corresponding to the maximum applied shear stress (from the locomotive wheels), before the plastic deformation limit is reached ; the probability of squat formation is consequently very low.

LIST OF PARTICIPANTS

FRA/ERRI International Conference on Rail Quality and Maintenance for Modern Railway Operation
June, 24-26, 1992

Mr. T.I. Adelegan
Nigerian Railway Corp.
P.O. Box 18337
Ibadan
Nigeria

Mr. D.A. Adelola
Nigerian Railway Corp.
P.O. Box 18337
IBADAN
Nigeria

Mr. D.R. Ahlbeck
Battelle
505 King Avenue
43085 Columbus
USA

Mr. S. Ahler
Danish State Railwat, DSB
Forbindelsesvej 3
Kobenhavn 2100
Denmark

Mr. A.A. Amiens
Nigerian Railway Corp.
P.O. Box 18337
IBADAN
Nigeria

Mr. R. Bartodziej
Huta Katourice SA
Dabrowa Gornicza
Poland

Mr. T. Benedek
Technical University of Budapest
H-1521 Budapest
Hungary

Mr. J.P. Bertrand
IRDIS
Voie Romaine BP 320
Mazieres-les-Metz
57214
France

Mr. J.P. Bettembourg
Unimetal Research
Direction Technique
57360 Amneville
France

Mr. J. Blodgett
Sperry Rail Service
46 Shelter Rock Road
06810 Danbury, CT
USA

P.E. Bold
Sheffield University/Mech. Eng.,P.O. Box,
Mappin St.
S1 4DU Sheffield
United Kingdom

Mr. T.J. Bos
N.V. Nederlandse Spoorwegen
Concordiastraat 67
3551 EM Utrecht
The Netherlands

Mr. A.F. Bower
Brown University
Division of Engineering
Providence R.I.
USA

Mr. R.Campbell
Loram Rail Ltd.
15 Redhills View
Glasgow G65 7LB
United Kingdom

Mr. D.F. Cannon
British Railways Board, RTC
Room 308B
Derby DE2 8YB
United Kingdom

Mr. F. Cantera
Viesa
P. Castellana 166
E-3-3-B
28046 Madrid
Spain

Mr. B.J. Cho
Kangwon Ind. Ltd.
#6, 2-KA,
Shinmoon-Ro,Chonro-Ku
Seoul 110-062
People Rep. of Korea

Mr. A. Chudzikiewicz
Warsaw University of Technology
Koszykowa 75
00-662 Warsaw
Poland

Mr. P. Clayton
Oregon Graduate Institute
19600 N.W.
Beaverton, Oregon
97006-1999, USA

Mr. H.Clement
R.A.T.P.
18 Rue St. Gothard
75 014 Paris
France

Mr.R. Conti
S. Gestion Qualite
MMRA BP 24
L-4801 Rodange
Luxembourg

Mr.J. Cooper
Speno International S.A.
22, Parc
Chateau-Banquet
1211 Geneva
Switzerland

Dr. D.L. Cope
British Rail - Inter City
RH 220 Euston House
24 Eversholt Street
NW1 1DZ London
United Kingdom

Mr. G.H. Cope
Mott MacDonald Consultants Ltd.
6, Beech Hurst Close
RH16 4AE Haywards Heath
United Kingdom

Mr. R.Y. Deroche
Unimetal
B.P. 140
57360 Amneville
France

Mr. R.A. Diest, van
Kloos Railway Systems B.V.
Postbus 3
2960 AA Kinderdijk
The Netherlands

Mr. H. Dol
N.V. Nederlandse Spoorwegen
Moreelsepark 1
3500 HA Utrecht
The Netherlands

Mr. M.C. Dubourg
INSA. - Batiment 113
20, Av. A. Einstein
69621 Villeurbanne cedex
France

Mr. C. Esveld
ECS
P.O. Box 331
5300 AH Zaltbommel
The Netherlands

Mr. E.S. Feilman
Loram Rail Ltd.
42 Rue de Fourqueux
78100 St. Germain-en-Laye
France

Mr. G. Fleischer
TU Magdeburg, Inst. fur Konstr.
Universitatsplatz 2,
PF 4120
0-3010 Magdeburg
Germany

Mr. B. Fournel
IRSID
34, rue de la Croix de Fer
78105 Saint Germain en Laye,
France

Mr. C.O. Frederick
British Railways Board, Brunel House,
Room 513,London Road, Derby DE2 8YB
United Kingdom

Mr. H. Fujimori
East Japan Railway Co., 2-10-1
Yurakucho,
Chiyoda-Ku,
100 - Tokyo, Japan

Mr. S.L. Grassie
Imperial College
Exhibition Road
SW7 2BX London
United Kingdom

Mr. H.D. Grohmann
Deutsches Reichsbahn
Am Sudtor
0-1802,Brandenburg-Kirchmoser
Germany

Mr. J. Gulyas
University of Miskolc
Miskolc-Egyetemvaros
Miskolc H-3515
Hungary

Mr. N.Harles
CFL
BP 1803
L-1018 Luxembourg
Luxembourg

Mr. T. Havady
Technical University Budapest
Mvegyetem RKP 3,
K.MF 22, H-1111
Budapest
Hungary

Mr. K.C. Ho
Kowloon-Canton
Railway Corp.
10/F, KCR House,
Shatin,Hong Kong,
New Territories
Hong Kong

Mr. H.J. Kaiser
Thyssen Stahl AG
Techn. Aus-u.
Weiterbildung
4100 Duisburg 11
Germany

Mr. Y.G.Kim
KHRC, 541, 5ga,
Namdaemonno
Jung-gu
Daewoo Building
Seoul
People Rep. of Korea

Mr. W.H. Hodgson
British Steel Track
Products, Moss Bay
CA14 5AE
Workington, Cumbria
United Kingdom

Mr. J.J. Kalker
Delft University of
Technology
Molenlaan 20
3055 EL Rotterdam
The Netherlands

Mr. C.H. Kim
Kangwon Ind. Ltd.
#6,2-KA,Shinmoon-Ro
Chongro-Ku
Seoul 110-062
People Rep. of Korea

Mrs. H. Ilias
Dipl.Ing
Marchstr. 12
1000 Berlin 10
Germany

Mr. J. Kalousek
National Res. Council
of Canada
3650 Wesbrook Mall
Vancouver V6S 2L2
Canada

Mr. J.Kisilowski
Warsaw University of
Technology
Koszykowa 75
00-662 Warsaw
Poland

Mr. T. Jackson
Irish Rail
50/60 North Wall
Dublin 2
Ireland

Mr. K.D. Kang
KHRC, 541, 5ga,
Namdaemunno
Jung-gu, Daewoo
Center Building,
Seoul
People Rep. of Korea

Mr. L. Kopsa
Czechoslovak
Railways
Nabrezi L. Svobody
12
110 15 Praha 1
Czechoslovakia

Mr. F. Jaeger
CFL
BP 1803
L-1018 Luxembourg
Luxembourg

Mr. A.X. Kappers
N.V. Nederlandse
Spoorwegen
Bitterschoten 14
3831 PC Leusden
The Netherlands

Mr. A. Korlaar
Nederlandse
Spoorwegen
Postbus 2025
3500 HA Utrecht
The Netherlands

Mr. A. Joeller
Voest-Alpine Stahl
Donawitz GmBH
Peter Tunnerstrasse
15
A-8700 Leoben
Austria

Mr. W. Karmowski
Cracow University of
Techn.
24 Warsaw Street
31-115 Cracow
Poland

Mr. Le Bihan
SNCF Tresorerie
21 Boulevard
Haussmann
F-75009 Paris
France

Mrs. C. Juckum
IRSID
34, rue de la Croix de
Fer
78105 Saint Germain
en Laye
France

Mr. L.M. Keer
Northwestern
University, CE
2145 N. Sheridon
Road, 60200 Evanston
Illinois
USA

Mr. D. Leykauf
Munich Technical
University
Baumbachstrasse 7
8000 Munich
Germany

Mr. A.L.P.Lopez
Lemos
CP-Est. de Sta
Apolonia
100 Lisboa
Portugal

Mr. R. Musil
CVUT Praha Fakulta
Strojni
Technicka 4
166 07 Praha 6
Czechoslovakia

Mr. A.B. Perlman
Tufts University
Dept. of Mech.
Engineering
02155 Medford, MA
USA

Mr. R.Lunden
Chalmers University
of Techn.
41296 Gothenburg
Sweden

Mrs. T. Nescheva
Bulgarien
kv. Sl. Atina -
VNVTU
Sofia
Bulgaria

Mr. J. Piotrowski
Warsaw University of
Technology
Sardynska 6 m 37
02-761 Warszawa
Poland

Mr. P.Lundvang
Danish State Railway,
DSB
Banegardspladsen 1F
8000 Arhus
Denmark

Mr. N. O'Donoghue
Irish Rail
50/60 North Wall
Dublin 2
Ireland

Mr.H.J.P. Pradier
ERRI
Oudenoord 500
3513 EX Utrecht
The Netherlands

Mr. J. Magiera
Cracow University of
Techn.
24 Warsaw Street
31-155 Cracow
Poland

Mr. M. Olzak
Warsaw University of
Technology
Nowowiejska 22/24
00-665 Warsaw
Poland

Mr.B. Prasil
SNCF Div. des Rails
17 Rue d'Amsterdam
75008 Paris
France

Mr. T. Martinek
CKD Lokomotivka
a.s.
Ceskomoravska 205
19002 Praha 9
Czechoslovakia

Mr. J.Orkisz
Cracow University of
Techn.
24 Warsaw Street
31-155 Cracow
Poland

Mr.L.F. Reid
Fatigue Technology
Ltd.
100 Andover Park W
Seattle, WA USA
98118
USA

Mr. M. Mirchev
Bulagarien
kv. Sl. Atina -
VNVTU
Sofia
Bulgaria

Mr. O. Orringer
Volpe Nat. Transp.
Center
Kendall Square
02142 Cambridge - MA
USA

Mr. J.H. Root
AECL Research
Chalk River
KO5 150
Canada

Mr. B. Mogard
Bankervet
S-781 55 Borlange
Sweden

Mr. M. Parente
CP Est. de
Sta-Apolonia
1100 Lisboa
Portugal

Mr.Y. Sato
East Japan Railway
24-26 Rue de la
Pepiniere
75008 Paris
France

Mr. J. Sauvage
Inrets
BP 34
Arcueil Cedex 94114
France

Mr. M. Siwek
Cracow University of
Techn.
24 Warsaw Street
31-115 Cracow
Poland

Mr. J. Stupnicki
Warsaw University of
Technology
Nowowiejska 22/24
Warsaw 00-665
Poland

Mr.J.C. Schaffner
Speno International
S.A., 22 Parc
Chateau-Banquet
1211 Geneva
Switzerland

Mr. T.B.B. Sorensen
Danish Institute
Railway, DSB
Forbindelsesvej 3
2100 Kobenhvn
Denmark

Mr. S. Suzuki
East Japan Railway
Co.,2-10-1
Yurakucho,Chiyoda-
Ku, 100 - Tokyo
Japan

Mr. W. Schoech
Speno International
S.A., 22 Parc
Chateau-Banquet
1211 Geneva
Switzerland

Mr. P.H. Sorlie
NSB Baneregion Sor
Stromso Torg 1
Drammen N-3006
Norway

Mr. R.P. Sweet
WS Atkins
Consultants Ltd.
Woodcore Grove,
Ashley Road
KI18 5BW Epsom,
Surrey, UK

Mr. R.E. Schramm
National Inst. of
Stand.& Techn.
325 Broadway
Division 853
Boulder Colorado
80303, USA

Mr. B.E. St. John
Permatrack Systems
Inc.
2668 Bellevue Avenue
V7V IE4 West
Vancouver, B.C.
Canada

Mr. Z. Swiderski
Railway Research
Institute, PKP
Chtopickiegostr. 50
04-275 Warszawa
Poland

Mr. H. Sehitoglu
Department of
Mechanical Enginee
University of Illinois
Urbana, Illinois 61801
USA

Mr. R.K. Steele
AAR Chicago
Technical Center
3140 South Federal
Street, Chicago
IL 60616, USA

Mr. J. Szelazek
Inst. of Fund.
Techn. Reserach
Swietokrzyska 21
Warsaw 00-049
Poland

Mr. S.B. Seo
KHRC
541,5ga,
Namdaemunno
Jung-gu Taewoo
Building, Seoul
People Rep. of Korea

Mr. D.H.Stone
Ass. of American
Railroads
3140 S. Federal Street
60616 Chicago, IL
USA

Mr. V.R. Terrill
Rail and Track
Consultant
32 Old Essex Road
01944-1205
Manchester, MA.
USA

Mr. Z.Shen
Chinese Academy of
Sciences
Southwest Jiatong U.
610031 Chengdu,
Sichuan
People's Rep. of
China

Mr. B. Stormink
N.V. Nederlandse
Spoorwegen
Moreelsepark 1
3500 HA Utrecht
The Netherlands

Mr. A.W. Thompson
Carnegie Mellon
University
Dept. Mat. Sci.
Engeneering
15213 Pittsburgh, PA
USA

Mr.F.W.T. Velderman
N.V. Nederlandse
Spoorwegen
Moreelsepark 1
3500 HA Utrecht
The Netherlands

Mr. D. Weinhauer
TU Magdeburg, Inst.
fur Konstr.
Postfach 4120
0-3010 Magdeburg
Germany

Mr. J. de Vre
Waggonfabrik Talbot
Julicherstrasse 215
D-5100 Aachen
Germany

Mr. K. Wick
Thyssen Stahl AG
Kaiser Wilhelm Str.
100
W 100 Duisburg 11
Germany

Mr. T.A. de Vries
Van Welzenes
Spoorbouw
Kilkade 53
3316 BC Dordrecht
The Netherlands

Mr. P.K. Wiersma
Nederlandse
Spoorwegen
Concordiastraat 67
3551 EM Utrecht
The Netherlands

Mr. P.J. Webster
University of Salford
M5 4WT Salford
United Kingdom

Mr. M.J. Wynne
Fatigue Technology
Ltd.
P.O. Box 601
Sutton Coldfield
B76-8LX
United Kingdom

Mr. A.J. Weel
N.V. Nederlandse
Spoorwegen
Postbus 2025
3500 HA Utrecht
The Netherlands

Mr. P.H. Zaalberg
N.V. Nederlandse
Spoorwegen
Moreelsepark 1
3500 HA Utrecht
The Netherlands