A Practical Approach to Controlling Rolling Contact Fatigue in Railways

Eric E. Magel, Kevin J. Sawley, Peter S. Sroba and Joe Kalousek,
National Research Council of Canada

Summary: Rolling contact fatigue (RCF) is a problem for most railways, heavy or light axle load, slow or high speed. A practical strategy for controlling RCF is discussed based on a multi-pronged approach of increasing component resistance to initiation and propagation of RCF, reducing applied stresses that cause RCF, and providing a maintenance regime for both controlling RCF and managing the risk associated with rail or wheel failure due to RCF. RCF is influenced by, and can be controlled by improvements to, the wheel profiles, rail profiles, wheel and rail metallurgy, gauge-face and top-of-rail friction coefficients, track geometry standards, and rolling stock characteristics. Mitigation can include wholesale changes to one parameter or selective and more modest changes to several parameters. Since RCF is a threshold phenomena, even a small reduction in the stresses, or modest increase in the strength of the component, can dramatically reduce the number of contact cycles that promote fatigue. For any given railway, the most appropriate approach to controlling fatigue will depend on its specific operating and maintenance strategies.

Index Terms: rolling contact fatigue, friction management, rail grinding, rail and wheel metallurgy, rail and wheel profiles.

1 Introduction

Wheels and rails generally fail by wear or rolling contact fatigue (RCF). Since wear and RCF are intimately related, it is necessary to consider both in any practical discussion of fatigue.

Wear is the gradual attrition of metal through stress and slip, whereby small particles are removed from the wheel and rail surface. Wheel removal due to wear is governed by limits set on flange height, flange width, rim thickness, and, on some railways, by wheel hollowing. Allowable wear on rails is limited by fatigue and fracture considerations, and is dictated by factors such as the rail section, axle load and wheel/rail dynamics.

Rolling contact fatigue defects form in response to millions of intense wheel-rail contact cycles that repeatedly over-stress the surface or subsurface material. With surface defects, cycles of high stress and high traction cause a thin layer of the steel surface to flow in the direction of the applied load. The surface layer strain-hardens and fractures, initiating a surface crack [1]. These cracks can grow under subsequent loads, link up and lead to shelling.

1.1 RCF Defects

The progression of surface fatigue on both the rail and wheel is illustrated in Figure 1. Depending on the specific operating conditions, small surface cracks can wear away as fast as they are generated, especially in particularly dry environments.

Besides cracking and shelling other RCF defects are squats, crushed heads and deep-seated shells on rails. On wheels, the shattered rim [2] is a subsurface initiated RCF defect.
**Figure 1**: Examples of rolling contact fatigue on wheels and rails.

**Squats**, or dark spots, are top-of-rail defects that appear to initiate from patches of microscopic surface martensite. Since the formation of martensite at wheel/rail contacts requires very high rates of slip, squats are most common on track with very high tractive efforts and usually high speeds.

Very common on European [3] and Japanese [4] high speed rail systems, squats have also been found on freight track with loaded up-grade passage, evidently due to slip of locomotive wheels under high traction.

**Crushed heads** are another surface initiated fatigue defect, typically about a quarter to half meter in length. They are characterized by increased frequency and growth of large surface cracks and plastic flow to one or both sides of the
running surface. The concentration of cracks is usually caused by a local stress raiser associated with improper rail grinding, lack of rail grinding, track irregularities, dirty steel or some combination of all of these factors. Grinding can prolong the life of crushed-head-affected rail by either removing, or moving the contact band away from, the stress raiser.

Only one of the many internal defects that occur in rails can be considered a rolling contact fatigue defect. While vertical and horizontal split heads and tache ovales have sometimes been classified as RCF defects, they are more properly dealt with as metallurgical defects since contact stresses have little direct influence on their formation. In contrast, the deep-seated shell (Figure 3 A&B) is a direct result of high contact loads at the extreme gauge-corner that cause the rail to collapse along a shear or "slip" line [5]. In steels with metallurgical imperfections, the deep-seated shell can initiate a transverse defect (Figure 3C). Since rolling contact stresses are only active near the surface, the transverse defect must be propagated by bending, residual and thermal stresses.

Deep-seated shells can be minimised through the use of harder steel and rail grinding to shift load from the extreme gauge corner.
1.2 Shakedown Theory and RCF

The initiation of rolling contact fatigue is most easily discussed using shakedown theory [6]. The combination of contact stress (P), normalized surface tractions (T/N) and shear strength of the steel (K) required to generate an increment of fatigue are summarized in the Shakedown Diagram (Figure 4). Contacts below lines A and C are benign. Those to the right and above lines A and C contribute to surface fatigue, and those bordered by lines A and B contribute to subsurface fatigue. Although there is debate about the value that should be applied to K, the shakedown diagram gives clear directions for improving system health. Changes that shift a distribution of contacts towards the lower left will reduce RCF. This shift is achieved by:

A) Minimizing normal contact stress: According to Hertz's elastic contact equations, contact stress is proportional to load raised to the $1/3^{rd}$ power. Hence doubling of wheel load will only increase contact stress by about 27%. Poor transverse profiles, produced for example by poor design, poor grinding or hollow wheels, can increase stress by a factor of 3 [7].

B) Minimizing tractions: Rail/wheel shear forces develop due to small relative slip between the rails and wheels. The amount of slip (known as creep) depends on the steering and traction demands. These creep forces, or tractions, cannot exceed the product of normal force and available adhesion (m). Traction can be minimized by modifying the properties of the interfacial layer (i.e. friction control) and controlling creepage through better profiles and flexible or steering trucks.

C) Using the highest strength, field proven materials available: For any given type of steel structure, resistance to RCF increases with hardness, with the shear yield-strength developed in the work-hardened rail surface layers being most relevant. For example, bainitic steels may have higher bulk strength than pearlitic steels, but laboratory tests indicate that bainitic steels work-harden less than pearlitic steels under rolling contact conditions. Thus, pearlitic steels likely develop work-hardened layers with greater strength. This may explain why tests with bainitic and martensitic steels, have produced conflicting results with respect to fatigue resistance.

2 Controlling RCF

In theory, RCF can be completely eliminated if stresses and tractions can be reduced sufficiently and/or steel strength raised sufficiently. In practice, the dynamics of the wheel/rail system and natural changes in transverse shape mean that RCF is virtually unavoidable in an uncontrolled environment, even a low axle-load one. The only exception occurs in systems with unacceptably high rates of wear.

To minimise, manage or eradicate RCF, several practical tools available are:

- improved wheel profiles
- improved rail profiles
- improved rail grinding and wheel re-truing practices
- friction management
- improved track quality
- improved suspension bogies.
- improved rail and wheel metallurgies

The tools implemented by a railway will necessarily be specific to that system's conditions and constraints, such that the best solution for one railway will rarely apply directly to another. The most cost-effective approach in general is to make small changes to several of the key elements, exploiting synergies where possible. The final solution will be confined by cost or operating boundaries, but the engineer should challenge the status quo or historical institutional limitations.

The following sections outline the benefits associated with each of the tools available to the railway for the control and treatment of RCF. Synergies are identified with the other tools.
2.1 Improved wheel profiles

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

- Improved steering in curves to reduce wear, L/V forces, and RCF. The Cartier Railway Company reduced wheel shelling by 60% by adopting a custom wheel profile [8]. Canadian Pacific Railway increased the life of coal fleet wheels by 18% using an improved wheel shape [9].
- Control of creep forces, by matching wheel profile curving ability (conicity) to the curving requirements. A wheel profile designed for the UK railways [10], called the WRISA2, showed a considerable improvement with respect to shakedown (Figure 5).
- Improved stability and reduced hunting by providing an appropriate tread slope over the tangent running band to control effective conicity in tangent track [11].
- Reduced damage to the field side of the wheel and the low rail as a result of lower tread wear rates and by lightly rolling off the field side of the profile to minimize hollowing.

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

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

2.2 Improved rail profiles

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

Besides reducing RCF of the rail, a family of rail profiles that contact the wheel at different running bands will provide a significant benefit to the wheel. Rail profiles designed to spread wear on the wheel slows the development of a false flange or geometrical stress raiser [12]. A properly designed system of wheel/rail profiles that controls stress and wear provides for durable, stable, and optimized wheel/rail performance.

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

2.3 Controlling metal removal through rail grinding and wheel machining

Early field studies suggest that crack initiation on rails takes place over 3-6 MGT [13]. Metallurgical analysis shows that newly initiated surface cracks are a fraction of a millimetre in length and propagate at an angle of 5 to 15 degrees to the surface. These newly initiated cracks generally propagate at a relatively slow rate, then much faster at the intermediate length of about 5-10 mm and depth of about 1-3 mm. At greater length and depth, where the contact stresses are less intense, crack propagation proceeds at a slower rate [14].

At the same time as these cracks are propagating into the surface, they are being truncated by either natural wear or machining processes such as wheel re-truing, abrasive brake shoes, or rail grinding. Regular re-truing of wheels to an optimized shape has increased wheel life by more than 50% in Australia [15], South Africa and Canada [8]. Abrasive brake shoes, such as those with cast iron inserts, can be effective in scrubbing metal from the wheel tread to remove incipient fatigue cracks as well as wheel-slide damage [16].

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

A more effective approach is to use lubrication to control wear and through rail grinding frequently remove a small amount of the rail surface material to truncate existing cracks and remove very shallow damage. The optimal strategy is to remove just the right amount of metal to control both surface and subsurface crack-initiation and to remove short cracks while the rate of propagation is still slow. This metal removal rate has been called “The Magic Wear Rate” [17].

Although different rail grinding strategies have evolved over the years, it is generally agreed that the ideal approach is to grind the rail preventively at The Magic Wear Rate [7]. Preventive grinding of rails to establish an optimized system of profiles and regularly remove incipient cracks has, with improved metallurgies, contributed to a two-fold increase in system rail life and a four-fold increase in system rail fatigue life over the last 25 years [18].

2.4 Friction management

Lubrication of the rail gauge-face/wheel-flange can reduce gauge-face/flange wear by 95 to 100% [19]. Lubrication can also contribute to reductions in RCF initiation by reducing tractions. However lubrication of the gauge-face decreases the steering moment that can be developed by a wheelset and increases the leading-axle angle-of-attack in curves [20]. The consequential increase in lateral creepage and lateral force increases top of low-rail wear, RCF initiation on both high- and low-rail top surfaces and the formation of deep seated shells on the high-rail gauge-corner. Furthermore, grease or oil contamination of surface cracks increases RCF propagation. Poor equipment or improper settings on lubricators may also cause grease to migrate to the TOR, thereby compromising traction and braking. Gauge-face lubrication, by preventing wear of the gauge-corner, combined with significantly longer grinding cycles has resulted in deep gauge-corner shelling on premium rails [18]. Thus, gauge-face lubrication that matches the grease application rate to type of grease, dispensing equipment and track conditions must be used in conjunction with a preventive rail-grinding program that regularly removes metal from the rail gauge-corner.

The benefits of top-of-rail friction control have been demonstrated for noise [21], corrugation suppression [22], reduced lateral forces and reduced rail wear [23]. In the case of RCF, both field tests and laboratory studies are currently underway to measure the extent of benefits. Significant benefits are expected according to the shakedown diagram: the shakedown limit for a friction level of 0.35 is 30-70 percent greater than that for dry rail (0.45-0.6).

2.5 Track geometry

Vertical and lateral track irregularities are often associated with rapid rates of RCF development in rail adjacent to the irregularity. Although dynamic impact forces are one obvious outcome of track errors, the effect on contact stress is minor. Of greater significance are the following:

- Under dynamic loading, the rail experiences crushing and gross plastic flow that leads to rapid profile deterioration. Loss of profile affects contact conditions and increases the probability of RCF. Some minor rail discontinuities can be removed by rail grinding to control the dynamic loads.

- Track irregularities cause wheels to displace laterally relative to the rail. In tangent track, this may lead to gauge-corner contact. High contact stresses, with strong longitudinal tractions and subsequent fatigue, are a common outcome [7]. Minor irregularities can be removed through rail grinding or otherwise accommodated by grinding the rail to shapes that will withstand these contact conditions. Track geometry correction will be needed to address most wide or narrow track gauge, cross-level or alignment errors.

- Track that is under- or over-elevated for the traffic that runs over it will encounter greater loads. When traffic is running under balanced speed, the result is higher wheel loads on the low rail which can cause rail rotation and RCF. In high cant-deficiency track, the high-rail mid-gauge is subject to strong longitudinal creepages and RCF cracking [24].

- Tight gauge in tangent track promotes gauge corner contact, truck hunting and RCF. Controlling wide gauge in curves is essential for reducing damage to the low-rail from hollow wheels. Besides the high contact stress, hollow wheels on wide-gauge track produce greater dynamic rail rotation, especially on poorly restrained rails that promotes unfavourable contact geometry.

- Plate-cut sleepers or poor fasteners will allow the rail to rotate dynamically with lateral loads. This rotation usually increases gauge-corner loading on the high rail and promotes RCF.

Rail grinding can dress weld discontinuities, profile rail to control hunting and remove corrugations. Other track maintenance procedures are required to remove discontinuities at bridges, switches and crossings, track pumping locations etc.

Although the rail fatigue cycles associated with track geometry problems are mirrored on the wheel, these cycles usually represent only a tiny fraction of the total contact cycles on the wheel, and the net contribution to wheel fatigue is small, while on the rail the damage is focused and very apparent.
2.6 Improved suspension bogies

A bogie suspension that is flexible in bending can take advantage of optimized profiles that provide favourable steering forces. The flexible suspension can reduce yaw angles in curves which will reduce RCF initiation on the rail in shallow curves and on the wheels. Reduced wheel shelling on the Canadian Pacific Railway has been attributed to the use of a flexible bogie [9]. However, a more flexible truck can also respond to unfavourable steering moments and increase the yaw angle, especially in the case of bogies that have been poorly maintained and are running with worn-out components. The adverse impact of high stiffness bogies has also been noted. A number of bogies have been developed to improve upon the limited curving performance of the standard three-piece bogie but without compromising stability in tangent track [25]. Additionally, there are retrofits available to the standard truck that have demonstrated measured reductions in wheel/rail forces and tangible reductions in wheel shelling.

2.7 Rail and wheel metallurgy

The resistance to rolling contact fatigue of steel is governed by its strength in shear (K), which in turn is proportional to its work-hardened hardness. The shakedown diagram shows that by increasing the shear strength K, higher contact stress can be accommodated without fatigue. Harder rail steels with optimized rail profiles will dramatically reduce both the probability of RCF initiation and the rate of propagation.

Based on its as-manufactured hardness, 400HB hypereutectoid premium steels can withstand contact stresses about 45% greater than standard rail steel (275HB) without yielding. The top row of Figure 6 shows that the effect of the harder steel is roughly a 60% reduction in the predicted rolling contact fatigue damage on the high rail of a 4-degree curve [26]. In combination with improved profiles, the benefit can be even more dramatic. The RE136 lb/yd rail profile, the RE141 lb/yd profile and the CPR-H are plotted for four different values of rail hardness. The CPR-H is an optimized profile ground onto the Canadian Pacific Railway curves at regular cycles. The harder steels reduce the number of contacts that exceed shakedown by considerably more than the change in hardness suggests, thereby dramatically reducing the probability of crack initiation (and the rate of propagation).

Figure 6: Predicted distribution of RCF damage on three different rail sections, for four different hardness values based on 300 measured worn wheels. The numbers on each plot represent the relative areas and peaks of the RCF distribution.

Alloy wheel steels have long been promoted as an effective means of controlling wheel RCF. In-service testing of an experimental alloy steel on two Canadian unit train operations, where wheel removal is almost entirely from mechanical (not skid-flat) shelling, found that there was a large improvement in wheel life - a near doubling of average life in one case. Operating conditions, including braking demands and type of bogies, impact the effectiveness of metallurgy in reducing RCF.

2.8 Inspection and defect detection

The risk of rail failure from internal flaws remains a large concern for railways. Railways manage this risk by using non-destructive inspection to find internal flaws. Accurate and reliable testing with ultrasonic systems requires the rail surface to be clean and free of moderate to severe RCF defects.

RCF of wheels is subject to several inspection criteria in North America. A visible shell of a certain size or two adjoining shells of a certain size are cause for wheel removal under interchange service. As well, wheel impact load detectors are widely used on the Class 1 railroads to detect shelled and out-of-round wheels that cause large dynamic forces. These impact detectors report to a centralized data collection system that submits warnings to the owner starting at 60 kips (267 kN), with wheel removal for all offenders exceeding 90 kips (400 kN). Removal of shelled wheels, besides preventing the rare derailment, also prevents damage to the rail (including broken rails) and track components.
3 A practical approach

One approach to successfully controlling and treating RCF is to apply the following five-stage process:

**Investigation:** Identify those RCF defects that incur the greatest expense and risk, including the costs of rail replacement, wheel re-truing, rail grinding and inspection. Determine if failures are clustered or system-wide.

**Constraint Identification:** Identify the constraints, such as lack of cooperation or financial, manpower, capacity or time limitations.

**Evaluation of Options:** Compare the relative benefits of the treatment against its cost. Determine whether the change is small or large compared with the current conditions, assess the potential benefit in that environment, and identify the cost of the change. Often, this process will culminate in the development of a business case that asks management to allocate funds to purchase materials or hire the required resources.

**Trial and Validation:** Measure (usually through field testing) the benefit of a particular treatment and compare against a baseline. Develop a more detailed business plan.

**Implementation:** Develop specifications, identify suppliers, design and implement training programs, and establish controls to measure and manage the success of the change.

4 Conclusions

Rolling contact fatigue is a difficult but solvable problem. A structured review of the wheel and rail profiles and metallurgy, track geometry, friction management, rail grinding practices, and bogie type will reveal the most appropriate candidates for change, as well as the need for improved inspection and detection.

For each available tool, it is necessary to compare the benefits of each against its costs, where both benefits and costs will be very specific to a particular railway. On the basis of the potential costs-benefits, a business case can be constructed. If the business case is accepted, the next step is usually to perform a field test to validate and document the expected improvements. Besides establishing arrangements with suppliers for products and servicing, broader system implementation will also require the development of programs for training, quality assurance and ongoing monitoring.

An effective program to mitigate rolling contact fatigue will take a multi-pronged approach that includes installing improved steels in curves, correcting the worst track geometry problems, improving the wheel profile where possible and preventively grinding rail to a family of shapes that includes different profiles for tangent, high and low rails. Improved suspension trucks can, in some cases, provide a significant reduction in rolling contact fatigue, as well as provide other benefits such as reduced lateral forces and fuel consumption. Friction management promises to be a powerful tool for reducing RCF, but field evidence to date is limited.

Although the needs and constraints will vary (sometimes dramatically) from one railway to another, a comprehensive review of the vehicle/track system should be undertaken to identify a series of modest, practical and often affordable modifications to existing materials and maintenance processes for controlling RCF and thereby extending rail and wheel life, and reducing risk.

5 References


16. Cobra® Treadguard™ Shoe Description and Application. White paper from Wabtec Corporation, undated


