Design of an optimised wheel profile for rail vehicles operating on two track gauges

J. SANTAMARIA*, J. HERREROS, E.G. VADILLO, N. CORREA

Department of Mechanical Engineering. University of the Basque Country UPV/EHU. 
Alameda Urquijo s.n., 48013 Bilbao, Spain

Please cite this paper as: Santamaria, J., Herreros, J., Vadillo, E.G., Correa, N. 
Design of an optimised wheel profile for rail vehicles operating on two track gauges. 
Vehicle System Dynamics, Vol. 51, pp. 54-73. 2013

NOTICE: This is an Author's Accepted Manuscript of an article published in 
VEHICLE SYSTEM DYNAMICS, Vol. 51, Nº 1, pp. 54-73 on January 2013 
[copyright Taylor & Francis], available online at: 
http://www.tandfonline.com/10.1080/00423114.2012.711478

DOI: 10.1080/00423114.2012.711478

* Corresponding Author. Email: javier.santamaria@ehu.es
Design of an optimised wheel profile for rail vehicles operating on two track gauges

J. SANTAMARIA*, J. HERREROS, E.G. VADILLO, N. CORREA

Department of Mechanical Engineering, University of the Basque Country UPV/EHU
Alameda Urquijo s.n., 48013 Bilbao, Spain

This paper sets out an optimum synthesis methodology for wheel profiles of railway vehicles in order to secure good dynamic behaviour with different track configurations. Specifically, the optimisation process has been applied to the case of rail wheelsets mounted on double gauge bogies, that move over two different gauges, which also have different types of rail: the Iberian gauge (1668 mm) and the UIC gauge (1435 mm). Optimisation is performed using Genetic Algorithms and traditional optimisation methods in a complementary way. The objective function used is based on an ideal equivalent conicity curve which ensures good stability on straight sections and also proper negotiation of curves. To this end the curve is constructed in such a way that it is constant with a low value for small lateral wheelset displacements (with regard to stability), and increases as the displacements increase (to facilitate negotiation of curved sections). Using this kind of ideal conicity curve also enables a wheel profile to be secured where the contact points have a larger distribution over the active contact areas, making wear more homogeneous and reducing stresses. The result is a wheel profile with a conicity that is closer to the target conicity for both gauges studied, producing better curve negotiation while maintaining good stability on straight sections of track. The paper shows the resultant wheel profile, the contact curves it produces, and a number of dynamic analyses demonstrating better dynamic behaviour of the synthesised wheel on curved sections with respect to the original wheel.

Keywords: Wheel profile optimisation, Genetic Algorithms, wheel-rail contact, double gauge vehicles.

1. Introduction

The geometry of wheel-rail contact is a key feature in resolving the dynamics of a rail vehicle. It is well known that the shape employed to design both the wheel profile in the rolling area and the railhead substantially conditions the vehicle's response from the point of view of a number of factors such as stability, absence of vibrations, curve negotiation ability, wear of wheels and rails and ride safety, among others [1,2]. Hence, therefore, the need to secure the best possible profiles for the intended ride conditions. This leads to a better dynamic response, increasing the vehicle's stability, for instance, or reducing wear on the profiles and the forces transmitted to the track on curved sections.

Designing a wheel profile that will produce good dynamic behaviour is no easy task. The vehicle's dynamics are highly sensitive to the geometry of the profiles, and although the same wheel profile may perform satisfactorily on a certain vehicle or track, it may nevertheless produce a poor response when changes are made to the ride conditions. There is an additional factor which introduces particular complications to the process of obtaining an optimised profile: the inevitable wear that will ensue on the profile itself as a result of
operation over the first few thousand kilometres. Wear is not only an adverse circumstance from the economic viewpoint (material is eliminated, and the wheels must be reprofiled or replaced periodically, and thus minimising wear in all cases is a constant concern); it also brings about changes in the shape of the profile originally obtained [3-5]. To a greater or lesser extent, this culminates in premature loss of the dynamic wheel-rail contact characteristics that the initial profile design phase attempted to establish. If wheel wear occurs rapidly, the profile which was obtained during the design phase and performed well will only prove useful over a short period of time, and consequently this will not be a good design. In addition to other parameters, therefore, any methodology which seeks to optimise rolling profiles for rail dynamics must take account of homogeneous wear on the profile.

Until a few years ago the design process was manual, and based on the designer's experience, but IT progress has now made it possible to use specific optimisation methods to try to adjust the profiles to specific ride conditions. The great sensitivity of contact curves to minor changes in rolling profiles, however, makes optimisation by means of traditional methods problematic in these cases. Conventional optimisation methods alone may not produce satisfactory results, and frequently lead to relative minima that may be far from the optimum solution. Genetic Algorithms (GAs) can help to avoid local minima being obtained, which barely improve the initial behaviour of the wheel, and constitute a particularly useful method for this type of problem [6-8].

In Spain, the gauge used for all intercity tracks has traditionally been 1668 mm (Iberian gauge), as distinct from most European systems, where the gauge is generally 1435 mm (UIC gauge). When high speed railway lines were introduced in 1992, however, it was decided that the new high speed trains would use the international 1435 mm gauge employed in the rest of Europe. The result is that Spain now operates conventional 1668 mm gauges and high speed 1435 mm gauges, both with a rail cant of 1:20 (which also differs from most European tracks, with rail cants of 1:40).

In order to put to good use the advantages of high speed tracks for conventional railway vehicles, at the present time vehicles are also in operation which can adapt to two different gauges and make one portion of the journey on the conventional track and the other portion on high speed track, thereby making best use of rail infrastructures. These vehicles use what are known as double gauge bogies, where the wheelset wheels change position and adapt to the new setup on the gauge changers built for this purpose. These special vehicles can reach operating speeds of up to 250 km/h on both track gauges.

This therefore introduces a situation in which the same wheelset moves over tracks with different gauges at high speed. Nor is the type of rail used on the two gauges the same: a 54 E 1 rail is used on the 1668 mm gauge, and a 60 E 1 rail for the 1435 mm gauge. Both factors can lead to major differences in wheel-rail contact conditions, making the design of the wheel profile particularly important in this case. In order to get the vehicle providing a good dynamic response for both sets of ride conditions, the wheel profile must be designed in due consideration of the two different contact situations.

This paper presents a synthesis methodology of wheel profiles optimised for two specific ride conditions (Table 1): a Spanish high speed railway track with a 1435 mm gauge, a 60 E 1 rail, and a rail cant of 1:20 (UIC case), and a conventional Spanish track with a 1668 mm gauge, a 54 E 1 rail, and a rail cant of 1:20 (Iberian case). In Table 1, wheel base refers to the lateral distance between the nominal rolling points on the two wheels of a wheelset. Unlike traditional methods, optimisation is carried out by minimising an objective function using Genetic Algorithms (GAs) together with conventional optimisation methods. The objective function is based on the least squares error between a target conicity curve and the
calculated conicity curve for a given variable setting. The target conicity chosen is an ideal conicity curve which reflects good performance of the vehicle on both straight sections and curves, low wear, realistic geometry of the profile, and uniform distribution of contact points on the wheel profile.

Table 1: Parameters corresponding to the two gauges

<table>
<thead>
<tr>
<th>Track gauge (mm)</th>
<th>Wheel base (mm)</th>
<th>Rail</th>
<th>Rail cant</th>
</tr>
</thead>
<tbody>
<tr>
<td>UIC case</td>
<td>1435</td>
<td>1500</td>
<td>60 E 1</td>
</tr>
<tr>
<td>Iberian case</td>
<td>1668</td>
<td>1734</td>
<td>54 E 1</td>
</tr>
</tbody>
</table>

The bibliography contains a number of previous papers concerning the optimisation of rolling profiles and the utilisation of GAs. One of the first was drawn up by Persson et al. [6], in which GAs were used to optimise the rolling profile of a railway wheel, using the minimisation of several adverse ride parameters as a criterion. Persson subsequently used a similar methodology to study the potential optimisation of a rail profile on a suburban train [7]. In 2007 Novales et al. [8] used similar methodology and software to optimise a rail profile for two different types of track: railway and light rail tracks. The objective function considered three ride quality parameters, evaluated through dynamic simulations: wear index, risk of derailment and contact stresses. In 2008 Shevtsov et al. [9] employed a conventional optimisation method to synthesise a wheel profile considering the Rolling Radii Difference function (RRD), and demonstrated the improvements secured for a particular case of operation. The RRD function has continued to be used as the optimisation basis in more recent works - for example, Markine [10] and Shen [11] - improving contact conditions at low computational cost. In 2009 Polach [12] demonstrated it was possible to design a wheel profile with the aim of adjusting to a suitable conicity curve and achieving a wide contact spreading over the wheel profile (thereby helping to prevent concentrated wear). This synthesis method is a direct process (with no optimisation) which calculates the geometry of the profile that creates a certain conicity function in a certain operating situation. In 2011 Dabin Cui et al. [13] used a conventional optimisation procedure to obtain a wheel profile which minimised a new objective function: weighted wheel/rail gap. The aim of this minimisation process is, as in [12], to secure a distribution of contact points over the wheel profile for a certain type of track. Very recently, Pålsson et al. [14] successfully applied the GAs method to optimisation of the geometry of a crossover to reduce wear. The geometry of the entire crossover is described on the basis of only 4 definition parameters, and optimisation is carried out considering both travel directions.

The main new feature of this work lies in the optimisation of a wheel profile for two different gauges, where the operating speed of the vehicle is very high in both cases, and thus its stability becomes an extremely important condition. Methods such as that which focuses exclusively on a study of wheel/rail gap, or those which concentrate mainly on obtaining uniform distribution of contact points over the profile, or which only take account of the minimisation of certain parameters not directly related to stability, produce solutions with equivalent conicities that are too high and cannot be used in this case. Considering equivalent conicity as an objective function makes it possible from the outset to impose the condition that the conicity values for minor wheelset displacements must be extremely low, guaranteeing the stability of the vehicle. The work shows how considering this function can improve the vehicle's negotiation of curves, maintaining equivalent conicity within a certain range of values. Optimisation for two different track conditions at once also makes it
impossible to use direct methods for designing the geometry, as set out in [12,15], and renders the optimisation process more costly.

Furthermore, all profile optimisation projects utilise moving interpolation points and spline curves or NURBS curves [16] to define the optimised geometry. This paper, however, uses an approach similar to that used in [14], but applied to the wheel profile, constructing the geometry with definition parameters that enable the number of variables in the problem to be reduced and optimisation efficiency to be increased.

The following sections of the paper provide an explanation of the selection of this objective curve, the optimisation methodology, and the improvements achieved with the new wheel profile obtained. For the purpose of these comparative analyses dynamic simulations were performed on a type vehicle fitted with both the currently used and the new wheel profiles, and studies conducted on the levels of wear on curves of different radii. The work presented was commissioned by the Spanish Ministry of Development, as part of a project to design a wheel profile with low track aggressiveness, and was continued on a project by the Spanish Research Ministry (MICINN).

2. Problem description

2.1 Current wheel profile

Figure 1 shows the wheel profile currently used by vehicles travelling on the two co-existing gauges. Since stability of the vehicle is of such importance in these cases, the profile must have low conicity across a broad range of lateral wheelset displacements. Since this must also occur on two types of gauge, the wheel tread, which is composed of a straight section with a 2.5% gradient, extends to areas in proximity to transition to the flange. This means that the contact point with the rail remains on this straight section over a considerable range of lateral displacements of the wheelset and for different types of rail, and that equivalent conicity is constant and low (matching the gradient of the straight section) within this range. This is all very much to the benefit of the vehicle's stability, which is not the case with other profiles, where the straight section of the wheel tread begins to curve well before it approaches the wheel flange.

Figure 1. Wheel profile in current use.
In the case of the wheel shown in Figure 1, it may also be observed that the profile is composed of a concatenation of successive straight sections and circumference arcs. Specifically, in the active zone the profile comprises a straight section on the wheel tread, two circular sections, and a straight section corresponding to the flange.

The counterpart to this wheel design, extremely stable regardless of the type of rail or gauge, emerges on curved sections. For a rail vehicle to negotiate a curve, there must be a certain difference in the rolling radii of the two wheels on the wheelset. If this radii difference does not emerge, the contact point is located over the wheel flange, leading to major sliding and higher risk of derailment. To correct the vehicle's curve negotiation, the necessary value of the radii difference must be obtained before the flange makes contact with the rail.

Figure 2 represents the contact conditions of the wheel in Figure 1 on both types of gauge. The graphs were calculated considering the wheel and the rail as rigid solids and with a zero yaw angle, since on these types of tracks with high operating speeds and large-radius curves, this angle is always very small and its influence on the geometric contact parameters is negligible. The upper figures show the distribution of contact points on profiles for different lateral wheelset displacements. As may be observed, in both cases the contact points are concentrated on a small strip of the wheel tread, giving rise to low, almost constant conicity. When the value of lateral displacement is very high and clearance is exhausted, the contact point shifts abruptly to the flange with barely any intermediate points on the wheel. The lower figures show the RRD function and the conicity function as a result of lateral displacement of the wheelset, for the two tracks. It will be observed that the RRD value increases very slowly and practically constantly over a large range of lateral wheelset displacements. It is not until the wheel flange makes contact with the rail (where \( y \approx 6.6 \) mm) that the value of the RRD function suddenly increases. The conicities show a similar pattern of behaviour. Conicity is calculated by means of expression (1), where \( r_l \) and \( r_r \) are the left and right rolling radii respectively, and \( y \) is the wheelset's lateral displacement.

\[
\lambda(y) = \frac{r_l - r_r}{2y}
\]  

(1)

All this indicates that the wheelset will undergo major lateral displacements when negotiating curves, and that the wheel flange will contact the rail for most curve radii, with the exception of very open curves. The difference observed in wheel-rail clearance between the two curves (the sudden increase in the function occurs at values showing an appreciably difference to \( y \)) is because the increased wheel back-to-back distance is not exactly the same as the increase in the gauge (Table 1), and because the rail is also different.

The considerable distance between the contact point on the flange and the preceding points constitutes an adverse factor from the point of view of railway dynamics. When the wheel flange makes contact with the rail, a double contact point situation will arise in which one contact area will be located over the wheel tread and another over the flange, at a considerable distance from each other. This entails a great deal of wear on the flange contact point, in addition to other negative effects. Moreover, this situation will arise for curve radii which are not very small, since the profile's low conicity is unable to generate an RRD value that will prevent contact with the wheel flange.
Design of an optimised wheel profile for rail vehicles operating...

From these results it may be concluded that the wheel profile analysed provides excellent stability on straight sections, but there is opportunity for improvement for negotiation of curved sections by attempting to increase the RRD function. Furthermore, the fact that the contact zone is reduced to a small region on the wheel profile will lead to uneven wear, with the result that the shape of the wheel will change rapidly and the dynamic characteristics of the nominal profile will not be maintained over time. The following sections investigate the possibility of producing a wheel profile that maintains major stability at all times and succeeds in increasing the value of the RRD function to improve performances on curves.

3. Optimisation methodology

3.1 Objective function

The objective function formulated in this paper is a curve with optimum equivalent conicity. Unlike other projects in which the reference curve is the RRD rolling radii difference curve [9-11], the conicity curve supplies information in high speed applications which is more useful in terms of stability, since this enables equivalent conicity values to be accurately set that are permissible for slight lateral displacements of the wheelset. The shape of the curve also enables it to be deduced whether the profile will produce higher values of RRD in connection with negotiation of curves, thereby maintaining its advantages. It should be pointed out that basing optimisation on this type of contact curve has a much lower computational cost than selecting optimisations based on dynamic simulations. Where dynamic simulations are included, calculation time can take weeks [8], whereas in the cases set out in this paper it is reduced to a matter of hours. The counterpart naturally lies in the
fact that the use of contact curves contemplates global optimisation of dynamic behaviour, whereas with dynamic simulations it is possible to specify or elaborate on the reduction of specific parameters (risk of derailment, contact stresses) on a given curve radius or a representative routing.

Figure 3 shows the conicity curve presented as the target for the optimisation process. The curve corresponding to the Iberian gauge is shown by way of example. The curve maintains a constant value of $\lambda = 0.025$ up to $y = 4$ mm. This conicity matches the inclination of the tread on the original wheel, and ensures good vehicle stability. As of 4 mm, conicity increases (to increase the value of RRD), though smoothly. The smoothness of the curve will lead to a gradual lateral displacement of the wheelset as it moves over the curve, a smooth movement of the contact area over the wheel profile, minimisation of double contact point situations, and greater vehicle stability. Finally, the target curve is brought to an end at the same point as the original conicity curve (Figure 2d), since this value now corresponds to a contact point located on the wheel flange. As will be observed below, the wheel flange remains unalterable during optimisation, and thus this optimised wheel conicity value will almost completely match that of the original wheel, and cannot be changed. The curve finally portrayed in figure 3 depicts a fit to a simple exponential function beyond 4 mm, producing the necessary smoothness and an ever greater increase in the function as it approaches the final point, which will secure good dynamic results.

In the case of the UIC gauge, flange contact arises with a slightly greater lateral displacement of the wheelset. This point on the curve cannot be changed during the optimisation process either, and thus the function shown in Figure 3 must be adapted when the UIC case is analysed. The modification merely consists of a displacement to the right of the curve, until the final point matches the value of $y$ at which contact is made with the flange in the case of the UIC gauge.

Although for the purposes of computation of the representative equivalent conicity of a wheelset the custom is to employ $y = 3$ mm (as stipulated in UIC Leaflet 518 [17]), this project requires low conicity up to 4 mm in order to ensure that the stability of the optimised wheel is practically the same as that of the original wheel. It is assumed that improvements in terms of curve negotiation will emerge in the case of wheelset displacements greater than 4 mm, increasing the conicity curve from this point onwards.

The objective function $f$ is calculated by comparing the optimum conicity curves in Figure 3 to the conicity curves obtained with the real wheel analysed, on the basis of Least
Mean Squares. The comparison is carried out twice for each wheel analysed, once for each gauge case. The value of $f$ arises from the weighted sum of the squared vertical differences between the two curves, as shown in Equation (2), where $\lambda_j^{\text{obj}}(y_i)$ and $\lambda_j^{\text{obt}}(y_i)$ are respectively the objective and obtained conicities for a lateral displacement $y_i$ in the case of gauge $j$, and $p(y_i)$ is a penalisation coefficient varying with $y$. For the purposes of calculating $f$, lateral displacements $y_i$ were considered every 0.2 mm up to the flange, and thus $m = 36$. The geometric contact model used to calculate the $f$ function is a rigid model, where only one contact point on each wheel-rail profile is considered for each lateral displacement of the wheelset, and this enables calculation time to be shortened considerably.

$$f = \sum_{j=1}^{3} \sum_{i=1}^{m} \left[ (\lambda_j^{\text{obj}}(y_i) - \lambda_j^{\text{obt}}(y_i)) \cdot p(y_i) \right]^2$$ (2)

In this type of optimisation it is essential for the sum of squared differences to be corrected using a weighting factor $p(y)$ (Equation (3)). Non-consideration of this weighting factor inevitable leads to solutions which are not valid from the practical viewpoint. This is because certain conicity curves which are a good fit for the reference curve do not represent good wheel profiles. This is the case, for instance, of a curve that is a good fit for the reference curve for $y > 3$ mm, but with conicity values that are moderately high for shorter displacements. Although the sum of squared differences may be low, the wheel is unsuitable for high speed operations since it does not have low conicity for small lateral displacements.

There is also a possibility of using relative error instead of the absolute differences between the curves, thereby making the errors obtained for small conicities greater than those obtained for large conicities, with no need for an additional penalty function. In this project, however, it was considered appropriate to use the weighted absolute differences with a suitable penalty function which, for instance, would permit variable penalisation to be deployed for the area of constant conicity. Here it must be pointed out that it is the conicities of lateral displacements in proximity to the centred position which exert the most influence on the vehicle’s stability [18,19]. The weighting function therefore penalises differences between curves for the $y < 4$ mm range decreasingly, so that penalisation is greater the lower the value of $y$. The penalisation, moreover, is only applied if the actual conicity does not lie within a certain range of values considered admissible. Although conicity must not increase above a certain value for stability to be guaranteed, nor is it advisable for it to be reduced too much since this may cause low-frequency vibrations on the vehicle [12]. This paper has considered conicity to be admissible between 0.022 and 0.05.

The weighting function acts within the $y > 4$ mm range, penalising the squared vertical differences to a greater extent if the conicity obtained is lower than the target conicity. The intention is to direct the optimisation process towards a wheel which can actually obtain a rolling radii difference above $y = 4$ mm. It was observed during this work that introducing this weighting coefficient greatly improved the results obtained. Likewise, it would also be possible to add a weighting function to penalise any solutions that are unsatisfactory from the point of view of curvature ranges or gradients required for the profile depending on the zone, though it was not necessary for this work.
3.2 Definition of the optimisation variables

Each wheel is represented by a vector with optimisation variables $x$. The $x_i$ variables must unequivocally define the wheel profile they represent. Normally these variables are allocated to a number of points on the wheel profile plane that are interpolated subsequently to generate the complete profile. For example, it is possible to select $n$ points on the $y$ axis and allocate a variable $z$ coordinate to each, and these coordinates constitute the vector of design variables [6-10,13]. The profile is unequivocally defined by performing interpolation of the points, with spline curves, for example.

This paper shows a novel approach for the definition of the wheel profile in the context of an optimisation algorithm: the curve is defined on the basis of the successive connection of rectilinear sections and circular arcs, which must meet some specific boundary conditions. Only the active section of the wheel profile is defined, the section to be optimised. As shown in Figure 4, the section being optimised runs from the straight section on the wheel tread to the straight section on the flange. The left-hand section plays no part in wheel dynamics because the contact point is not located within this area. The wheel's flange also plays a major role in terms of running safety. Both the flange angle and the total thickness are parameters which exercise a decisive influence on running safety [20]. The profile is therefore left with the same shape as the original wheel as of the straight section at 70º, thus retaining the properties of the original wheel against the risk of derailment.

![Figure 4. Wheel profile section to be optimised, and definition variables.](image)

The boundary conditions to be met are as follows (see Figure 4):

a) The first section is a straight section with a 2.5% gradient which starts at the same point as the original wheel tread.

b) The last section must arrive at an inclination of 70º to the first point on the straight section comprising the wheel flange.

The $x_i$ design variables are the curvatures of the circular arches that shape the profile (zero for straight sections) and the $y$ coordinates of the boundary points separating one arc from the next. Although one may use as large a number of sections as one likes, the utilisation of 3 intermediate sections with constant curvature (apart from the initial
rectilinear section with a 2.5% gradient) produces a good compromise between flexibility in terms of definition of the wheel and convergence in the optimisation process. Figure 4 represents this definition. To do this, it is necessary to define only 4 design variables: the curvatures of the first two sections (1/R₁ and 1/R₂), the point where the first section starts (y₀) and the point where the second section starts (y₁). The tangency condition on the flange imposes two conditions which assume that both the third curvature (1/R₃) and the point of separation between the second and third section (y₂) are set when the other 4 variables have been defined. The x vector is finally defined as shown in Equation (4).

\[ \mathbf{x} = \begin{pmatrix} y_0, y_1, \frac{1}{R_1}, \frac{1}{R_2} \end{pmatrix}^T \]  

(4)

There are two main alternatives for defining the profiles: using curves such as B-Splines, Bezier's, NURBS; or using sections of curves. Both have their advantages and peculiarities. With respect to definition of the profile using circumference arcs, which is the option used in this paper, the following observations may be made:

a) Appropriate flexibility has been obtained, and a wide range of possible solutions covered with 4 design variables. In the specific case of the GAs, 4 variables is a very suitable number for convergence and produces good results [14].

b) As in the case of using an interpolation with B-splines, the result obtained matches the final actual wheel profile, with no need to smooth out the curve subsequently or correct the position of certain points in order to eliminate irregularities or unrealistic areas.

c) The wheel profile is defined in a simple manner using the circumference arcs of which it is made up.

d) It is not necessary to implement any additional restrictions on the optimisation process such as, for example, restrictions concerning the relative position of the points, the variation in the profile's gradient or changes in the sign of curvatures, although not all possible combinations of variables guarantee tangency on the flange, and these solutions must be eliminated using a high penalty.

The disadvantage in comparison to the utilisation of interpolating points to define the profile lies in a possible loss of flexibility. This loss of flexibility, however, is only relative, since it is always possible to increase the number of defining sections by extending the range of solutions, and additionally it has been observed that with these 4 variables alone it is possible to obtain a spectrum of profiles that will suffice to meet the needs of an optimisation process of the type presented in this paper.

3.3 Optimisation method

The wheel is optimised through minimisation of the objective function \( f \). This function is characterised by the abrupt changes to its value when the wheel profile undergoes slight variations, and hence the utilisation of optimisation methods is applied with problems and difficulties. The deployment of Genetic Algorithms (GAs) proves particularly useful in this case. They have been successfully applied in other disciplines, and also in a number of related projects. GAs are based on Darwin's theory of evolution, in which the finest specimens and those that adapt best to their surroundings show higher levels of reproduction, and thus the positive characteristics are passed on to subsequent generations.

In the optimisation process described in this paper, each individual in the same generation is a certain wheel, represented mathematically by \( \mathbf{x} \), and aptitude is measured
logically through the objective function described in the preceding section. The optimisation problem is thus set out in the following terms:

\[
\begin{align*}
\text{Minimise:} & \quad f(x), \quad x \in \mathbb{R}^n \\
\text{Subject to:} & \quad A_i < x_i < B_i \\
\text{Where:} & \quad x = \{x_1, x_2, \ldots, x_n\}^T, \quad i=1,\ldots,n
\end{align*}
\]  

(5)

Upper and lower limits are specified for the vector with the design variables \( x \) in order to reduce the search space to realistic and feasible solutions.

GAs optimisation is performed in the MATLAB environment using the appropriate toolbox [21], using the \textit{ga} function. Satisfactory results were obtained using a linear crossover type and a mutation rate of 30\%. The intention in using such a high rate of mutation is that, with an objective function such as that established in this paper, the solution will not rapidly converge towards a relative minimum, but that solutions which have not been analysed will continue to be explored during each generation. The property of elitism is also employed, so that the best specimen in each generation is preserved and the solution is not lost.

During the optimisation process, some of the individuals generated (in the initial population randomly generated and also through crossovers and mutations) are not valid from the geometric viewpoint. In other words, there are values of \( x \) within the boundaries stipulated in Equation (5) with which a wheel profile cannot be built. This is because, for certain combinations of curvatures and boundary points, tangency cannot be secured at the point where the wheel flange commences. In these cases, a particularly high value is allocated to the objective function so that these wheels are eliminated from successive generations.

The final result was obtained using an initial population of 500 individuals, and calculating 12 successive generations. The greater computational cost is associated with evaluation of the objective function of each individual. Practically all calculation time is expended on these evaluations, which require a detailed geometric analysis of profiles in contact for a large number of lateral positions of the wheelset (in each individual and in each generation), and the computational cost of the other tasks is marginal. The objective function was programmed in FORTRAN, and an appropriate communication link between the two languages was designed. The result was a calculation time of approximately 4 seconds for the evaluation of each individual, using an Intel Core Duo T9900 at 3.06 GHz with 4 GB of RAM. The time required for a full calculation is approximately 6.5 hours, a short calculation period in comparison to other wheel optimisation processes. This is because using the conicity curve as an objective function eliminates the need to conduct a dynamic analysis of the vehicle, which is much more costly in computational terms.

When an optimised wheel profile has been obtained, the result may be refined slightly more by applying conventional (non-evolutionary) optimisation methods on the basis of the solution obtained using GAs. This ensures that the solution obtained is effectively a minimum, although obviously it is not possible to guarantee that this is the absolute minimum. MATLAB has its own algorithms for this kind of optimisation such as \textit{fminsearch} (non-linear minimisation of a vectorial function based on the Nelder-Mead method). In this case calculation time is not so long, and it took 38 minutes to evaluate 100 iterations.
4. Optimised wheel profile for two gauge cases

When the optimisation process has been carried out (first using GAs and subsequently the Nelder-Mead method), a solution wheel profile is obtained. The conicity curves of this wheel provides the best fit (also considering the weighting function) for the target conicity curves defined for the two gauges. Figure 5 shows the wheel profile obtained, and also the curvature radii of the arcs and the $y$ coordinates separating successive sections. This wheel profile has been patented by the authors in Spain [22]. It may be observed that the length of the straight section of the wheel tread has now been reduced with respect to the original wheel, and a section with a small curvature ($R_1 = 212$ mm) starts in a zone still at some distance from the flange. This will bring about an increase in the RRD function although, as will be observed, this does not affect stability.

Figure 5. Wheel profile obtained as a result of optimisation. Results in mm.

Figure 6 represents the equivalent conicity curves of the optimised profile for the UIC gauge and the Iberian gauge (shown as unbroken lines) in comparison to the theoretical target conicity (broken lines). The result obtained provides a reasonably good fit in both cases, and conicity gradually increases as lateral displacement increases. The value of conicity also virtually matches that of the original wheel for $y$ values of less than 4 mm. This constant value of $\lambda = 0.025$ for small lateral displacements ensures the same level of stability as the original wheel.

It should be pointed out here that, in comparison to the optimisation of a profile for a single type of track, in this case it may be observed that it is difficult to obtain an accurate fit of the conicity curves for both the track types considered. A better fit for one of the curves is indeed possible, as the authors have demonstrated in the case of the UIC gauge [23], although this entails impairing the fit of the other curve. Thus a compromise solution must be found for the best possible balanced approximation to each of the conicity curves. Another remarkable aspect is the shift observed between both the optimised curves and also between the target curves for each type of track. As already mentioned, this is because the clearance between rail and flange is lower in the Iberian gauge than in the UIC gauge.
Figure 6. Comparison of the equivalent conicity obtained with the optimised profile (unbroken lines) versus the target conicity (broken lines) for the UIC gauge and the Iberian gauge (rigid contact model).

Figure 7 compares the wheel profile obtained after optimisation to the original wheel profile. The transition zone between tread and flange shows lesser concavity of the optimised profile in comparison to the original profile.

Figure 8 represents the distribution of contact points on the optimised wheel profile for both gauges. A rigid contact model was used to generate the graphs. It may be clearly observed that the contact points have been distributed quite uniformly, thereby improving the vehicle's behaviour, as will be shown in detail following analysis of the results with the elastic contact model.

In order to ascertain the advantages of the synthesised profile using an elastic contact model and a commercial dynamic simulation software, a file was generated containing the data of the profile adapted to the wheel-rail contact module in the SIMPACK dynamic simulation package [24]. This module furnishes a proper modelling of the double contact point situations that may arise, taking account of the elasticity of the bodies, and the contact curves are smoothed out with respect to the rigid model.
Figure 8. Distribution of contact points on the optimised wheel profile for the a) UIC gauge and b) the Iberian gauge.

Figure 9 represents distribution of the contact points along the wheel profile, comparing the optimised profile to the original for both gauge cases. The figures were obtained using the SIMPACK elastic wheel-rail contact module. It may be observed that there is a substantial improvement in the distribution of contact points between wheel and rail and, most importantly for this work, on both gauges. This will entail a smoother ride, greater comfort and lower wheel wear, thereby preserving the initial shape of the profile over a longer lifespan. This aspect is extremely relevant in that, as wear occurs on a wheel profile, the initial design of the profile changes, and the change leads to dynamic responses in the vehicle which differ from those originally expected.

UIC gauge

Iberian gauge

Figure 9. Improvement in the distribution of contact points on the wheel profile for the UIC gauge and the Iberian gauge. The original wheel is on the left, and the optimised wheel on the right.
5. Dynamic simulations with the optimised wheel profile

5.1 Quasi-static results

There follow some dynamic results obtained with the optimised wheel in the preceding section. The objective is to ascertain that using the optimised wheel is advantageous with respect to the original wheel, taking account of aspects such as curve negotiation, wear, running safety, maximum contact stresses and Rolling Contact Fatigue (RCF).

For the purposes of the study, quasi-static analyses were performed on the operating conditions of a railway vehicle negotiating representative curves of different radii for both gauges: 10000 m, 7500 m, 4000 m and 2500 m for the UIC gauge, and 5000 m, 3500 m, 2000 m and 1500 m for the Iberian gauge. A number of parameters were calculated for each radius to indicate the aspects under examination.

In accordance with the previous results, as expected the behaviour of the wheel on curved sections showed a significant improvement due to the change in the situation of the contact points. Figure 10 compares the contact areas obtained as a result of quasi-static analysis of the wheel on a curve of radius 4000 m, on the UIC gauge. An operating speed of 250 km/h and non-compensated acceleration of 1 m/s$^2$ were considered. The DINATREN dynamic simulation tool developed by the authors was used for the analysis, and considered a 3-D elastic multi-point contact model [25,26]. The case shown in Figure 10 is taken by way of example since this is representative in terms of differences in the contact situation, and may be extrapolated to other curve radii analysed. It will be observed that two contact patches emerge on the two wheels. The contact area of the optimised wheel’s tread is larger than on the original wheel, and the second contact point is located at a greater distance from the flange. This leads to smoother curve negotiation, with lower wear on the profile in the zone around the flange. In Figure 10, $N_1$ and $N_2$ provide a qualitative representation of the normal forces associated with each contact zone.

![Figure 10. Comparison of the quasi-static contact conditions on a curve of radius 4.000 m on the UIC gauge. a) Original wheel b) Optimised wheel.](image)

Figure 10 compares, for all the curve radii considered, the results obtained in terms of the wear index and risk of derailment. The wear index $W$ is calculated on the outer wheel on the first wheelset on the vehicle's front bogie, the most representative wheel. To calculate the wear index, the friction force emerging on the contact is used, according to Equation (6), where $t$ represents creep force (calculated using FASTSIM) and $\nu$ the creepage. The risk of derailment is calculated by means of the quotient between the lateral force and the vertical force applied to the wheelset ($L/V$), according to Nadal's theory. In all cases non-compensated acceleration of 1 m/s$^2$ was considered, with maximum cant 140 mm and maximum speed 250 km/h.
\[ W = t \cdot v \]  

The results show that wear is reduced significantly, especially as the radius of the curve diminishes. The substantial wear reduction secured for the optimised wheel with respect to the original wheel is as much as 60% - 70% on the UIC gauge, and up to 80% on the Iberian gauge.

Figure 11. Reduction obtained in terms of wear index and risk of derailment with the optimised wheel, for a number of curve radii (quasi-static analyses), on both gauges.

With regard to the derailment coefficient, values were also reduced on the new profile for each curve analysed. For the UIC gauge, in the most unfavourable case (a curve with a radius of 2500 m) the value was lowered from 0.14 to 0.05, while in the case of the Iberian gauge it was reduced from 0.19 to 0.05 over curves with a radius of 1500 m. All cases at any rate are far from the maximum value of 0.8 stipulated in UIC Leaflet 518 [17], and thus wear reduction is more significant.

The fact that on some graphs the derailment coefficient does not increase uniformly as the curve radius decreases is because on these curves the yaw angle is directed towards the inside of the curve. This makes it difficult for the wheel to derail and, as the graph shows, the derailment coefficients are extremely small. It is on the curve of radius \( R = 1500 \) that the yaw angle is directed towards the outside of the curve, and the derailment coefficient increases as the curves become tighter.

Figure 12 compares the maximum stresses emerging in the contact areas for a number of curve radii and both gauges. The figures show, for each curve radius, the maximum stresses calculated according to Hertz’s theory in each of the two contact patches that may emerge. In this figure, \( s_1 \) refers to the contact area closest to the wheel tread, and \( s_2 \) to the contact area closest to the flange. If only one contact zone appears on the wheel for a curve radius, it
is shown as $s_1$. Likewise, for each curve radius the figures show the stresses calculated with the original wheel (left) and the optimised wheel (right). As may be observed, a substantial reduction of contact stresses is achieved in the optimised wheel, for all cases analysed.

![Figure 12. Reduction obtained in terms of maximum contact stresses with the optimised wheel, for a number of curve radii (quasi-static analyses), on both gauges.](image)

Another adverse phenomenon which may emerge on the wheels is RCF. This consists of surface fatigue affecting the material on the wheel and also the rail, due to repetitive cycles of shear stress, especially on curves. Depending on contact conditions, surface cracks may evolve into wear or an RCF phenomenon. The factors determining one condition or the other and the severity of this condition are tangential and normal forces, contact stresses, creepage and the presence of lubrication fluid.

Since the risk of RCF is not included in the objective function during the optimisations process, it is important to verify that the synthesised wheel does not increase levels of RCF in the normal operating conditions intended for the vehicle. To predict the existence of RCF, Ekberg [27] proposes a surface fatigue index $FI_{surf}$ according to Equation (7):

$$FI_{surf} = \mu - \frac{2\pi abk}{3F_z}$$  \hspace{1cm} (7)

where $\mu$ is the traction coefficient, $a$ and $b$ are the semi-axes of the Hertz contact ellipse, $F_z$ is the normal force and $k$ is the shear yield stress. The traction coefficient is defined as the quotient of the tangential force and the normal force on contact, as seen in Equation (8):

$$\mu = \frac{\sqrt{F_x^2 + F_y^2}}{F_z}$$  \hspace{1cm} (8)

where $F_x$ is the longitudinal force and $F_y$ is the lateral force on contact. The formula predicts the emergence of RCF if $FI_{surf}$ is positive. The normal force $F_z$, the traction coefficient $\mu$ and a small contact area would increase the risk of RCF.

Applying this theory to the case analysed here, Figure 13 shows a representative examples of the results obtained for a small-radius curve, comparing the two wheels for both the UIC and Iberian gauges. In these figures, $s_1$ and $s_2$ have the same meaning as in Figure 12, and 400 MPa was used as the definition of shear yield stress $k$. This corresponds to heat-treated high quality steel, although it is now possible to find premium steels with a greater $k$ value.
It may be observed in Figure 13 that on the original wheel the surface fatigue index is positive at the contact point close to the flange, since the traction coefficient is high and the contact patch is small. Not only does the optimised wheel reduce the index, but it even manages to render it negative through the new contact conditions, thereby decreasing the risk of RCF. The double contact point situation disappears in the case of the Iberian gauge, as was also observed in Figure 12. This reduction arises for other curve radii, and thus it may be concluded that the new wheel is also well suited in terms of a lower risk of RCF.

5.2 Time-integration results

To complete the dynamic study of the new wheel, time simulations are carried out to analyze the vehicle’s behaviour as it negotiates real curves on the routing, on both the UIC and Iberian gauges. The aim is to corroborate the improvement to the vehicle's dynamic response using the optimised wheel with respect to the original wheel for the intended ride conditions.

In the case of the UIC gauge, the vehicle describes a curve of radius $R = 7250$ m, with an incoming transition curve, a circular section and an exit transition curve of respective lengths 460 m, 7056 m and 460 m. A final straight section stabilises the vehicle after the curve. The track cant is $h = 0$ mm, and operating speed $V = 250$ km/h.

Figure 14 shows the results obtained over time for a number of indicators of dynamic response, in the case of both the original wheel and the optimised wheel. These results were obtained using SIMPACK, considering an elastic multi-point contact model and modelling the entire car.

It will be observed that the lateral displacements on the first bogie's wheelsets are smaller with the optimised wheel. It will also be noted that, the wear index is significantly lower on both the first and the second wheelset. It may also be observed that the first wheelset's yaw angle is higher with the optimised wheel, but the positive values on the graphs show that the wheelset is directed towards the inside of the curve, so this value is advantageous. The derailment coefficient is slightly lower for the optimised wheel, although in practice this does not entail any improvement in the operating conditions, since the values are extremely small and negligible in both cases.

In the case of the Iberian gauge, the vehicle describes a curve of radius $R = 960$ m, with an incoming transition curve, a circular section and an exit transition curve of respective lengths 160 m, 320 m and 160 m. Track cant is $h = 140$ mm, and operating speed $V = 150$ km/h, causing non-compensated acceleration $a = 1$ m/s$^2$. 
Figure 14. Results of dynamic simulation for a full vehicle negotiating a curve of $R = 7250$ m on the UIC gauge, performed with SIMPACK. The original wheel is shown as an unbroken line, and the optimised wheel as a broken line. a) Lateral displacement of first wheelset. b) Lateral displacement of second wheelset. c) Wear index on the left wheel of first wheelset. d) Wear index on the left wheel of second wheelset. e) Derailment coefficient of first wheelset. f) Yaw angle of first wheelset.

Figure 15 shows the results obtained in this case. The wheelset displacements on the front bogie are similar, slightly smaller with the optimised wheel, and also on the transitions. The wear graphs show that, as the vehicle moves over this curve, in the case of the original wheel a double contact point is maintained over time on the first wheelset's outer wheel. In the case of the optimised wheel, however, this only arises on the transitions. The result is a lower wear rate, and also on the contact point on the wheel tread. Total wear on the wheel is also shown in this case, the sum of the wear rates on the two contact areas, and this is substantially lower on the optimised wheel. Finally, both the yaw angle and the
Design of an optimised wheel profile for rail vehicles operating...

risk of derailment decrease when negotiating the curve, thereby improving its dynamic response.

Figure 15. Results of dynamic simulation for a full vehicle negotiating a curve of $R = 960$ m on the Iberian gauge, performed with SIMPACK. The results obtained for the original wheel are shown as an unbroken line, and those of the optimised wheel as a broken line. a) Lateral displacement of first wheelset. b) Lateral displacement of second wheelset. c) Wear index on the left wheel of first wheelset. d) Wear index on tread and flange contact of the left wheel of first wheelset. e) Derailment coefficient of first wheelset. f) Yaw angle of first wheelset.

6. Conclusions

This paper presents and develops an optimum synthesis methodology for rail wheel profiles that can provide a satisfactory dynamic response with different track configurations and
high operating speeds. The optimisation process is based on the application of Genetic Algorithms as a complement to conventional optimisation methods. The combination produces good results in view of the high sensitivity of the wheel-rail contact problem to minor changes in the wheel profile. The objective function chosen is based on an ideal conicity curve, designed so that values are very low for small lateral displacements of the wheelset, increasing smoothly as the displacements become greater. This methodology was applied to design a wheel profile suitable for two track gauges, and a solution wheel profile was obtained which, in relation to the original wheel, improves the dynamic response of the vehicle on curves and maintains high stability on straight sections. It could be interesting as future work to combine the procedure explained in this paper with an estimation of the worn profile evolution, to ensure the stability of the profile shape over time.

The main conclusions drawn from the paper are as follows:

- An objective optimisation function based on an ideal conicity curve offers certain advantages in terms of designing wheel profiles for high operating speeds. The system produces an extremely low computational cost in comparison to functions based on dynamic simulations (the entire optimisation process takes less than 5 hours) and, as was observed throughout the work, it ensures good stability on straight sections, proper curve negotiation, and homogeneous distribution of contact points on the profiles.

- Defining the wheel profiles on the basis of circumference arcs with continuity on the first derivative, where the design variables are the curvatures of the arcs and the boundary points between them, provides a good compromise between flexibility for definition of the profile and the number of variables to be considered (only 4 were considered in the example presented). This also prevents unrealistic profiles with changes of signs on the gradient or curvature, with no need to introduce additional restrictions, although not all combinations of the 4 variables secure tangency on the flange.

- Applying this methodology to the design of a wheel capable of moving at high speed over two track configurations (the UIC and Iberian gauges), where not only the gauge but also the type of rail varies, has produced a new wheel profile which represents a substantial improvement to the dynamic behaviour of the original wheel. Quasi-static and dynamic analyses of the vehicle operating with the new wheel demonstrate that the optimisation methodology, as proposed in this paper, works in an extremely satisfactory fashion, and may be used for other wheel designs with different specifications.

8. Acknowledgements

The authors thank the Spanish Research Ministry MICINN for their funding through contract TRA2010-18386 including the FEDER funds of the European Union, as well as the Basque Government for its financial assistance through IT-453-10 and for Research Grant BFI08.172. They also thank the financial help received from UPV/EHU through the training and research unit UFI11/29.

9. References