COMPARISON OF NON-ELLIPTIC CONTACT MODELS: TOWARDS FAST AND ACCURATE MODELLING OF WHEEL-RAIL CONTACT

Matin S. Sichani^{1)*}, **Roger Enblom**^{1,2)}, and **Mats Berg**¹⁾

1) Division of Rail Vehicles, Dept. of Aeronautical and Vehicle Engineering,
Royal Institute of Technology (KTH), Stockholm, Sweden

2) Bombardier Transportation, Västerås, Sweden

* Corresponding author e-mail address: matinss@kth.se

ABSTRACT

The demand to investigate and predict the surface deterioration phenomena in the wheel-rail interface necessitates fast and accurate contact modelling. During the past twenty years, there have been attempts to determine more realistic contact patch and stress distributions using fast non-iterative methods. The main aim of this work is to compare some of these state-of-the-art, non-elliptic contact models available in the literature. This is considered as the first step in introducing a fast accurate non-elliptic contact model that can be used on-line with vehicle dynamic analysis. Three contact models, namely STRIPES, Kik-Piotrowski and Linder methods are implemented and compared in terms of contact patch prediction, contact pressure and traction distribution. The paper discusses the gaps to be filled in terms of contact model validation and the results indicate the need for better contact pressure and patch estimation in certain contact cases.

1 INTRODUCTION

Accurate determination of the contact patch and calculation of stress distributions within this region has gained importance due to the increasing need for the prediction of wear and rolling contact fatigue. At the same time, on-line application of the contact model in multi-body-systems (MBS) codes imposes restrictions on computational cost of the contact model. Thus, there is always a trade-off between accuracy and time efficiency when it comes to solving the wheel-rail contact problem. Generally, in today's commercial MBS codes [1], the contact patch is considered to be elliptic based on the Hertz theory. Some models approximate a non-elliptic contact patch by one or more ellipses (equivalent elastic, multi-Hertzian). The tangential stresses and creep forces are then calculated using the FASTSIM algorithm [2]. The results are usually acceptable for the dynamic analysis of the vehicle-track interaction. However, the unphysical characterization of the contact patch and inaccurate

estimation of the stresses in certain cases leads to unrealistic wear and fatigue predictions. Therefore, there is a need for non-elliptic contact model that is almost as fast.

2 OVERVIEW of NON-ELLIPTIC CONTACT MODELS

Non-elliptic contact models can be divided into two categories; the models based on numerical methods such as the finite element method (FEM) or the boundary element method (BEM) and the models which predict the contact patch analytically. In this paper these two categories are named as *advanced numerical methods* and *fast non-elliptic methods*, respectively.

2.1 Advanced numerical methods

In advanced numerical methods the contact conditions are satisfied in discrete points of the bodies in contact. The bodies should therefore be discretized into

elements. The contact solution is highly dependent on the element size. To obtain smoother and accurate results a finer mesh is needed which requires more computational effort.

FEM The most advanced numerical method to use for solving contact problem is FEM. This is because it can include different sources of nonlinearity originated from material property or contact geometry. However, due to numerical issues, the contact conditions and the friction law should be linearized. The linearization should be treated carefully in order to achieve reliable results. Moreover, the need for a very fine mesh in the contact area makes it too expensive from a computational point of view. For a summary on wheel-rail contact modelling using FEM see [3].

BEM Solving rolling contact problem using BEM seems to be more common than using FEM. BEM based methods generally utilize half-space theory. It requires the contact patch dimensions to be much smaller than the dimensions of the bodies in contact. Furthermore, it assumes a flat (planar) contact patch. Using half-space theory, the deformations can be related to the stresses through influence functions. The deformation in one point is calculated by superposing the effect of tractions applied over the contact patch. The contact equations may be solved using different mathematical techniques. The CONTACT software [2] by Kalker utilizes variational method while Björklund and Andersson's [4] method, ContLab, uses matrix inversion.

Since the real contact patch is not known in advance and the surface deformation-traction relationship is dependent on an integral over the contact area, one should start with a so-called potential contact area and do iterations to arrive at the real contact patch and contact pressure distribution. Such an iterative procedure slows down the models based on BEM. Knothe and LeThe [5] also proposed a contact model based on BEM. This model seems to be faster since the patch is divided into strips rather than rectangular elements and the contact pressure distribution in the rolling direction is assumed to be elliptical in advance.

2.2 Fast non-elliptic methods

The computational expenses associated with advanced numerical methods motivate attempts to look for faster non-elliptic contact models. As mentioned in the preceding section, the iterative procedure to determine the contact patch is time consuming, hence, for a fast non-elliptic method the contact patch should be determined analytically. There are several different non-elliptic methods introduced in the literature which are based on a so-called virtual interpenetration. A survey of such models is given in [6].

2.2.1 Virtual interpenetration

In order to simplify the contact patch determination, the surface deformations are neglected ($\bar{u}_z = 0$) and it is assumed that the bodies can rigidly penetrate into each other. The achieved interpenetration zone is of course bigger than the real contact area but encircles it. It has been suggested that a smaller penetration (approach) value, δ_v , than the prescribed one, δ_0 , may result in an interpenetration zone which is close enough to the real contact area and can be regarded as the contact patch. Such a virtual penetration value can be defined in different ways. Based on experience, it should be about half the prescribed penetration value.

The exact contact condition at contact boundaries can be expressed as,

$$z(x, y) = \delta_0 - \bar{u}_z \quad (1)$$

where $z(x, y)$ is the separation function between the two surfaces in contact. In wheel-rail contact, one can assume a quadratic representation of the surfaces in the rolling direction. The profile combination in the lateral direction is usually presented by a discrete function of the lateral coordinate, $f(y_i)$. Hence, equation (1) can be rewritten as:

$$Ax^2 + f(y_i) = \delta_0 - \bar{u}_z \quad (2)$$

where A is the relative longitudinal curvature. The above equation can be approximated by virtual interpenetration as,

$$Ax^2 + f(y_i) = \epsilon \cdot \delta_v \quad (3)$$

where ϵ is the scaling factor. In order to evaluate the above simplification, a Hertzian contact condition is taken into consideration. If the Hertzian assumptions

are valid then the contact solution can be expressed using Hertz solution. Using Hertz, the contact patch is an ellipse with,

$$a = m_0 \sqrt{\frac{\delta_0}{r_0(A_0 + B_0)}} \quad (4)$$

$$b = n_0 \sqrt{\frac{\delta_0}{r_0(A_0 + B_0)}} \quad (5)$$

where, a, b are the semi-axes of the contact ellipse and m_0, n_0, r_0 are functions of relative curvatures at the point of contact, A_0, B_0 .

The scaling factor ϵ in virtual interpenetration-based methods can be defined in different ways. Below, four different approaches introduced in the literature are discussed.

Approach 1 In the Linder method [7], the scaling factor is a constant number $\epsilon = 0.55$. This leads to the following semi-axes,

$$A_0 x^2 + f(0) = \epsilon \delta_0 \Rightarrow x = a = \sqrt{0.55 \delta_0 / A_0} \quad (6)$$

$$b = \sqrt{0.55 \delta_0 / B_0} \quad (7)$$

for the contact ellipse in a Hertzian case.

Approach 2 In the Kik-Piotrowski method [8] the scaling factor is also a constant number $\epsilon = 0.55$. However, a shape correction strategy is applied to the patch later on. The correction is applied in a way to impose Hertzian semi-axes ratio to the patch while preserving its area. Thus,

$$a_c = \sqrt{ab n_0 / m_0} = 0.55 \delta_0 \sqrt{\frac{n_0 / m_0}{A_0 B_0}} \quad (8)$$

$$b_c = \sqrt{\frac{ab}{n_0 / m_0}} = 0.55 \delta_0 \sqrt{\frac{m_0 / n_0}{A_0 B_0}} \quad (9)$$

while,

$$\frac{a_c}{b_c} = \frac{n_0}{m_0}, a_c b_c = ab. \quad (10)$$

This strategy, however, does not result in the same Hertzian ellipse. It is because the corrected ellipse still has the same area as the interpenetration area which is different from the Hertzian ellipse area.

In the STRIPES method [9], the scaling factor is a function of relative radii of curvature at the point of contact $\epsilon = \epsilon(A_0, B_0)$. Since the virtual interpenetration will lead to an ellipse with different semi-axes ratio than a Hertzian ellipse, a correction has been applied to the local relative curvatures of the bodies in contact. Two correction methods presented in [9] are discussed as follows.

Approach 3 In this approach the local A curvatures are corrected. The scaling factor and corrected curvature are

$$\epsilon = \frac{n_0^2}{r_0(A_0 + B_0)} B_0 \quad (11)$$

$$A_{ci} = (n_0 / m_0)^2 B_i \quad (12)$$

where n_i, m_i, B_i are local values at coordinate y_i . The contact patch in Hertzian condition will then be,

$$a = \sqrt{\epsilon \delta_0 / A_{c0}} = m_0 \sqrt{\frac{\delta_0}{r_0(A_0 + B_0)}} \quad (13)$$

$$b = n_0 \sqrt{\frac{\delta_0}{r_0(A_0 + B_0)}} \quad (14)$$

which is exactly the same as in Hertz' solution.

Approach 4 In this approach both relative curvatures A, B are supposed to be corrected. The corresponding scaling factor and corrected curvatures are

$$\epsilon = \frac{n_0^2}{r_0(1 + (n_0 / m_0)^2)} \quad (15)$$

$$A_{ci} = \frac{(A_i + B_i)(n_0 / m_0)^2}{(1 + (n_0 / m_0)^2)} \quad (16)$$

$$B_{ci} = \frac{(A_i + B_i)}{(1 + (n_0 / m_0)^2)} \quad (17)$$

The contact patch in Hertzian condition will then be

$$a = m_0 \sqrt{\frac{\delta_0}{r_0(A_0 + B_0)}} \quad (18)$$

$$b = n_0 \sqrt{\frac{\delta_0}{r_0(B_0 + B_0(n_0 / m_0)^2)}} \quad (19)$$

which has the same length as a Hertzian ellipse but different width. The difference between A_0 and $B_0(n_0 / m_0)^2$ decreases when $\lambda = A_0 / B_0$ approaches unity or zero, hence the contact ellipse converges to

the Hertzian ellipse. For large values of λ the ellipses differ considerably.

2.2.2 FASTSIM adaptation to non-elliptic patches

The tangential part of the contact problem is generally treated using the FASTSIM algorithm which is based on Kalker's simplified theory. This is because FASTSIM is a non-iterative approach which outputs the surface tractions distribution rather than just the creep forces. The FASTSIM algorithm is originally developed for elliptic contact patches. However, it is not confined to elliptic contacts. In order to adapt FASTSIM to non-elliptic patches the flexibility parameters, L_i , should be calculated. This can be done in various ways. Several different techniques are discussed below.

Non-elliptic flexibility parameters Knothe and Le-The [10] have introduced a method to calculate flexibility parameters for a non-elliptic contact patch based on the results from elastic half-space linear solution for that non-elliptic contact. In order to derive the linear solution based on the elastic half-space theory, one needs to apply numerical integration to slip equations. The results are specific to each non-elliptic patch and therefore numerical integration should be done for every contact case. Thus, this approach is rather slow.

Equivalent ellipse In order to skip the calculation of flexibility parameters and employ the elliptical parameters, an imaginary equivalent ellipse may be assigned to the non-elliptic patch. Piotrowski and Kik [8] defined an equivalent ellipse for each separate contact zone by setting the ellipse area equal to the non-elliptic contact area and ellipse semi-axes ratio equal to length to width ratio of the patch. The estimated flexibility parameters may then be used to calculate a single flexibility parameter or they may be used individually force-creepage equations.

Local ellipses for each strip By discretizing the contact patch into strips, as done in the Linder method [7] and the STRIPES method [9], one may treat each strip differently. Instead of assigning an equivalent

ellipse to a whole patch, a virtual local ellipse may be assigned to each strip separately. The flexibility parameters may then be calculated based on the local ellipses and applied to each respective strip. Assigning a local ellipse to a strip may be done in various ways. Figure 1 depicts two different techniques used in the Linder model and STRIPES. In the Linder model, all local strip ellipses have the same lateral semi-axis. The longitudinal semi-axis is then calculated so that the strip fits into the resulting ellipse. In STRIPES, dissimilarly, the local ellipses semi-axes are calculated using Hertz solution by entering relative curvatures at the center of the strip as the input.

2.2.3 Computational cost

The methods based on virtual interpenetration need the prescribed penetration value as the input. However, in wheel-rail applications the wheel loads are given. Hence, the solution of the contact problem should start with an initial penetration value. The resulting vertical force calculated by the contact model may then not agree with the prescribed value. Therefore, the correct penetration value must be determined iteratively. The penetration value is related to the vertical force and

Fig. 1. Local ellipse assignment to contact patch strips. Local ellipses for two arbitrary strips based on (a) Linder (b) STRIPES methods.

contact area through contact stress distribution. Despite the iteration process needed, fast non-elliptic models are still faster than advanced numerical methods such as BEM based models. This is because there is no need for element discretization and evaluating pressure values in each element (as it is done in the CONTACT software).

3 RESULTS and DISCUSSIONS

In order to evaluate the approaches used in fast non-elliptic models, three different proposed models are considered. The Kik-Piotrowski model [8], The STRIPES model [9], and the Linder model [7] are numerically implemented based on how they are documented in the literature. Two different contact case studies are done.

Firstly, the non-elliptic contact case of S1002/UIC60 pair with 1:40 rail inclination for central wheelset position is analyzed. In this case, the CONTACT software [2] results are taken as the reference. Nevertheless, it should be kept in mind that CONTACT is bound to the half-space assumption and therefore the resulted contact patch is planar. The contact patch, normal stress distribution, stick-slip regions and tangential traction distribution and creep forces predicted by these models are aimed to be compared.

Secondly, the contact patch and contact pressure distribution for different wheelset lateral displacements are considered. In this case the results from the STRIPES method using two different correction strategies are compared with CONTACT solutions.

In the Kik-Piotrowski model, instead of calculating a weighted average of the flexibility parameters, three different flexibility parameters are calculated and used individually. The patch correction is also applied as it is described in the latest version [8]. The Linder model has no patch correction strategy.

The STRIPES method is implemented using the *A* curvature correction strategy (see Approach 2 in 2.2.1).

Moreover, a smoothing filter is applied to the local *B* curvatures as it is documented in [9] in order to achieve smooth contact patch boundaries. The STRIPES results documented in [6] for the same contact case with central wheelset position is slightly different from what is achieved here. This is mainly because the smoothing filter used in [6] is purely mathematical and tuned to better fit the reference results. Implementing STRIPES using *A&B* curvature correction without applying any smoothing filter is also done for the second study.

3.1 Wheelset central position

The first case study is done for the right wheel-rail pair S1002/UIC60 with 1:40 inclination for wheelset central position and given load 78500 *N*. This is a classical example of a non-elliptic tread contact. In this case, two separate ellipses seem to merge. Figure 2 shows the contact patch estimated by the three models compared with CONTACT results. The lateral axis is positive towards the field side. As it can be seen, all three methods predict a non-elliptic patch with relatively good accordance with the CONTACT results. However, the Linder model overestimates the patch. This is mainly because no correction is applied to the patch which is gained by virtual interpenetration. The STRIPES patch is closer to the CONTACT patch.

Considering other contact cases, it seems that the STRIPES correction method results in more accurate patches (taking CONTACT results as reference).

Fig. 2. Contact patch estimation by CONTACT, STRIPES, Kik-Piotrowski, and Linder models for wheelset central position.

Maximum contact pressure distribution over the patch width is illustrated in Figure 3. The CONTACT solution shows a peak at the left end of the patch; which is not captured with any of the fast methods. In fact, there are two separate peaks in the CONTACT solution, each belongs to an ellipse. In the fast methods the pressure is distributed rather uniformly across the contact patch. The Linder method predicts lower pressure levels due to the fact that its estimated contact area is bigger than the other methods.

The peaks in the STRIPES solution are smoothed out due to the smoothing filter applied to local B curvatures. One can get better results by modifying the smoothing filter (see Figure 22 in [6]).

The tangential stress distribution within the contact patch and slip-stick areas of these four contact models are illustrated in Figure 4 for a pure quasi-static spin 0.052 rad/m . A theoretical creep case is chosen in order to have the same input to all contact models. The contact patch is assumed to be planar in all models except for STRIPES. In STRIPES, each strip has a different contact angle and thus different spin value. For the other three models the spin value is constant across the patch and calculated using the contact angle at the rigid point of contact. This explains why the slip area in STRIPES is dissimilar to the others.

Fig. 3. Maximum pressure distribution along the lateral axis estimated by CONTACT, STRIPES, Kik-Piotrowski and Linder models.

Fig. 4. Tangential stress distribution and slip-stick zones in the contact patch for pure spin case. The red line encircles the slip region.

The contact angle changes about 4 degrees across the patch. This leads to 0.16 rad/m change in spin which is three times the spin value at the point of contact. This is why the tangential stress in the right end of the STRIPES patch is lower than the other methods and it gets higher on the other end. Such an unbalance gives rise to much higher longitudinal creep force.

3.2 Wheelset offset positions

As mentioned in the previous case, none of the fast methods could capture the peaks in the maximum pressure distribution. Using A curvature correction strategy (Approach 3) in STRIPES, it is necessary in some cases to filter local B curvatures in order to achieve a smooth contact patch. However, this smoothing also affects the pressure values and smears out the peaks in the pressure distribution as well. In order to avoid applying smoothing filters, STRIPES is re-implemented using $A\&B$ curvature correction strategy (Approach 4). In this case there is no need to apply filtering in order to have smooth patch boundaries. Nevertheless, the B curvature values must all be positive non-zero values. Hence, the whole local B curvature values are shifted above zero when it is needed. This is different from STRIPES, as explained by Ayasse and Chollet in [9], where negative local B curvature values are replaced by a minimal positive value.

In order to compare the two correction strategies in various non-elliptic cases the same wheel-rail pair is considered for different wheelset lateral displacements. Figure 5 contains graphs of contact patch and maximum pressure distribution for different wheelset lateral displacements. The results from the two aforementioned correction strategies of STRIPES are compared to CONTACT results.

It can be seen from Figure 5 that the contact patches resulted from different correction methods are relatively similar except for case (a) $\Delta y = -1\text{ mm}$. However, the pressure distribution graphs show that the Approach 4 correction (see 2.2.1) leads to better agreement with CONTACT results in terms of predicting maximum pressure within the contact patch. Case (b) in Figure 5, for instance, illustrates that avoiding smoothing enables STRIPES to capture the peaks in the pressure distribution for the case studied in the previous section.

Another important point to notice is that although the fast non-elliptic methods predict the contact patch for central wheelset position in a quite accurate way, there

are several contact cases in which the patch determination is not as accurate. STRIPES' contact patch deviation from CONTACT results is noticeable for cases (a), (c), and in particular (d) of Figure 5. The Kik-Piotrowski and the Linder models also predict patches similar to STRIPES. A closer look on these cases reveals that in all three cases, the local B curvature values drop to zero or less in one end of the patch. Figure 5 also depicts the evolution of the local B curvature within the contact patch for all cases. A negative B curvature value implies that the two bodies cannot come into contact at that point, while a zero value means the two bodies are in conformal contact. Since a positive local B curvature value is needed for each strip to calculate the maximum pressure and apply the corrections to the strip length one should correct the negative B curvature values.

4 CONCLUSIONS

The focus of this investigation is to compare fast non-elliptic contact models used in wheel-rail applications. The contact cases studied show that the contact patch prediction by fast non-elliptic methods needs to be improved, especially in contact cases where the local lateral curvature at the ends of the contact patch is around zero. Moreover, the prediction of contact pressure distribution requires to be improved as well. This will directly result in a better tangential solution.

The new curvature correction strategy applied to STRIPES and avoidance of smoothing leads to improved pressure distribution prediction. However, a better correction method which can represent the shear effect between neighboring strips is needed in order to get a better patch estimation.

The non-planar contact patch and effects of contact angle change is an important issue taken into account by STRIPES. It will be more of interest to investigate it in gauge-corner contact. A comparison of fast non-elliptic methods to FEM results for gauge-corner contact may lead to better understanding of the differences between these methods.

Fig. 5. Contact patch, maximum pressure distribution and local B curvature variation for different wheelset lateral displacements. Comparison between CONTACT results and two different correction strategies applied to STRIPES. Positive displacement moves the contact towards gauge corner.

Finally it worth mentioning that, in this work, the CONTACT software results are used as reference for evaluation as it is used by other researchers in similar investigations. However, the limitation of this method to half-space theory implies a restriction to use it in a gauge corner contact. Thus, the study cases in this paper are limited to tread contact. Measurement results or a detailed FEM analysis are needed to be used as reference rather than CONTACT solutions.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the financial support from Swedish Transport Administration (Trafikverket), Stockholm Transport (SL), Bombardier Transportation, The Association of Swedish Train Operators (Tågoperatörerna), and Interfleet Technology, making this research possible.

REFERENCES

- [1] Shackleton, P., Iwnicki, S., Comparison of wheel-rail contact codes for railway vehicle simulation: An introduction to the Manchester Contact Benchmark and initial results, *Vehicle System Dynamics*, 46, 1-2, 129-149, 2008.
- [2] Kalker, J.J., Rolling contact phenomena – linear elasticity. In: B.O. Jacobson and J.J. Kalker (Eds), *Rolling Contact Phenomena, CISM Courses and Lectures*, No. 411, Chap. 1, (Wien, New York: Springer), pp. 1–84, 2001.
- [3] Knothe, K., Wille, R., Zastra, B.W., Advanced contact mechanics-road and rail, *Vehicle System Dynamics*, 35, 4, 361-407, 2001.
- [4] Björklund, S., Andersson, S., A numerical method for real elastic contacts subjected to normal and tangential loading, *Wear*, 179, 117-122, 1994.
- [5] Knothe, K., Le-The, H., A contribution to the calculation of the contact stress distribution between two elastic bodies of revolution with non-elliptical contact area, *Computers & Structures*, 18, 6, 1025-1033, 1984.
- [6] Piotrowski, J., Chollet, H., Wheel-rail contact models for vehicle system dynamics including multi-point contact, *Vehicle System Dynamics*, 43, 7, 455-483, 2005.
- [7] Linder, Ch., *Verschleiss von Eisenbahnradern mit Unrundheiten*, Dissertation ETH Nr. 12342, 1997.
- [8] Piotrowski, J., Kik, W., A simplified model of wheel/rail contact mechanics for non-Hertzian problems and its application in rail vehicle dynamic simulations, *Vehicle System Dynamics*, 46, 2, 27-48, 2008.
- [9] Ayasse, J.B., Chollet, H., Determination of the wheel rail contact patch in semi-Hertzian conditions, *Vehicle System Dynamics*, 43, 3, 161-172, 2005.
- [10] Knothe, K., Le-The, H., A method for the analysis of the tangential stresses and the wear distribution between two elastic bodies of revolution in rolling contact, *Int. J. Solids Structures*, 21, 8, 889-906, 1985.