Experimental and numerical modelling of wheel-rail contact and wear

A. Rovira¹*, A. Roda¹, M.B. Marshall², H. Brunskill², R. Lewis²

¹Universidad Politecnica de Valencia, Dpto. de Ingenieria Mecanica. C. Vera s/n. E46022 Valencia. Spain
²University of Sheffield, Dep. of Mechanical Engineering. Mappin Street, S1 3JD Sheffield, UK

*Corresponding author: Tel. +34 96 387 7007 Ext. 76263, fax: +34 96 387 7629. E-mail address: anrocar@mcm.upv.es (A. Rovira).

Abstract

In the field of simulation of railroad vehicles, there are many numerical models to estimate the interaction forces between the wheel and rail. The main advantage of these models is that they can be used together with multi-body dynamics software to calculate the motion of a vehicle in real time. However, the result of these contact models is usually post-processed to estimate wear on the profiles and some hypotheses assumed by the contact models may be inadequate for wear analysis. This is the case when considering surface roughness, which is not introduced in the numerical models and makes wear prediction imprecise. In this work an experimental method based on the measurement of ultrasonic reflection is used to solve the contact problem, together with a FASTSIM (simplified theory of rolling contact) algorithm. This technique is suitable to deal with rough surfaces and gives a better approximation of the material behaviour. Wear is estimated by means of the energy dissipation approach (T•gamma). Two different contacts are investigated, using wheel and rail profiles coming from unused and worn specimens. In order to obtain realistic results, special care is taken when locating the specimens to reproduce the same contact that appears between the wheel and the rail in the track.

KEYWORDS: Wheel-rail contact; surface roughness; ultrasound; wear modelling

1 Introduction

The likelihood of railway wheel wear occurring is rising, this is as a result of new specifications imposed on railway wheels that are leading to an increase in the severity of the wheel/rail contact conditions [1].

It is important that the wheel wear process can be modelled, as this enables improvements to be made to wheel design and materials in order to keep wear at a minimum. Excessive wear needs to be avoided as this can affect the dynamic behaviour of the railway vehicles, which will reduce the ride comfort, augment the cost for maintenance and replacement of wheels, increase the vehicle-track interaction forces that may have an impact upon the potential for derailment, and reduce the integrity of the wheel material [2]. There will also be negative effects on the track, increasing maintenance cost of rails and the infrastructure. Understanding of the wear rates will also help with planning regrinding schedules.

Fig. 1. Typical Wheel Wear Modelling Methodology.

Design tools have been developed that integrate vehicle dynamics and wear modelling to predict the evolution of wheel profiles [2-7]. These tools all have the same basic structure as outlined in
Fig. 1, although different multi-body dynamics software packages are used as well as local contact analysis formulations and wear models.

All three modelling elements have their own drawbacks. Of particular concern are the contact mechanics methods used and the approach to wear modelling. Telliskivi and Olofsson [8,9] studied different contact mechanics methods used in wear simulations, and concluded that they significantly affected the accuracy. Further, the outputs from the numerical simulations of contacts do not take the roughness into consideration, which could be a significant cause of inaccuracy. In this work the limitations of this part of the wear modelling procedure are investigated by comparing numerical contact outputs with experimentally measured contact pressure distributions.

A non-destructive technique is used to explore the wheel-rail contact. As in a previous paper [10], the ultrasound technique is applied. A piece of rail and a piece of wheel are pressed together to get a contact. For a given normal force and known surface characteristics, the normal stress distribution can be calculated. This distribution together with creepage hypothesis is the input for the tangential problem, where the FASTSIM algorithm is used [11,12].

Wear is estimated through the energy dissipation in the contact patch (T-gamma approach). From the energy dissipation and depending on the wear regime [6], wear rate (mass loss) in the wheel and the rail is calculated. Finally, assuming that the wheel is rolling along the x axis, the damage on the profiles (mm per rolled metre) is obtained. Those magnitudes calculated for a real engineering surface are compared with the predicted values given by the numerical models generally used in railroad simulation.

2 Wheel-Rail Wear and Contact Modelling

The calculation of the forces arising in the wheel-rail contact, along with the estimation of wear, is divided into three steps (Fig. 2). First of all, the problem of a wheel normally loaded against a rail is solved, and the geometry of the contact patch and the associated normal pressure distribution is determined. There are a number of different ways of doing this, including experimental, analytical and numerical techniques. In this study, the normal problem will be solved in two different ways; using an experimental ultrasonic technique (Section 2.1), as well as by numerical methods (Section 2.2). The numerical approach applied here is typical of that applied by many researchers [2-7]. Measured contact profiles are loaded together and the area of interpenetration calculated. Following on from this an equivalent ellipse is fitted, and an assumed Hertzian contact determined.

![Fig. 2. Procedure followed to calculate the wheel-rail forces and wear estimation.](image)

The determined normal contact condition together with the kinematic situation under investigation is used to solve the tangential problem. In this case, the FASTSIM algorithm [11] as developed by Kalker (section 2.3) is used. Two different calculations are performed; one using the experimentally generated wheel rail contact data, with the other using the numerical results. The stress fields and the adhesion/slip areas will be compared for both calculations.
Finally, with the normal and tangential stress fields in conjunction with the kinematic information, wear of the profiles is estimated. When estimating the wear, an energy dissipation criterion will be used in the contact patches (see for example [4,6]), and is detailed in Section 6.

2.1 Ultrasound theory

Ultrasound (sound greater in frequency than 20 kHz) is widely used in the field of non-destructive testing to inspect components for the presence of cracks and other subsurface defects. More recently, the measurement of ultrasonic reflection from a rough surface contact has been used to determine contact pressure [10,13,14]. This technique has been applied to map the contact pressure distribution in a variety of machine element contacts, such as interference fits [15] and bolted joints [16], and has proven to be a successful non-intrusive technique.

The technique is based on the reflection of ultrasonic waves from a partially contacting interface (Fig. 3a). When viewed on a micro-scale real engineering surfaces consist of micro-asperities. If two surfaces are pressed together, asperities join together to form discrete junctions surrounded by air gaps. When an ultrasonic wave is incident at such an interface, provided that it is long in wavelength compared to the scale of the asperities, the interface as whole behaves as a reflector [17]. The ultrasonic signal is transmitted through the asperity contacts and reflected by the air gaps. In this way, a reflection coefficient can be defined for an interface, and is the fraction of ultrasound incident that is reflected from it [17].

![Fig. 3. Ultrasound and interfaces: (a) incident, transmitted and reflected waves. (b) Spring model representation.](image)

Kendall and Tabor [18] investigated this case, and proposed that the behaviour of the asperity contacts in response to the sound wave is analogous to that of a spring (Fig. 3b). The interface when described using the spring model has a stiffness $K$ (expressed per unit area), which is defined as the change in nominal contact pressure $p_{\text{nom}}$ required to cause unit approach $u$ of the mean lines of separation of the two surfaces. The reflection coefficient was related to the interfacial stiffness, and for two similar materials in contact reduces to:

$$|R| = \frac{1}{\sqrt{1 + (2K/\omega z)^2}}$$

where $\omega$ is the angular frequency ($= 2\pi f$) of the ultrasonic wave, and $z$ the acoustic impedance (the product of wave speed and density for the material). Drinkwater et al. [13] investigated the applicability of the spring model, and showed that the interfacial stiffness was strongly dependent on the interface load, and that the model could be applied successfully up to frequencies of 50 MHz. It was also noted that the interfacial stiffness is not only a function of the applied contact pressure, but also a function of the size, number, and distribution of the asperity contacts. Therefore, there is no single relation between stiffness and contact pressure, and a specific calibration is required for the contacting pair under investigation [19,20].
2.2 Contact Modelling and Reference Frames

Different numerical approaches exist to estimate the normal force for a given contact patch [21]. Hertz model [22] is a very fast and simple way to estimate the normal force in a contact. It is widely used in the field of railway simulation due to its simplicity [23,24]. The geometry of the bodies in the vicinity of the contact must be described by second degree polynomials. As a consequence, the contact area is an ellipse with semi-axes $a$ and $b$ in the rolling (longitudinal) and lateral directions, respectively. The values of $a$ and $b$ depend on the local curvatures of the bodies in the point of contact.

Together with its simplicity, the main advantage of the Hertz method is that if the hypotheses are completely fulfilled, there is an analytical solution for the contact problem. However, if the assumptions are not fulfilled, the result may not be precise. One of the biggest limitations is the shape of the contact area. It must be elliptic, and its shape depends on the local values of the curvatures of the two bodies in the point of contact. Two alternatives appear to overcome this drawback; the approximation of a non-elliptic contact zone by a set of ellipses as used by Pascal & Sauvage [25], and a method based on virtual interpenetration (Kik & Piotrowsky [26,27] and Ayasse & Chollet [28,29]).

The numerical method used in this paper to solve the normal contact problem is based on the virtual interpenetration. The profiles in the interface are allowed to interpenetrate, and the interaction force is a function of the indentation between the profiles. The area of contact is obtained from the interpenetration area as explained in [26]. In this work, an ellipse is fitted in the contact area and the normal force is calculated for this ellipse using the Hertz theory. This is known as the equivalent ellipse method. The computational effort using this method is very low which is very important in the field of railway simulation.

There is a fundamental difference between the contact ellipse calculated using Hertz and the equivalent ellipse. The former uses the radii of curvatures at the geometrical point of contact as the input data. As the wheel and rail profiles come from the digitalisation of real ones, interpolation and smoothing strongly influence the local geometry. Therefore, the calculated radii of curvatures at a point are biased by these procedures. Nevertheless, the global geometry is not influenced and, consequently, the equivalent ellipse method is little affected [21].

Fig. 4. Reference frames used: track reference frame $(xyz)$ and contact reference frame $(x_c, y_c, z_c)$.

To model the contact patches two frames of reference are used: one associated to the track, and another associated to the contact, as seen on Fig. 4. In the track reference frame $xyz$, $x$ is the longitudinal axis, indicating the forward direction, $y$ is parallel to the plane of the track and to the right, and $z$ is vertical and upwards. The $x$- and $y$-axes define the plane of the track.
There is a contact reference frame associated to each contact. It is defined from the track reference frame through a rotation around the longitudinal axis of the track of value \( \gamma_c \). \( \gamma_c \) is the average value of the conicity for the whole contact patch. The area of contact is placed inside a rectangle aligned with the \( x_c \)- and \( y_c \)-axes. The centre of the rectangle is taken as the origin of the reference frame.

### 2.3 Fastsim

The FASTSIM algorithm [11,12] is used to solve the tangential problem. It was developed by Kalker as a linear version of the full theory of elasticity implemented in CONTACT [30]. The basic idea of FASTSIM is that the relation between elastic displacements \( u \) and stresses \( p \) in the contact plane is linear, where \( L \) is called the flexibility coefficient:

\[
u = L p
\]

The problem is solved in two steps: First of all, the value of the flexibility coefficient for each area of contact is obtained. Secondly, for a given normal stress distribution and creepages in the contact, the algorithm is applied and the tangential stress distribution is determined.

In the first step, the whole area of contact is supposed to be in adhesion (this might be a restriction in the applicability of FASTSIM). The tangential pressure distribution is integrated, giving the values of \( F_x \) and \( F_y \) in the longitudinal and lateral directions, respectively. These values are equated to the solution of the full elastic problem given by CONTACT, and the value of the flexibility coefficient can be solved.

Solving the full elastic problem requires a high computational cost. If a regular figure is assumed, the magnitudes can be non-dimensionalised and the results can be tabulated. The natural choice for the shape of the area of contact is an ellipse. Given the ratio between the semiaxes of the ellipse \( a/b \) the problem is solved by interpolating in a table. Some other authors propose different regular figures to fit non-elliptic patches [31].

Kalker distinguishes three flexibility coefficients for an elliptic area \( (L_1, L_2, L_3) \) in order to express the linear relation in longitudinal and lateral displacements and spin rotation, respectively:

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

where \( c_{ij} \) are the values of Kalker’s coefficients for an elliptic area, and \( G \) the shear modulus of the material. The values of \( c_{ij} \) can be obtained from [30].

In the second step, the tangential stress distribution is calculated. Two inputs are required by FASTSIM: the normal stress distribution \( p_z \) and the creepages (longitudinal \( (\nu_x) \), lateral \( (\nu_y) \) and spin \( (\phi) \)). The normal stress distribution is usually calculated with Hertz’s model.

For a given element in the contact area, the kinematic equation for the stationary state is:

\[
s(x,y) = V \begin{bmatrix} v_x - \phi y \\ v_y + \phi x \end{bmatrix} - \frac{\partial u(x,y)}{\partial x}
\]

where \( s \) is the slip in the contact area, \( V \) the longitudinal velocity of the vehicle and \( v_x, v_y, \) and \( \phi \) the longitudinal, lateral and spin creepages, respectively.
The area of contact is partitioned into slices parallel to the $x$ axis (or longitudinal strips) as shown in Fig. 5. The algorithm FASTSIM divides each strip into the same number of elements of length $\Delta x$. Using the relation (2) and the values of the flexibility coefficients, the previous equation yields:

$$
\frac{s(x, y) \Delta x}{V} \frac{\Delta x}{L} = \left\{ \frac{v_x - \varphi y}{L_1} - \frac{\varphi_x}{L_3} \right\} \Delta x + [p(x - \Delta x, y) - p(x, y)]
$$

(5)

Assuming adhesion ($s = 0$) the tangential stress on an element can be calculated as:

$$
p^A(x - \Delta x, y) = p(x, y) - \left\{ \frac{v_x - \varphi y}{L_1} - \frac{\varphi_x}{L_3} \right\} \Delta x
$$

(6)

This formula is recursively applied, starting from the leading edge, where the tangential stress is assumed to be zero, to the trailing edge. The superindex $'A'$ indicates that is calculated under adhesion. This hypothesis is checked using the Coulomb’s law, where the coefficient of friction $\mu$ is considered a constant and known value

$$
\sqrt{(p_z^A(x - \Delta x, y))^2 + (p_y^A(x - \Delta x, y))^2} \leq \mu p_z(x - \Delta x, y)
$$

(7)

where $p_z$ is the normal pressure on the cell $(x-\Delta x, y)$. Otherwise the surfaces must slip and the tractive force follows the curve $\mu p_z(x, y)$. In the case of slip, the stress at the point $(x-\Delta x, y)$ equals:

$$
p_k^s = \mu p_z \frac{p_k^A}{|p|^4}
$$

(8)

with $k = x$ (longitudinal), $y$ (lateral), and the superindex $'S'$ stands for slip. Therefore, the tangential pressure at a cell is $p = p^A$ if there is adhesion or $p = p^S$ is there is slip.

The non-dimensional slip is defined as

$$
\gamma(x, y) = \frac{s(x, y)}{V}
$$

(9)

In the case of adhesion $\gamma = 0$, $\gamma$ can be calculated in the slip zone, using eqs. (5) and (6), as:

$$
\gamma(x - \Delta x, y) = \frac{s(x - \Delta x, y)}{V} = \frac{L}{\Delta x} \left(p^S(x - \Delta x, y) - p^A(x - \Delta x, y)\right)
$$

(10)

where $L = L_1$ for the equation in $x$, and $L = L_2$ for the equation in $y$.

The energy dissipated in the contact per unit area is used as a wear indicator. It is calculated for each element as
\[
\left( \frac{T_y}{A} \right)(x,y) = p(x,y)\gamma(x,y) = p_x(x,y)\gamma_x(x,y) + p_y(x,y)\gamma_y(x,y)
\]

(11)

For each value of \( T_y/A \) the corresponding wear rate \( (W_R) \) is calculated in \( \text{\( \mu \text{g/m/mm}^2 \)} \). This is achieved using the wear coefficients presented in Table 1, obtained for R8T wheel and 900A rail materials and extracted from [32]. Three regimes are possible, namely mild, severe and catastrophic regimes.

<table>
<thead>
<tr>
<th>Regime</th>
<th>( T_y/A ) (N/mm²)</th>
<th>Wear Rate ( (W_R) ) (( \mu \text{g/m/mm}^2 ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>K₁ (mild)</td>
<td>( T_y/A \leq 10.4 )</td>
<td>5.3 ( T_y/A )</td>
</tr>
<tr>
<td>K₂ (severe)</td>
<td>( 10.4 &lt; T_y/A \leq 77.2 )</td>
<td>55.0</td>
</tr>
<tr>
<td>K₃ (catastrophic)</td>
<td>( 77.2 &lt; T_y/A )</td>
<td>( 55.0 + 61.9 \times (T_y/A-77.2) )</td>
</tr>
</tbody>
</table>

### 3 Test Specimens and Geometry

#### 3.1 Wheel and rail relative position

The study of the wheel-rail contact is a complex problem because of the nonlinearity of the profiles. For a given set of wheel and rail profiles, the number, position, size and shape of the contact area depends on the relative position of the wheel-set with respect to the rail [24].

Fig. 6 shows the wheel/rail contact with respect to the frame of reference of the track. It corresponds to the centred position of the wheel-set. The rail is tilted at an angle equal to the rail slant of the sleeper. Both rail profiles are separated by a distance equal to the gauge, measured at a distance 14 mm below the top plane of the rails (this distance may be incremented for worn rails). The nominal contact point of the wheel is set at 70 mm from the inner face. Taking this point as an origin, both wheel profiles are mounted at fixed distance. The geometric parameters used to set the profiles in the track are specified in Table 2.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rail inclination</td>
<td>1:20</td>
</tr>
<tr>
<td>Gauge</td>
<td>1435 mm</td>
</tr>
<tr>
<td>Radius</td>
<td>450 mm (new)</td>
</tr>
<tr>
<td></td>
<td>420 mm (worn)</td>
</tr>
<tr>
<td>Distance between contact points</td>
<td>1500 mm</td>
</tr>
</tbody>
</table>
As shown in Fig. 7, the position of the wheel-set with respect to the track is modelled with a set of four coordinates: lateral ($y$) and vertical displacement ($z$) of the centre of mass of the wheel-set, and roll ($\phi$) and yaw ($\psi$) angles (rotation around the longitudinal and vertical axes respectively). Lateral and vertical displacements are measured in the track reference frame $xyz$. Yaw and roll rotations are the Euler angles in the sequence $Z-X$. This is a typical choice in railroad simulation.

3.2 Specimens

Two pairs of profiles were used to analyse new and worn wheel/rail contacts; the new profiles were in a factory as machined condition, and the worn profiles taken from heavily worn in-service components (these are the same specimens used previously in [10]). As shown in Fig. 8, the specimens were cut from actual wheel and rail sections. The rail specimen contained the head of the profile, whereas the wheel specimen was cut containing the tread and the flange fillet.

Surface roughness values were measured from the samples and are stated in Table 3. The new set of profiles is still complete with machining marks (turning), giving a rough surface. A combined roughness has been calculated for each pair, and unsurprisingly the lower value is found for the worn components. In this case, wear has especially reduced the worn rail roughness.
Table 3. Wheel-rail specimen surface roughness values.

<table>
<thead>
<tr>
<th>Wheel-rail</th>
<th>Specimen component</th>
<th>Mean surface roughness $R_a$ (μm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>New</td>
<td>Wheel ($w$)</td>
<td>2.07 (transversal) 1.80 (longitudinal)</td>
</tr>
<tr>
<td></td>
<td>Rail ($r$)</td>
<td>11.15</td>
</tr>
<tr>
<td>Worn</td>
<td>Wheel ($w$)</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>Rail ($r$)</td>
<td>1.10</td>
</tr>
</tbody>
</table>

3.3 Profiles

As shown in Figure 9, the wheel and rail profiles for the new and worn sets have been digitised using a Miniprof device. Clearly visible in the Figure, is the effect wear has had on the profile of the worn specimens. In the case of the worn rail, wear has reduced the width of the head giving a wider gauge and allowing for a greater lateral displacement of the wheel-set without flanging. The rail crown is also visibly flattened. For the worn wheel, its profile is hollowed, giving a lower conicity on the tread and a false flange on the outer face. These effects are critical, as the shape of the profiles will condition the position and the shape of the contact area.

Fig. 9. Comparison of new and worn profiles. The flange and back face of the wheel are plotted as a reference (thin line)

4 Experimental Method and Results

4.1 Apparatus, Specimen Alignment and Methodology

The experimental set-up is shown schematically in Fig. 10, and consisted of a 10 MHz focusing transducer, a UPR (ultrasonic pulser-receiver), a digital oscilloscope, a control PC, a scanning table (automated with x, and y-stepper motors) and a load frame.

Fig. 10. Schematic of the ultrasonic pulsing receiving apparatus and scanning system.

As shown in the Figure, the wheel and rail specimens are loaded together in the frame, and the ultrasonic transducer positioned in a distilled water bath above them (air bubbles in the water would scatter the sound signal). The transducer has a known focal length of 76.2 mm in water, and is positioned in the water bath so as to focus the ultrasonic signal onto the interface of...
interest. It should be noted that this is achieved by taking into consideration refraction effects at the boundary between the upper surface of the rail specimen and the water bath [10]. Further to this function, the water also couples the ultrasound to the specimens; as ultrasonic sound waves are rapidly attenuated in air.

A control signal is emitted from the PC, and in turn triggers the UPR to electrically excite the transducer. In response to the excitation, the transducer emits an ultrasonic sound wave; it also receives back the reflected signal from the interface. The signal emitted by the transducer is a multi-frequency waveform with a centre frequency of 8.8 MHz. Both the emitted and reflected sound signals are displayed on the oscilloscope and the averaged peak to peak amplitude of the reflected signal downloaded. Upon completion of the measurement at a given point, the PC controlled scanning table moves the transducer to the next point in a pre-determined sequence until the complete interface is scanned. In this work, a scanning increment of 0.25 mm was used in both the x and y directions, and five averages of the peak to peak amplitude were recorded at each point.

As well as being partially transmitted through the interface, the emitted ultrasonic signal is also attenuated by the material bulk. In order to calculate the reflection coefficient, these two effects must be de-coupled. A reference scan is taken without the wheel specimen present. As the interface is now metal-air, the entire signal is reflected, and the returned signal is only diminished by attenuation. Therefore, by dividing a given data point by its corresponding reference value, reflection coefficients can be calculated. Scans of reflection coefficient were then converted to interfacial stiffness, and finally contact pressure using a calibration appropriate to the interface under examination [10].

![Fig. 11. Experimental set-up.](image)

Fig. 11 shows a more detailed view of the previously described load frame. When compared to previous studies [10], greater emphasis was placed on re-creating actual wheel/rail contact conditions. The frame consists of an upper fixed plate, and a lower actuated floating plate. The rail specimen is located in a groove on the upper plate, and the wheel specimen on the floating plate. Profiles were set in the test rig in the same position as they are on the track when the train is in operation. The relative position of the wheel with respect to the rail is given by the set of coordinates \( \{y,z,\phi,\psi\} \) (Section 3.1, Fig. 7). A grid was marked on the floating plate, and was used to set the lateral displacement \( y \) and the yaw rotation \( \psi \) for the contact. Rotation around the \( x \) axis was set by the wedge; this value comprised of the roll angle of the wheel-set \( \phi \) and the rail slant \( \phi_r \). It should also be noted that in this study only the centred position of the wheel-set is studied \((y = \phi = \psi = 0)\), and therefore the angle of the wedge is \( \alpha_w = \phi \). With the specimens
positioned, load was then applied to the specimens in line with the contact, so as not to cause rotation of the specimens (this was checked using an orbit surface level on the floating plate).

### 4.2 Ultrasonic Results

Fig. 12 shows the ultrasonically measured contact pressure distributions for the new and worn profiles. These scans were taken for a load of 50 kN. Integration of the measured pressure profiles was performed, and gave an error of less than 5% when compared to the known applied load highlighting the validity of the technique. Both the contact patches are globally similar in size. However, the result recorded from the new profiles is visibly fragmented, and has higher peak pressure values when compared to the worn profiles (1800 MPa as opposed to 1350 MPa).

![Fig. 12](image)

Fig. 12. Normal stress distributions obtained with the ultrasound technique for new (left) and worn (right) profiles.

This effect is best explained by considering the roughness of the long range waviness and machining marks present on the new wheel specimen. They reduce interface conformity and cause a fragmentation of the pressure distribution. This result is consistent with previous studies where new, worn, and sand damaged wheel/rail specimens have been considered [10]. As the wheel and rail wears, the contact becomes more conformal, both on a long range and an asperity scale, and the stress distribution is closer to the one expected for perfectly smooth surfaces.

### 5 Contact Modelling

In this section the normal problem is solved numerically following the procedure detailed in Section 2.2. The input values for the contact problem are the coordinates \( y, \phi \) and \( \psi \) (for the centred wheel-set) and the normal force \( F_N = 50 \text{ kN} \) in this case. For a given set of coordinates \( \{y,z,\phi,\psi\} \) the interpenetration area is calculated as in [24], and a value of the indentation, \( \delta \), obtained. The contact area is then estimated using the interpenetration area via the relation \( \delta_0 = k \cdot \delta \) (a value of \( k = 0.55 \) as in [26] is used). An ellipse is fitted to the determined area of contact, with the geometrical semi-axes \( a \) and \( b \) defined. The tangential problem is then calculated for the equivalent ellipse by applying FASTSIM. The use of the equivalent ellipse method represents a major advantage in the application of FASTSIM, as the values of the flexibility coefficients are tabulated (2) resulting in a reduction in the computer processing power required.
Fig. 13. Areas of contact obtained numerically for new (left) and worn (right) profiles.

Fig. 13 shows the areas of contact obtained numerically using the interpenetration method, plotted on the track reference system. The plan view of the contact in the $x$-$y$ axes is also plotted in the same scale. The profiles used to solve the contact problem are the ones in Fig. 9. The wheel is generated from the revolution of the profile around the main wheel-set axis, whereas the rail is generated as a longitudinal extrusion.

Fig. 14. Areas of contact for new (left) and worn (right) profiles in the contact reference frame.

The ‘$x$’ is the area of contact calculated numerically.

In Fig. 14, the area of contact obtained numerically is plotted together with the normal stress distribution obtained experimentally, in the contact reference frame. The numerical problem is solved taking as an input the 2D profiles in the interface. This might explain the small differences between the contact patch obtained experimentally using the ultrasound technique and the contact patch obtained numerically.

The shapes of both areas of contact seem to be close to an ellipse, especially for the worn profiles. To evaluate this characteristic a parameter of ellipticity is defined. The area of contact is framed inside a rectangle with its axis parallel to the coordinate system of the contact (see section 2.2). Then, an ellipse is fitted to the rectangle, with $2a$ and $2b$ the dimensions of the ellipse in the longitudinal and lateral directions, respectively.

With $A_c$ being the value of the area of contact and $A_e$ the area of the equivalent ellipse, the parameter of ellipticity $\varepsilon$ is defined as:

$$\varepsilon = 1 - \frac{(A_c \cup A_e) - (A_c \cap A_e)}{A_c}$$  \hspace{1cm} (12)

The numerator of the fraction represents the difference between the area of contact and the fitted ellipse (striped area in Fig. 15). This value is divided by the area of the equivalent ellipse to get a non-dimensional parameter. The parameter of ellipticity $\varepsilon$ equals 1 for a perfect ellipse.
If the value of the parameter $\varepsilon$ is sufficiently large, the area of contact can be considered an ellipse and be calculated using Hertzian hypotheses. Otherwise, non-Hertzian methods should be used with the increase in computational time. In any case, Kalker in [12] suggests ellipcsisation of the contact area for the application of vehicle simulation.

The characteristic parameters of the contact patches are detailed in Table 4. The area of contact is greater for the experimental results (5% greater for new profiles and 15% for worn ones). This is calculated adding the elements with positive pressure. Each element represents a discrete position in the scanning window; therefore, the pressure in each element is an average value. Consequently, the size of the area of contact is an estimate and depends on the resolution taken in the scanning window (scanning step). For a smaller scanning step, the size of the area of contact is expected to be closer to the size of the area of contact obtained numerically up until the resolution of the ultrasonic transducer is reached.

Table 4 also shows the semi-axes of the equivalent ellipse. With these values, the parameters of ellipticity are calculated for the numerical/experimental and new/worn contacts. $\varepsilon$ is greater for the worn contact, because the shape of the contact patch is closer to an ellipse. This is also seen in Fig. 14. A prove of the validity of the assumption of the equivalent ellipse is the agreement shown in Fig. 14. The procedure used to find the area of contact which exerts a normal force equal to 50 kN involves iteration. If the solution of the numerical models converges to the area of contact obtained from the ultrasound measurements, the assumptions taken are valid.

**Table 4.** Characteristic parameters for the contact patches for new (a) and worn profiles (b).

<table>
<thead>
<tr>
<th>AREA OF CONTACT</th>
<th>NUMERICAL (Interpenetration)</th>
<th>EXPERIMENTAL (Ultrasound tech.)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(a) New profiles</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Area of contact $A_c$ (mm$^2$)</td>
<td>77.96</td>
<td>81.62*</td>
</tr>
<tr>
<td>Semiaxes eq. ellipse $(a \cdot b)$ (mm)</td>
<td>5.75 · 4.03</td>
<td>6.25 · 4.125</td>
</tr>
<tr>
<td>Param. of ellipticity $\varepsilon$</td>
<td>0.84</td>
<td>0.84</td>
</tr>
<tr>
<td>Max. pressure $p_{\text{max}}$ (MPa)</td>
<td>1030**</td>
<td>1800</td>
</tr>
<tr>
<td><strong>(b) Worn profiles</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Area of contact $A_c$ (mm$^2$)</td>
<td>70.73</td>
<td>84.13*</td>
</tr>
<tr>
<td>Semiaxes eq. ellipse $(a \cdot b)$ (mm)</td>
<td>5.65 · 4.09</td>
<td>6.00 · 4.25</td>
</tr>
<tr>
<td>Param. of ellipticity $\varepsilon$</td>
<td>0.96</td>
<td>0.89</td>
</tr>
<tr>
<td>Max. pressure $p_{\text{max}}$ (MPa)</td>
<td>1033**</td>
<td>1350</td>
</tr>
</tbody>
</table>

* Estimated value adding the elements with positive pressure
** Value calculated for the equivalent ellipse using the Hertz theory.

The normal stress distribution for the contact patch obtained numerically is calculated using the Hertz model for the equivalent ellipse. This is plotted together with the normal stress distribution obtained experimentally for new and worn contacts (Fig. 16). The maximum pressure obtained by means of ultrasonic measurements is greater than Hertz maximum pressure.
(74% for new profiles, and 30% for worn profiles). This is probably caused by the rough surface topography not considered by Hertz. This effect is less important for worn contacts where the profiles become more conformal. Furthermore, there is a lack of symmetry comparing the measurements with the Hertzian distribution. This might be produced because the profiles do not have a constant radius of curvature around the point of contact.

![Graph](image)

Fig. 16. Comparison of the normal stress distribution obtained through ultrasonic measurements and predicted by the Hertz theory for new profiles (up) and worn profiles (down).

### 6 Tangential problem and wear analysis

Once calculated the solution of the normal problem and prescribing some kinematic conditions, the tangential problem can be solved. For each wheel-rail pair, two different inputs will be used:
- The normal stress distribution given by Hertz model,
- and the normal stress distribution coming from ultrasound measurements.
FASTSIM is applied to the data coming from the experimental measurements. It will allow investigating the effect of surface roughness on the tangential stress distribution and wear.

For the normal stress distribution coming from ultrasound measurements, an alternative input is used. The elastic coefficients $L_i$ are calculated for the equivalent ellipse, using the equations (3). This error is assumed to be small because the parameter of ellipticity $\varepsilon$ is close to 1.

To calculate the tangential problem and carry out wear analysis, the same discretisation of the experimental measurements is used. It is composed of a grid of elements, with each element of size $\Delta x \times \Delta y$ (Fig. 17). Therefore, not all the strips have the same number of elements. A longitudinal strip is made by a row of elements aligned with the $x_c$ axis. Three cases are possible:

- The element does not belong to the contact area $p_{ij}^y = 0$
- The element is in adhesion $\sqrt{(p_{ij}^x)^2 + (p_{ij}^y)^2} < \mu p_{ij}^y$
- The element is in slip $\sqrt{(p_{ij}^x)^2 + (p_{ij}^y)^2} = \mu p_{ij}^y$

where $i,j$ are the coordinates of the element in the scanning window. If an element $ij$ does not belong to the contact area, $p_{ij}^x$ and $p_{ij}^y$ are zero. $p_{ij}^x$ and $p_{ij}^y$ are the tangential stresses in the longitudinal direction and in the lateral direction, respectively.

> Fig. 17. Discretisation of the area of contact coming from ultrasound measurements.

Wear is modelled using the criterion of energy dissipation. It is assumed that wear rate is related to work done at wheel-rail contact [6]. The values of energy dissipated per unit area ($T \gamma / A$) are correlated with the wear rate (mass loss per unit area and rolled metre) (see Table 1 for wear coefficients).

For both new and worn set of profiles and with the normal stress distribution coming from ultrasound measurements and obtained numerically, three different calculations are performed. First of all, the solution of the tangential problem and wear estimation in the contact patch are presented, for both new and worn profiles. Each magnitude is mapped in the contact patch to compare the results for the data coming from ultrasound measurements with the results for the Hertz distribution. Secondly, those parameters are integrated for the contact area and its evolution is plotted for increasing values of creepage. Finally, assuming that the wheel is rolling following the $x$ axis and the relative position of the wheel respect to the rail is kept constant, damage on the profiles are calculated under three defined scenarios. For all the calculations, the coefficient of friction is supposed to be known and constant in the whole area of contact. A value $\mu = 0.4$ is used.
6.1 Results for moderate creepage

A first set of results is presented for moderate creepage: a value for longitudinal creepage ($\nu_x$) of 0.3% is assumed. As the relative position of the wheel with respect to the rail corresponds to the centre position, with no yaw angle, no lateral creepage is expected ($\nu_y = 0.0\%$). Finally, the spin creepage is the corresponding to the wheel conicity ($\varphi = 5.0 \ (\text{1/m})\%$).

As a solution of the tangential problem and wear analysis the following magnitudes are plotted for each element in the contact patch (Figs. 18 and 19 for new and worn profiles, respectively):

- Adhesion and slip regions
- Tangential stress distribution in the longitudinal direction $p_x \text{ (MPa)}$
- Wear rate $W_R \text{ (\mu g/m/mm}^2\text{)}$

As a result of the application of FASTSIM, the leading edge is in adhesion and the trailing edge in slip. The frontier is smooth for the Hertziand pressure distribution. For the ultrasound measurements, this line is irregular and shows some overlapping patches of adhesion in the slip area. There is a correspondence between the normal stress distribution and the fragmentation of the adhesion/slip regions: where there is a high peak in the normal pressure (Fig. 16), due to Coulomb’s law, it allows higher traction forces; saturation is not reached and those elements are in adhesion. Conversely, when there is a valley in the normal stress distribution, no high values of traction are allowed and those elements are under slip condition.

![Contact Patch](image1.png)

![Contact Patch, Solution of Tangential Problem (Hertz)](image2.png)

![Tangential Pressure (Longitudinal) for new profiles](image3.png)

![Wear Rate for new profiles](image4.png)

Fig. 18. Surface plots of different parameters from the tangential problem from data coming from ultrasound measurements and from the Hertz model, for new profiles and moderate creepage ($\nu_x = 0.3\%$).

For the longitudinal stress surface plot ($p_x$), the shape obtained is quite smooth for both situations, coming from ultrasound measurements and from the Hertz model, for worn and new profiles. On the one hand, in the adhesion region the stress shows a monotonic increase since it
reaches the saturation given by the value $\mu p_0$. This peak value is about 27% greater for the distribution coming from the measured contact patches than for the distribution coming from Hertz calculation (for both new and worn profiles). On the other hand, in the slip region, the tangential stress field is proportional to the normal stress distribution, and it shows the same peaks that appeared in the normal stress distribution. It makes the tangential stress distribution coming from measurements to have a less regular shape; this was expected due to the fragmentation of the contact patch.

![Surface plots of different parameters from the tangential problem from data coming from ultrasound measurements and from the Hertz model, for worn profiles and moderate creepage ($v_f = 0.3\%$).](image)

**Fig. 19.** Surface plots of different parameters from the tangential problem from data coming from ultrasound measurements and from the Hertz model, for worn profiles and moderate creepage ($v_f = 0.3\%$).

To represent the damage on the contact patch, the wear rate is mapped. In the adhesion region (composed by the leading edge and the high peaks of the normal stress distribution), according to the wear criterion, there is no wear; it only appears in the slip zone (trailing edge of the contact patch, and valleys in the normal stress distribution). Locally, there is a big difference between the expected wear for the measured data and the data generated with Hertz model. This suggests that the application of Hertz to a contact may not predict local effects of wear in the profiles. For the data coming from ultrasound measurements there are high wear peaks with values up to 30.0 $\mu$g/m/mm$^2$ surrounded by valleys with no wear rate (there is adhesion). Those peaks will cause a high local wear on the wheel and rail profiles. For data coming from Hertz model, there is a smooth distribution of wear rate with the maximum around 9 $\mu$g/m/mm$^2$. This pattern appears for both new (Fig. 18) and worn profiles (Fig. 19), being more severe for new profiles.

This observation highlights a benefit of the ultrasonic technique. When the surfaces are new and unworn, the interface is not yet conformal as running in has not occurred. The application of the technique at this point allows judgments to be made as to where the initial high rates of wear are likely to occur, based on the initial surface topography. This then provides valuable information
for predicting the evolution of the contact patch. Although for the worn in contact some points of high wear rate remain, on average values are reduced, meaning the Hertzian approach offers a far closer approximation of the contact conditions.

However, the widespread use of the experimental technique would require the scanning of each individual wheel/rail contact and to this end is not realistic. Although, the experimental results presented in this paper may help to understand the actual wear mechanisms present for real wheel and rail surfaces during the running-in phase, and the shape of contact that ultimately prevails for a given initial surface finish.

It is also interesting to point out that, for the moderate creepage scenario, despite the high peaks appeared for data coming from measurements, the entire contact patch for the new and worn contacts and calculated from the Hertz model and from ultrasound measurements is in the mild regime so wear rate is proportional to the energy dissipated in the contact. For greater values of creepage, the full contact patch would be in slip, and a transition from mild to severe regime is expected. The values of the high peaks of wear rates would saturate with a value of 55.0 μg/m/mm².

### 6.2 Global parameters of contact for range of increasing creepage values

In this section, the distributions of the magnitudes calculated before are integrated over the contact patch. These global magnitudes give an overall idea about the contact area, but in order to get accurate results, the magnitudes must be considered locally.

A global value for the energy dissipated in the contact patch is calculated averaging the local \( Tγ/A \) from (11) in the slip region, being \( A_s \) the slip region and \( \Delta x, \Delta y \) the dimensions of each element of the contact patch.

\[
\frac{Tγ}{A} = \frac{1}{A_s} \sum A \left( \frac{Tγ}{A} \right) \Delta x \Delta y \tag{13}
\]

This value can also be used to calculate a global wear rate using the wear law in Table 1.

Lateral and spin creepages are kept constant (\( ν_z = 0.0 \% \), \( φ = 5 (1/m) \% \)) while longitudinal creepage is changed, simulating a range of increasing values of traction effort. The value for the measured contact is compared with the expected value calculated with the Hertz model. The relative variation between them is also presented (plotted on secondary axes). Results are presented in Fig. 20-left for new profiles and Fig. 20-right for worn ones.

The magnitudes plotted are:
- Percentage of the slip region over the total area of contact (%)
- Traction force \( F_z \) (kN)
- Global wear rate \( W_R \) (μg/m/mm²), which is calculated as an average of the local wear rate over the slip region.

The first plot shows the influence of the increasing creepage on the size of the slip region. It is interesting to point out that for small creepages (0.2-0.3 %), the slip region is very small and the discretisation of the contact patch may not be adequate because the size of the elements may be too big compared to the size of the slip region. This might explain the differences between the two methods for very low creepages. Above this threshold, the curves show the same tendency, and the differences between them decrease for increasing values of creepage. For \( ν_z > 1.0 \% \) the variation is less than 5 %. Beyond \( ν_z = 1.4 \% \) this difference is negligible.
Analysing the traction force, the curves show a good agreement using Hertz vs. measured data. This suggests that the Hertz theory is suitable for simulation. For a creepage greater than 0.4 %, the difference is estimated to be less than 5%. This difference vanishes for increasing creepage.

The last plot is relative to the global wear in the contact patch. The value of \(T \gamma / A \Delta x \Delta y\) for each cell is summed for the area of contact and averaged over the slip region. When the previous curves are saturated, these ones have a linear behaviour with the creepage, due to the linear increase of the non-dimensional slip \(\gamma\) for the constant value of \(T\). For the maximum longitudinal creep of 2 %, the maximum expected wear rate is 25-30 \(\mu g/m/mm^2\) which means that all the range is in the mild regime (Table 1). For increased loads and, therefore, stresses in the contact patch, a transition to the severe regime is expected. That would result in a horizontal line in the plot. It is interesting to point out that for the studied contact patches, the predicted global wear rate is greater for the Hertzian model. The variation between the estimated wear for Hertz and for measured data, for both new and worn contacts, is estimated between 10-15% for creepages greater than 1.2 %.

**Fig. 20.** Comparison of the values for global parameters of the contact area for increasing creepage for new profiles (left) and worn profiles (right).
6.3 Calculation of damage in the profiles for low, moderate and high creepage

Finally, in this section, the damage on the profiles is estimated as the wear depth created in each cell per rolled metre. Assuming that the wheel is rolling in the same position respect to the rail, it is used the factor $\Delta x/(2 \pi R)$ to average the damage of each element of the contact area in the whole circumference of wheel running radius $R$, being $\Delta x$ the length of the cell (equal for all elements of the contact patch). The damage caused in each cell is calculated as

$$D_{cell} = \frac{W}{\rho} \frac{\Delta x}{2 \pi R}$$  \hspace{1cm} (14)

where $\rho$ is the density of the material. The summation of the averaged damage of all the elements in a strip gives the estimated damage for each strip of the profile. As the contact is in the wheel tread, it is taken an average radius $R$ for the entire contact patch.

Three situations are studied for each contact: one with low creepage, another with moderate creepage and finally with high creepage. The different scenarios used to calculate wear are listed in Table 5. The contact area under a slip condition is calculated to be about 7 % for low creepage, 50 % for moderate creepage and 98 % for high creepage, both for worn and new profiles. In Fig. 21 the accumulated damage for the two sets of profiles and the three scenarios is shown.

Table 5. Scenarios for wear calculation.

<table>
<thead>
<tr>
<th>Creepage</th>
<th>Low Creepage</th>
<th>Moderate Creepage</th>
<th>High Creepage</th>
</tr>
</thead>
<tbody>
<tr>
<td>$v_c$ (%)</td>
<td>0.05 %</td>
<td>0.3 %</td>
<td>1.0 %</td>
</tr>
<tr>
<td>$v_y$ (%)</td>
<td>0.0 %</td>
<td>0.0 %</td>
<td>0.0 %</td>
</tr>
<tr>
<td>$\phi$ (1/m)</td>
<td>5.0 %</td>
<td>5.0 %</td>
<td>5.0 %</td>
</tr>
</tbody>
</table>

Fig. 21. Damage produced by the contact patch on the profile in mm per rolled metre for low, moderate and high creepage conditions (note that the scale is different for each plot).
It is interesting to analyse the shape obtained for measured distributions compared to hertzian distributions. For the hertzian pressure distribution the expected damage follows a smooth curve whereas for the pressure distribution coming from measurements, damage on the profile has a more irregular shape. This is the consequence of the local effects of wear rate on the contact patch, caused by the presence of asperities. This damage pattern can only be obtained with data coming from ultrasound measurements. For worn profiles, the prediction of damage given by Hertz is closer to the shape obtained for the measured data. This is due to the effect of the asperities, which are worn out. As a general conclusion, the estimation of damage for the profiles offered by the ultrasound measurement is slightly greater than the one expected for the hertzian case.

It is worth commenting on the shape of the damage predicted for the hertzian distribution under the low creepage scenario. As the size of the elements is too big compared to the size of the slip region, there are a few elements in the slip region and no smooth solution is obtained as in the other two scenarios.

Another characteristic to highlight is the nonlinearity of the creepage with the accumulated damage (note the different order of magnitude for each scenario). This is associated to the saturation of the contact patch (see Fig. 20, upper plots): for low creepage, only approximately 10% of the area is in slip; in the next step, for moderate creepage, 50% of the area is in slip and finally for high creepage, above 95% of the contact patch is in slip. After full saturation and for greater values of slip, the transition from mild to severe regime in the contact patches is expected.

Finally, it is interesting to comment on the effect of spin. While longitudinal creepage is increased, spin is kept constant. Therefore its relative influence is only noticed in the low creepage scenario which makes the damage distribution non-symmetrical.

7 Conclusions

In this paper, a methodology to study the wheel-rail contact for real surfaces has been proposed. Two techniques have been used: one experimental using the measurement of the ultrasonic reflection, and another numerical based on the Hertz model. The tangential problem was calculated using FASTSIM for data coming from both procedures. The results were compared and the main points discussed are:

- When two rough surfaces are pressed together, the load is supported by the asperities in contact. This effect is observed in the normal stress distribution obtained experimentally, which has high peaks corresponding to roughness within the contact. This is not found in the distribution obtained numerically, because the surfaces are assumed to be smooth. The numerical method gives better results for worn contacts than for the new ones, because the asperities are partially worn out.

- To study the tangential problem a kinematic state is prescribed. The contact patch is divided into adhesion and slip regions. For a real surface these regions are fragmented and do not show a smooth frontier.

- Data coming from experimental measurements reproduce the real contact where the load is discretely supported by asperities (in both new and worn profiles). Therefore, the wear distribution obtained from experimental measurements is a non-continuous function showing high peaks of wear rate. However, the modelisation with Hertz theory predicts a smooth wear rate distribution. Thus, care must be taken when analysing some wear estimations coming from a numerical model which does not take into account the effect of surface roughness.
Although analysing the contact locally there is a big deviation between measured data and Hertz theory, the difference of the average magnitudes of the contact patches was found to be small. This means that the numerical method is suitable for simulation as it is typically used, but special care must be taken when analysing wear.

The wear life of a component highly depends on its initial topography, and the wear mechanisms of the surface will similarly vary. Therefore, different initial conditions for the same running environment of the component would influence its overall life.

Finally, there is a strong limitation on the use of the experimental method. To use this method each contact patch must be scanned with a good resolution and it requires a lot of time. Nevertheless, the use of experimental data is useful to understand the mechanisms involved in the contact of rough surfaces and wear.

Acknowledgements

The corresponding author gratefully acknowledges the cooperation of C. Hardwick and Portec Rail Inc. for supplying the Miniprof device. This research was supported by Universitat Politecnica de Valencia (Spain).

References


