Detailed wheel/rail geometry processing with the conformal contact approach

Edwin Vollebregt[1] (ID)

Abstract This paper proposes a new way of considering wheel–rail contact in multibody systems simulation that goes beyond the traditional planar constraint and elastic approaches. In this approach, wheel–rail interaction is modelled as a force element with pressures and shear stresses distributed over a contact area that may be curved, supporting conformal contact situations. This by-passes the selection of the contact reference location and reference angle, which are delicate aspects of planar contact approaches.

The idea is worked out introducing the curved reference surface as the new backbone for the computations, instead of the tangent plane used previously in planar contact approaches. The steps are described by which the curved reference is constructed in CONTACT, using generic facilities for markers, grids, and coordinate transformations, by which generic wheel/rail configurations can be analyzed in a fully automated way.

Numerical results show the capabilities of the new method for measured, worn profiles, suppressing discontinuities in the forces when multiple contact patches split or merge. A further application concerns the evaluation of strategies used in planar contact approaches. There we find that the tangent plane's inclination is of the biggest importance. This should be defined in an averaged way to achieve maximum correspondence to the more detailed curved contact approach.

Keywords Wheel–rail contact · Wheel–rail profile · Contact geometry · Contact forces · Conformal contact

1 Introduction

The development of efficient, stable, and accurate models that account for wheel–rail contact is of great interest for vehicle developers in the industry, for vehicle operators and infrastructure managers and also for researchers. The railway system is a complex one, with strong and subtle interactions between the different components of the vehicle, within the track

✉ E. Vollebregt
edwin.vollebregt@cmcc.nl

1   Vtech CMCC, Rotterdam, The Netherlands

Fig. 1 Schematization of the
wheel–rail rolling contact forces:
acting mainly as a stiff, variable
spring in normal direction,
Hertzian or non-Hertzian, and
like a variable spring and damper
tangentially, that breaks loose if a
large force is required [13]
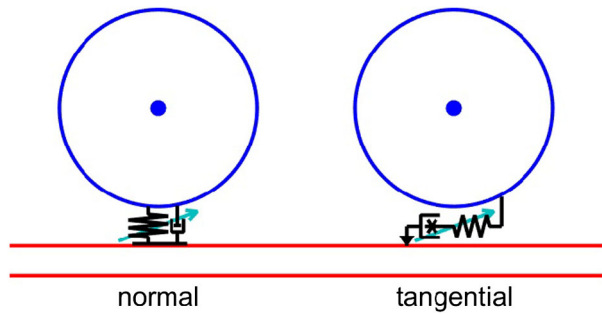
system, and in the wheel/rail interface. This complexity is addressed by vehicle dynamics
simulation, providing a well-established tool for the design and analysis of many aspects of
railway systems, such as ride comfort, running stability and safety, virtual homologation,
and the prediction of wear and rolling contact fatigue [1–7]. Wheel–rail interaction is a key
aspect in these simulations: using steel-on-steel contacts, very high forces are transmitted in
small areas, which may change quickly with the contact circumstances, such as the contact
location, creepage, and the available friction. This leads to the problem of wheel/rail contact
modelling for multibody system simulation.

The multibody modelling approach sets a framework, where a model consists (typically)
of discrete elements like rigid bodies and force elements. Force elements are then connected
to the bodies at markers, i.e. at special points where the kinematics on the one hand and the
forces and moments on the other hand are exchanged. Within this context, wheel–rail con-
tacts are usually modelled as a specific type of force element. This receives the kinematical
quantities (translational and rotational displacements and velocities) from the markers, at
which it is connected to the bodies representing the wheel and the rail. The force element
provides an effective means for modular programming, permitting the contact problem to be
solved in isolation, treated as a black-box from the side of multibody dynamics.

In the implementation of a force element for multibody system dynamics, the wheel–rail
contact problem is generally divided into three consecutive stages [2, 8–12]: (1) the con-
tact geometry stage, identifying the contact location and corresponding tangent plane; (2)
the kinematic stage, determining the undeformed distance and the creepages, i.e. the rela-
tive motion of wheel and rail surfaces, at the contact position; and (3) the contact mechanics
stage, computing the resultant normal and tangential forces. The background lies in the phys-
ical phenomena involved in wheel–rail contact, as schematized in Fig. 1. These phenomena
differ widely between normal and tangential directions, introducing a strong anisotropy in
the contact problem [13]. The third stage is divided further into two sub-stages, for the nor-
mal and tangential contact solution. For instance combining a Hertzian approach for the
normal contact with Fastsim or Polach's method for the tangential problem [14, 15].

The tangent plane provides a convenient interface between the quantities and coordinates
of the multibody system approach, on one side, and of rolling contact mechanics, on the
other. It provides the coordinate directions needed to define the contact problem: (a) the so-
called undeformed distance function, for the normal contact problem, and (b) the rigid slip
velocity, i.e. surfaces' relative motion, for the tangential one. These directions are used next
to express the solution: (c) the contact pressures, acting in normal direction, and (d) frictional
stresses, acting tangentially. That is, the tangent plane provides a further means for modular
programming, partitioning the wheel/rail contact into the preparations and post-processing,
versus the actual solution. The actual solution is then performed in isolation, in terms of
coordinates that are natural for the normal and tangential contact problems.
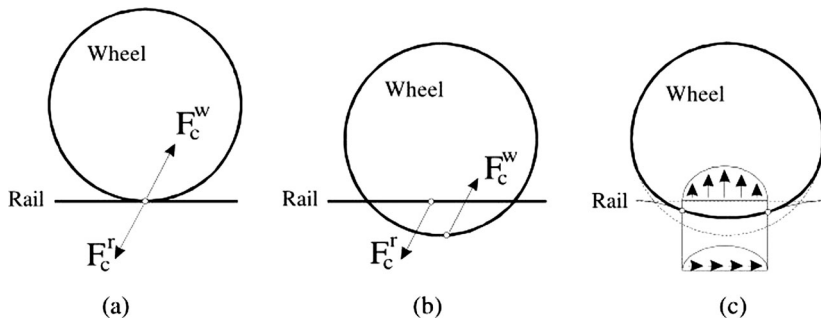
Fig. 2 Methods for the integration of wheel/rail contact in multibody systems simulation: (**a**) constraint method, (**b**) elastic method, (**c**) real, distributed configuration. Reprinted from [26, Fig. 1], with permission from Elsevier

However, the tangent plane does not allow one to capture contact problems fully. The method runs into difficulties for asymmetric, non-Hertzian problems, where the tangent plane can be defined in different ways, with different inclinations [16–18]. Which one represents the true contact interactions in the best way? Further complications arise in two-point contact situations, where the interaction between the two problems needs to be accounted for [19, 20], and where discontinuities occur when the main contact jumps from one point to the other [21], and in conformal contacts, where the curvature of the contact patch is inherent in the problem specification [22–25]. How can the true gap and rigid slip best be mapped from the curved surfaces onto a single plane?

Here we present the "curved reference surface", called curved reference for short, as an alternative backbone for the wheel/rail contact solution. This builds on the generalised curvilinear coordinates $[x, s, n]^T$ that were introduced for conformal contact situations [22, 23]. We show how these ideas are worked out in the automated analysis of the contact geometry problem in CONTACT. A key element is to use a form of object orientation, using "markers" to capture positions and orientations precisely, and "grids" to manipulate with curves and surfaces, with spline interpolation [26]. This software foundation is extended for the curved reference and curved contact surface. Next, the undeformed distance is computed relative to the local normal direction. The rigid slip velocity is computed separately at each potential contact position, by-passing the creepages that simplify the actual situation on the basis of a reference position that is open to question. Curvature of the contact surface in rolling direction is incorporated using a so-called "mid-plane correction", to make the rigid slip behave in the desired way, avoiding rigid slip at places where it is not expected. The contact problem is solved in the curved coordinate system, after which the resulting tractions are rotated and aggregated to the resultant total forces and moments.

The method using a curved reference surface provides a natural extension of the elastic approach as discussed by Shabana et al. [27], illustrated in Fig. 2 (b), which placed strong emphasis on the contact location and the contact forces as the primary concepts in the calculation. Acknowledging that contact arises on a contact area with finite extensions, a distributed force element may be considered for the solution of the contact problem. This lets us account for the local, distributed kinematics of the problem rather than using approximate values obtained from a single contact location, and defines the actions on the wheel and rail by the contact stresses, as shown in Fig. 2 (c), rather than using the contact forces as the central concept. The curved reference provides a means to approximate this real contact configuration. The roles of the contact location and contact forces are much reduced;

these are used in the software implementation, as a means to summarize the situation in a convenient way, as needed by the multibody simulation. Different from the original elastic approach, the contact point may then be selected after the contact solution, and does not affect the solution anymore.

The content of this paper is structured as follows. First, Sect. 2 gives a concise description of rolling contact problems. This shows the planar contact approach embedded in the wider mathematical context and indicates the places where extensions are made for the curved reference. The extensions are discussed in Sect. 3: presenting the overall structure of the computations and the detailed choices made in the proposed method. Numerical results are presented in Sect. 4, applying the new method to two-point and conformal contacts, and as a reference for benchmarking planar contact approaches. After this, the paper ends with conclusions and discussion.

2 Formulation of the contact problem

Here we describe the mathematical formulation of rolling contact problems, laying the foundations for the extensions presented in consecutive sections. We start from the viewpoint of a dynamical system, to show the embedding of contacts in a multibody formulation. This aims to give a better understanding of the undeformed distance and rigid slip functions, i.e. the quantities that are at the heart of the contact problem. These are related to the bodies' positions and deformations as seen from the multibody perspective.

Two coordinate systems are used: $[x, y, z]^T$ for a global system, with z the overall vertical direction normal to the track plane, positive downwards, and $[x, s, n]^T$ for local coordinates used to formulate the contact problem.

2.1 Undeformed reference state

We consider two elastic bodies, particularly a rail and a wheel, that undergo rolling with sliding motion. They are indicated by numbers a, with $a = 1$ for the rail, the upper body (z positive downwards), and $a = 2$ for the wheel. Following the Arbitrary Lagrangian Eulerian (ALE) descriptions given by Nackenhorst [28] and Donea et al. [29], we first identify the geometries of the bodies separately, without being in contact, named the initial state. Particles of the bodies are referred to by their coordinates $\mathbf{X}$ in the coordinate system that is used, with $\mathbf{X} \in \Omega_{ini}^{(a)}$, where the domain $\Omega_{ini}^{(a)}$ represents body a in its initial (undeformed) state.

Next, the bodies are brought together such that contact is formed. Here, too, we identify the geometry without considering deformations, the so-called undeformed reference state. This can be formalized using mappings $\mathbf{\Psi}^{(a)}$ for bodies $a = 1, 2$ that describes the reference position, $\mathbf{x}_{ref}$, of each particle $\mathbf{X}$ at the current time instance t.

$$\mathbf{\Psi}^{(a)} : \Omega_{ini}^{(a)} \times [t_0, t_{final}) \to \mathbb{R}^3, \quad \mathbf{x}_{ref} = \mathbf{\Psi}^{(a)}(\mathbf{X}, t), \quad \Omega_{ref}^{(a)}(t) = \mathbf{\Psi}^{(a)}(\Omega_{ini}^{(a)}, t). \quad (1)$$

These mappings are used to describe rigid body translation and rotation. An example concerns a wheelset that is moved from an arbitrary initial position to the current position $\mathbf{x}_{ws}$ and orientation $\mathbf{v}_{ws}$ (Euler angles). The reference position of each particle is obtained as

$$\mathbf{x}_{ref}(\mathbf{X}, t) = \mathbf{x}_{ws}(t) + \mathbf{R}(\mathbf{v}_{ws}(t))\mathbf{X}. \quad (2)$$

Here $\mathbf{X}$ is the position of a particle relative to the center of mass and $\mathbf{R}$ is the 3×3 rotation matrix corresponding to the orientation relative to the initial one.

Considering flexible multibody dynamics, the overall flexible bending of rails, axles, wheels, occurs at a much slower rate than the elastic deformation in and around the contact areas. "Global" deformations are therefore kept separate from the "local" contact solution. This is formalized by incorporating global deformations in the reference state. The mapping could be decomposed into parts, separating rigid body motion from flexible bending, as used in [30]. Here we focus on the combined result of these different motions.

2.2 Elastic displacements

The third state that is considered is the actual configuration at time t, including (local) deformation. This is described by further mappings $\boldsymbol{\Phi}^{(a)}$.

$$\boldsymbol{\Phi}^{(a)} : \{\Omega_{ref}^{(a)}(t), t\} \to \mathbb{R}^3, t \in [t_0, t_{final}), \quad \mathbf{x}_{act} = \boldsymbol{\Phi}^{(a)}(\mathbf{x}_{ref}, t). \tag{3}$$

This allows one to introduce local elastic displacements of all particles as the difference between actual and reference states:

$$\mathbf{u}^{(a)} : \{\Omega_{ref}^{(a)}(t), t\} \to \mathbb{R}^3, t \in [t_0, t_{final}),$$

$$\mathbf{u}^{(a)}(\mathbf{x}_{ref}, t) = \boldsymbol{\Phi}^{(a)}(\mathbf{x}_{ref}, t) - \mathbf{x}_{ref}(t). \tag{4}$$

The particles' velocities are obtained as the material time derivative of their actual positions. By Eq. (4), this may be split as follows:

$$\mathbf{v}_{act}^{(a)}(\mathbf{X}, t) = \frac{D\mathbf{x}_{act}^{(a)}}{Dt} = \frac{D\mathbf{x}_{ref}^{(a)}}{Dt} + \frac{D\mathbf{u}^{(a)}}{Dt} \tag{5}$$

The reference part comes from rigid body motion and (slow) overall flexible bending, which are assumed known in the solution of contact problems. The other part describes the change of elastic displacements of particles due to the contact. These are an important concern in frictional contacts with rolling and sliding.

2.3 Considerations on the normal direction

The bodies are brought into contact and, as a result, stresses $\boldsymbol{\sigma}$ (tensor), strains $\boldsymbol{\epsilon}$ (tensor) and displacements $\mathbf{u}$ (vector) arise in the bodies and at their surfaces.

The usual boundary conditions for a single body are to prescribe displacements $\mathbf{u}^{(a)} = \bar{\mathbf{u}}^{(a)}$ on one part of the boundary $\Gamma_u^{(a)} \subset \partial\Omega^{(a)}$ and stresses $\boldsymbol{\sigma}^{(a)}\mathbf{n}^{(a)} = \bar{\mathbf{p}}^{(a)}$ at another part $\Gamma_\sigma^{(a)} \subset \partial\Omega^{(a)}$. This does not work at contacts because the two bodies form each others boundary conditions. Neither $\bar{\mathbf{u}}$ nor $\bar{\mathbf{p}}$ is known, so that these cannot be prescribed in advance.

In the study of contact mechanics, the stresses $\bar{\mathbf{p}}$ acting at the outside of a continuous, deformable body, are often called "tractions" or "surface tractions" [13, 31]. That is, pressures are also called "normal tractions", and what are called "tractions" in other fields of study are here referred to as "tangential tractions". This stems from the central role of surface stresses in contact interactions, by which it is convenient to have a short term for their designation.

Contact conditions are introduced to complete the specification of the problem of the bodies' deformation. This starts from regions $\Gamma_c^{(a)} \subset \partial\Omega^{(a)}$ that include all points on the two bodies that may get in contact. One body may then be designated as the master body, using its local surface normals $\mathbf{n}^{(a)}(\mathbf{x})$ to identify corresponding points on the other body. Having established such a relation, the so-called gap function can be defined, as well as the mutual velocity of the two points.

The main difficulties of such a scheme become clear from an example, considering two soft equal spheres that are pressed together. A flat contact area is formed that may be relatively large compared to the size of the spheres. In this example, large deformations occur that introduce considerable differences between the surface normals $\mathbf{n}_{ref}$, $\mathbf{n}_{act}$ in reference and actual configurations. This complicates the formalization of the gap function: should this be measured in the initial or in the final direction? Furthermore, this also complicates the mapping of the two surfaces onto each other. How can we locate the corresponding point on the other body if the surface normals change considerably by deformation?

Generic solutions to these problems have been established; see for instance [32–34] for strategies used in finite-element methods. Instead, we require that the normal and tangential directions are defined a priori to the contact solution: identical for all surface positions when using a planar contact approach, or space-varying when using the conformal method.

2.4 Introducing a curved reference surface

In formulating the contact problem, we are particularly interested in the surface quantities. According to the anisotropy of contact problems (Fig. 1), this needs the local normal and tangential directions. But the actual surface normals, after deformation, become known only after the contact solution. This then calls for iteration, building the solution in progressively finer detail, or approximation, ignoring certain effects on the surface normals.

Assuming that elastic displacements and displacement gradients are small, the surface normals cannot change much between reference and actual configurations [35, p. 3]. This allows one to use the surface normals of the reference configuration as a basis for the contact solution. Further, if the contact area is small compared to the local dimensions of the contacting bodies, that is, if the contact is concentrated, the contact will be effectively flat, contained in a plane, permitting a planar approximation. This latter assumption will now be relaxed, forming a curved reference surface on the basis of the undeformed configuration.

A curved reference surface is introduced in between the wheel and rail surfaces, with surface parameters s_1 and s_2 (cf. [2]):

$$\mathbf{x}_{cref} : \mathcal{S} \rightarrow \mathbb{R}^3, \ \mathcal{S} \subset \mathbb{R}^2, \ \mathbf{x}_{cref} = \mathbf{x}_{cref}(s_1, s_2). \tag{6}$$

At each position $\mathbf{x}_{cref}$, the tangent directions are found using

$$\mathbf{x}_{,1} = \frac{\partial \mathbf{x}_{cref}(s_1, s_2)}{\partial s_1}, \ \ \mathbf{x}_{,2} = \frac{\partial \mathbf{x}_{cref}(s_1, s_2)}{\partial s_2}, \tag{7}$$

and the local normal to the surface is

$$\mathbf{n}(s_1, s_2) = \frac{\mathbf{x}_{,1} \times \mathbf{x}_{,2}}{|\mathbf{x}_{,1} \times \mathbf{x}_{,2}|}. \tag{8}$$

It is assumed that the curved reference lies close to the surface where the actual contact occurs, and that its normal vector $\mathbf{n}$ is close to the surface normals of rail and wheel, before or after deformation. The curved reference is then used in place of the surface in which the actual contact area is contained, and the normal vector $\mathbf{n}$ of Eq. (8) is used instead of the surface normals of rail and wheel. The method may be applied with different strategies for the construction of the curved reference. The better this approaches the actual contact, the smaller will be the errors of approximation.

Next, the boundaries $\partial \Omega_{ref}^{(a)}$ are projected onto the curved reference surface. This gives (s_1, s_2)-coordinates for each boundary point on the wheel and the rail. This mapping can be inverted in the contact area and its surroundings, by which each (s_1, s_2) maps to a single

point on the rail and wheel surfaces.

$$\mathbf{x}_{ref}^{(1)} = \mathbf{x}_{ref}^{(1)}(s_1, s_2), \quad \mathbf{x}_{ref}^{(2)} = \mathbf{x}_{ref}^{(2)}(s_1, s_2). \tag{9}$$

By this projection, quantities like the displacements $\mathbf{u}^{(a)}$ or the tractions $\mathbf{p}^{(a)}$ that are defined on the actual surfaces may be considered as functions defined on the curved reference surface instead.

2.5 Contact conditions

We consider the positions and velocities for two particles that are in contact with each other in the actual configuration. Firstly, their distance $\mathbf{x}_{ref}^{(1)} - \mathbf{x}_{ref}^{(2)}$ in the reference state prescribes the difference of their elastic displacements $\mathbf{u} := \mathbf{u}^{(1)} - \mathbf{u}^{(2)}$ in the actual state. This yields a constraint in the normal direction, concerning the gap or undeformed distance of the profiles, with negative values for interpenetration.

$$\text{undeformed distance}: h_n := \mathbf{n}^T(\mathbf{x}_{ref}^{(1)} - \mathbf{x}_{ref}^{(2)}), \tag{10a}$$

$$\text{deformed distance}: e_n := \mathbf{n}^T(\mathbf{x}_{act}^{(1)} - \mathbf{x}_{act}^{(2)}) = h_n + \mathbf{n}^T\mathbf{u}. \tag{10b}$$

Concerning the velocities, we may measure and compare the velocity differences in the reference and actual states. These are called the rigid slip velocity and actual slip velocity, denoted by $\mathbf{w}, \mathbf{s}$. Here we are interested primarily in the components in tangential directions:

$$\text{rigid slip velocity}: \mathbf{w}_a := [\mathbf{t}_x, \mathbf{t}_s]^T(\mathbf{v}_{ref}^{(1)} - \mathbf{v}_{ref}^{(2)}), \tag{11a}$$

$$\text{actual slip velocity}: \mathbf{s}_a := [\mathbf{t}_x, \mathbf{t}_s]^T(\mathbf{v}_{act}^{(1)} - \mathbf{v}_{act}^{(2)}) = \mathbf{w}_a + [\mathbf{t}_x, \mathbf{t}_s]^T\frac{\mathrm{D}\mathbf{u}}{\mathrm{D}t}. \tag{11b}$$

A subscript $[]_a$ is added to indicate that these are absolute dimensioned values (e.g. [mm/s]) rather than relative dimensionless values ($[-], [\%]$). Equation (11b) may be scaled by an overall rolling velocity V, replacing absolute slips by relative slips $\mathbf{w}_r, \mathbf{s}_r$. This brings in that many rolling contacts are insensitive to the velocity V, producing the same tractions as long as the same relative rigid slip is maintained. This is mostly a matter of presentation, the same solutions may be computed with or without such scaling.

Using these preparations, the contact conditions may be stated in their customary form [31, 35, 36]:

$$\text{normal problem}:$$

$$\text{in exterior area } E: \quad e > 0, \ p_n = 0, \tag{12a}$$

$$\text{in contact } C = H \cup S: \quad e = 0, \ p_n \geq 0, \tag{12b}$$

$$\text{tangential problem}:$$

$$\text{in exterior area } E: \quad \mathbf{s}_a \text{ free}, \ \mathbf{p}_t = \mathbf{0}, \tag{12c}$$

$$\text{in adhesion area } H: \quad \|\mathbf{s}_a\| = 0, \ \|\mathbf{p}_t\| \leq \mu p_n, \tag{12d}$$

$$\text{in slip area } S: \quad \|\mathbf{s}_a\| > 0, \ \mathbf{p}_t = -\mu p_n \mathbf{s}_a/\|\mathbf{s}_a\|. \tag{12e}$$

Here $\mathbf{p} = \mathbf{p}^{(1)}$ stands for the surface tractions acting *on* body 1, $\mathbf{p}^{(2)} = -\mathbf{p}$, with p_n the component in (local) normal direction and $\mathbf{p}_t = [p_x, p_s]^T$ the tangential parts. Equations (12a) and (12b) state that the normal pressure is compressive and that there is no interpenetration

between the bodies. Equations (12d) and (12e) correspond to Coulomb's law of dry friction. That is, the tangential tractions cannot exceed the local traction bound μp_n, slip occurs only where the traction bound is reached, and if there is slip, the tractions act in the opposite direction.

The coefficient of friction μ is often considered as a known input to the simulation, optionally with independent values at the top and side of the rail. The values used depend on circumstances like contamination (e.g. wear debris, clay, leaves) and applied agents such as friction modifiers and flange lubrication. On-going research is considering whether this model suffices and if so, how μ comes about, modeling the effects of third body layers, roughness, fluids, temperature, etc. on the creep forces [37–42].

2.6 Planar contact approach

Three more steps remain to complete the elastic, planar contact approach for contacts in rail vehicle dynamics simulation: (1) defining the tangent plane for the computations, i.e. filling in the normal direction, (2) defining the creepages, thereby specifying the rigid slip, and (3) solution or approximation of the resulting, coupled elasticity problems. The extension to conformal contacts is presented in Sect. 3.

The identification of an appropriate tangent plane was discussed in previous publications [17, 18, 26]. This begins with an analysis of the wheel and rail geometries in their undeformed states, locating the regions where contact may occur. For each region one can look at the so-called initial contact position, the point with maximum interpenetration, and define the contact plane tangent to the surfaces at this position. This is the simplest initial contact method. One extension is to set the inclination for the tangent plane in some averaged way. Another alternative is to base the tangent plane off a different reference location, typically using some form of weighted averaging [16, 26]. This is called the weighted center method.

Weighted averaging can be worked out in different ways: using the horizontal interpenetration area, the vertical interpenetration area at the contact locus, the 3d interpenetration volume, using estimated contact pressures, and so on. Here we employ weighting using linear or quadratic functions of the vertical gap values $g(x, y)$ as shown in Eqs. (13) and (14) [26],

$$[x, y, \alpha]_{cwgt1} = \iint_{g<0} [x, y, \alpha] \, g(x, y) \, dx \, dy \Big/ \iint_{g<0} g(x, y) \, dx \, dy, \qquad (13)$$

$$[x, y, \alpha]_{cwgt2} = \iint_{g<0} [x, y, \alpha] \, g^2(x, y) \, dx \, dy \Big/ \iint_{g<0} g^2(x, y) \, dx \, dy. \qquad (14)$$

The different methods composed from this are defined in Table 1. The motivation for using squared gaps is that they concentrate more on the bigger gap values. This appears to correlate better to the contact pressures than the linear gap values.

The rigid slip velocity $\mathbf{w}_a$ of Eq. (11a) is a field that describes the tangential velocity difference of the two surfaces at every point in contact. This field is usually simplified, considering the velocity difference at "the" contact location. Three creepages ξ, η, ϕ are then defined for longitudinal, lateral, and spin motions. Longitudinal and lateral creepages are found as the scaled velocity difference at the contact position, while spin creepage represents the scaled relative rotation about "the" normal direction. Generic formulas may be derived from the velocity of points on rigid bodies, using the time derivative of Eq. (2); see for instance [11, 12] or [43]. These formulas may also be expanded using the wheelset dimensions and orientation, to get creepages in terms of railway-specific terms, e.g. [43–45].

Table 1 Computational methods considered in the statistical evaluation of planar contact approaches in Sect. 4.2.3

nr.	method	type	reference position	reference angle	weighting
0.	curved	curved reference	weighted center	weighted	quadratic
1.	crv-icp	curved reference	initial contact	sampled	–
2.	wgt-loc-1	planar contact	weighted center	sampled	linear
3.	wgt-loc-2	planar contact	weighted center	sampled	quadratic
4.	wgt-wgt-1	planar contact	weighted center	weighted	linear
5.	wgt-wgt-2	planar contact	weighted center	weighted	quadratic
6.	icp-loc	planar contact	initial contact	sampled	–
7.	icp-wgt-1	planar contact	initial contact	weighted	linear
8.	icp-wgt-2	planar contact	initial contact	weighted	quadratic

When using creepages, the scaled rigid slip velocity is presented in a compact, yet approximate way, i.e.

$$\mathbf{w}_r(\mathbf{x}) = \begin{bmatrix} \xi - \phi s \\ \eta + \phi x \end{bmatrix}. \tag{15}$$

Here $\mathbf{x} = [x, s]^T$ is a position in the tangent plane, defined relative to the local origin used, i.e. the contact reference position [26]. A complete derivation was presented by Kalker [31, §1.7.3]. This shows that small terms are neglected compared to the complete rigid slip of Eq. (11a).

Historically, the notion of rigid slip did not appear until the end of the 1970-ies. No symbol was used for this in older work by Kalker, writing out formulas like $s_x = v_x - \phi y + \partial u/\partial x$ [46, Eq. (1.6a)], using v_x instead of ξ. This notation was maintained by Johnson in 1985, e.g. writing $\xi_x - \psi y/c$ in [47, Eq. (8.3a)]. This work used the symbol $\mathbf{s}$ for the actual slip.

The term rigid slip seems to have been introduced with the variational theory for rolling contacts [48, 49]. This placed more emphasis on the (undeformed) reference state and its elastic deformation, and for rolling, on the motion of separate particles instead of the bodies as a whole. New symbols were introduced, inconsistently, denoting the rigid and actual slips as $\mathbf{r}, \mathbf{s}$ in [50], while using $\mathbf{s}, \mathbf{w}$ in [51]. The background of the latter is found in a paper in German [52], where rigid slip is translated as "*starre Schlupf*" $\mathbf{s}$ and actual slip as "*wahre Schlupf*" $\mathbf{w}$. Kalker used this notation up to about the paper introducing Fastsim [53]. The roles of $\mathbf{s}$ and $\mathbf{w}$ have been interchanged thereafter, denoting the rigid slip as $\mathbf{w}$, as followed here in this paper.

2.7 Elasticity in the contact problem

The description of the contact mechanics problem is completed by presenting the laws for the material behavior. Adopting small displacements theory, these are generically described as

$$\rho^{(a)}\ddot{\mathbf{u}}^{(a)} = \nabla \cdot \boldsymbol{\sigma}^{(a)} + \mathbf{f}^{(a)}, \tag{16a}$$

$$\boldsymbol{\sigma}^{(a)} = \mathsf{C}^{(a)} : \boldsymbol{\epsilon}^{(a)}, \tag{16b}$$

$$\boldsymbol{\epsilon}^{(a)} = \frac{1}{2}\left[\nabla\mathbf{u}^{(a)} + (\nabla\mathbf{u}^{(a)})^T\right]. \tag{16c}$$

Here $\rho^{(a)}$ is the mass density, $\mathbf{f}^{(a)}$ represents body forces, and $\mathsf{C}^{(a)}$ is a fourth-order stiffness tensor. Equation (16a) is the equation of motion (Newton), Eq. (16b) describes the material behavior according to Hooke's law, and Eq. (16c) is the linearised strain–displacement relation.

These equations are assumed to hold everywhere in the bodies' interiors $\mathbf{x}_{act} \in \Omega_{act}^{(a)}$, where suitable boundary conditions are required on the boundaries $\partial\Omega^{(a)}$. Considering a single body, this gives rise to initial and boundary value problems, amenable to finite-element solution. In-depth information is provided by Laursen [32], for instance, including the extensions to large deformations and to include contacts into the formulation.

Instead of approaching the elasticity equations (16a)–(16c) directly, using finite elements, it pays off to first reformulate the problem. One main approach for this is to use a boundary element method. This method is based on the concept of influence functions $\mathbf{A}(\mathbf{x}, \mathbf{x}')$, which establish a relation between the displacements $\mathbf{u}$ at surface position $\mathbf{x} = [x, s, 0]^T$ and a tractions $\mathbf{p}(\mathbf{x}')$ at another surface position $\mathbf{x}' = [x', s', 0]^T$. For linearly elastic behavior, the full solution $\mathbf{u}$ can then be composed by superposition:

$$\mathbf{u}^{(a)}(\mathbf{x}) = \int_{\mathbf{x}' \in C} \mathbf{A}^{(a)}(\mathbf{x}, \mathbf{x}')\mathbf{p}^{(a)}(\mathbf{x}')\, dx'\, ds' \tag{17}$$

This method considers stresses and deformations at the surfaces only, avoiding the discretization of the interiors of wheel and rail bodies.

The influence functions depend on the geometry and material parameters of the two bodies. They can be computed numerically and stored in a file [22, 25], or can be calculated analytically in certain specific cases. The latter is particularly convenient for concentrated contact problems that permit using the halfspace approach. Considering homogeneous, isotropic, linearly elastic materials, and neglecting inertial effects, which are assumed to be small compared to the contact stresses [54], the influence functions can be calculated exactly from the formulas of Boussinesq and Cerruti [35, 55].

The equations are discretized in CONTACT using a rectangular potential contact area Γ_c that needs to encompass the actual contact area C, discretized into $m_x \times m_s$ elements of size $\delta_x \times \delta_s$. Subscripts i, j are used for grid point numbers in x- and s-directions. For the sake of simplicity, the indices i, j are merged into a single index $I = i + m_x \cdot (j - 1)$.

The continuous tractions $\mathbf{p}$ are approximated by piecewise constant or piecewise bilinear functions [55]. Purpose-built solvers were constructed for the resulting numerical problems, i.e. NormCG for the normal problem [56] and TangCG, SteadyGS [57, 58] for the tangential problem. These are based on Conjugate Gradients and the Gauss–Seidel method, extended with active set strategies for dealing with the contact conditions (12a)–(12e), and using FFTs for fast computation of structured matrix-vector products.

2.8 Heuristic approaches

The boundary element approach implemented in CONTACT yields a complete, "exact" treatment of wheel–rail contacts at sufficiently flat sections the wheel and the rail. With the increasing speed of computers, this can now be integrated in multibody dynamics simulation, as shown by recent publications [42, 59, 60]. This provides benefits in a number of special situations where a high level of detail is needed, such as computations of locomotive traction power including the effects of third body layers, flange-climb derailment with a

conformal geometry, wheel-flats or corrugation, studies of wear and RCF damage that need detailed contact stresses, or validation of more simplified contact approaches.

However, even though CONTACT may be thousands of times faster than using the finite-element method, it remains rather slow for vehicle dynamics simulations considering longer stretches of track, where millions or billions of contact problems must be solved. Fast approximations are therefore preferred for day-to-day work, in most cases, as long as one is not too critical on the results.

Many methods have been presented for the modelling of the contact forces [14, 15], which work by simplifying the full contact model in one way or another. Normal contact was analyzed by Hertz in a seminal paper [61], providing a detailed solution for frictionless contacts for surfaces that have constant curvatures in the vicinity of the contact [31, 47, 62]. Under these assumptions, an elliptical contact arises with a semi-ellipsoidal pressure distribution. This has long served as a reasonable approximation in railway vehicle dynamics simulation, and is provided still in several simulation packages today. One main reason is that the Hertzian approximation brings the tangential contact problem into a standard form, simplifying its approximation [14].

Since real wheels and rails have varying curvatures along the width of the profiles, there has been much work aiming to extend the normal contact to non-Hertzian geometries and non-elliptical contact patches. Heuristic methods were proposed using virtual interpenetration, multi-Hertzian and semi-Hertzian strategies [9, 15, 16, 20, 63–66]. These generally work by assuming a correspondence between the actual contact area and the interpenetration area, and a pressure distribution analogous to the Hertizan pressures. Iterative solution of the elasticity equations, as done by NormCG, is thereby avoided and replaced by a much reduced set of equations, dependent on the heuristic that is used. Various extensions were published in recent years; see e.g. [12, 67–70], addressing for instance the effect of a non-zero yaw angle. More work needs to be done on how good these methods perform, in terms of speed and accuracy, in a wide range of circumstances.

The tangential creepage and creep force are related to each other in a complex, non-linear way. In vehicle dynamics simulation, this is approached with different heuristics by Vermeulen and Johnson [71], Shen–Hedrick–Elkins [72], Polach [73, 74], Bosso et al. [75], or by using Fastsim [36, 53], or by building a table on the basis of CONTACT's results [76]. Most of this work is based on an assumption of an elliptical contact. Solution of the elasticity equations is avoided by adopting a formula, based on certain theoretical considerations (Vermeulen–Johnson, Shen–Hedrick–Elkins, Bosso et al.), or by adopting a simplified material law instead of Eq. (16a)–(16c) (Fastsim), or by analytical approximation of the results of the Fastsim approach (Polach).

Only few works considered the creep forces in non-Hertzian conditions. An extension of Fastsim is provided by the FaStrip algorithm [77], but this does not seem to be picked up widely. The table-based approach is also extended to non-Hertzian circumstances [78, 79]. Fastsim itself can be employed if a strategy is defined to get the flexibilities needed. The main strategy here is using an equivalent ellipse [66], an alternative route is presented by Alonso and Giménez [80]. Further work is needed to develop and test algorithms for non-Hertzian circumstances.

3 Extensions for conformal contacts

This paper presents the extension of the above approach for planar contact by combination with the ideas for conformal contacts as presented in [22, 23]. This revolves around
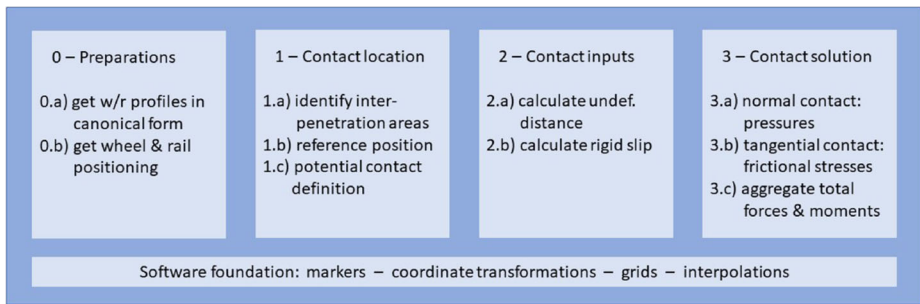
0 – Preparations	1 – Contact location	2 – Contact inputs	3 – Contact solution
0.a) get w/r profiles in canonical form	1.a) identify inter-penetration areas	2.a) calculate undef. distance	3.a) normal contact: pressures
0.b) get wheel & rail positioning	1.b) reference position	2.b) calculate rigid slip	3.b) tangential contact: frictional stresses
	1.c) potential contact definition		3.c) aggregate total forces & moments

Software foundation: markers – coordinate transformations – grids – interpolations

Fig. 3 Steps used for the wheel/rail contact solution

the added value provided by the generalized curvilinear coordinates $[x, s, n]^T$ that are extended here to the curved reference surface. These coordinates replace the Cartesian contact coordinates $[x_{cp}, y_{cp}, z_{cp}]^T$. The curved reference replaces the tangent plane as the central backbone for the computations, decoupling actual contact problem from the preparations and post-processing, and decoupling the normal from the tangential contact solution. The computations are generalized further to the 3-d contact search procedure as presented in [26] and are automated using the mechanisms presented in there.

Section 3.1 presents the overall solution approach. Section 3.2 introduces the curved reference, while Sect. 3.3 describes its computation. The extension of undeformed distance calculations and rigid slip velocities are presented in Sects. 3.4, 3.5 and 3.6, the latter introducing a new mid-plane correction. Section 3.7 finally addresses the aggregation of resulting forces and moments.

3.1 Overall solution procedure

In a separate paper [26], we presented the analysis of contact geometry using the planar contact approach. This is built on an object oriented-like software foundation for markers, grids and coordinate transformations. The main steps of the method are summarized in Fig. 3, with steps 0.b to 2.a illustrated graphically in Fig. 4.

The method starts by reading wheel and rail profiles from various sources, within their own coordinate systems, making corrections and preparations to bring them into a canonical form. Figure 4 (a) displays placing of these profiles in a common coordinate system, the track frame (tr), using markers $\mathbf{m}_{rw(tr)}$ and $\mathbf{m}_{rr(tr)}$ that connect the profiles to the bodies representing the wheel and the rail, governing their kinematics. This uses geometrical and kinematical information from a half wheelset on a half track or a half pair of rollers. The rail is considered prismatic in rolling direction, whereas wheels and rollers are assumed bodies of revolution. Track irregularities and rail and wheelset flexibilities are included as provided by the multibody model.

The location of contact patches is determined using the interpenetration in overall vertical direction (Fig. 4 (b)). A full 3-d contact search is performed including effects of the yaw angle, building the wheel surface in wheel coordinates first, then transformed to the track system. Interpenetration areas can be determined using a brute force grid-based approach or using a refined version of the contact locus [26]. Interpenetration areas are described by the interval $[y_{sta}, y_{end}]$ in Fig. 4 (b) and by the length $[x_{sta}, x_{end}]$ (not shown). These may be larger but not smaller than the actual interpenetration area, as required by the potential contact area. In the software, the name "contact patch" is used as a synonym for "independent contact problem", storing all corresponding information.
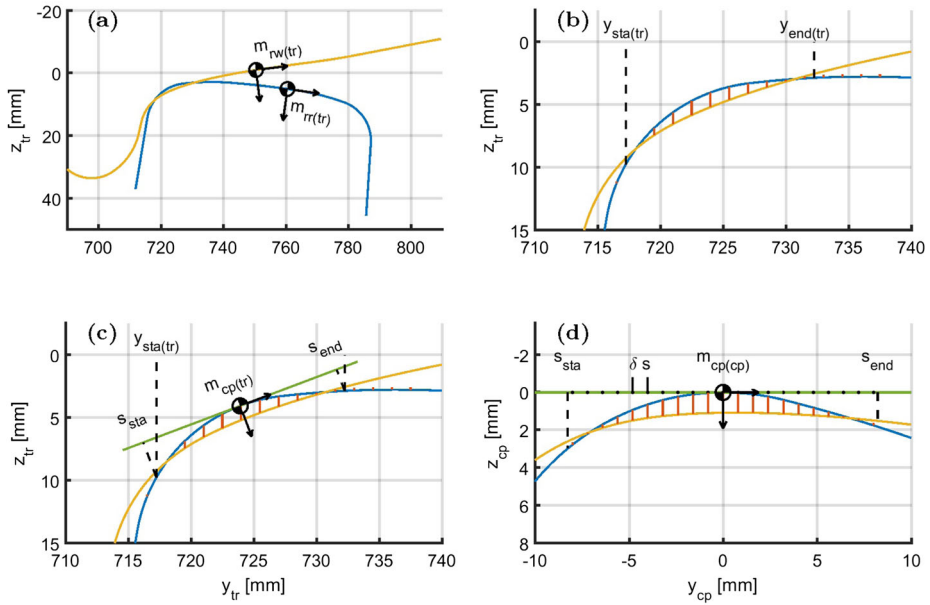
Fig. 4 The contact location problem and potential contact definition (planar contact). (**a**) For a given wheel and rail, at given positions, including the yaw angle (not shown), (**b**) determine the regions where interpenetration occurs. This uses the overall vertical direction, and a 3-d search procedure [26]. (**c**) Defining contact local coordinates and an appropriate potential contact area. (**d**) Discretization and undeformed distance calculation in normal direction

The contact reference position and reference angle are computed using one of the approaches of Sect. 2.6. This yields the contact marker $\mathbf{m}_{cp(tr)}$ and contact plane as illustrated in Fig. 4 (c). The interval $[y_{sta}, y_{end}]$ is projected onto this plane to get the corresponding extent $[s_{sta}, s_{end}]$ in the contact coordinate system. Adjacent contact patches may be taken together when the distance between them is less than a threshold, using a blending approach to mitigate jumps in the numerical results at perturbation of the input values [26]. From there, the discretization is formed, the undeformed distance and creepages are computed on each contact patch (Fig. 4 (d)), and the contact problems are solved using the halfspace approach, using the discretization and solvers as discussed in Sect. 2.7.

3.2 Introducing the curved reference surface

A curved reference surface is now introduced for conformal contacts as illustrated in Fig. 5, using the curved coordinate s in the tangential direction and the curved coordinate n in the normal direction. A main characteristic of this reference surface is that it uses a local construction, producing separate curved reference grids for contact patches that are computed separately, instead of one curved reference for all contact patches together. The reason for this is illustrated in Fig. 6: the construction of a single, global curved reference is complicated for geometries that have multiple, unconnected interpenetration zones with bigger positive gaps in between. In such cases we cannot maintain correspondence between arc lengths $\tilde{s}$, s_r, s_w of the different surfaces, requiring a more complex administration.

A second characteristic of our curved reference is that we consider the curvature in lateral direction only. Actual contact surfaces can be curved in longitudinal direction too, for

Fig. 5 Curved reference surface (green, dotted) for a contact patch, with varying normal and tangential directions
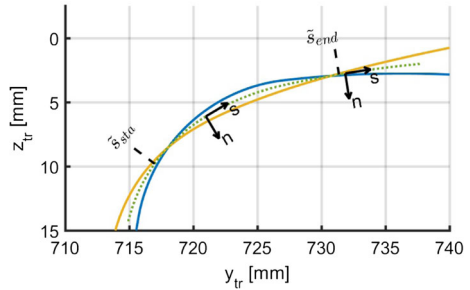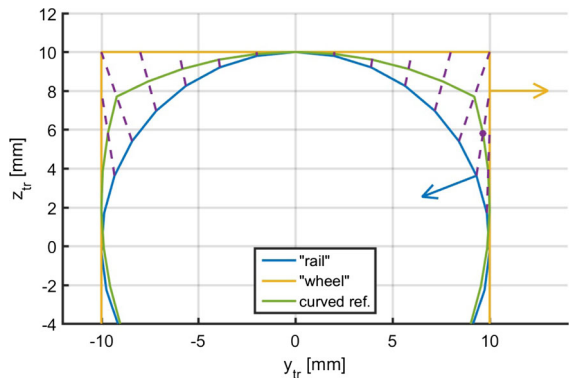

Fig. 6 Curved reference plane: the averaging of $(\mathbf{x}_r + \mathbf{x}_w)/2$ goes wrong if two surfaces separate and join again

instance a raceway in a ball bearing, or for a pin rolling in a hole. This curvature in longitudinal direction affects the rigid slip in wheel–rail contacts, as will be explained more in detail below. Yet to consider this would make our software more complicated. A prismatic curved surface can be represented by a 1-d grid (curve) in the Oyz-plane, and the transformation to Cartesian coordinates involves just a single rotation. This makes intermediate and final results of the program easier to understand and work with in further computations.

A beneficial result of using locally defined curved references is that they fit nicely in the overall structure as shown in Fig. 3. The steps for the planar approach are all reused, with relatively minor extensions to accommodate conformal contacts. The steps are modified as follows:

1.c' The potential contact area is defined on the curved reference surface.
2.a' The interpenetration of the undeformed surfaces is measured in the local normal direction of the reference surface, i.e. using the local orientation of the normal instead of a constant normal direction defined for the whole contact area;
2.b' The apparent slipping of the undeformed surfaces (rigid slip) is measured along the local tangent vectors instead of using a single tangent plane for the entire contact area;
3.a', 3.b' The true elastic deformations of the wheel and rail can be modelled using influence functions for non-planar bodies, e.g. computed by finite-element methods, instead of using the half-space approximation. This extension is skipped here. The development of influence functions is discussed in [81, 82].
3.c' The surface tractions are defined using the local normal and tangent directions, and this is accounted for in the accumulation of the total forces and moments.
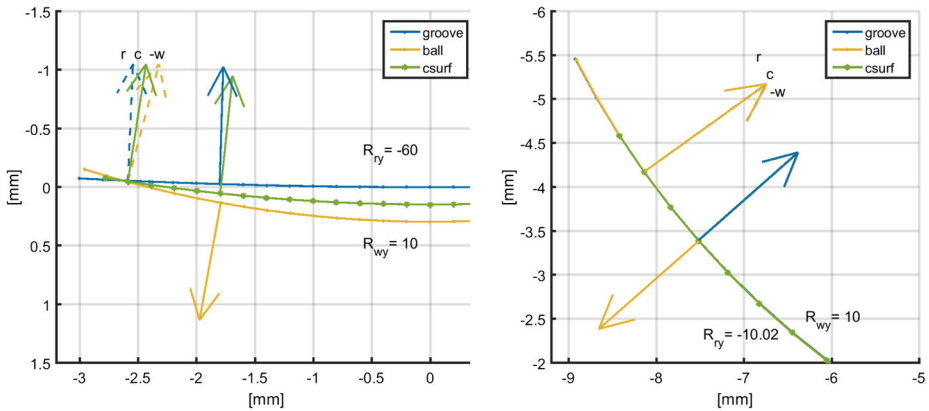
Fig. 7 Computed curved reference surface and surface normals for a circular profile with $R_{wy} = 10$ mm on a concave, grooved profile with $R_{ry} = -60$ mm (left) and $R_{ry} = -10.02$ mm (right)

3.3 Step 1.c', Defining the curved contact surface

The curved reference serves as the potential contact area for the contact solution. Using rectangular discretization elements, this surface should have fixed grid sizes δx and δs, with coordinates $\mathbf{x}_l$ following the curved rail and wheel surfaces. Using the rail as the basis for the coordinate system, this grid will be defined by the positions $\{x_i\} \times \{y_j, z_j\}$. Corresponding information can then be stored using 1-d grids and grid-functions as defined in [26, §2.6].

A simple approach for the curved reference would be to simply follow the rail profile. This may work well as long as the wheel and rail profiles stay in close proximity to each other, but introduces slight deviations in the surface normals. This is illustrated in Fig. 7 for circular profiles. Note that the left figure shows an extreme situation, with a contact width of 5 mm on a ball diameter of 20 mm. For steel bodies, such a contact width may be realistic only if the rail profile is grooved much stronger. In that case the surface normals will stay much closer together, as shown in the right part of the figure.

A more refined approach is used that is based on the dominant effects in the elastic deformations. For equal material laws and equal material parameters, the actual contact lies halfway between the undeformed wheel and rail, and the surface normal is averaged as well. This ignores the effect of the geometry of the solids on the deformations, which is also ignored when using the halfspace approach.

In case of material dissimilarity, the notion "in between" the two surfaces may be refined using appropriate scaling. The direct effect of pressures p_n on normal displacements scales with $(1 - \nu)/G$. This is incorporated in the mean surface using

$$\mathbf{x}_m = \theta^{(1)}\mathbf{x}^{(1)} + (1 - \theta^{(1)})\mathbf{x}^{(2)}, \text{ with} \tag{18a}$$

$$\theta^{(1)} = \frac{(1 - \nu^{(1)})G^{(2)}}{(1 - \nu^{(1)})G^{(2)} + (1 - \nu^{(2)})G^{(1)}} \approx \frac{u_n^{(1)}}{u_n^{(1)} + u_n^{(2)}}. \tag{18b}$$

If one body is rigid, this makes the curved reference follow the surface of this body. This may be relevant for instance in the simulation of a linear guiding system or a roller coaster, when using different materials for the rail and the rollers.

A straightforward implementation of this idea was presented in [22], using the chord distances s_r and s_w along the rail and wheel profiles, evaluated at regular steps $j \cdot \delta s$ relative

to a common reference, then computing the mean surface $\mathbf{x}_m = (\mathbf{x}_r + \mathbf{x}_w)/2$. This method starts from wheel and rail profiles defined in the same plane, i.e. ignoring the yaw angle. This procedure has now been extended to the 3-d contact search procedure that is used in CONTACT.

The extended approach starts from the contact reference position $\mathbf{x}_{cp(tr)}$, in track coordinates, as defined by the contact search procedure. The rail profile is sampled at regular steps $s_{cp(r)} + j \cdot \delta s$, for a suitable range of j to encompass the whole interpenetration area. The difficult part is then to find the opposite points on the wheel profile, recognizing that the wheel particles found at x-coordinate $x_{cp(tr)}$ do not lie in the principal wheel profile, at $x_{(w)} = 0$ in the wheel profile coordinate system. This is resolved by an iterative calculation:

– Set $s_{j(w)} = s_{cp(w)} + j \cdot \delta s$ and set initial estimate $x^0_{j(w)} = x_{cp(w)}$ for iteration $k = 0$,
– Compute the wheel surface positions $\mathbf{x}^k_{j(w)}$ in wheel coordinates, at surface parameters $(s_{j(w)}, x^k_{j(w)})$.
– Transform the positions from wheel to track coordinates using the generic "to global" operation.
– Determine the mismatch $\Delta_j = x^k_{j(tr)} - x_{cp(tr)}$, set $x^{k+1}_{j(w)} = x^k_{j(w)} - \Delta_j$, $k = k + 1$, and iterate as needed.

This builds strongly on the software foundation as described in [26] for manipulating with profiles and surfaces (grids), conversion between different coordinate systems, and easy spline interpolation. For instance, tangent vectors along the wheel and rail surfaces are obtained as $[0, \partial y/\partial s, \partial z/\partial s]^T$, directly from the spline representation. These are rotated 90° to form the surface normals.

It may be argued that the steps δs in the wheel surface get shortened a little by the yaw angle, dependent on the local surface inclination. This was found to be negligible in our calculations, and is neglected in the curved reference surface.

3.4 Step 2.a', Undeformed distance in local normal direction

A convenient way to compute the undeformed distance for conformal contacts is to re-use the existing computations for planar contacts and correcting for the change in surface orientation. Denoting variables for planar contact by an overbar, such as the reference angle $\bar{\delta}$, this gives

$$h_j = \bar{h}_j \cdot \cos(\delta_j - \bar{\delta}). \tag{19}$$

This is illustrated in Fig. 8, left, for a configuration where $\bar{\delta} = 0$, $\delta_j = -60°$, with $\cos(\delta_j - \bar{\delta}) = 0.5$. The interpenetration in local normal direction is shorter than the distance in vertical direction. The size of this effect is computed by cosine of the enclosed angle. This is accurate as long as the interpenetration $\bar{h}_j$ is small, such that the angle δ does not vary much in the considered region.

Figure 8, right, shows a test-case used for the validation of these computations. A deep groove is used with curvature $r_{groove} = -10.01$ mm and bottom point at $[0, -0.02]^T$. A ball with radius $r_{ball} = 9.99$ mm is placed at $[0, -9.97]^T$, putting the lowest point at $[0, +0.02]^T$, at 0.04 mm interpenetration. The case is constructed such that the planar contact approach uses Oxy as the tangent plane, and the curved reference is circular with radius $r = 10$ mm and center at $[0, -10]^T$. This allows all points and distances to be evaluated manually and check the computations.
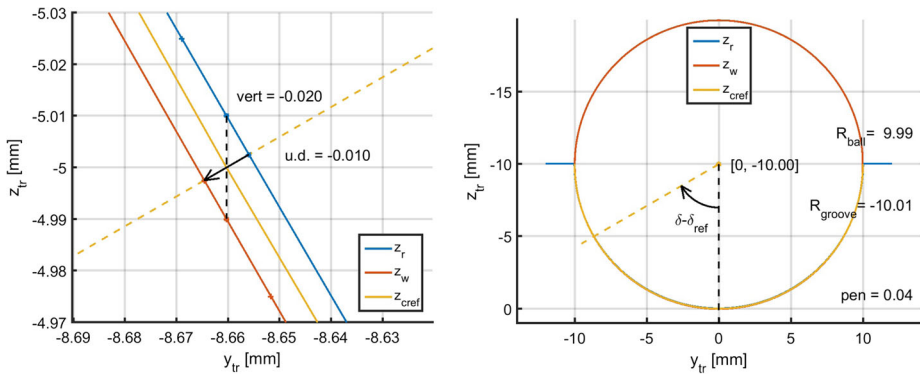
Fig. 8 Geometry for a ball in a conforming groove, analytical test-case used to develop and validate the undeformed distance calculation of Eq. (19)

3.5 Step 2.b', Grid-based rigid slip computation

In the traditional planar contact approach, the creepages ξ, η, ϕ of Eq. (15) are obtained from the rigid body velocities of the rail and wheel bodies. This works using the material time derivatives of the reference configuration, conform Eqs. (1) and (2), using rigid body kinematics. This is worked out in the wheelset translational velocity $\mathbf{v}_{ws}$, the angular velocity $\boldsymbol{\omega}_{ws}$, obtained from the rates of change of Euler angles, and optionally, the rate of change of global deformations, and the corresponding variables for a roller. Two material points P and Q are selected at the rail and wheel surfaces at the contact reference position [26]. Their velocities $\mathbf{v}_{ref,P}, \mathbf{v}_{ref,Q}$ are obtained using standard formulas for a point on a rigid body. The creepages are then obtained as the scaled velocity difference.

The new approach that is used for the curved reference works largely the same as this previous approach. The main difference is that the reference velocities are computed at all contact grid points, $\mathbf{p}_I = [x_I, s_I, n_{Ir}]^T$, $\mathbf{q}_I = [x_I, s_I, n_{Iw}]^T$, sampling the wheel and rail surfaces, instead of using the origin only.

This grid-based approach provides further insights in the creepages ξ, η, ϕ. Namely, we recover the results of the traditional approach with creepages when using the grid-based approach with modified grids $\mathbf{p}_I'' = \mathbf{q}_I'' = [x_I, s_I, 0]^T$. From this we learn that:

– for a wheel on a rail, the creep calculation amounts to approximating the wheel by a purely conical surface;
– for a ball on a plane, the creep calculation amounts to approximating the ball by a cylindrical surface.

These approximations may be understood by the analysis presented by Kalker [31, §1.7.3]. The lateral curvatures of ball, wheel and roller are represented in Kalker's analysis in the conformity terms $(-1)^{a+1} B^{(a)}(s)$, which are ignored in the simplified form of Eq. (15). The traditional creep calculation may thus be improved upon by retaining the conformity terms.

3.6 Step 2.b', Introducing the mid-plane correction

A second extension to the rigid slip is that the normal and longitudinal directions are redefined temporarily to account for the surface inclination in longitudinal direction:

$$\mathbf{v}_{mid} = \frac{1}{2} \left(\frac{\mathbf{v}_P}{\|\mathbf{v}_P\|} + \frac{\mathbf{v}_Q}{\|\mathbf{v}_Q\|} \right), \tag{20a}$$
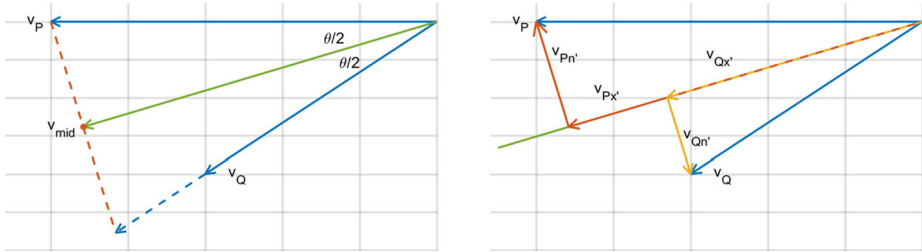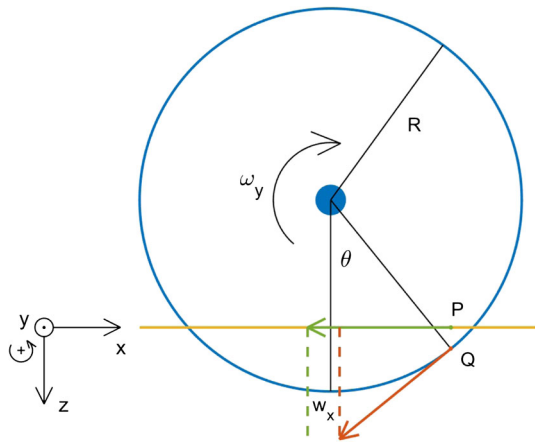
Fig. 9 Left: construction of midplane in longitudinal direction conform Eq. (20b). Right: velocity vectors $\mathbf{v}_P, \mathbf{v}_Q$ expressed in local (x', n')-coordinates

Fig. 10 Considering the velocities of points P and Q in the undeformed rail and wheel surfaces, rigid slip $w_x = v_{Px} - v_{Qx}$ arises due to the wheel surface inclination θ

$$\mathbf{t}'_x = -\frac{\mathbf{v}_{mid}}{\|\mathbf{v}_{mid}\|}, \quad c = (\mathbf{t}'_x, \mathbf{t}_x), \quad s = (\mathbf{t}'_x, \mathbf{n}), \tag{20b}$$

$$\mathbf{t}'_x = \begin{bmatrix} c \\ 0 \\ s \end{bmatrix}, \quad \mathbf{n}' = \begin{bmatrix} -s \\ 0 \\ c \end{bmatrix}, \quad \mathbf{R} = \begin{bmatrix} c & 0 & -s \\ 0 & 1 & 0 \\ s & 0 & c \end{bmatrix}, \tag{20c}$$

$$\mathbf{v}'_P = \mathbf{R}^T \mathbf{v}_P, \quad \mathbf{v}'_Q = \mathbf{R}^T \mathbf{v}_Q. \tag{20d}$$

These equations are illustrated in Fig. 9. They are obtained using $\mathbf{t}_x, \mathbf{n}$ as the global system, making $\mathbf{t}'_x, \mathbf{n}'$ the local system and making (20d) to-local operations. The rigid slip is then computed in the temporary coordinate system, e.g. $w_x = v_{Px'} - v_{Qx'}$, as illustrated in Fig. 9.

This temporary change compensates an artificial effect that comes from using a prismatic curved reference surface, linearly extruded in the rolling direction. This effect is illustrated in Fig. 10: in a contact-following coordinate system that is moving with velocity V, the longitudinal velocities show up as $v_{Px} \equiv -V$ and $v_{Qx} = -V \cos(\theta)$. This introduces rigid slip $w_x = -V(1 - \cos(\theta))$, non-zero when there is a difference in surface inclination. Equation (20b) puts the tangent direction $\mathbf{t}'_x$ half-way between $\mathbf{v}_P$ and $\mathbf{v}_Q$. Using this we get $v_{Px'} = v_{Qx'} = -V \cos(\theta/2)$ and get no rigid slip anymore in the situation of Fig. 10.

This unwanted rigid slip is described by the terms $\frac{1}{2} R_{ax}^{-1} x^2$ in Kalker's analysis. Kalker stated that these terms are very small and can hence be neglected [31, p. 42]. Here we argue

differently: the terms are not negligibly small by themselves, but will cancel each other when measured in normal rather than vertical direction.

3.7 Step 3.c', Aggregating tractions into resultant forces

For the aggregation of total forces, we consider the tractions $\mathbf{p} = \mathbf{p}^{(1)}$ acting on the grid locations $\mathbf{x}_I$ and in the coordinate directions of the curved reference surface. They are rotated to a common, Cartesian coordinate system by a generic curved-to-Cartesian operation. This can be the local coordinate system introduced for planar contact analysis, adhering to the existing flow of computations, reusing the existing code where possible. The rotated tractions may be denoted as $p_{I\bar{n}} = p_{\bar{n}}(\tilde{x}_I, \tilde{s}_I)$ for a traction in overall $\bar{n}$-direction at grid point I of the curved reference surface.

The resultant forces acting on the upper body, the rail, are obtained by integration:

$$F_{\bar{n}}^{(1)} = \int_{\tilde{C}} p_{\bar{n}}(\tilde{x}, \tilde{s}) \, d\tilde{x} \, d\tilde{s} \;=\; \sum_I p_{I\bar{n}} \, \delta x \, \delta s, \tag{21}$$

likewise for $F_{\bar{x}}$ and $F_{\bar{s}}$. These forces act in the positive directions of the (planar) contact reference coordinates. The forces exerted on the wheel are obtained by "action = –reaction".

The moments of the tractions $\mathbf{p}$ with respect to the contact reference position are computed similarly, using the same transformation. The contribution of a single point is obtained from the lever arm times surface tractions:

$$\mathbf{M} = \mathbf{x} \times \mathbf{p} = \begin{bmatrix} \bar{x} \\ \bar{s} \\ \bar{n} \end{bmatrix} \times \begin{bmatrix} p_{\bar{x}} \\ p_{\bar{s}} \\ p_{\bar{n}} \end{bmatrix} = \begin{bmatrix} \bar{s} \, p_{\bar{n}} - \bar{n} \, p_{\bar{s}} \\ \bar{n} \, p_{\bar{x}} - \bar{x} \, p_{\bar{n}} \\ \bar{x} \, p_{\bar{s}} - \bar{s} \, p_{\bar{x}} \end{bmatrix}. \tag{22}$$

The total moments are then obtained by integration over the contact area, e.g.

$$M_{\bar{x}}^{(1)} = \int_{\tilde{C}} (\bar{s} \cdot p_{\bar{n}}(\tilde{x}, \tilde{s}) - \bar{n} \cdot p_{\bar{s}}(\tilde{x}, \tilde{s})) \, d\tilde{x} \, d\tilde{s} \;=\; \sum_I (\bar{s}_I \, p_{I\bar{n}} - \bar{n}_I \, p_{I\bar{s}}) \, \delta x \, \delta s \tag{23}$$

The components $M_{\bar{x}}$ and $M_{\bar{s}}$ are zero when using planar contact ($\bar{n} \equiv 0$) with a symmetric pressure distribution. The component $M_{\bar{n}}$ is called the spin moment.

To aggregate the total forces and moments of different contact patches requires that they are brought together in a common coordinate system. Obvious choices for actions on the rail are to use orientations of rail or track systems, while forces and moments on the wheel are interpreted more easily in the wheelset or wheel coordinate frames. The fictitious location of the aggregate forces may be shifted by adding compensating moments. For contact patch i, we can compute the equivalent force and moment acting at the rail marker as

$$\mathbf{F}_{i(tr)} = \mathbf{R}_{i(tr)} \mathbf{F}_{i(cp)}, \quad \mathbf{M}_{i@cp(tr)} = \mathbf{R}_{i(tr)} \mathbf{M}_{i@cp(cp)}, \tag{24a}$$

$$\mathbf{M}_{i@rr(tr)} = \mathbf{M}_{i@cp(tr)} + \left(\mathbf{x}_{i(tr)} - \mathbf{x}_{rr(tr)} \right) \times \mathbf{F}_{i(tr)}. \tag{24b}$$

The total forces can be shifted similarly to any point P. The resultant couple can be made zero if the total force and moment vectors are perpendicular to one another. Else, we can split the moment in parallel and perpendicular parts and annihilate the latter by shifting:

$$\mathbf{M}^{\parallel} = \frac{(\mathbf{M} \cdot \mathbf{F})}{\|\mathbf{F}\|} \frac{\mathbf{F}}{\|\mathbf{F}\|}, \quad \mathbf{M}^{\perp} = \mathbf{M} - \mathbf{M}^{\parallel}. \tag{25}$$

When a force $\mathbf{F}$ is shifted from its original location $\mathbf{0}$ to $\mathbf{x}_{av}$, then this shift is taken into account in the moment by an additional term $\mathbf{x}_{av} \times \mathbf{F}$. Using the vector triple product,

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}, \tag{26}$$

we may eliminate $\mathbf{M}^{\perp}$ as follows:

$$\mathbf{x}_{av} \times \mathbf{F} = -\mathbf{M}^{\perp} \quad \rightarrow \quad \mathbf{F} \times (\mathbf{x}_{av} \times \mathbf{F}) = \mathbf{F} \times (-\mathbf{M}^{\perp}) \tag{27a}$$

$$\rightarrow \quad (\mathbf{F} \cdot \mathbf{F})\mathbf{x}_{av} - (\mathbf{F} \cdot \mathbf{x}_{av})\mathbf{F} = -\mathbf{F} \times \mathbf{M}^{\perp} \tag{27b}$$

$$\overset{(\mathbf{F} \cdot \mathbf{x}_{av}) = 0}{\rightarrow} \quad \mathbf{x}_{av} = -\frac{\mathbf{F} \times \mathbf{M}^{\perp}}{(\mathbf{F} \cdot \mathbf{F})} = \frac{\mathbf{M}^{\perp} \times \mathbf{F}}{(\mathbf{F} \cdot \mathbf{F})} = \frac{\mathbf{M} \times \mathbf{F}}{(\mathbf{F} \cdot \mathbf{F})}. \tag{27c}$$

The location $\mathbf{x}_{av}$ may be placed at the rail surface by shifting along the force direction $\mathbf{F}/\|\mathbf{F}\|$, with no effect on the total force and moment. This involves the intersection of a spline curve with a straight line, which is solved by an iterative calculation. This finally defines an average location where the total forces may be conceived, collapsing the surface tractions with the smallest possible moment. This reduces the errors when moments are neglected in a multibody simulation.

4 Results

This section presents numerical results computed with the new method and accompanying results of planar contact methods as listed in Table 1. All results are computed using the CONTACT program, version 20.2 [13], using stand-alone contact analysis. Two sets of test-cases are presented, using worn profiles obtained from one North American railroad in Sect. 4.1, and using a bigger collection of measured wheel profiles, obtained from another North American railroad, in Sect. 4.2.

The first area of application for the curved reference lies in the study of conformal and two–point wheel–rail contacts. This is illustrated in Sect. 4.1, with Sect. 4.1.1 discussing how discontinuities in the resultant forces may be avoided using the new method, and Sect. 4.1.2 addressing the refined calculation of the rigid slip velocities and their effect on the creep forces. The second application concerns the benchmarking and tuning of planar contact approaches. This is shown in Sect. 4.2, presenting a statistical evaluation of planar methods compared to the new curved reference approach.

4.1 Two-point and conformal wheel/rail contacts

An interesting case for two–point and conformal contacts was obtained thanks to Dr. Kevin Oldknow [Simon Fraser Univ., pers.comm]. This concerns a small collection of wheel and rail profiles, gathered during field work for a North American heavy haul railway in 2005 and recognized as test-cases with conforming shapes. Rail profiles were measured for three different curves with nominal radii 257 m, 282 m, and 388 m. Wheel profiles were collected from revenue service wagons with standard North American three-piece bogies and a gross load of 130 tonnes. Profiles were measured using MiniProf wheel profilometers for wheels with low, mid and high mileage.

Here we focus on the 257 m radius curve interacting with worn wheels with mid and high mileage. Figure 11 shows how the profiles of the high rail and of the high-mileage wheel produce a distinctly conformal contact. Test-cases for the contact analysis for the
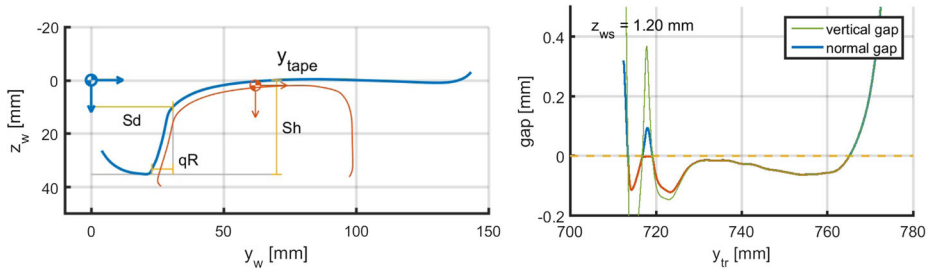
Fig. 11 Left: conformal worn profiles of the high-mileage wheel and high rail for the 257 m radius curve. Right: resulting gap function for positioning considered in Fig. 12
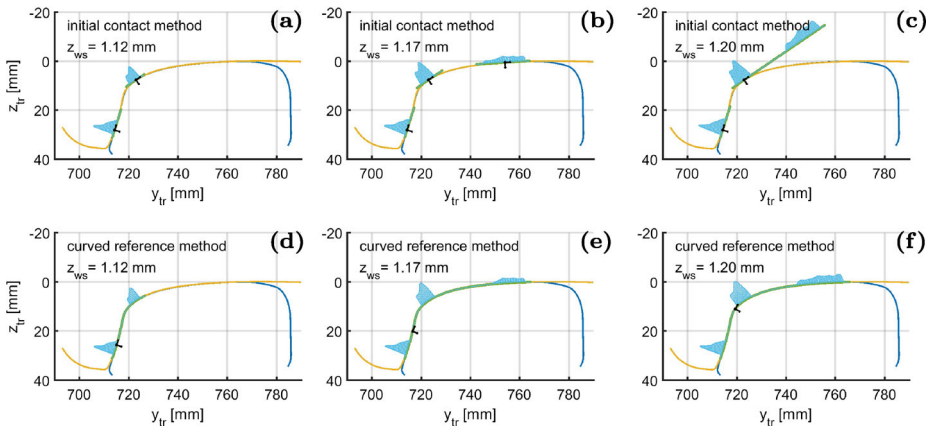

Fig. 12 Contact configurations for the high-mileage wheel at lateral position $y_{ws} = 14.85$ mm and increasing vertical positions z_{ws}. (**a**)–(**c**): computed using the initial contact method "icp-loc"; (**d**)–(**f**): corresponding results for the curved reference method

high rail are created using representative values obtained from geometry and kinematic data: nominal wheel load 160 kN for a loaded wagon, angle of attack $\psi = 5$ mrad, roll angle $\phi = -1.3$ mrad. Full flange contact occurs when the wheel profile is shifted by $y_{ws} = 14.85$ mm from the nominal position, owing to the reduced flange thickness $S_d = 30.9$ mm and wear of the rail. The coefficient of friction is set to $\mu = 0.35$ at the top of the rail (surface inclination $\alpha \geq -20°$) and $\mu = 0.20$ at the gauge face ($\alpha \leq -30°$) with a linear transition in between, modelling the use of flange lubrication.

4.1.1 Combined computation of contact patches

A difficult feature for the planar contact approach lies in the separate or combined computation of adjacent contact patches. This is illustrated in Figs. 12 and 13 for the planar contact method "icp-loc" of Table 1, which uses the initial contact position and the corresponding contact angle. The high-mileage wheel is considered at fixed velocities $v = 15$ m/s, $\omega_{ws} = -32.1$ rad/s (nominal rolling radius $r_{nom} = 444$ mm). One, two or three contact patches may be found in this configuration, at the wheel flange, the flange root, and the tread of the wheel, varying the vertical displacement z_{ws} of the wheelset, at fixed lateral displacement $y_{ws} = 14.85$ mm. The interpenetration areas on the flange root and tread merge
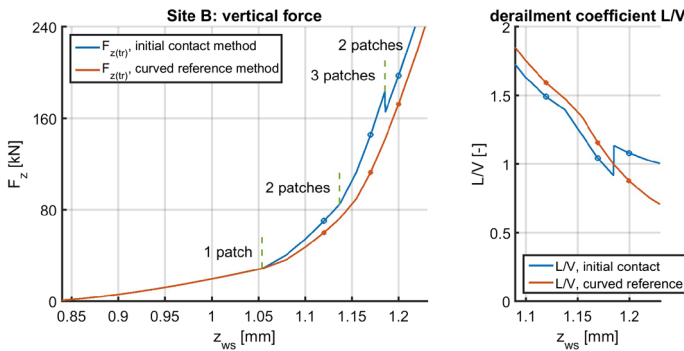
Fig. 13 Left: graph of vertical force versus vertical position for the configurations of Fig. 12, at fixed lateral displacement. Right: derailment coefficients $F_{y(tr)}/F_{z(tr)}$

when the wheel is moved down beyond $z_{ws} = 1.18$ mm, such that two contact patches are computed from thereon (compare Figs. 12 (b) and (c)). This gives a drop of 18.1 kN in the vertical forces computed and an accompanying change in the lateral forces, increasing L/V from 0.92 to 1.13, as shown in Fig. 13.

Two different effects occur when two contact patches are taken together in the planar contact approach:

1. The contact reference angle changes suddenly from two values for the separate patches to a single value when they are combined.
2. The contact pressures of the two patches suddenly start interacting with each other. This generally increases the computed elastic deformations, reducing the pressures needed at fixed interpenetration.

The effect of the contact angle can be lessened to some extent using an averaged value instead of the value sampled at the initial contact position. This works especially for the case of Fig. 12 (c) where patches on the flange root and tread of the wheel are combined. Bigger difficulties remain if the contact patch on the flange needs to be combined with the one on the flange root, due to the bigger changes in contact angle.

The effects of interactions are seen in Fig. 13 at $z_{ws} \leq 1.18$ mm. As the contact is nearly flat within each of the three contact patches, close correspondence could be expected between the planar and conformal approaches. Yet the conformal method finds considerably lower pressures and forces compared to the planar method. This is the result of taking the interactions between different patches into account instead of neglecting them. This is achieved by "greedy" combination of patches in the conformal approach. Interpenetration regions are combined there that lie less than $d_{sep} = 30$ mm from each other, resulting in a single patch in Figs. 12 (d)–(f).

These discontinuities in the total forces pose a trade-off between reliability and robustness for the planar contact approach. Different strategies are followed by vehicle dynamics softwares. VAMPIRE, for instance, tends to avoid jumps in the results by supporting at most one contact on the tread of the wheel and at most one at the flange, at the expense of the trustworthiness of the results in more challenging situations. SIMPACK and Universal Mechanism, on the other hand, support multiple contact patches, appearing, disappearing, and merging as the simulation evolves. This goes at the expense of some robustness and performance, using time step refinement and permitting some irregularities in the results when jumps occur.
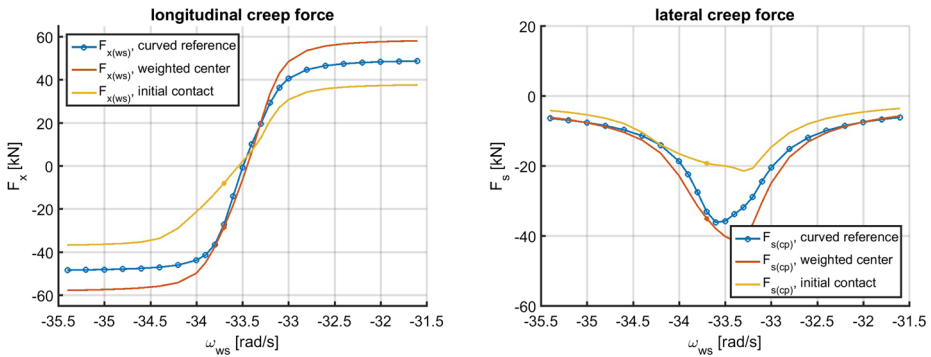
Fig. 14 Tangential creep forces F_x, F_s for the mid-mileage wheel at lateral position $y_{ws} = 10.65$ mm, vertical load $F_{z(tr)} = 160$ kN, and varying angular velocities ω_{ws}, computed by different methods
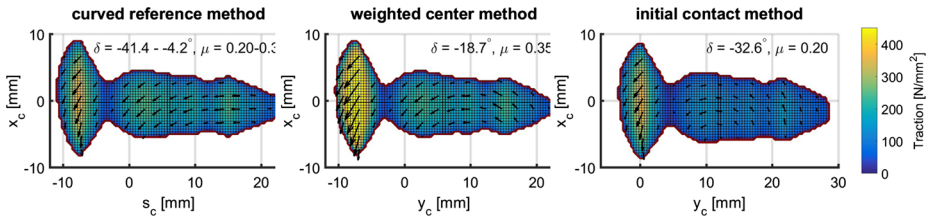

Fig. 15 Tangential tractions $[p_x, p_y]^T$ computed by different methods for the mid-mileage wheel at $y_{ws} = 10.65$ mm, $F_z = 160$ kN, $\omega_{ws} = -33.7$ rad/s

The benefit of the conformal contact approach is to provide smoothness by using the local contact angles, for the undeformed distance as well as the rigid slip computation. Jumps regarding the contact interaction can then be mitigated using greedy combination of patches, or using a blending approach in which the interaction is gradually turned on when contact patches get closer to one another [26].

4.1.2 Effects of the rigid slip computation

The effects of different computational methods on the frictional forces are examined with a test for the mid-mileage wheel. The wheelset is placed at lateral position $y_{ws} = 10.65$ mm ($S_d = 36.3$ mm) and roll angle $\phi_{ws} = 1.3$ mrad, pressed down to a vertical load $F_{z(tr)} = 160$ kN. The forward velocity and yaw angle are kept fixed at $v = 15$ m/s and $\psi_{ws} = 5$ mrad. The angular velocity ω_{ws} is varied in a wide range to explore how this affects the creep forces. Results are shown in Figs. 14–16 for the methods icp-loc and wgt-loc-2 of Table 1.

Figure 14 shows considerable differences between the creep forces computed using different methods. A first origin for this lies in the coefficient of friction employed in the different calculations. In the planar contact methods, a constant value is used that is determined on the basis of the contact reference angle. The initial contact method finds contact angle $\delta = -32.6°$, by which the case is considered as flange contact with $\mu = 0.20$. The weighted center method, on the other hand, finds $\delta = -18.6°$ and sets $\mu = 0.35$ for tread contact. The conformal contact method, using a variable coefficient of friction, on the basis of the actual inclination of the curved reference surface, produces $\mu = 0.2$ on the left side and $\mu = 0.35$ on the right side of the contact area.
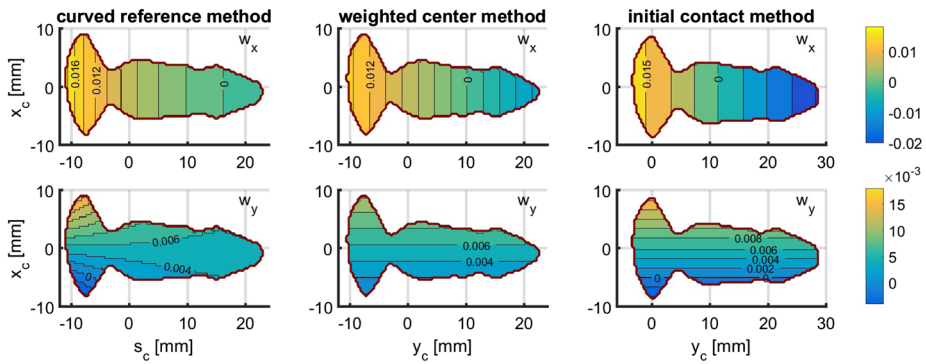
Fig. 16 Rigid slip functions $[w_x, w_y]^T$ computed by different methods for the mid-mileage wheel at $y_{ws} = 10.65$ mm, $F_z = 160$ kN, $\omega_{ws} = -33.7$ rad/s. Left: local kinematics, evaluated at each surface position. Middle/right: kinematics evaluated at the contact reference marker with linear extrapolation
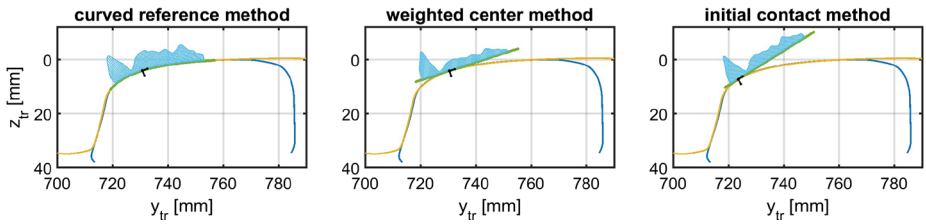

Fig. 17 Contact configurations for the mid-mileage wheel at lateral position $y_{ws} = 10.65$ mm and vertical load $F_{z(tr)} = 160$ kN

Looking in detail, we find that the different coefficients of friction do not fully explain the different shapes of the curves: a slower transition from negative to positive F_x for the initial contact method, and attaining the lowest lateral creep-force values at different angular velocities ω_{ws}. These differences are analyzed further by considering one specific case, at $\omega_{ws} = -33.7$ rad/s, in Figs. 15–17.

The slower transition found for the initial contact method is attributed mainly to the rigid slips $[w_x, w_y]^T$ that are used (Fig. 16). These are understood with the aid of the contact situations as illustrated in Fig. 17. The curved reference method computes rigid slip velocities based on the actual wheel and rail surface shapes, that is, using the local kinematics at each surface position. The planar methods approximate this using the kinematics at a single position, i.e. the contact reference marker. At their respective contact reference positions, both planar contact methods thereby produce the same rigid slip as the conformal method. Yet the methods differ in the rigid slips computed at other positions. Using Eq. (15), these rigid slips vary linearly in the planar methods. The coefficient ϕ is bigger for the initial contact method corresponding to the bigger contact angle. For the conformal method, the rigid slips vary more rapidly on the left side, and slower on the right, in proportion to the inclination of the contact reference surface. This corresponds to the conformity terms $B^{(a)}(s)$ as discussed in Sect. 3.5.

A similar effect plays a role in the lateral creep forces F_s. Further, the coefficient of friction plays a role in the results for the weighted center method, with bigger stresses $p_y < 0$ in the range -10 mm $\leq y_c \leq -5$ mm (Fig. 15). Thirdly, we find that the aggregate results plotted for the curved reference method are affected by the contact pressures. This concerns

the force $F_{\bar{s}}$ for the contact reference orientation, as discussed in Sect. 3.7. Depending on the actual orientation of the contact pressures, positive contributions are found on the left and negative on the right of the reference position. This may be perceived as a downside of the curved reference surface: the results are not interpreted as easily as for the planar contact approach.

4.2 Statistical evaluation of planar contact approaches

The results shown above suggest that the initial contact approach is more susceptible to jumps in forces than the weighted center approach, and that it may be better to use an averaged contact angle than sampling the angle at the reference position. So far, however, this is just anecdotal evidence, with no indication on the generality of the results. Here we present statistics for a large number of cases to get more balanced information.

4.2.1 Data-set for freight cars with worn profiles

Data are obtained from project work for CSX, a US class I railroad, performed by Sentient Science, NRC Canada, and Vtech CMCC on rail life and rail life extension. The overall simulation approach starts by selecting a curve and collecting traffic data: the number of trains, types of wagons, loads, speeds, wheel profiles, as well as track data, e.g. geometry, friction management practices. These data are analyzed and classified, e.g. distinguishing different quality levels [83]. Loading ensembles are then defined that represent the actual situation, maintaining the ranges and variability of input parameters, however, with a much reduced number of loadings. Simulations with VAMPIRE are used to assess the vehicle dynamics in steady curving, using VAMPIRE's internal contact algorithms. The contact situation is then analyzed anew, in more detail, using CONTACT, and the resulting contact stresses are fed into Sentient's calculation of wear and grinding. This is repeated many times in a cycle, representing an overall duration of 25 years or so, until the rail profile is worn down to a predefined threshold and the life is exhausted.

Here we consider the data obtained for a 5 degree right hand curve (radius 349 m), near Cartersville, GA. A thousand freight cars are simulated, a mixture of flat, hopper and tank cars, with heavily worn wheel profiles, according to the 20% of wheels with the lowest quality index. Speeds vary between 7 and 40 mph (11 to 64 km/h), and weights between 50 000 and 300 000 lbs (222 to 1334 kN) for empty and full wagons. Two axles are considered for each car, for the leading bogie. Four coefficients of friction are set for each axle, two for each wheel, taken from normal distributions with mean value $\bar{\mu} = 0.35$ and standard deviation $\sigma_\mu = 0.16$ at the top of the rail (surface inclination $\alpha \leq 20°$) and $\bar{\mu} = 0.20$, $\sigma_\mu = 0.0167$ at the gauge face ($\alpha \geq 30°$). Two rail profiles are considered: tentative new profiles for the high and low rails of the curve.

CONTACT calculations are performed once as a post-processing step after each VAMPIRE simulation. This gives a considerable speed up compared to full integration of CONTACT in the multibody simulation, owing to the fast heuristic contact algorithms implemented in VAMPIRE, executed many times to find the steady curving situation. Calculations with CONTACT are then done with wheelset kinematics held fixed as computed by VAMPIRE, including axle positions, orientations and velocities. One exception concerns the vertical position. Each wheel is raised or lowered as needed to maintain the vertical force $F_{z(tr)}$ obtained from VAMPIRE. This addresses the differences in the relation between position and vertical force between CONTACT and VAMPIRE's contact methods. Agreement in the positions is considered much less important than agreement in the forces, here, because the positions have no direct influence on the wear calculation, whereas the forces govern the stresses that do have a major effect on the wear calculation.
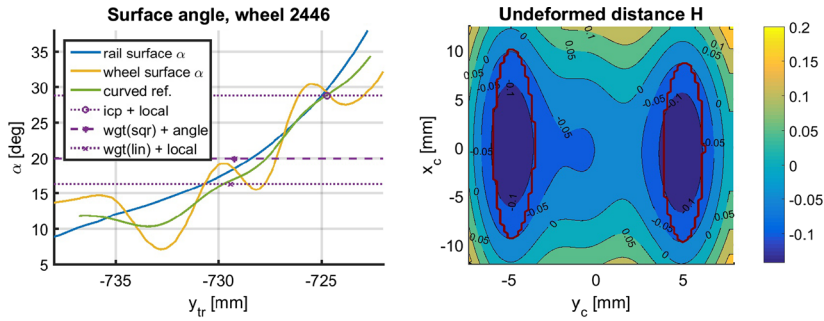
Fig. 18 Contact geometry for wheel 2446 of the group 5 data: wheel surface inclination α and undeformed distance function h
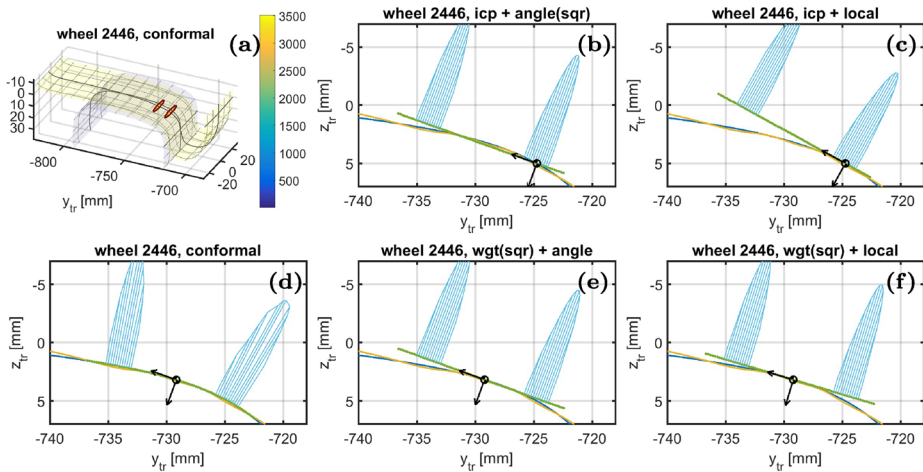

Fig. 19 Effect of the choices for contact reference position and reference angle for wheel 2446 of the group 5 data

4.2.2 Detailed inspection for wheel 2446

Results are computed for all wheels using different methods and compared in different ways. The wheels that showed the biggest differences were investigated further. Figures 18 and 19 show the results for wheel 2446, which are particularly insightful on how the different strategies work out for the contact reference point and reference angle.

The wheel profile that is used exhibits some grooves and ridges across the profile. This manifests itself as oscillations in the surface inclination (Fig. 18, left) and gives rise to disjoint valleys in the undeformed distance function (right). At a vertical force of 152.7 kN, these valleys are contained in one interpenetration area with two local minimum values. This results in two actual contact areas with nearly elliptical shapes, with some interactions between the two patches.

Figure 19 shows the contact planes and markers used with the resulting pressure distributions. This shows clearly how the initial contact method places the contact reference at the deepest bump (b)–(c), while the weighted center method goes somewhere in the middle (e)–(f). Next, the effect of an averaged angle is seen clearly comparing pictures (b) and (c)

Table 2 Results for different methods for wheel 2446 of the group 5 data, with the biggest differences to the conformal method marked in italic

method	ref.angle [deg]	F_x [kN]	F_s [kN]	F_y [kN]
conformal	19.9	−41.4	−26.0	−27.7
wgt(sqr) + angle	20.0	−38.3	−27.5	−26.2
wgt(lin) + angle	19.4	−38.6	−27.3	−24.7
wgt(sqr) + local	*16.7*	−39.3	−26.5	*−18.1*
wgt(lin) + local	*16.4*	−39.4	−26.4	*−17.3*
icp + angle(sqr)	20.0	−39.1	−26.7	−27.0
icp + local	*28.8*	*−32.8*	*−32.3*	*−47.2*

for the initial contact method. The weighted center + angle method (e) uses the same angle as (b) but raises the contact plane slightly due to the different reference position. This gives a slight difference in rolling radius between the weighted center and initial contact methods when using the averaged contact angle.

Figure 18 (left) shows the inclinations of various tangent planes using dashed and dotted lines. The dashed line for the weighted averaging methods approximates the average value for the curve for the curved reference surface. The approximation used by the icp + local method is rather poor because it samples the actual inclinations away from the center. The value used by wgt(lin) + local appears to underestimate the averaged value. This is explained by the convex shape of the actual curve, which is a typical feature of the rail gauge corner.

Table 2 shows the total forces obtained from different methods. For icp + local we see that over-estimating the contact angle results in a bigger lateral creep force F_s, reducing F_x. Furthermore, the contact angle lets F_n contribute stronger to F_y. For wgt + local, under-estimating the contact angle, F_s is reduced slightly compared to wgt + angle while F_x is increased. The main change comes again from the propagation of F_n on F_y. The wgt + angle and icp + angle methods seem to perform a little worse in F_x and F_s but compensate through the F_n contribution. Squared weights improve the contact angle and the ultimate correspondence of the results.

4.2.3 Evaluation of methods

Many different planar contact methods can be constructed using different strategies for defining the contact reference position and reference angle. The question is: which of these methods comes closest to the true contact situation? This is assessed for the methods as listed in Table 1 with the curved reference method as referee, i.e. the best method currently available at our disposition.

The results of different calculations are compared mainly through the signed derailment coefficients L/V. These are sensitive to the contact angle and contact location, while ignoring the different weights of different wagons. Figure 20 shows the comparison for a number of methods, with the wheels sorted by the L/V values computed by the curved reference method. This shows that there is a bigger spread in the computed results for wheels with negative L/V values. These are mostly wheels on the left side, the high rail, where we find multiple contacts and contacts with bigger contact angles.

Table 3 shows the statistics of the differences of each method compared to the curved reference method. This takes each line of Fig. 20 in turn, computes the difference for each wheel compared to the reference method, and then computes the mean value, mean of absolute values, root-mean-square, and maximum absolute values.
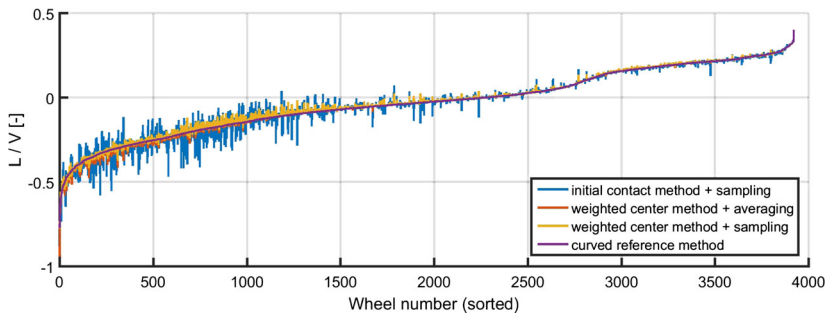
Fig. 20 Comparison of L/V-values computed for the statistical evaluation of planar contact methods

Table 3 Statistics of the differences of computed L/V-values for different methods compared to the results for the curved reference method. (*) excluding cases with $L/V < -1.0$

tested method	mean(dif)	mean(abs(dif))	rms(dif)	max(abs(dif))
curved reference with initial contact	0.0000	0.0003	0.001	0.013
weighted center, sampled, linear	0.0061	0.0088	0.016	0.129
weighted center, sampled, squared	0.0046	0.0067	0.012	0.089
weighted center, average, linear	0.0008	0.0064	0.013	0.155
weighted center, average, squared	0.0003	0.0051	0.011	0.170
initial contact, sampled (*)	−0.0011	0.0126	0.028	0.247
initial contact, average, linear	0.0005	0.0064	0.013	0.158
initial contact, average, squared	0.0001	0.0051	0.011	0.178

The first line of Table 3 concerns the comparison of different variants of the curved reference method, using different ways to define the reference position. This is used as an anchor for the curved reference surface and its grid discretization. The table shows that there is a slight dependence on the way how the contact reference position is selected. These differences diminish when the grid is refined.

The second part of the table shows results for the weighted center method. This appears to give somewhat biased results (nonzero mean difference), when sampling the contact angle at the reference position. This is seen in Fig. 20 as well, with more yellow above than below the results for the curved reference method. This comes from the typical shapes of rail profiles, with low curvature at the top of the rail and curvature increasing towards the gauge corner. The weighted center method places the reference position centered in the contact patch, on average. The surface inclination that is found there is then typically a little lower than the average inclination over the width of the contact.

The third part of the table concerns the initial contact method. Here we found 7 wheels with L/V-values below -1.0, related to big contact angles at the flange of the wheel. These are filtered out to avoid their big influence on the statistics. After this, we still find nearly twice bigger differences than for the weighted center method with sampled contact angles. These differences are removed when switching to an averaged contact angle. We conclude that the contact angle is the most sensitive parameter in the planar contact approach. Weighted averaging appears to give better values than local sampling, and the squared gaps of Eq. (14) work a little better than the linear gaps used in Eq. (13).

4.2.4 Performance

Considering the performance, each of the methods needs about 450 s on a commodity laptop for the computations for 4 000 wheels. About 20 s is spent on the geometrical analysis, 60 s on the normal contact problem, 310 s on the tangential problem, and 60 s on in- and output. In this time, each wheel is analyzed 4–5 times to satisfy the requested vertical force, using about 25 ms per case with prescribed positions. The resolution used is $\delta x = \delta s = 0.5$ mm, by which 60–1120 elements end up in contact, with an average of 325. The main difference between the curved reference and planar methods lies in the geometrical part. The curved reference method needs 30% more time there, corresponding to an increase of about 1% of the total time.

5 Conclusions and discussion

This paper addressed the role of the tangent plane in the planar contact approach –defining the normal and tangential directions needed for the undeformed distance, creepage, pressures, and slip– and discussed the complications in its definition, prior to the calculations, and using a single inclination to approximate the entire contact. The curved reference surface is put forward as an alternative backbone for detailed wheel/rail contact computations. This builds on the generalized coordinates $[x, s, n]^T$ introduced for conformal contact situations that provide different normal and tangential directions at different locations. This facilitates the use of local kinematics instead of approximate, global values, by-passing the creepages, evaluating the rigid slip velocities in a more detailed way. The steps are described by which the curved reference is constructed, using generic facilities for markers, grids, and interpolations, enabling the conformal method to be used in a fully automated way.

A main characteristic of the curved reference surface that is used in CONTACT is that it uses a local construction. Separate curved reference grids are produced for different contact patches, to avoid the complications that arise in global curvilinear coordinates for surfaces that separate by a bigger distance and then come back together. A beneficial effect of this choice is that it permits re-using the steps for planar wheel/rail contact solution with relatively minor extension.

The computation of the curved reference surface is based on the chord distances s_r and s_w along the wheel and rail profiles. Earlier calculations on this are extended for the 3-d contact search procedures used in CONTACT, including yaw. The surface is represented by a planar curve in the Oyz-plane of the track frame, which is extruded in the direction of rolling. This makes the curved reference easier to understand and work with because orientations like the local normal can be described using a single rotation. This simplified view creates an artifact in the rigid slip calculation that is resolved locally using the so-called midplane correction.

One application of the curved reference surface is to provide reference solutions for the evaluation of simpler contact approaches. This is worked out for the tangent plane used in the planar contact approach. A data-set is used for a thousand freight cars with heavily worn wheel profiles, with statistical evaluation. The tangent plane's inclination appears to be the most influential aspect, which needs to approximate the average value over the contact region. After this, the reference position for the local coordinate system can be defined equally well with the initial contact position or using weighted averaging.

The planar contact approach seems to give quite acceptable results in most cases, but may break down in particular circumstances. This happens mainly for measured, worn profiles where multiple contact patches are forming and growing together. Discontinuities are found

in the resultant contact forces, originating from sudden changes in the contact reference angle and in the interactions between different contact patches. In such cases, the planar contact approach is restricted by the use of a constant inclination for the tangent plane, while the curved reference accommodates variations in inclination naturally. This permits for more smooth and robust contacts in multibody simulation.

The benefits of the curved reference may be exploited also in heuristic contact approaches, for instance using the Kik–Piotrowski method in combination with Fastsim. Discontinuities may be reduced there as well, and creepages may be avoided using the more detailed rigid slip computation. This is an interesting topic for further study. Another way forward is to improve the planar contact approach with the aid of the curved reference surface. One could improve the calculation of creepages for instance by adding a conformity term. A third direction for future work concerns truly conformal contacts: assessing the consequences of our simplifications, extending the influence coefficients used beyond the halfspace approach, and comparing the results to those obtained from finite-element calculations. A final challenge remains to get better understanding of the coefficient of friction: how it comes about, and how influential factors like fluids and third body layers can be included in our modeling.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Iwnicki, S.D. (ed.): Handbook of Railway Vehicle Dynamics CRC Press, Boca Raton (2006)
2. Shabana, A.A., Zaazaa, K.E., Sugiyama, H.: Railroad Vehicle Dynamics: A Computational Approach. CRC Press, Boca Raton (2008)
3. Weidemann, C.: State-of-the-art railway vehicle design with multi-body simulation. J. Mech. Syst. Transp. Logist. **3**, 12–26 (2010)
4. Polach, O., Böttcher, A., Vannucci, D., Sima, J., Schelle, H., Chollet, H., Götz, G., Garcia Prada, M., Nicklisch, D., Mazzola, L., Berg, M., Osman, M.: Validation of simulation models in the context of railway vehicle acceptance. Proc. Inst. Mech. Eng., F J. Rail Rapid Transit **229**, 729–754 (2015). https://doi.org/10.1177/0954409714554275
5. Zhai, W., Han, Z., Chen, Z., Ling, L., Zhu, S.: Train–track–bridge dynamic interaction: a state-of-the-art review. Veh. Syst. Dyn. **57**, 984–1027 (2019)
6. Escalona, J.L., Aceituno, J.F.: Multibody simulation of railway vehicles with contact lookup tables. Int. J. Mech. Sci. **155**, 571–582 (2019)
7. Escalona, J.L., Aceituno, J.F., Urda, P., Balling, O.: Railway multibody simulation with the knife-edge-equivalent wheel–rail constraint equations. Multibody Syst. Dyn. **48**, 373–402 (2020)

8. Kortüm, W.: Review of multibody computer codes for vehicle system dynamics. Veh. Syst. Dyn. **22**(suppl. 1), 3–31 (1993)
9. Kik, W., Piotrowski, J.: A fast approximate method to calculate normal load at contact between wheel and rail and creep forces during rolling. In: Zobory [84]
10. Pombo, J., Ambrósio, J., Silva, M.: A new wheel–rail contact model for railway dynamics. Veh. Syst. Dyn. **45**, 165–189 (2007)
11. Zaazaa, K.E., Schwab, A.L.: Review of Joost Kalker's wheel–rail contact theories and their implementation in multibody codes. In: Proceedings of the ASME 2009 International Design Engineering Technical Conferences, September, pp. 1–12 (2009)
12. Magalhães, H., Marques, F., Liu, B., Antunes, P., Pombo, J., Flores, P., Ambrósio, J., Piotrowski, J., Bruni, S.: Implementation of a non-Hertzian contact model for railway dynamic application. Multibody Syst. Dyn. **48**, 41–78 (2020)
13. Vollebregt, E.A.H.: User guide for CONTACT, Rolling and sliding contact with friction. Technical Report 20-01, version 20.2, Vtech CMCC (2020). https://www.cmcc.nl/documentation
14. Vollebregt, E.A.H., Iwnicki, S.D., Xie, G., Shackleton, P.: Assessing the accuracy of different simplified frictional rolling contact algorithms. Veh. Syst. Dyn. **50**(1), 1–17 (2012). https://doi.org/10.1080/00423114.2011.552618
15. Meymand, S.Z., Keylin, A., Ahmadian, M.: A survey of wheel–rail contact models for rail vehicles. Veh. Syst. Dyn. **54**, 368–428 (2016)
16. Vollebregt, E.A.H., Weidemann, C., Kienberger, A.: Use of "CONTACT" in multi-body vehicle dynamics and profile wear simulation: initial results. In: Iwnicki, S.D. (ed.) Proceedings of the 22nd International Symposium on Dynamics of Vehicles on Roads and Tracks, Manchester, pp. 1–6 (2011). IAVSD
17. Vollebregt, E.A.H.: Comments on "The Kalker book of tables for non-Hertzian contact of wheel and rail". Veh. Syst. Dyn. **56**(9), 1451–1459 (2018). https://doi.org/10.1080/00423114.2017.1421767
18. Piotrowski, J., Bruni, S., Liu, B.B.: Reply to comments on "The Kalker book of tables for non-Hertzian contact of wheel and rail" by E.A.H. Vollebregt. Veh. Syst. Dyn. **56**, 1460–1469 (2018). https://doi.org/10.1080/00423114.2018.1437274
19. Piotrowski, J., Kalker, J.J.: The elastic cross-influence between two quasi-Hertzian contact zones. Veh. Syst. Dyn. **17**, 337–355 (1988)
20. Piotrowski, J., Chollet, H.: Wheel–rail contact models for vehicle system dynamics including multi-point contact. Veh. Syst. Dyn. **43**, 455–483 (2005)
21. Aceituno, J.F., Urda, P., Briales, E., Escalona, J.L.: Analysis of the two-point wheel–rail contact scenario using the knife-edge-equivalent contact constraint method. Mech. Mach. Theory **148**, 103803 (2020)
22. Vollebregt, E.A.H., Segal, A.: Solving conformal wheel–rail rolling contact problems. Veh. Syst. Dyn. **52**(suppl. 1), 455–468 (2014)
23. Vollebregt, E.A.H.: Conformal contact: corrections and new results. Veh. Syst. Dyn. **56**(10), 1622–1632 (2018). https://doi.org/10.1080/00423114.2018.1424917
24. Blanco-Lorenzo, J., Santamaria, J., Vadillo, E.G., Correa, N.: On the influence of conformity on wheel–rail rolling contact mechanics. Tribol. Int. **103**, 647–667 (2016)
25. Blanco-Lorenzo, J., Santamaria, J., Vadillo, E.G., Correa, N.: A contact mechanics study of 3D frictional conformal contact. Tribol. Int. **119**, 143–156 (2018)
26. Vollebregt, E.A.H.: Detailed wheel/rail geometry processing using the planar contact approach. Vtech CMCC, May 2020
27. Shabana, A.A., Zaazaa, K.E., Escalona, J.L., Sany, J.R.: Development of elastic force model for wheel/rail contact problems. J. Sound Vib. **269**, 295–325 (2004)
28. Nackenhorst, U.: The ALE-formulation of bodies in rolling contact - Theoretical foundations and finite element approach. Comput. Methods Appl. Mech. Eng. **193**, 4299–4322 (2004)
29. Donea, J., Huerta, A., Ponthot, J.Ph., Rodríguez-Ferran, A.: Arbitrary Lagrangian-Eulerian methods. In: Stein, E., de Borst, R., Hughes, J.R. (eds.) Encyclopedia of Computational Mechanics. Wiley, Chichester (2004). Chap. 14
30. Kaiser, I.: Refining the modelling of vehicle-track interaction. Veh. Syst. Dyn. **50**(suppl), 229–243 (2012)
31. Kalker, J.J.: Three-Dimensional Elastic Bodies in Rolling Contact. Solid Mechanics and Its Applications, vol. 2. Kluwer Academic, Dordrecht (1990)
32. Laursen, T.A.: Computational Contact and Impact Mechanics, Fundamentals of Modeling Interfacial Phenomena in Nonlinear Finite Element Analysis. Springer, Berlin (2002)
33. Benson, D.J., Hallquist, J.O.: A single surface contact algorithm for the post-buckling analysis of shell structures. Comput. Methods Appl. Mech. Eng. **78**, 141–163 (1990)
34. Wriggers, P.: Computational Contact Mechanics, 2nd edn. Springer, Heidelberg (2006)
35. Kalker, J.J.: Rolling contact phenomena - linear elasticity. In: Jacobson, B., Kalker, J.J. (eds.) Rolling Contact Phenomena. CISM Courses and Lectures, vol. 411, pp. 1–85. Springer, New York (2000)

36. Vollebregt, E.A.H., Wilders, P.: FASTSIM2: a second order accurate frictional rolling contact algorithm. Comput. Mech. **47**(1), 105–116 (2010). https://doi.org/10.1007/s00466-010-0536-7
37. Chen, H., Ishida, M., Namura, A., Baek, K.S., Nakahara, T., Leban, B., Pau, M.: Estimation of wheel/rail adhesion coefficient under wet conditions with measured boundary friction coefficient and real contact area. Wear **271**, 32–39 (2011)
38. Spiryagin, M., Polach, O., Cole, C.: Creep force modelling for rail traction vehicles based on the Fastsim algorithm. Veh. Syst. Dyn. **51**, 1765–1783 (2013)
39. Vollebregt, E.A.H.: Numerical modeling of measured railway creep versus creep-force curves with CONTACT. Wear **314**, 87–95 (2014)
40. Six, K., Meierhofer, A., Müller, G., Dietmaier, P.: Physical processes in wheel–rail contact and its implications on vehicle-track interaction. Veh. Syst. Dyn. **53**(5), 635–650 (2015)
41. Magel, E.E.: A survey of wheel/rail friction. Technical Report DOT/FRA/ORD-17/21, Federal Railroad Administration, September 2017
42. Vollebregt, E.A.H., van der Wekken, C.D.: Advanced modeling of wheel–rail friction phenomena. Technical Report TR19-11, VORtech, November 2019. FRA project
43. Elkins, J.A.: Prediction of wheel/rail interaction: the state-of-the-art. Veh. Syst. Dyn. **20**, 1–27 (1992)
44. Garg, V.K., Dukkipati, R.V.: Dynamics of Railway Vehicle Systems. Academic Press, Ontario (1984)
45. Wickens, A.H.: Fundamentals of Rail Vehicle Dynamics: Guidance and Stability. Swets and Zetlinger, Lisse, The Netherlands (2003)
46. Kalker, J.J.: On the rolling contact of two elastic bodies in the presence of dry friction. PhD thesis, Delft University of Technology (1967)
47. Johnson, K.L.: Contact Mechanics. Cambridge University Press, Cambridge (1985)
48. Duvaut, G., Lions, J.L.: Les Inéquations en Mécanique et en Physique. Dunod, Paris (1972)
49. Kalker, J.J.: Variational principles of contact elastostatics. J. Inst. Math. Appl. **20**, 199–219 (1977)
50. Kalker, J.J.: The computation of three-dimensional rolling contact with dry friction. Int. J. Numer. Methods Eng. **14**, 1293–1307 (1979)
51. Kalker, J.J.: Survey of wheel–rail rolling contact theory. Veh. Syst. Dyn. **8**(4), 317–358 (1979)
52. Kalker, J.J.: Über die Mechanik des Kontaktes zwischen Rad und Schiene. ZEV-Glas. Ann. **102**(7/8), 214–218 (1978)
53. Kalker, J.J.: A fast algorithm for the simplified theory of rolling contact. Veh. Syst. Dyn. **11**, 1–13 (1982)
54. Wang, G., Knothe, K.: The influence of inertia forces on steady-state rolling contact. Acta Mech. **79**, 221–232 (1989)
55. van der Wekken, C.D., Vollebregt, E.A.H.: Numerical calculation of the elastic field in a half-space using bilinear elements. Math. Mech. Solids **24**, 3537–3553 (2019)
56. Vollebregt, E.A.H.: A new solver for the elastic normal contact problem using conjugate gradients, deflation, and an FFT-based preconditioner. J. Comput. Phys. **257, Part A**, 333–351 (2014)
57. Zhao, J., Vollebregt, E.A.H., Oosterlee, C.W.: A fast nonlinear conjugate gradient based method for 3D concentrated frictional contact problems. J. Comput. Phys. **288**, 86–100 (2015)
58. Vollebregt, E.A.H.: Improving the speed and accuracy of the frictional rolling contact model "CONTACT". In: Topping, B.H.V., Adam, J.M., Pallarés, F.J., Bru, R., Romero, M.L. (eds.) Proceedings of the 10th International Conference on Computational Structures Technology, pp. 1–15. Civil-Comp Press, Stirlingshire (2010)
59. Kaiser, I., Poll, G., Voss, G., Vinolas, J.: The impact of structural flexibilities of wheelsets and rails on the hunting behaviour of a railway vehicle. Veh. Syst. Dyn. **57**, 564–594 (2019)
60. Spiryagin, M., Vollebregt, E., Hayman, M., Persson, I., Wu, Q., Bosomworth, C., Cole, C.: Development and computational performance improvement of the wheel–rail coupling for heavy haul locomotive track damage simulation studies. Veh. Syst. Dyn. (2020). https://doi.org/10.1080/00423114.2020.1803371
61. Hertz, H.: Über die Berührung fester elastischer Körper. J. Reine Angew. Math. **92**, 156–171 (1882)
62. Ayasse, J.B., Chollet, H.: Wheel–rail contact. In: Iwnicki [1], pp. 85–120, Chap. 4
63. Pascal, J.P., Sauvage, G.: The available methods to calculate the wheel/rail forces in non-Hertzian contact patches and rail damaging. Veh. Syst. Dyn. **22**, 263–275 (1993)
64. Ayasse, J.B., Chollet, H.: Determination of the wheel rail contact patch for semi-Hertzian conditions. Veh. Syst. Dyn. **43**, 159–170 (2005)
65. Linder, C.: Verschleiss von Eisenbahnrädern mit Unrundheiten. PhD thesis, Nr. 12342, ETH Zürich (1997)
66. Piotrowski, J., Kik, W.: A simplified model of wheel/rail contact mechanics for non-Hertzian problems and its application in rail vehicle dynamics simulations. Veh. Syst. Dyn. **46**(1–2), 27–48 (2008)
67. Sichani, M.Sh., Enblom, R., Berg, M.: Comparison of non-elliptic contact models: towards fast and accurate modelling of wheel–rail contact. Wear **314**, 111–117 (2014)
68. Liu, B., Bruni, S., Vollebregt, E.A.H.: A non-Hertzian method for solving wheel–rail normal contact problem taking into account the effect of yaw. Veh. Syst. Dyn. **54**(9), 1226–1246 (2016)

69. Sun, Y., Zhai, W., Guo, Y.: A robust non-Hertzian contact method for wheel–rail normal contact analysis. Veh. Syst. Dyn. **56**(12), 1899–1921 (2018)
70. Zhu, B., Zeng, J., Zhang, D., Wu, Y.: A non-Hertzian wheel–rail contact model considering wheelset yaw and its application in wheel wear prediction. Wear **432–433**, 202958 (2019)
71. Vermeulen, P.J., Johnson, K.L.: Contact of nonspherical bodies transmitting tangential forces. J. Appl. Mech. **31**, 338–340 (1964)
72. Shen, Z.Y., Hedrick, J.K., Elkins, J.A.: A comparison of alternative creep force models for rail vehicle dynamic analysis. Veh. Syst. Dyn. **12**(1), 79–83 (1983)
73. Polach, O.: A fast wheel–rail forces calculation computer code. Veh. Syst. Dyn. Suppl. **33**, 728–739 (1999)
74. Polach, O.: Creep forces in simulations of traction vehicles running on adhesion limit. Wear **258**, 992–1000 (2005)
75. Bosso, N., Zampieri, N.: A novel analytical method to calculate wheel–rail tangential forces and validation on a scaled roller-rig. Adv. Tribol. **2018**(1–11), 7298236 (2018)
76. Kalker, J.J.: Book of tables for the Hertzian creep-force law. In: Zobory [84], pp. 11–20
77. Sichani, M., Enblom, R., Berg, M.: An alternative to FASTSIM for tangential solution of the wheel–rail contact. Veh. Syst. Dyn. **54**(6), 748–764 (2016)
78. Piotrowski, J., Liu, B.B., Bruni, S.: The Kalker book of tables for non-Hertzian contact of wheel and rail. Veh. Syst. Dyn. **55**, 875–901 (2017). https://doi.org/10.1080/00423114.2017.1291980
79. Marques, F., Magalhães, H., Liu, B., Pombo, J., Flores, P., Ambrósio, J., Piotrowski, J., Bruni, S.: On the generation of enhanced lookup tables for wheel–rail contact models. Wear **434–435**, 202993 (2019)
80. Alonso, A., Giménez, J.G.: Tangential problem solution for non-elliptical contact areas with the FastSim algorithm. Veh. Syst. Dyn. **45**(4), 341–357 (2007)
81. Zhao, J., Vollebregt, E.A.H., Oosterlee, C.W.: Extending the BEM for elastic contact problems beyond the half-space approach. Math. Model. Anal. **21**(1), 119–141 (2016)
82. Blanco-Lorenzo, J., Vollebregt, E.A.H., Santamaria, J., Vadillo, E.G.: Approximating the influence coefficients of non-planar elastic solids for conformal contact analysis. Tribol. Int. (2020). https://doi.org/10.1016/j.triboint.2020.106671
83. Magel, E.E., Oldknow, K.: Quality indices for managing rail through grinding. In: Li, Z., Núñez, A. (eds.) Proceedings of the 11th International Conference on Contact Mechanics and Wear of Rail/Wheel Systems, pp. 658–667. Delft University of Technology, The Netherlands (2018)
84. Zobory, I. (ed.): Proceedings of the 2nd Mini Conference on Contact Mechanics and Wear of Wheel/Rail Systems Technical University of Budapest, Hungary (1996)