An Application of the Finite Element Method to Curve Fitting

A. Samartin
L. Moreno

(1) Professor
(2) Assistant Professor. Dept of Structural Analysis. E.T.S. de Ingenieros de Caminos, Canales de Puertos Santander, Espania.

ABSTRACT
An application of the Finite Element Method (FEM) to the solution of a geometric problem is shown. The problem is related to curve fitting i.e. pass a curve trough a set of given points even if they are irregularly spaced. Situations where curves with cusps can be encountered in the practice and therefore smooth interpolating curves may be unsuitable. In this paper the possibilities of the FEM to deal with this type of problems are shown. A particular example of application to road planning is discussed. In this case the functional to be minimized should express the unpleasant effects of the road traveller. Some comparative numerical examples are also given.

INTRODUCTION
The main purpose of this paper is to show the possibility of using other curves than the traditional ones (straight lines, circles and clothoides) in order to define the longitudinal axis of a road. The paper has an interdisciplinary character and the proposed curves are simply the interpolation functions of the FEM and they are specified by minimizing a given functional.

The finite element method was first used more than twenty five years ago [1] in the solving of a structural problem. Since then it has been developed to a considerable degree to the point that now it constitutes an important tool for the studying of problems in the field of mathematics, enabling a large variety of physical situations to be dealt with. The mathematical foundations of the method are well established [2], and numerous variations and formulations are possible: weighting function techniques (Galerkin, collocation, etc.), semianalytical procedures (finite strips, layers and finite prisms, etc.), as the well known boundary element method, are just some of the examples of the numerous possibilities which exist. See [3] for a summary of such examples. In the present work, a concrete application is described, namely applying the method in the planning of roads.

FORMULATING THE PROBLEM
From a mathematical point of view the horizontal projection of the axis of the road can be established as follows

The coordinates of N points, \( P_i(x_i,y_i) \), are given with reference to a global coordinate system \( x,y \) (figure 1). These points are ordered according to the forward direction of the axis (i=1,2,3,... N).

A continuous curve with continuous slope and curvature which passes through the above-mentioned points needs to be found. This curve may need to satisfy in addition a number of other "boundary" conditions, in such a way that it commands an entry slope, exit slope or slope at a midway point, or values of the curvature \( \kappa \) at some arbitrary points \( P_k \).

APPLICATION OF THE FINITE ELEMENT METHOD (FEM)
The segment \( i \) is defined by the two extreme points \( P_i \) and \( P_{i+1} \), and according to Figure 1, the following can be written:

\[ l_i = \frac{1}{2} \left( (x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2 \right) \]

\[ \cos \alpha_i = \frac{x_{i+1} - x_i}{l_i}; \quad \sin \alpha_i = \frac{y_{i+1} - y_i}{l_i} \]

The required curve \( y=y(x) \) which passes through the points \( P_i \) and is continuous \( C^2 \) will be restricted by the following "smoothing" condition:

\[ y(x) \text{ minimizes the functional } \int ds \]

Where the curvilinear integral extends to \( y=y(x) \), \( ds \) is the differential of arc and \( R \) the radius of curvature.

The functional (1) to be minimized represents only a possibility among several ones. Other functionals can include higher order derivatives (first order derivative of the curvature) expressing the unpleasant effects of the road in the traveller. The mathematical treatment of these functionals by the FEM is similar to the one given here. See [4].

As it is known the finite element technique enables problem (1) to be solved by expressing the solution \( y=y(x) \) a sum of piece-wise functions. In this case the following considerations are valid:

For the segment \( i \), the adimensional local coordinates \((\xi,\eta)\) of the chord, shown in Figure 2, are adopted and related to the global coordinates by means the following expressions:

\[ \xi = \frac{1}{2} (x_i + x_{i+1}) \cos \alpha_i + \frac{1}{2} (y_i + y_{i+1}) \sin \alpha_i \]

\[ \eta = \frac{1}{2} (x_i + x_{i+1}) \sin \alpha_i - \frac{1}{2} (y_i + y_{i+1}) \cos \alpha_i \]

the inverse transformation of which is:
Figure 1. Polygon

\[ x = \frac{1}{2}x_i + \frac{1}{2}x_{i+1} + l_i(\cos \alpha_i - \sin \alpha_i) \]  
\[ y = \frac{1}{2}y_i + \frac{1}{2}y_{i+1} + l_i(\sin \alpha_i + \cos \alpha_i) \]

where \( l_i \) is the length of the segment \( i \).

As is usual in the FEM the required curve \( y = y(x) \) may be found within the section \( P_i^1 \) in the following form:

\[ n = (N_1, N_2, N_3, N_4) \]

where \( \theta_0, \theta_1, \theta_2, \theta_3 \) are the slopes and curvatures of the points \( \alpha_0, \alpha_1, \alpha_2, \alpha_3 \) corresponding to \( i \) and \( i+1 \) respectively in the case of Figure 2. As can be noticed, a local numbering has been introduced in the segment \( P_i^1 \).

The functions \( N_i, N_{i+1} \) correspond to the interpolation functions or shape functions and they normally adopt polynomials of the abscissa.

In order to linearize the problem, it is assumed that the points \( P_i \) and \( P_{i+1} \) are sufficiently close and then the curvature can be approximately expressed by the second derivative of the abscissa with respect to the ordinate, i.e.

\[ \frac{d^2}{dx^2} \ll 1 \]

In this case the interpolation functions are fifth order hermite polynomials:

\[ N_1 = \frac{1}{5}(5 - 7\xi^2 + 6\xi^3 + 3\xi^4 - 3\xi^5) \]  
\[ N_2 = \frac{1}{10}(5 + 7\xi^2 - 12\xi^3 + 6\xi^4 + 3\xi^5) \]

It can be shown that these interpolation functions satisfy the following equations:

\[ \frac{dN_1}{d\xi}|_{\xi=1} = 1 ; \quad \frac{dN_2}{d\xi}|_{\xi=1} = 0 \]

\[ \frac{d^2N_1}{d\xi^2}|_{\xi=1} = 0 ; \quad \frac{d^2N_2}{d\xi^2}|_{\xi=1} = 0 \]

\[ N_1(-1) = 0 ; \quad N_2(-1) = 0 \]

\[ N_1(1) = 0 ; \quad N_2(1) = 0 \]

These properties characterize the shape functions \( N_i, R_i \), and have allowed its analytical determination. These functions are shown in Figure 3.
Equation (4) can be written more conveniently in the following way:

\[ \eta = \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \]  \hspace{1cm} (7a)

with:

\[ \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} \frac{d a_1}{d a_0} \\ \frac{d^2 a_1}{d a_0^2} \end{bmatrix} \]  \hspace{1cm} (7b)

\[ \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = \begin{bmatrix} \frac{d a_1}{d \eta} \\ \frac{d^2 a_1}{d \eta^2} \end{bmatrix} \]  \hspace{1cm} (7c)

The continuity conditions which need to be satisfied by the curve \( y = y(x) \) are of the type \( C^2 \), i.e. the first and second derivatives with respect to the same axes must be the same at the joints which are considered as the extremes of the adjacent elements. That is to say, according to Figure 4 the following can be written:

\[ \frac{c_{i-1}}{2} + m_i = \frac{c_i}{2} + m_i = \lambda_i \]  \hspace{1cm} (8a)

\[ \frac{c_{i-1}^2}{8} - \frac{m_i}{4} = \frac{c_i^2}{8} - \frac{m_i}{4} = \tilde{\lambda}_i \]  \hspace{1cm} (8b)

where:

\[ m_i = \frac{c_{i-1} + c_i}{2} \]  \hspace{1cm} (8c)

\( m_i \) and \( \tilde{c}_i \) are the slope and curvature of the extreme \( \alpha \) of the section \( j(\alpha = 1, 2; j = i-1, i) \).

The parameters \( \lambda_i \) and \( \tilde{\lambda}_i \) must be selected in such a way that the functional (1) is minimized, i.e.:

\[ E = E(\eta) = \frac{1}{2} \int \left( \frac{d^2 a_1}{d \eta^2} \right)^2 d \eta = \frac{1}{2} \int \frac{1}{\alpha_1} \frac{1}{\alpha_2} \frac{1}{\alpha_3} \frac{1}{\alpha_4} \left( \frac{c_{i-1}}{2} \right)^2 \]  \hspace{1cm} (9)

where \( \eta_i \) corresponds to the ordinate at the interval \( P_i P_{i+1} \), given by the expression (4) and \( I \) is the number of sections (equal to \( N-1 \) in the case of an open polygon).

The contribution to (9) of a generic element \( i \) gives:

\[ E = \frac{1}{2} \int \left[ \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \right]^T \left[ \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \right] \]  \hspace{1cm} (10)

that is:

\[ E_i = \frac{1}{2} \int \left( \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \right)^T \left[ \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \right] d \eta \]

with:

\[ \left[ \begin{array}{cc} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{array} \right] d \eta \]

matrix of dimension (2x2) and where the second derivative with respect to \( \eta \) is indicated by a double accent.
the minimum of which is reached for the condition below:

\[
\frac{\partial E}{\partial u_i} = 0 \quad \text{for } i = 1, 2, \ldots, N \text{ (joints)}
\]

The following system is obtained:

\[
\begin{bmatrix}
 E_{11} & E_{12} & \cdots & E_{1j} \\
 E_{21} & E_{22} & \cdots & E_{2j} \\
 \vdots & \vdots & \ddots & \vdots \\
 E_{N1} & E_{N2} & \cdots & E_{Nj}
\end{bmatrix}
\begin{bmatrix}
 u_1 \\
 u_2 \\
 \vdots \\
 u_N
\end{bmatrix} = \begin{bmatrix}
 y_1 \\
 y_2 \\
 \vdots \\
 y_N
\end{bmatrix} = 0 \quad (10)
\]

which corresponds to the situation of open polygon which is a normal situation in road practice. \((N = 1 + \ell)\).

Obviously for the linear system of equations (10) with \(2N\) unknowns there exists the trivial solution: \(u_i = 0\) for all \(i\), which corresponds to a straight line.

In order to obtain the final solution, distortions are introduced at each section \(i\), with the following values at the extremes:

\[
\begin{align*}
\frac{\pi}{2} - m_1 & = \frac{\pi}{2} - m_{i+1} \\
\frac{\pi}{2} - m_2 & = \frac{\pi}{2} - m_{i+1}
\end{align*}
\]

That means the "initial forces or equivalent nodal forces" \(\vec{P}\) which appears in each section, and their expression in

\[
\begin{bmatrix}
 \vec{P}_1 \\
 \vec{P}_2 \\
 \vdots \\
 \vec{P}_{N-1}
\end{bmatrix} = \begin{bmatrix}
 \vec{P}_{11} & \vec{P}_{12} & \cdots & \vec{P}_{1j} \\
 \vec{P}_{21} & \vec{P}_{22} & \cdots & \vec{P}_{2j} \\
 \vdots & \vdots & \ddots & \vdots \\
 \vec{P}_{N-11} & \vec{P}_{N-12} & \cdots & \vec{P}_{N-1j}
\end{bmatrix}
\begin{bmatrix}
 u_1 \\
 u_2 \\
 \vdots \\
 u_{N-1}
\end{bmatrix} = \begin{bmatrix}
 \vec{u}_1 \\
 \vec{u}_2 \\
 \vdots \\
 \vec{u}_{N-1}
\end{bmatrix} = \begin{bmatrix}
 \vec{m}_1 \\
 \vec{m}_2 \\
 \vdots \\
 \vec{m}_{N-1}
\end{bmatrix} \quad (i = 1, 2)
\]

and quantities of the (kinematic variables) slope and curvature.

In this way the final system of linear equations, which enables the unknown \(u\) to be determined, is obtained:

\[
\begin{bmatrix}
 E_{11} & E_{12} & \cdots & E_{1j} \\
 E_{21} & E_{22} & \cdots & E_{2j} \\
 \vdots & \vdots & \ddots & \vdots \\
 E_{N-11} & E_{N-12} & \cdots & E_{N-1j}
\end{bmatrix}
\begin{bmatrix}
 u_1 \\
 u_2 \\
 \vdots \\
 u_{N-1}
\end{bmatrix} = \begin{bmatrix}
 y_1 \\
 y_2 \\
 \vdots \\
 y_{N-1}
\end{bmatrix} = \begin{bmatrix}
 \vec{u}_1 \\
 \vec{u}_2 \\
 \vdots \\
 \vec{u}_{N-1}
\end{bmatrix} = \begin{bmatrix}
 \vec{m}_1 \\
 \vec{m}_2 \\
 \vdots \\
 \vec{m}_{N-1}
\end{bmatrix} \quad (11)
\]

with \(I = N-1\), the number of sections and \(N\) the number of joints of the polygon.

Once the system (11) has been solved, the values of the slope and curvature at the extremes of each section \(i\), \((\alpha, \beta, \gamma)\), can be deduced. By means of the interpolation equations (4) and the transformation equations (3), the coordinates \((x, y)\) of the different points of the required curve \(y = y(x)\) can also be deduced.
EXAMPLES OF APPLICATION

In order to test the approximation of the method, several simple cases have been studied.

The first six cases, the data correspond to the coordinates of four points \( P_i \) situated along a circle of radius \( R = 100 \) and equally spaced the angle \( \alpha = \frac{2\pi}{4} = \pi/2 \) (Figure 5). In some cases, the values of the angles and/or curvatures are also specified in the two extreme points \( P_i \) and \( P_4 \), i.e., the entry and exit joints. The results obtained from the FEM are compared to the values of the given circle. The numerical sensibility of the FEM is studied by analyzing successive cases corresponding to different values of the angle \( \alpha = 10^\circ, 20^\circ \) and \( 30^\circ \).

Another example corresponds to the one shown in the Figure 6. It is a transition curve composed by two clothoides of parameter \( A = 100 \) and length \( L = 100 \). The coordinates \( x_i, y_i \), slope \( \theta_i \) and curvature \( \kappa_i \) of the two points \( P_i \) and \( P_j \) are specified and only the coordinates at \( P_i \) and \( P_j \). The points \( P_2 \) and \( P_3 \) are the entry and exit points. In order to check the efficiency of the method two intermediate points are situated randomly along the transition curve. Then, five extreme cases have been studied and each of them is defined by the distances \( L_1 \) and \( L_2 \) of the points \( P_2 \) and \( P_3 \). The results and the comparison with the values obtained from the transition curve are shown in Table II.

<table>
<thead>
<tr>
<th>CASE</th>
<th>( \alpha = \alpha_1 + \alpha_2 = 30^\circ )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0000 0.0000 0.0000 0.0000 0.0000</td>
</tr>
<tr>
<td>2</td>
<td>0.3667 0.0667 0.6667 0.6667 0.0667</td>
</tr>
<tr>
<td>3</td>
<td>0.3666 0.1667 0.1667 0.6667 0.3667</td>
</tr>
<tr>
<td>4</td>
<td>0.3666 0.1667 0.1667 0.6667 0.3667</td>
</tr>
<tr>
<td>5</td>
<td>0.3666 0.1667 0.1667 0.6667 0.3667</td>
</tr>
</tbody>
</table>

FEM: Specified values of the slopes and curvatures in the two extreme points.

FEMC: Specified values of the slopes in the two extreme points.

FCMC: Specified values of the curvatures in the two extreme points.

PMC: Free values of the slopes and curvatures in all the points.

Figure 5. Example of application 1.

is reached when the slope and curvature are specified at the entry and exit points. This conclusion still is valid even for the largest separation angles \( (30^\circ) \) and different angles \( (10^\circ, 20^\circ \) and \( 30^\circ \)). If only the slope is specified in the entry and exit points the above concordance still holds but not too closely. However, if no values of slope and curvature are given at the extreme points of the axis, the differences between the results increases. The explanation for difference has to be found from the fact it may be possible to find a curve different to the circle passing through the four points \( P_i \), \( i = 1, 2, 3, 4 \) that produces a smaller value to the functional (1) than the circle.

The results and the comparison with the values obtained from the transition curve are shown in Table II.

Some differences between them are observed. The reason should be found from the fact that the clothoides perhaps is not the curve minimizing the functional (1). The scarce number of points used in all these cases to define the curve can also explain these differences. However when the points are located along the curve with some engineering judgement, for example the two intermediate points near the inflexion point or contact between clothoides, these differences are dramatically reduced (case 5).

If the number of points to define the transition curve is
increased to six (Figure 7), the results obtained are given in the Table III and the increase in the accuracy is observed.

From the two above examples it can be deduced that the FEM allows to define unequivocally the curve of the road axis, subject to be continuous C^2 and minimizing a functional of the type (1). In this way, it is not necessary to use the traditional special curves such as straight lines, circles and clothoïdes, and therefore a more wide freedom for the road designer is reached. The axis can be defined using the FEM simply by the nodal coordinates of some special points (joints) and some extra constraints in the slope and/or curvature. The coordinates of intermediate points between two consecutive points can be obtained from the use of the shape or interpolation functions (6).

EXTENSIONS OF THE METHOD

It is understood that the technique which has just been described can be extended to deal with more complex cases which include conditions of slope and curvature specified at one or various points of the polygon, the procedure is very simple and is exactly the same as for structural situations where boundary conditions are introduced automatically and in a general way for calculations to be carried out by computer. See reference [5] for an interesting description of this point.

Obviously, simultaneous treatment of the horizontal plane and elevation can be carried out following similar criteria to those indicated, in such a case the functional to be minimized may be the following:

\[
\int \left( \frac{1}{R} + \lambda^2 \frac{1}{T^2} \right) ds
\]

(12)

where 1/R and 1/T correspond to the mainly bending and torsional curvatures, and \( \lambda \) a parameter which can be specified according to the conditions of use of the road. The previous functional (12) may be, and is probably more adequately decomposed into the following:

\[
\int \left( \frac{1}{R} + \frac{3}{R^2} + \frac{\lambda^2}{T^2} \right) ds
\]

with \( R_g \) and \( R_v \) the curvatures in the horizontal and vertical planes of the curve of the axis.

Finally, although the connexion with the planning of roads is less obvious, the method can be used successfully along the lines indicated above for the representation of surfaces (e.g. land representation), surfaces which passing through a series of points \( P_i(x_i,y_i,z_i) \) are found to be conditioned by simple continuity requirements, including the first derivative or even the curvature. The most important aspect of the FEM both for this type of problem and for other problems, involves the selection of shape or interpolation functions, since the composing and solving of the system of equations (11) is standard and can be found in any general matrix programmes of structures, for example: SAP, STRUDL, ANSYS, HASTRAN, etc. Figure 8 shows the possibility of C^2 elements which might be used or the more general problem of smooth and continuous surface representation. If only the continuity of the slope is required, the selection of triangular and compatible elements is very wide, and these elements correspond to all the compatible elements of bending of plates (refer to the specialised literature
above all two recent publications[6] and[7].

15 dof O $w$, $w_x$, $w_y$
    $w_{xx}$, $w_{xy}$, $w_{yy}$
    $w_{xxx}$, $w_{xyy}$, $w_{xyy}$$w_{yyyy}$
1 dof $n$ (normal derivative)
1 dof $n_0$ (normal derivative)
1 dof $w$ (function)

Figure 8. Continuity $C^2$ triangular element.

CONCLUSIONS

The FEM is an important tool to transform continuous problems into discrete ones in Structural Analysis. Applying the method to the solving of problems in other fields of Science and Technology has produced impressive results. The present work has shown by means of a simple example applied to road design that the FEM offers significant possibilities in the solution of practical curve and surface fitting problems.

REFERENCES

6. J. Udras del Valle and A. Samartín, "Una contribución al Estudio de Elementos Finitos en Flexión de..."