

SECOND ORDER NURBS INTERPOLATION OF REAL AFFINE AND PROJECTIVE PLANE CURVES

Luca Ghezzi

November 2001

Abstract

A method is presented to interpolate real affine plane curves with second degree NURBS (Non Uniform Rational Bezier Splines). The curve into object is initially partitioned into an arbitrary number of arcs. Then each arc is approximated with a conic section passing thru its extremes, thru a third intermediate point and being tangent to the arc at the extremes. Each arc is then parameterized by a second degree NURBS whose coefficients are computed in order to exactly fit the conic. Finally, the resulting NURBS arcs are joined together to form a unique global curve. The method is capable to manage also non-simple (i.e. self-intersecting), non-regular (i.e. piece wise differentiable) curves. A generalization of the method is also presented, capable to manage curve arcs containing improper points. This procedure, fully deploying NURBS interpolation capabilities, is theoretically set in the framework of real projective geometry.

1 Introduction

The problem of interpolating a generic plane line occurs in many practical circumstances. Essentially, the problem is that of finding a curve belonging to some determined family and being able to be analytically manipulated, in order to fit another curve, more complex or with an unknown parameterization, or a series of sampled points. In what follows, a method will be presented to interpolate plane curves by means of rational algebraic curves constituted by arcs of conics suitably joined together. This method is particularly aimed to the subsequent analytical description of profiles of plane objects, in view of further numerical processing for various applications, such as the solution of minimal distance problems (to be applied to contact modeling in multi-body and/or finite element codes) or simply for graphical purposes.

In case interpolation is carried out by means of high order polynomials, the possibility exists that little changes in the set of points to be interpolated will cause great changes in the fitting curve, which is not desirable. This sort of “instability” is a consequence of the increasing versatility of high order polynomials, which may go thru small spaced sampled points by assuming also very contorted shapes. For this reasons it is usually preferable to use interpolating curves constituted by many low order arcs, instead of few high order arcs. One of the main advantages of the method hereby discussed is its stability, due to the choice of conic arcs as basic components of the final global interpolating curve. As a matter of fact, conics are polynomials with a very low degree (i.e. 2) and the aforementioned drawbacks are prevented. Another possibility, would be to use piece wise cubic interpolation. The choice of conics instead of cubics is essentially due to keep the procedure as simple as possible.

The method hereby presented to interpolate real affine plane curves is based on piece wise quadratic interpolation. This means that the curve to be interpolated, called the global curve, is suitably partitioned into sub-arcs, called local arcs. The number and dislocation of local arcs can be freely chosen according to the level of precision to be attained. An approximating conic arc then substitutes every local arc, in such a way that the former is tangent to the latter in its extreme points and an additional intermediate passage point exists common to both. Every conic arc is then parameterized as a Non Uniform Rational B-Spline (NURBS). Then, these local interpolating curves are joined together in a suitable way in order to produce a unique global interpolating curve.

This procedure is suitable for real affine curves only, so that no improper points are allowed. Nevertheless, an extension of the method has been developed so to have the possibility to deal with real projective curves. The procedure functioning is analogous to the affine one, the only difference being the way NURBS curves are defined and conceived in the projective rather than in the affine plane. Particularly, in the former case a planar curve is attained after having embedded the real projective plane into tri-dimension space. This way of proceeding allows extending to the real projective plane the entire theoretical framework of spatial NURBS curves, together with the peculiar features of a projective manifold. The outcome is a powerful method enabling the user to get rid of the restrictions of affine curve parameterization. As a matter of fact, improper points are allowed, so that it is no needed worrying about parallel lines intersecting in their improper point nor to simulate “infinite valued” coordinates with “big numbers”. Apart from the obvious enhancements from the numerical standpoint together with the theoretical consistence (and elegance) of this second method, an effectively wider class of curves is managed, being possible to have arcs going to and coming from points placed “at infinity”.

This work is articulated as follows: in sect. **2**, notation and conventions are established and the **problem of finding the conic** satisfying the above mentioned interpolation conditions is discussed and analytically posed. Particularly, sufficient conditions are formulated and proved in order to obtain arcs suitable for later NURBS parameterization. Two local reference frames are also defined and discussed in detail, together with the relevant affinity and projectivity needed to pass from a generic global reference frame to the local one. Particularly, the second local reference frame is crucial for the analytical formulation of the whole theory and for the proofs of the main theorems. In sect. **3**, **NURBS theory** is briefly introduced as regards real affine curves, particular attention being paid to the usage to be done in this work, namely to second degree curves. In sect. **4**, the problem of finding **NURBS parameterization of conics** is deeply discussed, and all of the possible exceptions are handled and solved. The main theory is specialized to the simple, but frequently recurring, cases of a line segment and a circumference arc. Sect. **5** deals with the **extension to the real projective plane**. Particularly, an in depth analytical discussion is carried on to understand the geometrical meaning of such generalized NURBS curves. In sect. **6**, a method is presented to join together the local NURBS curve into a **global interpolating NURBS curve**. Such a method is effective both for affine and projective real curves. A practical example is provided for both the categories of curves, showing the potentialities of quadratic NURBS interpolation. Finally, in sect. **7** some general qualitative hints are given about the **interpolation strategy**.

2 Determination of the local fitting conic

A generic conic in the real projective plane $\mathbb{P}^2(\mathbb{R})$ is the zero set of a quadratic form

$$f(P) = P^T A P \tag{1}$$

where P is a vector collecting the three homogeneous coordinates and A is a symmetric matrix:

$$P = \begin{pmatrix} x \\ y \\ u \end{pmatrix} \quad A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix} \tag{2}$$

Expanding (1) one gets

$$f(P) = a_{11}x^2 + a_{22}y^2 + a_{33}u^2 + 2a_{12}xy + 2a_{23}yu + 2a_{13}xu \tag{3}$$

Considering the (real) projective space $\mathbb{P}^5(\mathbb{R})$, a conic is a point with homogeneous coordinates $(a_{11}, a_{22}, a_{33}, a_{12}, a_{23}, a_{13})$. Therefore, a conic is univocally determined provided that 5 (independent) linear conditions are specified. This is equivalent to intersecting 5 hyper-planes (each of them is defined by a linear equation) in $\mathbb{P}^5(\mathbb{R})$. If, and only if, those hyper-planes are linearly independent, then their equations are satisfied only by the points belonging to a homogeneous line, i.e. the solution is unique up to a constant factor. Some caution need to be taken in the choice of the linear conditions to be applied, because the solution could not be real. In other words the conic could be belonging to $\mathbb{P}^5(\mathbb{C})$ but not to $\mathbb{P}^5(\mathbb{R})$. This method, intended to interpolate real curves, will always be referred to curves going thru real

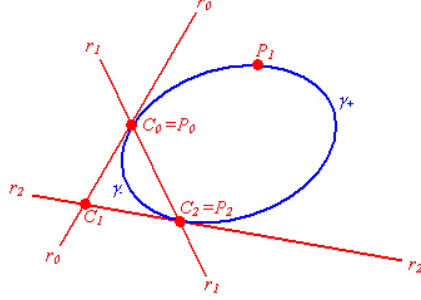


Figure 1: Notation

points and tangent to real lines, so that no non-real conics will be introduced. Moreover, in view of the successive parameterization of the conic, some further cautions need to be taken, as later detailed.

The conic we are looking for is passing thru 3 prescribed points, namely P_0 , P_1 and P_2 , and in 2 of this 3 points (say P_0 and P_2) it is tangent to 2 prescribed lines (r_0 and r_2 , respectively). In the following, we will show that these 5 conditions are linear and linearly independent, and we will show a simple condition on the 3 points and the 2 tangent lines so to have conics able to be parameterized according to the method hereby illustrated.

Before that, we will describe the notation to be used throughout the rest of this proceeding (see Figure 1). The tangent to the conic in P_0 will be called r_0 and, analogously, the tangent to the conic in P_2 will be called r_2 . We will call r_1 the line going thru P_0 and P_2 . In addition to the 3 passage points P_0 , P_1 and P_2 , 3 more points C_0 , C_1 and C_2 , named control points, will be defined for later usage. We will systematically place C_0 and C_2 coincident with P_0 and P_2 , respectively. C_1 will be the intersecting point of the two tangents r_0 and r_2 . In view of this, the naming convention for the 3 lines assume a precise logic: r_1 is the polar to the conic for C_1 and, analogously, r_0 and r_2 are the polars to the conic for P_0 and P_2 , respectively. Shortly, r_i is the polar to the conic for C_i for $i = 0,1,2$. The conic itself will be called γ . We will also call γ_+ the arc of conic starting from P_0 , ending at P_2 and going thru P_1 , while the complementary arc will be named γ_- .

First of all, we will consider the condition that a point $P_i(x_{P_i} : y_{P_i} : z_{P_i})$ ($i \in \{0,1,2\}$) belongs to the conic, i.e. that its projective coordinates satisfy eqn. (1)

$$P_i^T A P_i = 0 \quad (i \in \{0,1,2\}) \quad (4)$$

We can rewrite this condition in its extended form:

$$a_{11}x_{P_i}^2 + a_{22}y_{P_i}^2 + a_{33}u_{P_i}^2 + 2a_{12}x_{P_i}y_{P_i} + 2a_{23}y_{P_i}u_{P_i} + 2a_{13}x_{P_i}u_{P_i} = 0 \quad (i \in \{0,1,2\}) \quad (5)$$

Obviously this condition is linear. Moreover, since the two control points C_0 and C_2 coincide with the two passage points P_0 and P_2 , the following relations (useful when finding the NURBS parameterization of a conic) also holds:

$$C_i^T A C_i = 0 \quad (i \in \{0,2\}) \quad (6)$$

As regards the tangency condition, first we will write the equation of the line tangent to a conic described by $f(x : y : u) = 0$ in $P_i(x_{P_i} : y_{P_i} : z_{P_i})$ ($i \in \{0,2\}$):

$$\left(\frac{\partial f}{\partial x}\right)_{P_i} x + \left(\frac{\partial f}{\partial y}\right)_{P_i} y + \left(\frac{\partial f}{\partial u}\right)_{P_i} u = 0 \quad (i \in \{0,2\}) \quad (7)$$

Taking into account of eqn. (3) we get:

$$\begin{aligned} (a_{11}x_{P_i} + a_{12}y_{P_i} + a_{13}u_{P_i})x + (a_{12}x_{P_i} + a_{22}y_{P_i} + a_{23}u_{P_i})y + \\ (a_{13}x_{P_i} + a_{23}y_{P_i} + a_{33}u_{P_i})u = 0 \quad (i \in \{0,2\}) \end{aligned} \quad (8)$$

This line has to coincide with a prescribed line, namely r_i ($i \in \{0,2\}$). Apparently, this would mean that two conditions have to be imposed, since a line is completely described by 3 homogeneous parameters (for

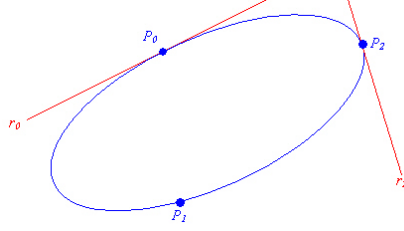


Figure 2:

example the 3 coefficients of x , y and u terms). However, the line described by eqn. (8) is, by construction, automatically going thru point P_i , so that only 1 (and not 2) linear condition has to be imposed. We can accomplish to this task by imposing that line (8) goes thru a given point $Q_i(x_{Q_i} : y_{Q_i} : z_{Q_i})$ of r_i . The only constraint on Q_i is not being coincident with P_i . Substituting the homogeneous coordinates of a generic point Q_i into eqn. (8) we get:

$$a_{11}x_{P_i}x_{Q_i} + a_{22}y_{P_i}y_{Q_i} + a_{33}u_{P_i}u_{Q_i} + a_{12}(x_{P_i}y_{Q_i} + y_{P_i}x_{Q_i}) + a_{23}(y_{P_i}u_{Q_i} + u_{P_i}y_{Q_i}) + a_{13}(x_{P_i}u_{Q_i} + u_{P_i}x_{Q_i}) = 0 \quad (i \in \{0, 2\}) \quad (9)$$

Eqn. (9) can be viewed also as $Q_i^T A P_i$ i.e. as the scalar product between vector Q_i and the vector resulting from matrix multiplication $A P_i$. However, $P_i = C_i$ for $i = 0$ and $i = 2$; besides, Q_i can be substituted by any point belonging to r_i , such as C_1 , which is a point of both the tangents r_0 and r_2 . Therefore, the following relations also hold

$$C_1^T A C_i = 0 \quad (i \in \{0, 2\}) \quad (10)$$

This, together with (6), will be very useful in finding the NURBS parameterization of a conic. It is also worth stressing that the condition that the polar to the conic for C_1 goes thru C_i will lead to exactly the same relation (10), in perfect agreement with the duality principle in Projective Geometry.

Of course also the tangency condition is linear, and then the system made up by the 3 eqns. (5) and the 2 eqns. (9) in the 6 homogeneous unknowns (a_{11} , a_{22} , a_{33} , a_{12} , a_{23} , a_{13}) is linear. Since the unknowns are homogeneous parameters, in order to have exactly one conic, the linear system should have ∞^1 solutions, which equivalent to say that the coefficient matrix rank should be 5. Since this matrix is 5×6 -dimensional, its rank is necessarily lower or equal to 5 (and then the existence of a solution must not be questioned). The rank is exactly 5 (and then the solution is unique) if, and only if, the 5 linear conditions (5-9) are linearly independent, which is not evident a priori. We will demonstrate that actually this is the case (and then that conic satisfying those conditions is unique) with the following

Theorem 2.1 *There is one, and only one, conic going thru three non collinear points P_0 , P_1 , P_2 , being tangent in P_0 to a line r_0 not containing neither P_1 nor P_2 , and being tangent in P_2 to a line r_2 not containing neither P_1 nor P_0 (see Figure 2).*

Proof The existence of at least one conic satisfying the 5 linear conditions algebraically expressing the geometric conditions is straightforward, as already explained. As regards the uniqueness of the conic, it necessary and sufficient to show that the 5 linear conditions are independent. Let us assume, as an absurd hypothesis, that the aforementioned 5 linear conditions are linearly dependent. If so, at least 1 of them is a consequence of the other 4. Then, any conic satisfying such 4 conditions will automatically satisfy the fifth one, too. Two distinct cases are possible. In the first case, the fifth condition is one of the two tangency conditions, say that the conic has to be tangent in P_2 to r_2 . Therefore, all conics going thru P_0 , P_1 , P_2 and being tangent in P_0 to r_0 should also be tangent to r_2 . Now, let us consider (see Figure 3) the degenerate conic constituted by the line going thru P_0 and P_1 and the line going thru P_2 and P_0 . Notice that P_0 is a singular double point for such a conic, so that the intersection multiplicity with any line, like r_0 , going thru P_0 is 2. Besides, since, by hypothesis, $P_2 \in r_0$ and $P_0 \in r_2$, the line thru P_0 and P_2 is not a component of the degenerate conic, so that P_2 is a regular point for such a conic and the intersection multiplicity with r_2 is 1. Therefore, this conic is tangent to r_0 but not to r_2 and we

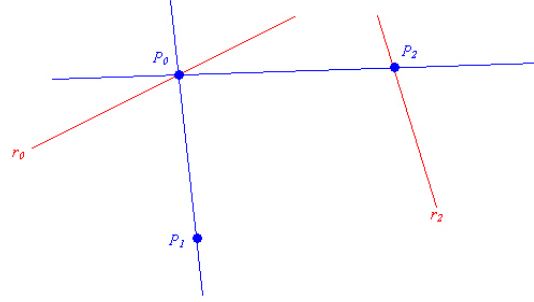


Figure 3:

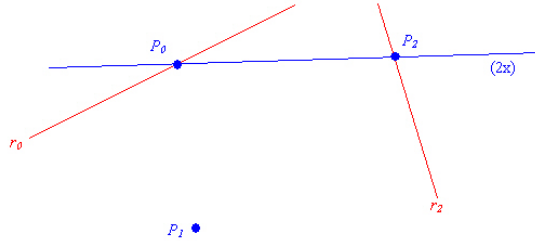


Figure 4:

have come to an absurd conclusion. In the other possible case, the fifth condition is the passage thru P_1 . Therefore, all conics going thru P_0 and P_2 , being there tangent to r_0 and to r_2 , respectively, should also pass thru P_1 . Now, let us consider (see Figure 4) the degenerate conic constituted by the double line thru P_0 and P_2 . Notice that each point of this degenerate conic is a singular double point, so that the intersection multiplicity between the conic and any line is 2. Therefore, this conic is tangent to r_0 and r_2 , but it does not pass thru P_1 and we have come to an absurd conclusion. This completes the proof. ■

It is evident that all the conics satisfying the conditions of the theorem are irreducible (i.e. non-degenerate). However, it is possible to increase the generality of the conditions to include degenerate conics and eventually to loose the uniqueness of solution. More precisely, if P_1 belongs to r_0 (see Figure 5), then this line will cut the conic three times: twice in P_0 and once in P_1 . Then r_0 will be a component of the conic, the latter being split into r_0 itself and r_2 . In this case we have got a unique conic degenerated into 2 distinct lines.

In view of the following NURBS parameterization, we will systematically avoid this particular case. More precisely, instead of trying to parameterize the arc from (with reference to Figure 5) P_1 to P_2 passing thru the singular point, we will consider two arcs, viz. the one from P_1 to the singular point and then the one from the singular point to P_2 . These two arcs are actually rectilinear and we will interpolate them as two conics, both degenerated into two double lines.

If P_0, P_1, P_2 were aligned along r_0 (see Figure 6), then any conic constituted by r_0 itself and any other line thru P_2 would be acceptable. In this case, we will arbitrary re-establish uniqueness by considering line r_0 doubly counted as the actual conic. This will also allow to treat the case of a line (plane algebraic curve of degree 1) as a particular case of a conic (plane algebraic curve of degree 2). This way, it will be possible to pass continuously from a line to a conic (or vice versa) by displacing point P_1 from (or to) the line thru the other two points.

Up to now we have written a set of 5 linear and linearly independent conditions, univocally defining a conic. Now we want to simplify as much as possible this set of conditions in order to get easily to the final result with the maximum analytical awareness on the various steps needed. More precisely, we will rewrite conditions (5-9) in a particular reference frame minimizing the analytical effort to obtain the conic. To this purpose, two reference frames will be considered, together with the coordinate transformations needed to switch from the actual global reference to the particular local one, and vice versa.

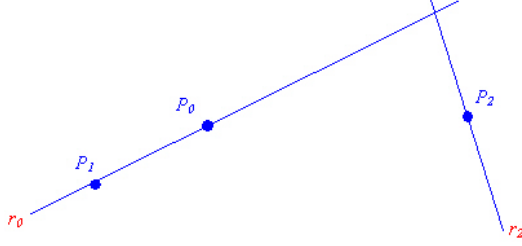


Figure 5:

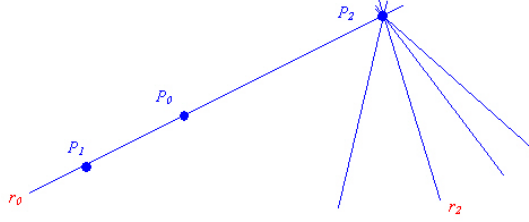


Figure 6:

Let us start by considering a local reference frame (see Figure 7, in black) such that the origin coincides with P_0 and the y -axis coincides with r_0 . It is always possible, by means of a roto-translation, to pass from the global reference (see Figure 7, in green) to the local reference. By construction we have:

$$\begin{aligned} P_0(0 : 0 : 1) &\Rightarrow x_{P_0} = 0 \quad y_{P_0} = 0 \quad u_{P_0} = 1 \\ r_{0\infty}(0 : 1 : 0) &\Rightarrow x_{Q_0} = 0 \quad y_{Q_0} = 1 \quad u_{Q_0} = 0 \end{aligned} \quad (11)$$

We can use condition (5) for $i = 0$ and find $a_{33} = 0$. Similarly, we can use condition (9) for $i = 0$ and find $a_{23} = 0$. Therefore, the interpolating conic equation reduces to

$$f(x : y : u) = a_{11}x^2 + a_{22}y^2 + a_{12}xy \quad (12)$$

It is immediate to notice that the three partial derivatives of this curve vanish when they are evaluated in point $P_0(0 : 0 : 1)$:

$$\begin{cases} \left(\frac{\partial f}{\partial x} \right)_{(0:0:1)} = (2a_{11}x + 2a_{12}y)_{(0:0:1)} = 0 \\ \left(\frac{\partial f}{\partial y} \right)_{(0:0:1)} = (2a_{22}y + 2a_{12}x)_{(0:0:1)} = 0 \\ \left(\frac{\partial f}{\partial u} \right)_{(0:0:1)} = 0 \end{cases} \quad (13)$$

Then P_0 is a (double) singular point for the conic. Therefore, the two lines (or the only line, doubly counted) composing the conic must go thru P_0 . This case is possible only if r_2 is going thru P_0 . The other line composing the conic depends on the position of P_1 : if P_1 is aligned with P_0 and P_2 , then r_2 is a double line; if not, the conic is degenerated into two distinct lines. In the first case, i.e. whether the conic degenerates into a double line, the equation is readily obtained by squaring the homogeneous equation of the line thru P_0 and P_2 . In the other case, i.e. whether the conic is reducible and reduced, we could obtain its equation by multiplying the homogeneous equations of the two lines involved. This is however unnecessary, since we have agreed not to treat directly this case but to decompose it into two cases of the previous kind.

Since all the possible exceptions have been handled, we can surely assume that the conic to be found is not degenerate and we can set $a_{13} = 1$ and then solve for a_{11} , a_{22} , and a_{12} the following linear system:

$$\begin{pmatrix} x_{P_1}^2 & y_{P_1}^2 & 2x_{P_1}y_{P_1} \\ x_{P_2}^2 & y_{P_2}^2 & 2x_{P_2}y_{P_2} \\ x_{P_2}x_{Q_2} & y_{P_2}y_{Q_2} & x_{P_2}y_{Q_2} + x_{Q_2}y_{P_2} \end{pmatrix} \begin{pmatrix} a_{11} \\ a_{22} \\ a_{12} \end{pmatrix} \begin{pmatrix} -2x_{P_1}u_{P_1} \\ -2x_{P_2}u_{P_2} \\ -x_{P_2}u_{Q_2} - x_{Q_2}u_{P_2} \end{pmatrix} \quad (14)$$

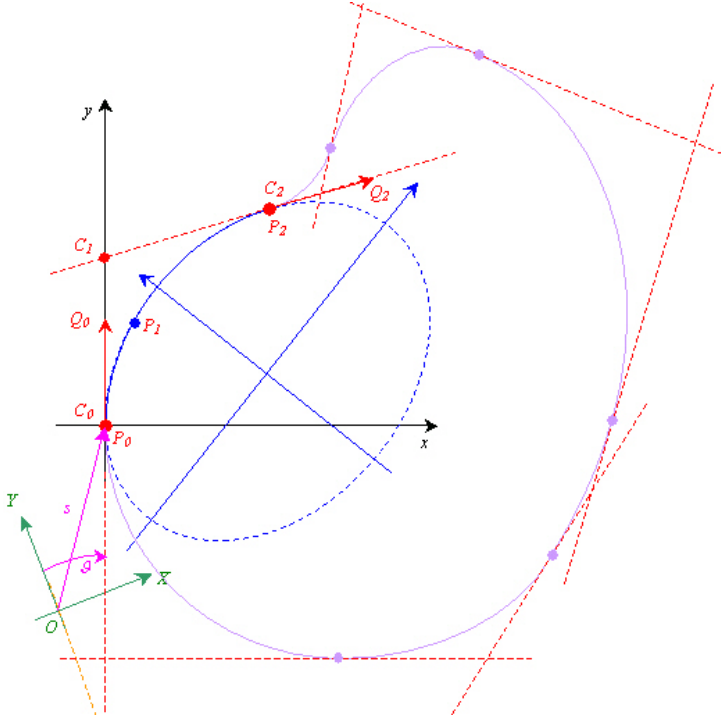


Figure 7: Real projective plane curve interpolation with conics

Figure 7 shows how the method has to be applied to interpolate a real affine plane curve. First of all, the global curve is partitioned into local sub-arcs (violet arcs). Passage points and tangent lines are then found (dotted red lines). With reference to a particular sub-arc (the blue ellipse, for example), a local fitting conic is found and then parameterized as a NURBS curve, as later detailed. Of course, in case the global curve is piece wise conic, then the NURBS curve is an exact parameterization of it and no approximation is introduced. In the picture, the roto-translation from the global to the local reference frame is also shown.

We will use this local coordinate system with reference to NURBS parameterization of $\mathbb{A}^2(\mathbb{R})$. The transformation needed to switch from the global reference frame to the local one is a plane roto-translation, (i.e. a special case of affinity) which is a very simple geometrical operation to think of. However, there is a particular reference frame offering a remarkably simple and deep insight as regards looking at and finding a solution to the problem of the interpolating conic. More precisely, let us consider a coordinate transformation such that r_0 and r_2 are mapped into the two coordinate axis, whilst r_1 is mapped into the improper line. Figure 8 shows an idealized sketch of this mapping showing in the same color a plane region and its relevant mapped image).

More precisely, let us assume for the 3 lines r_0 , r_1 and r_2 the following equations in homogeneous coordinates:

$$r_i \quad : \quad r_{i0}x + r_{i1}y + r_{i2}u = 0 \quad (i \in \{0, 1, 2\}) \quad (15)$$

Then, the above mentioned mapping can be represented by means of the following change of coordinates:

$$\begin{cases} x' = \lambda_0(r_{00}x + r_{01}y + r_{02}u) \\ y' = \lambda_2(r_{20}x + r_{21}y + r_{22}u) \\ u' = \lambda_1(r_{10}x + r_{11}y + r_{12}u) \end{cases} \quad (16)$$

where the presence of parameters $\lambda_i \in \mathbb{R}(i \in \{0, 1, 2\})$ is due to the fact that the 3 homogeneous coordinates x, y, u are defined up to a constant factor. We can set a value for λ_0 by reasoning this way. We expect the image of C_0 to be the improper point of the y' axis in the new coordinate system, i.e. it will have all coordinates equal to zero, except for y' . Provided that it is non-vanishing, the actual value of the y' coordinate is not meaningful, since it is only needed to assign the “vertical” direction.

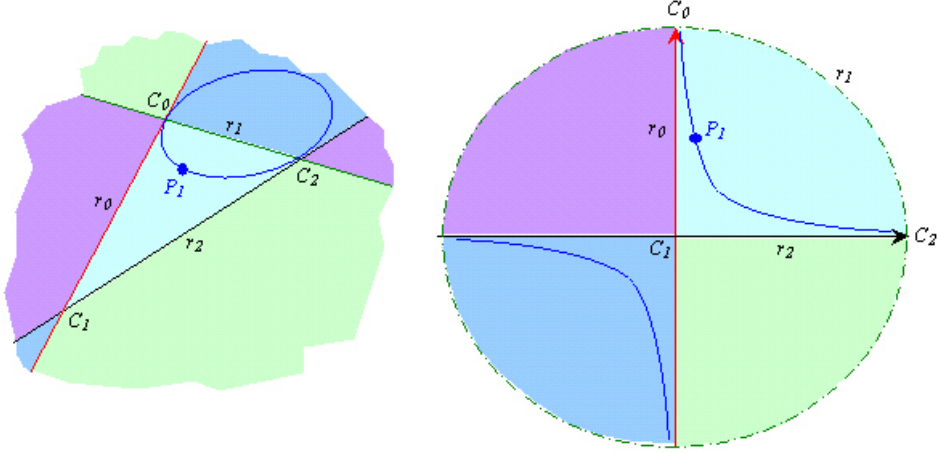


Figure 8: Graphical representation of the projectivity from the global to the local reference frame

Nevertheless, we want it to be exactly 1, for a reason that will be clear soon later. We can accomplish to this task by setting parameter λ_0 equal to the reciprocal of the value assumed by the equation of r_0 when calculated in correspondence of C_0 . As a matter of fact, the other two parameters λ_1 and λ_2 will not interfere, since the expressions they are multiplied for vanish when valued in correspondence of C_0 . Then C_0 will be mapped exactly into $(0 : 1 : 0)$. Similarly, λ_1 and λ_2 can be determined by reasoning the same way on C_1 (that has to be mapped exactly into $(0 : 0 : 1)$) and C_2 (that has to be mapped exactly into $(1 : 0 : 0)$), respectively. Summing up, we have

$$\lambda_i = \frac{1}{r_{i0}x_{C_{2-i}} + r_{i1}y_{C_{2-i}} + r_{i2}u_{C_{2-i}}} \quad (17)$$

For simplicity, we will refer to (16) in the following form

$$\begin{cases} x' = t_{00}x + t_{01}y + t_{02}u \\ y' = t_{20}x + t_{21}y + t_{22}u \\ u' = t_{10}x + t_{11}y + t_{12}u \end{cases} \quad (18)$$

where, obviously, $t_{ij} = \lambda_i r_{ij}$. Mapping (18) can also be expressed in compact matrix form as

$$P' = TP \quad (19)$$

where T is given by

$$T = \begin{pmatrix} t_{10} & t_{11} & t_{12} \\ t_{20} & t_{21} & t_{22} \\ t_{00} & t_{01} & t_{02} \end{pmatrix} \quad (20)$$

(please pay attention to the numbering of coefficients t_{ij} inside matrix T)

This transformation is a mapping from the real projective plane into itself. More precisely, it is a projectivity T and (like any projectivity) it can be represented by means of a transition matrix $T \in GL(3, \mathbb{R})$. By definition, T is non singular and then the projectivity can be inverted (i.e. it is a surjection on $\mathbb{P}^2(\mathbb{R})$). We will call T (respect. T^{-1}) both the direct (respect. inverse) projectivity and the direct (respect. inverse) transition matrix. Due to the linearity of the transformation, two distinct points must have distinct images (i.e. the transformation is an injection on $\mathbb{P}^2(\mathbb{R})$). Summing up, T is a bijection on $\mathbb{P}^2(\mathbb{R})$. Moreover, since T is linear, it is also a diffeomorphism from $\mathbb{P}^2(\mathbb{R})$ to itself.

In view of this, let us comment in some depth the choice of parameters λ_i . Just to fix ideas, let consider line r_0 and points C_2 and P_1 . Eqn. (15) defining r_0 splits $\mathbb{A}^2(\mathbb{R})$ (and not $\mathbb{P}^2(\mathbb{R})$) into two half planes. For all the points of each half plane, the equation of r_0 is either always positive or always negative. However, it is not a priori predictable which half plane is the “positive” one and which is the “negative”

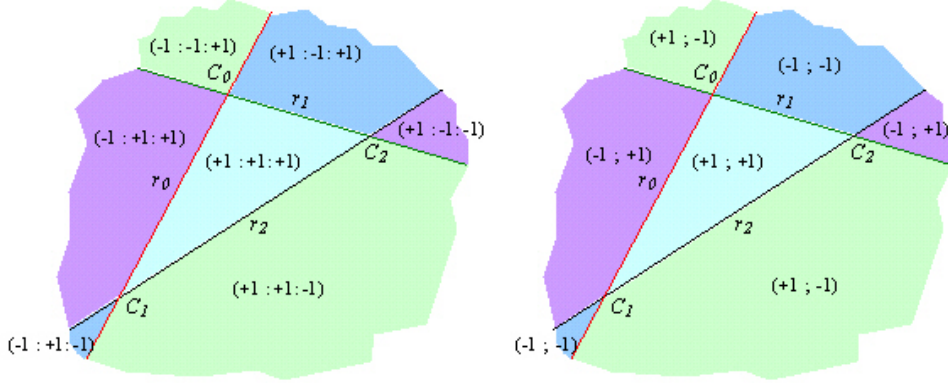


Figure 9: Values of ψ_i (left) and φ_i (right)

one, since it all depends on the choice of coefficients r_{00} , r_{01} and r_{02} , as they are a homogeneous triple. When we set the value of λ_i in order that the image of C_2 is “exactly” $(1 : 0 : 0)$, we are precisely choosing the half plane containing C_2 as the positive one. Therefore the image of P_1 will have a positive x' coordinate, since P_1 rests in the same positive half plane. Another way to look at this is to consider directly eqn. (16-17). When we evaluate the x' coordinate of the image of P_1 , we obtain the ratio of the same equation defining line r_0 , evaluated in P_1 and in C_2 . Since both the two points lie in the same half plane, we will always obtain a positive result, regardless which half plane is “positive” and which “negative”. Exactly the same applies to the other lines and homogeneous coordinates, as obvious.

We are now going to reformulate the previous reasoning in a more formal way in order to acquire a powerful tool to reference easily to lines and regions bounded by lines in $\mathbb{P}^2(\mathbb{R})$. To this end, let us define the following sets

$$\begin{aligned} W^2 &= \{-1, 0, +1\} \times \{-1, 0, +1\} \\ W^3 &= \{-1, 0, +1\} \times \{-1, 0, +1\} \times \{-1, 0, +1\} \end{aligned} \quad (21)$$

i.e. the double and triple cartesian product of set $\{+1, 0, -1\}$. This will be useful in defining an application

$$\Psi : \mathbb{P}^2(\mathbb{R}) \rightarrow W^3 \quad (22)$$

such that

$$\psi_i(P(x : y : u)) = \text{sgn} \left(\frac{r_{i0}x + r_{i1}y + r_{i2}u}{r_{i0}x_{C_{2-i}} + r_{i1}y_{C_{2-i}} + r_{i2}u_{C_{2-i}}} \right) \quad (i \in 0, 1, 2) \quad (23)$$

The meaning of this function is easy to get: once the homogeneous coefficients have been chosen for line r_i , we will have $\psi_i(P) = +1$ for all points P such that the equation of r_i yields a result having the same sign as the evaluation of r_i in C_{2-i} . Analogously, it will be $\psi_i(P) = -1$ for all points giving an opposite sign as C_{2-i} and $\psi_i(P) = 0$ for points belonging to line r_i .

To complete our formalism we need to introduce also an application

$$\Phi : \mathbb{P}^2(\mathbb{R}) \setminus r_1 \rightarrow W^2 \quad (24)$$

such that

$$\varphi_i(P(x : y : u)) = \frac{\psi_i(P(x : y : u))}{\psi_1(P(x : y : u))} \quad (i \in 0, 2) \quad (25)$$

What changes from (22) to (24) is that the second takes the result of ψ_1 as a reference to “normalize” the result of the other two applications. From a certain point of view, this is analogous to the operation to be made when shifting from a projective plane to an affine one. The practical effect of using (24) instead of (22) is that two elements like $(+1 : +1 : -1)$ and $(-1 : -1 : +1)$, which are distinct in W^3 , correspond to the same class of equivalence $(+1; -1)$ in W^2 . Of course, it is not possible to divide for ψ_1 when point \mathbb{P} belongs to r_1 , since the divisor would be vanishing. The analogue of this fact is the loss of a line (the so-called improper line) when shifting from the real projective plane to a real affine plane.

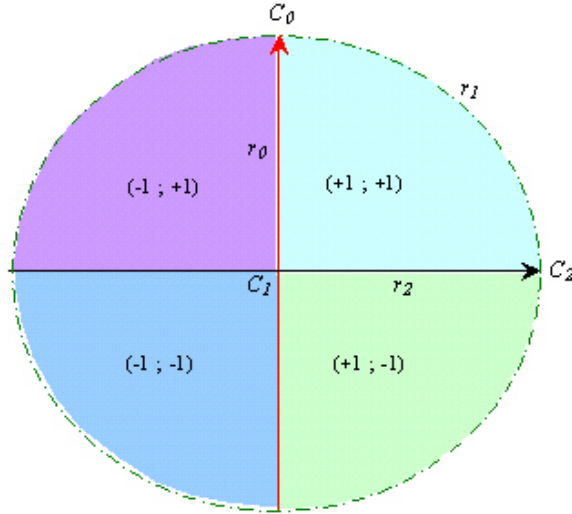


Figure 10:

We want to apply our previous conclusions to the graphical analysis of mapping T . The three lines divide $\mathbb{P}^2(\mathbb{R})$ into 4 triangles, i.e. regions bounded by 3 line segments. Since linear, T maps triangles into triangles. We agree to consider those triangles as open sets, i.e. deprived of their edges. The color scheme in the left and right part of Figure 8 is such to show the effects of this mapping, in the sense that any triangle and its image are colored the same. It is easy to understand what triangle is the image of what just observing the coloring scheme of its edges (actually, the projectivity has been built starting from this point). As an example, the light cyan triangle containing P_1 is bounded by a red segment, followed by a black and a green one in counterclockwise order. Then the mapped triangle will be the upper-right one (i.e. the North East one), since T preserves the color scheme.

Notice that each triangle can be associated to one, and only one, element of the discrete set W^2 having both entries different from zero and that this element is the same even after the transformation T has been applied. Formally, we can simply state the same with the following

Theorem 2.2 *The three lines r_0 , r_1 and r_2 divide $\mathbb{P}^2(\mathbb{R})$ into 4 triangles. All points P , and no others, of a given triangle are characterized by a unique value $\Phi(P)$ and this value is invariant to T , i.e. $T(\Phi(P)) = \Phi(P)$.*

Proof Since the three (not coincident) lines r_0 , r_1 and r_2 are pair wise intersecting (in C_0 , C_1 , C_2), they divide $\mathbb{A}^2(\mathbb{R})$ into exactly 7 regions. Each region lies in one half plane with respect to each line and then all points belonging to a given region have the same values of ψ_0 , ψ_1 and ψ_2 (see Figure 9-left). In $\mathbb{P}^2(\mathbb{R})$ the 7 regions reduce to 4 triangles, since improper points have to be considered. Particularly, any line going thru C_i and intersecting r_i have one, and only one improper point, belonging to one, and only one, triangle. The coloring scheme shows how the triangles are arranged in $\mathbb{P}^2(\mathbb{R})$. It is straightforward to see that the values of φ_0 and φ_2 are the same inside a given triangle, as shown in Figure 9-right.

As regards projective invariance, it is easy to notice that it holds due to the particular choice of scalar multipliers λ_i (see eqn. (17)). Particularly, the co-domain of T is such that all points with positive (respect. negative) x' coordinate are those with positive (respect. negative) sign for φ_2 , and the same holds with respect to y' and φ_0 (see Figure 10). ■

Let us come to the reason for having chosen such a local reference frame, namely the ease in finding the equation of the conic. To this end, in the new local coordinate system let us consider the pencil of conics passing thru P_0 and P_2 and being there tangent to the y' and x' axis, respectively. We can write the equation of such a pencil by linearly combining any two conics satisfying such conditions. For convenience, we will use the degenerate conic consisting of the two coordinate axis, i.e. $x'y' = 0$, and the

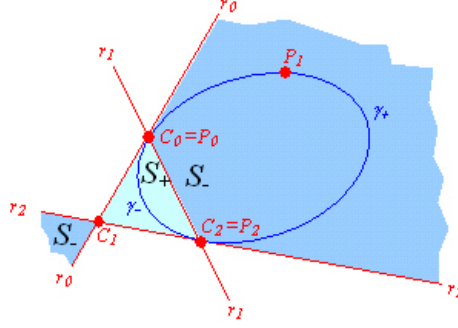


Figure 11: Definition of sector S as the union of its two subsets S_+ and S_-

degenerate conic constituted by the doubly counted line thru P_0 and P_2 (the improper line), i.e. $u'^2 = 0$. The equation of the pencil of conics is then given by

$$\lambda x' y' + \mu u'^2 = 0 \quad (26)$$

where λ and μ are two homogeneous parameters. The unique conic in the pencil passing for $P_1(x'_{P_1} : y'_{P_1} : u'_{P_1})$ is such that

$$\lambda = -\mu \frac{u'^2_{P_1}}{x'_{P_1} y'_{P_1}} \quad (27)$$

and, consequently, the equation of such a conic is given by

$$\frac{x' y'}{x'_{P_1} y'_{P_1}} = \frac{u'^2}{u'^2_{P_1}} \quad (28)$$

For point P_1 we have $x'_{P_1} \neq 0$ (because $P_1 \notin r_0$), $y'_{P_1} \neq 0$ (because $P_1 \notin r_2$), $u'_{P_1} \neq 0$ (because $P_1 \notin r_1$). Notice also that the conic assumes the form of an hyperbole in the local reference frame, even though it could be other conics, when projected back in the global reference frame. The matrix associated to conic (28) is given by.

$$A' = \begin{pmatrix} 0 & \lambda/2 & 0 \\ \lambda/2 & 0 & 0 \\ 0 & 0 & \mu \end{pmatrix} \quad (29)$$

Starting from the conic equation in the local reference frame, then recalling the linear bound between point coordinates in global versus local reference frame (eqn. (19)) and finally rearranging terms, one easily obtains the conic equation in the global reference frame:

$$0 = P'^T A' P' = (TP)^T A' (TP) = P^T (T^T A' T) P = P^T A P \quad (30)$$

Notice that, generally, T^T is distinct from T^{-1} , since we are facing a more general projectivity than a simple roto-translation. The conic matrix in the global reference frame is then bound to the one in the local reference frame by the following simple relation:

$$A = T^T A' T \quad (31)$$

One more good reason for choosing this local reference frame is the possibility of understanding easily what happens to the conic in case the passage points and the tangents are arranged in some way, including pathological ones. To the purpose, let us consider the sector S of $\mathbb{P}^2(\mathbb{R})$ containing all the points with either strictly positive or strictly negative values of φ_j :

$$S = \{P \in \mathbb{P}^2(\mathbb{R}) \mid \varphi_0(P)\varphi_2(P) = 1\} \quad (32)$$

We will divide sector S into 2 subsets by line r_1 and call S_+ (respect. S_-) the subset of S containing points with strictly positive (respect. negative) values of φ_j :

$$S_+ = \{P \in S \mid \varphi_0(P) = \varphi_2(P) = +1\} \quad (33)$$

$$S_- = \{P \in S \mid \varphi_0(P) = \varphi_2(P) = -1\} \quad (34)$$

According to their definition, S , S_+ and S_- do not contain points belonging to the 3 lines r_0 , r_1 and r_2 , including points P_0 and P_2 . Figure 11 shows the aforementioned regions with reference to a generic global coordinate system. In case the local one is chosen (in the usual way), S_+ is simply the first quadrant, whilst S_- is the third one.

We are now ready to state the following theorem, implicitly resting on the projective invariance of function Φ .

Theorem 2.3 *The interpolating conic γ (deprived of the two points P_0 and P_2) is contained inside sector S if, and only if, P_1 belongs to sector S . Moreover, γ_+ without its extremes is entirely contained in either S_+ or S_- , depending on whether P_1 belongs to either S_+ or S_- , respectively*

$$\begin{aligned} P_1 \in S & \Leftrightarrow \gamma \setminus (P_0 \cup P_2) \subseteq S \\ P_1 \in S_+ & \Leftrightarrow \gamma_+ \setminus (P_0 \cup P_2) \subseteq S_+ \\ P_1 \in S_- & \Leftrightarrow \gamma_+ \setminus (P_0 \cup P_2) \subseteq S_- \end{aligned} \quad (35)$$

Proof Without any loss of generality, it is possible to assume that the current reference frame were the local one, eventually obtained thru projectivity (19). As regards the necessity of all condition, it is obvious that, if P_1 does not belong to S , S_+ , S_- , then the arc of the conic containing P_1 cannot be contained into S , S_+ , S_- , respectively. On the other side, as regards the sufficiency of the condition, let us consider the equation of γ given by (30), hereby rewritten.

$$\frac{x'y'}{x'_{P_1}y'_{P_1}} = \frac{u'^2}{u'^2_{P_1}} \quad (36)$$

We recall that, since P_1 belongs to S , $x'_{P_1} \neq 0$ (because $P_1 \notin r_2$), $y'_{P_1} \neq 0$ (because $P_1 \notin r_0$), $u'_{P_1} \neq 0$ (because $P_1 \notin r_1$). Besides, if $P_1 \in S$, it follows $x_{P_1} > 0$ and $y_{P_1} > 0$, so that, being the right member necessarily non-negative, it must be $x'y' \geq 0$ for each point of γ . Therefore, we can assume that, (except for P_0 and P_2) any given point of the interpolating conic will have both x' and y' coordinates either strictly positive or strictly negative. In the first case such point lies in S_+ while in the second case they lie in S_- . Whatever the case, they lie in S . More precisely, γ is a hyperbole referred to its own asymptotes and therefore one of its two branches is in the first quadrant (i.e. in S_+), being the other in the third quadrant (i.e. in S_-). Of the two branches, γ_+ is the one going thru P_1 . Hence γ_+ is entirely in S_+ or S_- depending on the position of P_1 . This completes the proof of the theorem. ■

The utility of the previous theorem will be clear when discussing about the NURBS parameterization of conics. As a matter of fact, the theorem provides a sufficient condition for choosing a passage point P_1 in such a way that the relevant conic arc lies in a prescribed region, viz. S_+ or S_- . This in turn is a sufficient condition for being able to parameterize the conic as a NURBS curve. The point is that the arc to be represented as a NURBS must be continuous and cannot go thru an “infinite” (meaning improper) point and then come back to “finite” (meaning proper) points. For example, it would be impossible to parameterize two arcs of hyperbole joined at “infinity”. Thanks to the theorem, in case the conic arc does not have improper points, it is easy to see that γ_+ is not a troublesome curve, as regards this problem. This is because either γ_+ lies in a bounded region (like S_+ as depicted in Figure 11) or it lies in an unbounded region (like S_- as depicted in Figure 11) without having the possibility to reach infinite points (because the curve is by hypothesis without improper points). In case the conic has improper points, a more general method will be presented in sect. 5. Also in this case the previous theorem will be helpful.

Once the three lines r_i have been chosen, we can also wonder what happens if the passage point P_1 is chosen in some pathological manner. If P_1 belongs to the line thru P_0 and P_2 , then the conic we are looking for is such line doubly counted. Hence its equation will be simply $u'^2 = 0$ in the local coordinate

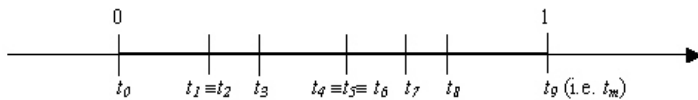


Figure 12: A possible partition of interval $[0, 1]$

system. On the other hand, if P_1 belongs at least to one of the two tangents, the conic degenerates in the two tangents themselves and in the local coordinate system its equation is simply $x'y' = 0$. After the conic is projected back in the global reference frame thru (30), we could have easily the equation of the conic even in the global coordinate system. However, we remind that the case of a conic degenerated into two distinct lines will not be faced when turning to NURBS parameterization, as already detailed.

3 NURBS definition

3.1 B-spline Basis Functions

In the real affine plane $\mathbb{A}^2(\mathbb{R})$ let us introduce a list of $n + 1$ points C_0, C_1, \dots, C_n , termed *control points*, whose coordinates will be given by (x_{C_i}, y_{C_i}) for $i = 0, 1, \dots, n$. Then, let us introduce a scalar real parameter t , ranging (for convenience only) from 0 to 1. For each control point C_i , let us define a family of piece wise polynomials defined on t :

$$N_{i,p} \quad : \quad [0, 1] \rightarrow \mathbb{R} \quad (37)$$

This functions, named *B-splines basis functions* (or, shortly, basis functions), constitute a set depending on two indexes i and p , the former being associated to the enumeration of control points (and then ranging from 0 to n), the latter being the degree of the polynomials composing the basis function itself. For example, for control point C_0 we have $N_{0,0}$, which is made up of degree 0 polynomials (i.e. it is piece wise constant), $N_{0,1}$, which is made up of degree 1 polynomials (i.e. it is piece wise linear), $N_{0,2}$, which is made up of degree 2 polynomials (i.e. it is piece wise quadratic), and so on for the other basis functions of the family associated to C_0 . The same applies to all of the other control points.

To define basis functions, we need to introduce a partition (see Figure 12) of the closed interval $[0, 1]$ by means of a sequence of values t_k such that $0 = t_0 \leq t_1 \leq t_2 \leq \dots \leq t_m = 1$. The $m + 1$ parameter values t_k will be termed *knots*, and the sub-intervals amidst them as knot spans. It is allowed that two or more knots coincide (and then the relevant knot span collapses into one point). If knots are equally spaced, the knot sequence is said to be uniform, otherwise it is said to be non-uniform.

We are now ready to recursively define the B-spline basis functions as follows:

$$N_{i,0}(t) = \begin{cases} 1 & \text{if } t_i \leq t \leq t_{i+1} \\ 0 & \text{otherwise} \end{cases} \quad (38)$$

$$N_{i,j}(t) = \frac{t-t_i}{t_{i+j}-t_i} N_{i,j-1}(t) + \frac{t_{i+j+1}-t}{t_{i+j+1}-t_{i+1}} N_{i+1,j-1}(t) \quad (1 \leq j \leq p)$$

In case a knot is repeated, a division by zero occurs when evaluating basis functions by means of the previous formula. In such a case, it is intended that the relevant term does not appear in the basis function. As an example, if $t_0 = t_1$, the computation of $N_{0,1}$ will lead to the division by $t_{0+1} - t_0 = t_1 - t_0 = 0$. According to our convention, we will discard the whole term and so $N_{0,1}$ will not be composed by $(t - t_0)/(t_1 - t_0)N_{0,0}$.

In what follows, the basis functions will be computed as regards a particular case that will be used later to find out the NURBS parameterization of a conic arc. To this purpose, we will use a $m + 1 = 6$ knots spaced as shown in Figure 13-top, i.e. the first 3 coinciding with $t = 0$ and the last 3 coinciding with $t = 1$. The computation of the basis functions up to the second order is straightforward by means of definition (38) and the results are noted down in Figure 13-bottom. The sum of the basis functions with the same degree is identically equal to 1, as evident by summing column wise in Figure 13. This property is called *partition of unity*.

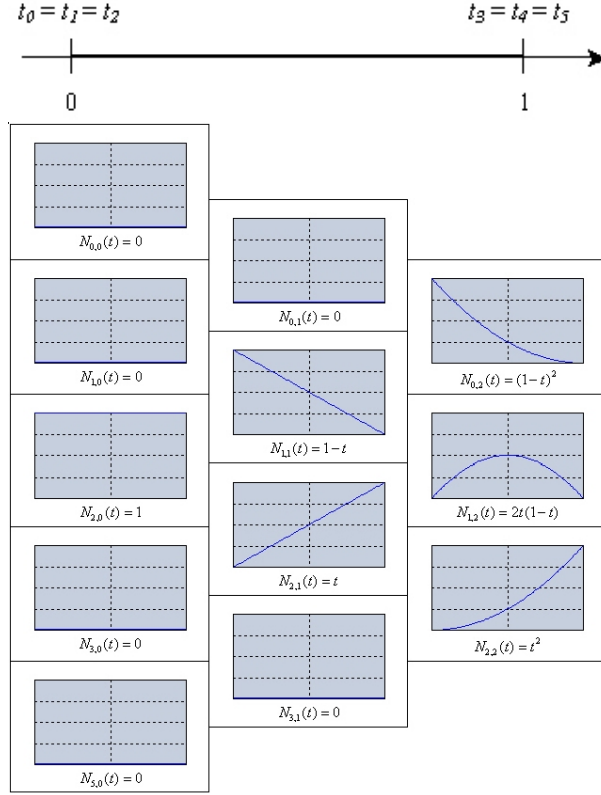


Figure 13: Bezier basis functions for a second degree NURBS

3.2 B-spline curves

Let us assume that the curve to be interpolated is entirely constituted by proper points, i.e. not belonging to the improper line of $\mathbb{P}^2(\mathbb{R})$. This is virtually equivalent to restrict our analysis to $\mathbb{A}^2(\mathbb{R})$. Recalling the set of $n + 1$ control points $C_0, C_1, \dots, C_n \in \mathbb{A}^2(\mathbb{R})$, we will call *control polyline* the polyline composed by line segments $C_0C_1, C_1C_2, \dots, C_{n-1}C_n$. The previously defined (see sect. 2) three control points C_0, C_1 and C_2 are precisely those to be used with reference to conic arc parameterization. We will also associate each control point C_i to a B-spline basis function $N_{i,p}(t)$, where p is the degree of polynomials to be used for interpolation. It can be easily shown that the number of control points, the number of knots and the polynomial degree must be such that

$$m = n + p + 1 \quad (39)$$

in order to have well posed definitions. We are now ready to define a curve in a parametric form as follows

$$P(t) = \sum_{i=0}^n N_{i,p}(t)C_i \quad (40)$$

The curve parameterized by eqn. (40) is said to be a *B-Spline curve*. Expanding the compact vector notation one gets

$$\begin{cases} x(t) = \sum_{i=0}^n N_{i,p}(t)x_{C_i} \\ y(t) = \sum_{i=0}^n N_{i,p}(t)y_{C_i} \end{cases} \quad (41)$$

B-spline curves have many important properties, which can be found, together with their proof, on the many texts dedicated to this subject. We will recall only what is necessary to understand what follows. If we set the first $p + 1$ knots equal to 0 and the last $p + 1$ knots equal to 1, what we get is a so called clamped B-spline, viz. it goes thru the first and the last control point and is there tangent to the control

polyline. For example, let us set $p = 2$, $t_0 = t_1 = t_2 = 0$ and $t_3 = t_4 = t_5 = 1$. Then the resulting B-spline will go thru C_0 being there tangent to the line thru C_0 and C_1 and it will also go thru C_2 being there tangent to the line thru C_2 and C_1 .

As regards the passage thru C_0 , we can argument as follows. Due to the particular choice of the first $p + 1$ knots, it follows that, keeping into account the definition of the basis functions, the first $p - 1$ with 0 degree are vanishing (see $N_{0,0}$ and $N_{1,0}$ in Figure 13). Then the first $p - 2$ basis functions with degree 1 will be vanishing in their turn (see $N_{0,1}$ in Figure 13) and so on up to $N_{0,p}$, which will be equal to 1 when valued in correspondence of $t = 0$ (see $N_{0,2}$ in Figure 13). Therefore, due to the partition of unit property of B-spline basis functions, all the others should be equal to zero when valued in correspondence of $t = 0$ (see $N_{1,2}$ and $N_{2,2}$ in Figure 13). As a result, for $t = 0$ summation (40) will yield exactly C_0 . The same holds symmetrically for the last control point. A similar argumentation applied to the first derivative of the B-spline will justify the assertion on tangency to the control polyline at the extreme control points.

3.3 NURBS curves

With reference to B-spline curves, we can think of the control points as elements of $\mathbb{P}^2(\mathbb{R})$ and express them in homogeneous coordinates. Since we are assuming that all control points and curve points are proper, their u coordinate will not be vanishing. We can set it equal to 1, or we can choose to multiply all coordinates of control point C_i for a constant factor w_i , named the i -th *weight*:

$$C_i = \begin{pmatrix} x_{C_i} \\ y_{C_i} \\ 1 \end{pmatrix} = \begin{pmatrix} w_i x_{C_i} \\ w_i y_{C_i} \\ w_i \end{pmatrix} \quad (42)$$

In view of this, eqn. (41) can be rewritten as

$$\begin{cases} x(t) = \sum_{i=0}^n w_i N_{i,p}(t) x_{C_i} \\ y(t) = \sum_{i=0}^n w_i N_{i,p}(t) y_{C_i} \\ u(t) = \sum_{i=0}^n w_i N_{i,p}(t) \end{cases} \quad (43)$$

In the particular case that all weights are equal (not necessary to 1) eqn. (43) gives the same curve as eqn. (41). However, If different weights are chosen for different control points, then we have a new curve. We can go back to affine coordinates by dividing the first two projective coordinates by the third one, obtaining

$$\begin{cases} X(t) = \frac{x(t)}{u(t)} = \frac{\sum_{i=0}^n w_i N_{i,p}(t) x_{C_i}}{\sum_{i=0}^n w_i N_{i,p}(t)} \\ Y(t) = \frac{y(t)}{u(t)} = \frac{\sum_{i=0}^n w_i N_{i,p}(t) y_{C_i}}{\sum_{i=0}^n w_i N_{i,p}(t)} \end{cases} \quad (44)$$

or, more compactly and referring to C_i as the vector storing the two affine coordinates of the i -th control point,

$$P(t) = \frac{\sum_{i=0}^n w_i N_{i,p}(t) C_i}{\sum_{i=0}^n w_i N_{i,p}(t)} \quad (45)$$

The curve parameterized by eqn. (45) is said to be a *NURBS curve*, acronym of *Non Uniform Rational B-Spline*. As previously mentioned, B-spline curves are a subset of NURBS curves, namely the subset of non-rational NURBS.

NURBS curves inherit most of B-spline properties. Particularly, if the first and last weight, viz. w_0 and w_n , are equal to 1 and the first and last $p + 1$ knots are coincident, then the resulting NURBS is clamped. The reason is exactly the same for which the analogous property holds for B-spline curves. We can see this if we evaluate in zero the three summations in (43), obtaining $x(0) = x_{C_0}$, $y(0) = y_{C_0}$ and $u(0) = 1$. The same holds at the other extreme and for the tangents.

4 NURBS parameterization of a single conic arc

The problem is now, given a conic arc γ_+ in the real projective plane $\mathbb{P}^2(\mathbb{R})$ (found according to the theory previously exposed or in any other way), to parameterize it with a NURBS curve. γ_+ goes thru 3 points P_0, P_1, P_2 and is tangent in P_0 to line r_0 and in P_2 to line r_2 . The complete conic γ containing γ_+ is given by $f(P) = P^T A P = 0$. We also assume that no point belongs to the improper line $u = 0$ and, for convenience purposes, that the reference frame is centered in P_0 and oriented so that the y -axis coincident with r_0 . Now we want to find, if any, the NURBS curve that reproduces exactly this conic. We assume that the three passage points have been chosen so that point P_1 belongs to sector S . As a consequence, we are sure that γ_+ is a conic arc entirely contained in S_+ or in S_- and not intersecting the improper line of the plane. This suffices to guarantee the possibility to parameterize γ_+ by means of second order NURBS curve, i.e. thru quotients of second degree polynomials.

The procedure to find this NURBS is as follows. First, let us choose the control points C_0, C_1, C_2 so that $C_0 = P_0, C_2 = P_2$ and C_1 coincides with the intersection point between lines r_0 and r_2 . Thus, the control polyline is made up of two segments cut from the two tangents r_0 and r_2 to the conic for C_1 . We will also assume that r_0 and r_2 are not parallel, so that C_1 does not belong to the improper line $u = 0$. In case the conic arc is such that the two tangents are parallel, it can always be split into two sub-arcs with non-parallel tangents. Since we have 3 control points and a second order NURBS, we need 6 knots. Let us choose $t = 0$ and $t = 1$, both counted three times, so that we will have $t_0 = t_1 = t_2 = 0$ and $t_3 = t_4 = t_5 = 1$. As a consequence, the three basis functions will be $N_{0,2} = (1-t)^2$, $N_{1,2} = 2t(1-t)$ and $N_{2,2} = t^2$, as already noted. As regards weights, we will choose $w_0 = w_2 = 1$, so that the NURBS curve will be tangent to its control polyline, and then to r_0 and r_2 , in P_0 and P_2 . The last requirement to accomplish is the passage thru point P_1 . We will assign w_1 a value such that the outcoming NURBS goes thru P_1 .

Recalling eqn. (43), the NURBS curve is parameterized by

$$P(t) = \sum_{i=0}^2 w_i N_{i,2}(t) C_i = w_0 N_{0,2}(t) C_0 + w_1 N_{1,2}(t) C_1 + w_2 N_{2,2}(t) C_2 \quad (46)$$

where $C_i = (x_{C_i} : y_{C_i} : 1)$ are the homogeneous coordinates of the control points. Since each point on the NURBS curve must also be a point of the conic, it must be

$$f(x(t) : y(t) : u(t)) \equiv 0 \quad \forall t \in [0, 1] \quad (47)$$

Taken into account eqn. (46) and the symmetry of matrix A , condition (47) can be formulated as

$$P(t)^T A P(t) = w_0^2 N_{0,2}^2(t) C_0^T A C_0 + w_1^2 N_{1,2}^2(t) C_1^T A C_1 + w_2^2 N_{2,2}^2(t) C_2^T A C_2 + 2w_0 w_1 N_{0,2} N_{1,2} C_0^T A C_1 + 2w_1 w_2 N_{1,2} N_{2,2} C_1^T A C_2 + 2w_2 w_0 N_{2,2} N_{0,2} C_2^T A C_0 = 0 \quad (48)$$

Since $C_0 = P_0$ and $C_2 = P_2$ belong to the conic by construction, $C_0^T A C_0 = C_2^T A C_2 = 0$. Moreover, since C_1 belongs both to the tangent r_0 to the conic in $C_0 = P_0$, and to the tangent r_2 to the conic in $C_2 = P_2$, then $C_0^T A C_1 = C_1^T A C_2 = 0$, so that we have simply to satisfy

$$w_1^2 N_{1,2}^2(t) C_1^T A C_1 + 2w_0 w_1 N_{0,2}(t) N_{2,2}(t) C_0^T A C_2 = 0 \quad \forall t \in [0, 1] \quad (49)$$

Now, substituting the basis functions in eqn. (49) together with weights $w_0 = w_2 = 1$ (which ensures that r_0 and r_2 are tangent to the conic, as we assumed when we simplified eqn. (48) to (49)) we come to

$$2t^2(1-t)^2(w_1^2 2C_1^T A C_1 + C_0^T A C_2) = 0 \quad \forall t \in [0, 1] \quad (50)$$

It is then possible to make the NURBS curve be the desired conic for any given value of parameter t if, and only if,

$$w_1 = \varepsilon \sqrt{-\frac{1}{2} \frac{C_0^T A C_2}{C_1^T A C_1}} \quad (51)$$

where ε is equal to +1 or -1 according the following reasoning. Eqn. (50) is satisfied by both the positive and by the negative determinations of the squared expression. This is perfectly explainable if we consider

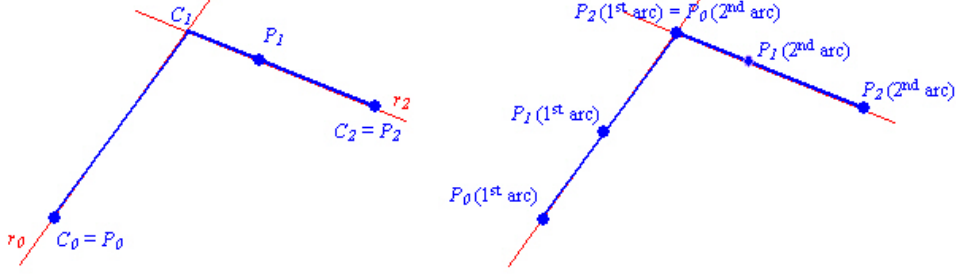


Figure 14: How to manage the case of a reduced and reducible conic by splitting the arc into two sub-arcs

that conic γ is made up of two sub arcs, namely γ_+ and γ_- . Thanks to NURBS theory, it can be proved that the choice of a positive weight corresponds to the choice of the sub arc contained in the convex hull of the three control points C_0 , C_1 and C_2 , i.e. inside the small pale cyan triangle in Figure 11. On the other hand, the choice of the opposite negative weight is equivalent to the choice of the complementary arc. Summing up, if P_1 (and then γ_+) belongs to S_+ (respect. S_-), then $\varepsilon = +1$ (respect. $\varepsilon = -1$). This is automatically performed if we use one of the following equivalent definitions for ε :

$$\varepsilon = \varphi_0(P_1) = \varphi_2(P_1) \quad (52)$$

The presence of a ratio in (51) can make some question arise about the possibility of finding an acceptable value for the weight of P_1 . However, if $C_1^T A C_1 = 0$ then C_1 belongs to the conic, the latter being then necessarily degenerated into the two lines r_0 and r_2 . We will discard this case since it can be managed with two line segments. In other words, in case P_1 is taken on one of the two tangents (or at the intersection of the two tangents), the conic arc will be subdivided into two rectilinear arcs, as shown in Figure 14.

In order to compute with greater ease the value of w_1 , we use the local reference frame such that $C_0 = P_0 = (0 : 0 : 1)$, $C_1 = (0 : y_{C_1} : 1)$ and $C_2 = (x_{C_2} : y_{C_2} : 1)$, where $x_{C_2} = x_{P_2}$, $y_{C_2} = y_{P_2}$. This is not against generality, since it is always possible, thru a plane roto-translation, to choose this particular reference frame. The y coordinate of C_1 can be found by intersecting the line thru P_2 and Q_2 (i.e. r_2) with the y axis (i.e. r_0), easily obtaining

$$y_{C_1} = \frac{x_{P_2} y_{Q_2} - x_{Q_2} y_{P_2}}{x_{P_2} - x_{Q_2}} \quad (53)$$

It is easy to see that $x_{P_2} - x_{Q_2}$ vanishes if, and only if, r_2 is vertical, but this case has been excluded since r_0 and r_2 cannot be parallel. A simple computation allows rewriting eqn. (51), as

$$w_1 = \varepsilon \sqrt{\frac{-a_{13} x_{C_2}}{2a_{22} y_{C_1}^2}} \quad (54)$$

It is important to notice that, whilst (51) is far general, (54) is valid only when the reference frame is as indicated. However, since weights are invariant to affine transformations, (54) need not to be modified when roto-translating back to the original reference frame.

A separate management is devoted to conics degenerated into a double line. As already deeply detailed, this case occurs when the three passage points P_0 , P_1 and P_2 lie on the same line. In such a case we can neglect the tangents, since no other condition is needed to find out the NURBS parameterization but the coordinates of those three points. This implicitly means that any choice of the tangent is acceptable. In order to get to a NURBS parameterization of a line segment, many reasonable choices are possible. Ours will be to keep on using a second degree NURBS curve, both for a uniformity reason, and in order to allow for easy modifications to the curve. As a matter of fact, we will choose as control points the usual couple of extreme passage points P_0 and P_2 and, as a third one, we will use P_1 . In case the NURBS has to be modified for any reason, an available control point C_1 exists and by slightly moving it out of the line thru the other two will result in non-degenerate conics. This can be helpful in case a strictly convex

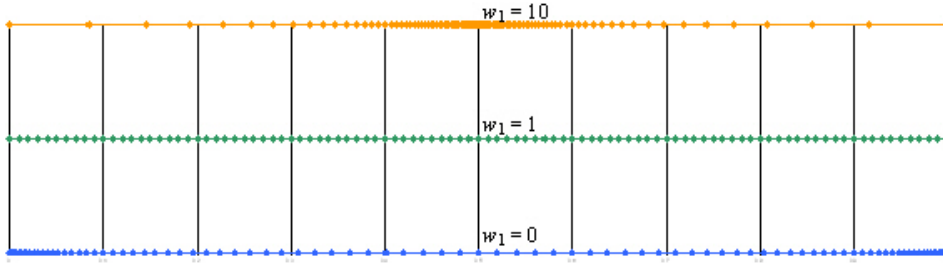


Figure 15: Effect of different choices for weight w_1

or concave curve is needed, for example with respect to the solution of a minimal distance problem, whenever strict convexity could be needed to avoid multiplicity of solution.

Generally speaking, the best choice is to take P_1 midway from the other two, so that when the outgoing NURBS curve will be sampled, the same number of sampled points will be available both before and after P_1 in a more or less symmetric pattern. In some occasion, however, it could be useful to place P_1 closer to one of the extremes, in order to have more sampled points in a given portion of the arc.

The need for sampling NURBS is bound to the usage of the NURBS itself, ranging from graphical plotting to more delicate numerical applications. As regards the former, no problem exists, since we are talking about a straight line. On the other side, it could be useful to refine the set of sampled points in some place if more accuracy is needed when dealing with such a place. As an example, we can think to the solution of a minimal distance problem between two bodies whose boundaries are parameterized as NURBS curves. If a narrow zone can be easily envisaged to be containing the solution point, a more precise solution is got whether the sampled points are denser in that zone. Than P_1 can be in some way helpful as regards this problem.

Coming back to the problem of parameterizing a segment as a NURBS, we will keep the same assumptions and conventions as for other conics. For convenience, we can choose to perform a roto-translation so that the local reference frame is centered in C_0 and with the y -axis aligned with the line thru the three control points. Then the equation of the conic is simply $x_2 = 0$, i.e. $a_{11} = 1$, being the only non-vanishing coefficient. Hence formula (54) for computing the weight of the intermediate control point becomes undetermined.

This is perfectly explainable as follows. It is well known (and quite easy to prove) that any NURBS curve with positive weights lies in the convex hull of its control points, i.e. inside the smallest convex set containing the control points. In our case, the three control points are collinear, so that their convex hull is simply the line segment having the two most extreme control points. Notice that this is independent on the choice of weights. This explains why (54) is undetermined: any positive value of w_1 will give a line segment. On the other hand, a negative value of w_1 will not be allowable, since it will give the portion of line external to the desired segment (which, by the way, is a segment with exactly the same dignity if the line is considered as projective one).

However, the choice of the weight is not completely neutral. As a matter of fact, even though the geometrical support of the curve (i.e. its shape) is the same, its analytical parameterization (i.e. the velocity with which the curve is spanned by parameter t) changes. The effects of this are evident when sampling the NURBS for equally spaced values of the parameter. Specifically, it can be easily seen that, for example, $w_1 = 0$ will slow down the curve close to the extremes, making them denser with sampled points; w_1 equal to a “big” positive number will have the opposite effect; $w_1 = 1$ will produce an equilibrated, uniform parameterization (see Figure 15). Once again, the best solution is that meeting the user needs.

4.1 NURBS parameterization of a circumference arc

We conclude this section by specializing the theory to a particularly important case, i.e. that of a circumference arc. Figure 16 shows a circumference together with the local reference frame. Let R be the radius and 2α the angular width of arc γ_+ (in blue). Point C_0 coincides with the origin of the local

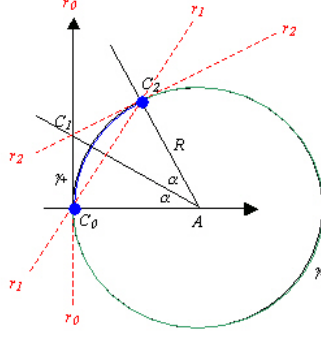


Figure 16: NURBS parameterization of an arc of circumference

reference frame. Hence the x coordinate of C_2 is readily computed to be $R(1 - \cos 2\alpha)$. The y coordinate of C_1 is equal to $R \tan \alpha$. The conic γ is a circumference centered in $A(R; 0)$ and with radius R . Then its equation is easily found to be $x^2 + y^2 - 2Rx = 0$. It follows that $a_{13} = R$ and $a_{22} = 1$.

Substituting into (54) one gets

$$w_1 = \varepsilon \sqrt{\frac{1 - \cos 2\alpha}{2 \tan^2 \alpha}} = \varepsilon \cos \alpha \quad (55)$$

Notice that $\varepsilon = +1$ gives the blue arc, whereas $\varepsilon = -1$ gives the green one. For $\alpha = \pi/4$ (a quarter circumference) we get $w_1 = 1/\sqrt{2}$. It is worth stressing that for $\alpha = \pi/2$ (a half circumference) the two tangents r_0 and r_2 are parallel and the y coordinate of C_1 is unbounded. In this case the arc have to be furtherly partitioned into (for example) two halves. In the next section a more general method will be presented able to manage control points located on the improper line and then overcoming this annoying drawback.

5 Extension to curves with improper points

We are now willing to extend the method so that a wider class of curves can be treated, essentially dropping the restrictive assumptions we exposed in sect. 4. Our aim is to include improper control points (i.e. belonging to the improper line $u = 0$) and then to allow for conic arcs of any kind, including hyperbole with the two improper points, or arcs with parallel tangents r_0 and r_2 . The outcome will be a method containing the previous one as a particular case.

Let us come back to (42), expressing the three homogeneous coordinates of i -th control point. Contrarily to what assumed in sect. 4, now we admit any point of $\mathbb{P}^2(\mathbb{R})$ to be a control point, even those with $u = 0$. Since the three coordinates $(x : y : u)$ are homogeneous, we can always multiply them by a weight, obtaining

$$C_i = \begin{pmatrix} w_i x_{C_i} \\ w_i y_{C_i} \\ w_i u_{C_i} \end{pmatrix} \quad (56)$$

The same apparatus of basis functions, control points, weights and knots will be preserved as in sect. 3. We can proceed similarly to (40) and define a curve having the following parameterization:

$$P(t) = \sum_{i=0}^n N_{i,p}(t) C_i \quad (57)$$

Vector eqn. (57) is equivalent to the following set of 3 scalar equations:

$$\begin{cases} x(t) = \sum_{i=0}^n w_i N_{i,p}(t) x_{C_i} \\ y(t) = \sum_{i=0}^n w_i N_{i,p}(t) y_{C_i} \\ u(t) = \sum_{i=0}^n w_i N_{i,p}(t) u_{C_i} \end{cases} \quad (58)$$

The curve parameterized by eqn. (58) is an extension to the real projective plane of a NURBS curves, meaning that improper points are now allowed and, particularly, control points can be also located on the improper line $u = 0$.

We can also look at the curve given by this parameterization as a “traditional” (meaning in the spirit of sect. 3) NURBS curve embedded in 3D space instead of in 2D plane. More precisely, we can think of the 3 homogeneous coordinates $(x : y : u)$ in $\mathbb{P}^2(\mathbb{R})$ as 3 affine coordinates in $\mathbb{A}^3(\mathbb{R})$; then we can add a fourth homogeneous coordinate v and pass to $\mathbb{P}^3(\mathbb{R})$, obtaining $(x : y : u : v)$. Similarly to (42), the coordinates of control points will be given by

$$C_i = \begin{pmatrix} x_{C_i} \\ y_{C_i} \\ u_{C_i} \\ 1 \end{pmatrix} = \begin{pmatrix} w_i x_{C_i} \\ w_i y_{C_i} \\ w_i u_{C_i} \\ w_i \end{pmatrix} \quad (59)$$

Similarly to (41), the NURBS curve having the given set of control points and weights and the given knot vector will have the following parameterization:

$$\begin{cases} x(t) = \sum_{i=0}^n w_i N_{i,p}(t) x_{C_i} \\ y(t) = \sum_{i=0}^n w_i N_{i,p}(t) y_{C_i} \\ u(t) = \sum_{i=0}^n w_i N_{i,p}(t) u_{C_i} \\ v(t) = \sum_{i=0}^n w_i N_{i,p}(t) \end{cases} \quad (60)$$

Coming back from $\mathbb{P}^3(\mathbb{R})$ to $\mathbb{A}^3(\mathbb{R})$ we will get

$$\begin{cases} \frac{x(t)}{v(t)} = \frac{\sum_{i=0}^n w_i N_{i,p}(t) x_{C_i}}{\sum_{i=0}^n w_i N_{i,p}(t)} \\ \frac{y(t)}{v(t)} = \frac{\sum_{i=0}^n w_i N_{i,p}(t) y_{C_i}}{\sum_{i=0}^n w_i N_{i,p}(t)} \\ \frac{u(t)}{v(t)} = \frac{\sum_{i=0}^n w_i N_{i,p}(t) u_{C_i}}{\sum_{i=0}^n w_i N_{i,p}(t)} \end{cases} \quad (61)$$

Now we want to come back from $\mathbb{A}^3(\mathbb{R})$ to $\mathbb{P}^2(\mathbb{R})$. In other word, we have to remember that the three coordinates $(x/v : y/v : u/v)$ are homogeneous, i.e. defined up to a constant factor. The smart choice of this constant factor is $v(t)$ point wise (“constant” is referred to the three homogeneous coordinates of the same point, meaning that it must be the same for them but being possible that different points are expressed with different “constant” factors, one for each point). As a consequence, multiplying the three homogeneous coordinates $(x(t)/v(t) : y(t)/v(t) : u(t)/v(t))$ by $v(t)$ we get from (61) to (58).

The geometrical meaning of the previous formulae is as follows. The points of $\mathbb{P}^2(\mathbb{R})$ are in one-to-one correspondence with those lines in $\mathbb{A}^3(\mathbb{R})$ going thru the origin. The correspondence is established by identifying the three homogeneous coordinates of a given point in $\mathbb{P}^2(\mathbb{R})$ with the three director cosines of a line in $\mathbb{A}^3(\mathbb{R})$. Such lines cut the surface of a sphere centered in the origin of $\mathbb{A}^3(\mathbb{R})$ into couples of antipodal points. Each couple of antipodal points is then in one-to-one correspondence with a point in $\mathbb{P}^2(\mathbb{R})$. Hence any curve in $\mathbb{P}^2(\mathbb{R})$ can be associated to a curve on the surface of sphere embedded in $\mathbb{A}^3(\mathbb{R})$. We can also decide to choose only one point as a representative for a given couple of antipodal points, so that a topological model for $\mathbb{P}^2(\mathbb{R})$ is a hemispherical surface, for example the northern hemisphere of the

aforementioned sphere. As a consequence, (61) is the parametric equation of a spatial curve having the particular property of being contained in a hemispherical surface. Generally speaking, the geometrical meaning of the passage from the real projective plane $\mathbb{P}^2(\mathbb{R})$ to the real affine plane $\mathbb{A}^2(\mathbb{R})$ is that of a projection. More precisely, we can choose a plane in $\mathbb{A}^3(\mathbb{R})$, for example the horizontal one being tangent to the sphere in its north pole. Then we can project points lying on the hemisphere onto the plane by using the center of the sphere (i.e. the origin of $\mathbb{A}^3(\mathbb{R})$) as the center of the projection. The plane the hemisphere is projected on is a bi-dimension sub-space embedded in tri-dimension space $\mathbb{A}^3(\mathbb{R})$ and it can be associated to the real affine plane $\mathbb{A}^2(\mathbb{R})$. It is clear that antipodal points have the same projection on the real affine plane and points lying on the equator are projected “at infinity” on the tangent plane. Hence the equator is what we call the improper line, to be dropped when passing from projective to affine. Of course, no conceptual difference exists at all between proper and improper points on the sphere, i.e. in $\mathbb{P}^2(\mathbb{R})$. Moreover, many other tangent planes can be chosen in lieu of the horizontal one, producing different affinizations of the projective plane. Once a tangent plane has been chosen, the relevant improper line is the equatorial circumference parallel to it.

From the theoretical standpoint, we are still facing a NURBS curve, even if in the framework of a different space. As a matter of fact, the object we are dealing with is a spatial NURBS and then it is endowed of all the important properties of NURBS curve. Just to exemplify, the convex hull property, continuity and differentiability properties, the effects of multiple knots, the clamping conditions, and the many other features we used throughout sect. 4 still hold without the necessity to proof them once again.

As regards the final step from $\mathbb{P}^2(\mathbb{R})$ (were improper points can be defined) to $\mathbb{A}^2(\mathbb{R})$ (where improper points have to be dropped), it is obvious that the final outcome will be the lacking of “infinite” points. Nevertheless, they are not completely lost, since they can be restored in any moment by looking at the relevant projective curve. Moreover, the latter can be affinized in any way, according to the needs of the user. In other words when one passes from the real projective to the real affine plane he has to drop a line, namely the improper line, but it is always possible to choose any line as the improper one. Hence, a different affinization can be achieved by changing the projection plane or, conversely, by operating projectivities. Of course, many different projections can be performed, in order to have a complete vision of the real projective curve by means of many real affinizations or we can also look at the projective curve depicted on the surface of a sphere (or of a hemisphere, if preferred). We can pass from $\mathbb{P}^2(\mathbb{R})$ to $\mathbb{A}^2(\mathbb{R})$ affine coordinates by dividing the first two projective coordinates by the third one, obtaining:

$$\begin{cases} X(t) = \frac{x(t)}{u(t)} = \frac{\sum_{i=0}^n w_i N_{i,p}(t) x_{C_i}}{\sum_{i=0}^n w_i N_{i,p}(t) u_{C_i}} \\ Y(t) = \frac{y(t)}{u(t)} = \frac{\sum_{i=0}^n w_i N_{i,p}(t) y_{C_i}}{\sum_{i=0}^n w_i N_{i,p}(t) u_{C_i}} \end{cases} \quad (62)$$

It is evident that the affinization (62) of the projective curve (58) will exclude improper points. It is however easy to implement an algorithm capable to avoid the latter and to keep the proper ones only, being also capable to switch from one affine branch to the next by “jumping” the improper connection point.

As regards the parameterization of the conic arc γ_+ , we will refer to the local coordinate system obtained thru projectivity T in sect. 2. In such a reference frame the control points have coordinates $C_0 = (0 : 1 : 0)$, $C_1 = (0 : 0 : 1)$, $C_2 = (1 : 0 : 0)$. We will go on using the same knot vector $(0, 0, 0, 1, 1, 1)$ and weights $w_0 = w_2 = 1$, so to clamp the NURBS curve at extremes of its control polyline. Being the knot vector unchanged with respect to sect. 4, the basis functions are also the same. As a consequence we can rewrite (58) as follows:

$$\begin{cases} x(t) = t^2 \\ y(t) = (1-t)^2 \\ u(t) = w_1 2t(1-t) \end{cases} \quad (63)$$

The only unknown parameter left is the weight of the intermediate control point, to be determined once again in order that the NURBS curve is exactly the conic arc. Substituting (63) in the equation (30) of the conic yields

$$t^2(1-t)^2(u_{P-1}^2 - 4w_1^2 x_{P_1} y_{P_1}) = 0 \quad (64)$$

This equation can be satisfied identically provided that

$$w_1 = \varepsilon \sqrt{\frac{u_{P-1}^2}{4x_{P_1}y_{P_1}}} \quad (65)$$

Eqn. (65) gives the weight to be assigned to the intermediate control point. Since P_1 has to be chosen in S , then $x_{P_1}y_{P_1} > 0$, so that by necessity the denominator is non-vanishing and the radicand is non-negative. As regards the sign ε of the weight, a similar reasoning can be carried out as in sect. 4 and the value of ε can be determined by means of (52).

6 NURBS interpolation of the global real projective plane curve

Up to now, we have found the equation of a single conic satisfying five linear conditions and then we have found a NURBS parameterization of such an arc in order to reproduce exactly the same conic. Now we want to join together the NURBS parameterizations of all the conic arcs constituting the interpolation of the original plane curve, in order to obtain just one NURBS curve. To this aim, we will not change neither the degree p of the NURBS nor weights w_i . We will only build a new knot vector (said global knot vector) and a new control polyline (said global control polyline) starting from those relevant to each conic arc (and thus said to be local control polylines).

As regards the global control polyline, we will simply merge the extreme nodes of local control polylines. More precisely, the first control point of a given local control polyline will be merged with the last control point of the preceding local control polyline. Only one of the two merged control point will survive, whereas the other will be lost, so that no double control points will exist. It is worth noticing that the tangents need not to be aligned in merged control points, since the curve to be interpolated can also be only piece wise (instead of continuously) differentiable. As already mentioned, each control point will keep the same weight. This is not source of troubles since for each local control polyline the intermediate control point is not merged and the merged control points have always weight 1. In case of closed curves, the last and the first control point will have the same coordinates and also the same weight, but they will not be merged.

The global knot vector has to be built in such a way to preserve the tangency of the NURBS to the control polyline in all merged control points. This can be achieved by putting a triple knot at the beginning ($t = 0$) and another triple knot at the end ($t = 1$), and by putting N double knots uniformly spaced in between the extremes, where N is the number of conic arc diminished by 1.

A simple example will explain how to manage it all. The curve to be represented is the boundary of a rectangle with half circles glued onto two edges. Figure 17 shows the final NURBS curve (and sample points on it) together with its control points and control polyline. Since this curve is really made up of conic arcs, its NURBS parameterization is indeed exact. Table 1 lists all control points coordinates and weights, also specifying how the global control polyline has been deduced from the local ones relevant to the many conic arcs, viz. the upper and lower horizontal segments and 4 quarters of circumference. Table 2 lists the global knot vector components. The global curve is constituted by 6 local conic arcs; hence there are $N = 6 - 1 = 5$ double knots equally spaced; two triple knots, one at the beginning and one at the end complete the vector.

The reason why such a procedure works is not difficult to be understood. The triple knots placed at the extremes make the NURBS to be clamped to its control polyline, because the interpolation degree p is still equal to 2. The other (double) knots are equally spaced and this ensures that the many conic arcs are the image of equally wide portions of interval $[0, 1]$. Moreover, the fact those knots are double makes the NURBS be tangent to the control polyline in correspondence of them, i.e. in correspondence of the passage from one conic arc to the other.

A complete proof of the last statement is quite tedious and beyond the scopes of this little work. Nevertheless, it can be readily deduced from some known facts about NURBS theory. Particularly, it can be proved that basis function are C^{p-k} continuous in correspondence of a knot having multiplicity k . In our case $p = k = 2$, so that C^0 continuity is granted where two conic arc are joined (whereas inside each conic arc the NURBS is infinitely differentiable). It can also be proved that the higher a knot multiplicity, the less the number of non-zero basis functions in correspondence of such a knot. More

| <i>notes</i> | <i>x-coordinate</i> | <i>y-coordinate</i> | <i>weight</i> | <i>notes</i> |
|-----------------------|---------------------|---------------------|---------------|-----------------------|
| upper | 0 | 10 | 1 | =last control point |
| horizontal | 25 | 10 | 0 | interm. control point |
| segment | 50 | 10 | 1 | NE |
| interm. control point | 60 | 10 | $1/\sqrt{2}$ | quarter |
| SE | 60 | 0 | 0 | circumference |
| quarter | 60 | -10 | $1/\sqrt{2}$ | interm. control point |
| circumference | 50 | -10 | 1 | lower |
| interm. control point | 25 | -10 | 0 | horizontal |
| SW | 0 | -10 | 1 | segment |
| quarter | -10 | -10 | $1/\sqrt{2}$ | interm. control point |
| circumference | -10 | 0 | 1 | NW |
| interm. control point | -10 | -10 | $1/\sqrt{2}$ | quarter |
| =first control point | 0 | 10 | 1 | circumference |

Table 1: Global control polyline

| t_0 | t_1 | t_2 | t_3 | t_4 | t_5 | t_6 | t_7 | t_8 | t_9 | t_{10} | t_{11} | t_{12} | t_{13} | t_{14} | t_{15} |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|----------|----------|----------|----------|----------|----------|
| 0 | 0 | 0 | 1/6 | 1/6 | 2/6 | 2/6 | 3/6 | 3/6 | 4/6 | 4/6 | 5/6 | 5/6 | 1 | 1 | 1 |

Table 2: Global knot vector

precisely, if the multiplicity of a knot is k , then there are at most $p - k + 1$ non-zero basis functions at this knot. Hence, in our case we have only 1 non-zero basis function in correspondence of each double knot, so that only one control point has non-zero coefficient. As a result of the partition of unity property, the NURBS is forced to go thru such a control point, i.e. the point where two conic arc are joined. A more refined argumentation, referencing to the NURBS derivative, will also proof the tangency to the legs of the control polyline.

Summing up, we have a degree 2 NURBS curve going thru the same points and tangent to the same tangents as the assembly of conic arcs. In other words the NURBS curve is a piece wise conic, each conic satisfying 4 out of 5 linear conditions defining the “correct” conic. Then, maintaining the same choice for the weights of the intermediate control points (those not merged) ensures that the NURBS is exactly fitting the conic arc assembly.

As already mentioned, the possibility exists to have non-smooth points if the tangents to different conic arcs are not coincident. Of course, wherever such a condition is satisfied the NURBS is at least C^1 continuous. Also polygons may be easily parameterized by means of this method, being all their edges to be managed as segments of (double) lines. As regards line segments, the convexity properties that had permitted to represent them when dealing with the local fitting conic are inherited by the global NURBS. This is due to a general property of NURBS theory known as strong convex hull property and stating that any NURBS curve point belongs to the convex hull of a limited set of control points. More precisely, if $t \in [t_i, t_{i+1}]$, then it can be proved that the relevant point $P(t)$ lies in the convex hull of control points $C_i, C_{i-1}, \dots, C_{i-p+1}, C_{i-p}$. In our case, if t belongs to the interval mapped onto the line segment, then its image belongs to the convex hull of the three collinear control points, i.e. it lies onto the segment. Analogous conclusions about line segments and weights influence can be drawn starting from this fact, as already detailed about the local fitting conic.

The procedure for joining together local fitting curves to produce the global curve can be applied both to the method for $\mathbb{A}^2(\mathbb{R})$ curves and to the more general, extended method for $\mathbb{P}^2(\mathbb{R})$ curves. The reason is, as usual, that the latter is still a method for producing NURBS, and no other properties have been used for justifying the joining procedure, but the “universal” NURBS properties.

Figure 18 shows an example of real projective plane curve with improper points, obtained thanks to the generalized method. The curves is a rectangular hyperbole $y = 1/x$ joined with a circle with radius equal to $1/\sqrt{2}$ and tangent to the hyperbole in (1;1). Since it is piece wise conic, the curve is exactly parameterized by the second degree NURBS chosen according to what previously explained. The curve has been partitioned into 6 sub arcs, indicated in Figure 18, together with their control points and the

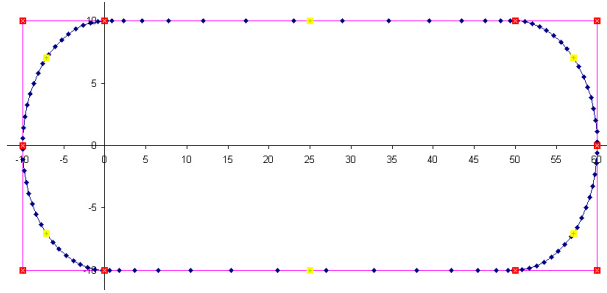


Figure 17: The NURBS curve (blue line with diamonds) together with its control polyline (fuchsia plain line), control points (red squares) and the other passage points (yellow squares)

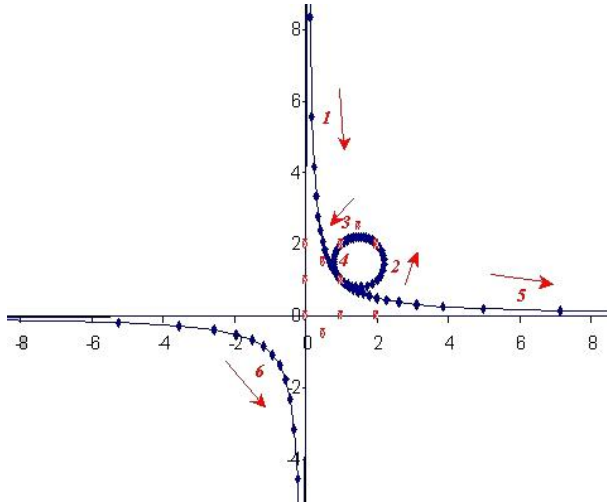


Figure 18: A real projective plane curve parameterized as a NURBS curve (generalized method)

sense it is spanned by the parameter. In case of improper control points, a proper point has been placed indicating the direction of the line whose improper point is the one into object. For example, the red square SE of the origin indicates the NW-SE direction. This is the direction associated to the improper point resulting from the intersection of the parallel, NW-SE directed tangents to sub arc no. 2. As regards this arc, notice that a half circumference has been exactly represented without the necessity to split it into 2 sub-arcs. The latter solution, compulsory in case of the non generalized method, has been used as a reference for the two quarter circumference arcs no. 3 and 4. Notice also that arcs no. 1 and 5 have one improper control point, whereas arc no. 6 has two improper control points.

7 Final remarks

The only approximation involved with this method is the choice of the partition of the global curve. Obviously, at least for regular curves it is always possible to approximate as close as desired any curve, provided that a suitably numerous set of arcs is chosen to be converted into conic arcs.

Furthermore, the method can handle also non-simple and non-smooth curves. In case of self-intersecting curves it is trivially sufficient to partition it in such a way that no arc is self-intersecting, as illustrated below (Figure 19.a). The only regularity conditions required are continuity and piece wise differentiability of the curve to be interpolated. In case of non regular curves, i.e. non continuously differentiable curves, it is once again sufficient to adopt a suitable curve partitioning in such a way that each arc is regular, as shown below (Figure 19.b).

We have already mentioned that, when sampling a NURBS curve, it could be preferable to have

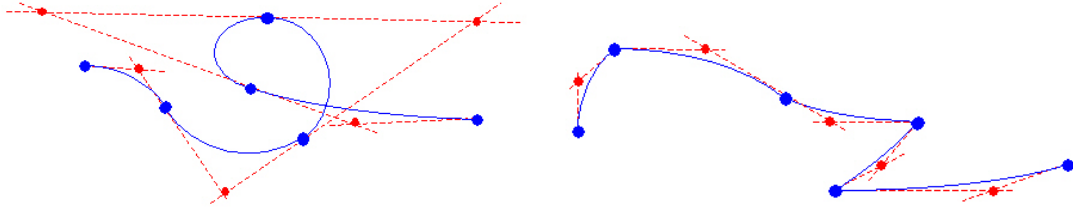


Figure 19: a) A non-simple curve, b) A non regular (continuous, piece wise differentiable) curve

closer sampled points in particular zone, depending on the application the NURBS is devoted to. A wise partitioning of the original curve can bring some benefits. Particularly, if the NURBS is sampled in correspondence of values of parameter t equally spaced, then, due to the particular choice of the knot vector, each sub-arc will receive more or less the same number of sampled points. If the opposite is desirable, one can work out a more rational choice of values for parameter t . However, this could not be possible in case the sampling routine is managed by an external code. In such a case, it is always possible to partition the curve in such a way that smaller arcs are placed in the more delicate zone. Then, since the number of samples is virtually constant, shorter arcs are denser with samples.