Approximation of Involute Curves for CAD-System Processing

Fumitaka Higuchi¹, Shuuichi Gofuku¹
Takashi Maekawa^{1,*}, Harish Mukundan², Nicholas M. Patrikalakis²

Yokohama National University, Department of Mechanical Engineering Digital Engineering Laboratory, Tokiwadai, Hodogaya, Yokohama 240-8501, Japan

² Massachusetts Institute of Technology, Department of Mechanical Engineering Design Laboratory, Cambridge, MA 02139-4307, USA

YNU Digital Engineering Laboratory Memorandum 05-1 Copyright ©2005 Yokohama National University, All rights reserved

December 21, 2006

Abstract

In numerous instances, accurate algorithms for approximating the original geometry is required. One typical example is a circle involute curve which represents the underlying geometry behind a gear tooth. The circle involute curves are by definition transcendental and cannot be expressed by algebraic equations, and hence it cannot be directly incorporated into commercial CAD systems. In this paper an approximation algorithm for circle involute curves in terms of polynomial functions is developed. The circle involute curve is approximated using a Chebyshev approximation formula [11], which enables us to represent the involute in terms of polynomials, and hence as a Bézier curve.

In comparison with the current B-spline approximation algorithms for circle involute curves, the proposed method is found to be more accurate and compact, and induces fewer oscillations.

Keywords: circle involute curves, involute gears, Chebyshev approximation formula, Bézier curves

^{*}Corresponding author. E-mail address: maekawa@ynu.ac.jp

1 Introduction

An evolute and its involute, are defined in mutual pairs. The evolute of any curve is defined as the locus of the centers of curvature of the curve. The original curve is then defined as the involute of the evolute. The simplest case is that of a circle, which has only one center of curvature (its center), which is a degenerate evolute. The circle itself is the involute of this point.

A circle involute curve (involute of a circle) is however defined as the locus of an end of a taut string as it unwinds from a base circle as depicted in Figure 1(a). A circle involute represents a path traced by a point on a line as the line rolls without slipping on the circumference of a circle. A rigorous mathematical definition is provided in Section 2. This property is extensively used in the design of spur gears, where the gear tooth surface is generated by extruding a circle involute curve along the direction which is orthogonal to the plane containing the involute curves.

Cycloids, generated by a point on the rim of a circle rolling on a straight line, are one of the first regular curves used for gear-tooth profiles. Although cycloid curves are rarely used for gear-tooth profiles today, they are used for impellors of pressure blowers [3]. Another interesting profile curve is a trochoid, which is the generalization of the cycloid defined as the locus of a point at a fixed distance from the center of a circle rolling on a straight line. The gear-tooth profiles for gear pumps, consisting of an inner gear and an outer gear, often employ trochoids. They are used as a fluid pump such as a fuel pump for automobile engines. The applications of cycloidal and trochoidal curves for gears are rather limited when compared with circle involute curves.

One of the most important characteristics of an involute gear is that it will transmit a uniform angular motion to the second gear, regardless of the changes of the distance between the centers of the two base circles [3, 8]. The tooth profile as shown in Figure 1(b) represents a commonly used gear tooth, and is an involute of a circle.

Key issues to be kept in mind while designing a gear profile are of high positional accuracy, affordable precision, consistent performance and zero backlash. Detailed gear design terminology can be found in [3]. As it will be shown in Section 2, a significant limitation of the circle involute curve is that it is transcendental and cannot be expressed by algebraic equations. This prevents circle involute curves from being directly incorporated into commercial CAD systems leading to decreased accuracy and performance degradation during gear manufacture. Further, this difficulty is multiplied if one needs to manipulate these approximated curves, for example during the introduction of fillets, or during finite element discretization.

A method proposed by Barone [2] to overcome the above issues involves evaluating and further interpolating a number of discrete points on the circle involute curve, as well as first and second derivatives evaluated from the tangent and the curvature of the involute curve using B-spline curve. However, due to the nature of interpolation, it is inevitable that oscillations occur between interpolating points. Such oscillations could result in a not-so-smooth tooth surface causing increased wear and tear leading to performance losses.

In this paper, we propose to use a Chebyshev approximation formula to approximate the transcendental equation of the circle involute curve into a polynomial in the monomial form. This is followed by representing the approximated polynomial in the monomial form to an equivalent Bernstein form (Bézier curve), which is numerically more stable

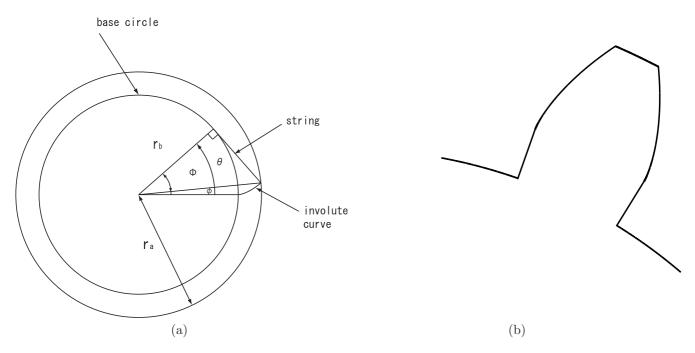


Figure 1: (a) Definition of a circle involute. (b) A gear tooth profile.

under perturbations [9]. When compared with the method proposed by Barone [2], the approximated curve based on our method oscillates with both lesser amplitude and lesser frequency.

The paper is structured as follows. In Section 2, the mathematical definition a circle involute curve and its differential geometry properties are presented. The Chebyshev approximation formula is introduced in Section 3, and is applied in Section 4 to obtain the Bézier curves. Section 5 illustrates our method with examples and comparisons with industry standards. Finally in Section 6 we conclude our work, and produce recommendations for further research.

2 Definition and Differential Geometry of Circle Involute Curves

If we denote the equation of a circle of radius r_b and centered at the origin by

$$x(\theta) = r_b \cos \theta, \quad y(\theta) = r_b \sin \theta,$$
 (1)

then the mathematical expression for its involute is derived as [7]

$$x(\theta) = r_b(\cos\theta + \theta\sin\theta), \quad y(\theta) = r_b(\sin\theta - \theta\cos\theta).$$
 (2)

Clearly, the equation (2) representing the involute of the circle is transcendental. Transcendental equations in general cannot be expressed as algebraic equations.

The first and second order derivatives of the circle involute curves are readily derived as

$$\dot{x}(\theta) = r_b \theta \cos \theta, \quad \dot{y}(\theta) = r_b \theta \sin \theta \,, \tag{3}$$

$$\ddot{x}(\theta) = r_b(\cos\theta - \theta\sin\theta), \quad \ddot{y}(\theta) = r_b(\sin\theta + \theta\cos\theta), \tag{4}$$

Accordingly the arc length, unit tangent vector and curvature of the circle involute curve are evaluated as:

$$s(\theta) = \int_0^\theta \sqrt{\dot{x}^2 + \dot{y}^2} d\theta = \frac{r_b \theta^2}{2} ,$$
 (5)

$$\mathbf{t}(\theta) = \left(\frac{\dot{x}}{\sqrt{\dot{x}^2 + \dot{y}^2}}, \quad \frac{\dot{y}}{\sqrt{\dot{x}^2 + \dot{y}^2}}\right) = (\cos \theta, \sin \theta), \tag{6}$$

$$\kappa(\theta) = \frac{\dot{x}\ddot{y} - \dot{y}\ddot{x}}{(\dot{x}^2 + \dot{y}^2)^{\frac{3}{2}}} = \frac{1}{r_b\theta} . \tag{7}$$

Equation (7) clearly confirms that the locus of center of curvatures of the circle involute is the circle itself. From equation (3), it is apparent that the circle involute curve is singular at $\theta = 0$, however, the unit tangent (6) is well defined as θ cancels out in the normalization process. Also we can observe that the curvature is infinite at $\theta = 0$, and rapidly decreases as we move away from the base circle. Therefore the curve near the base circle is very difficult to produce accurately, and should be avoided whenever possible [3]. The ending point of the involute curve along the tooth profile is determined by computing the intersection point between the circle involute curve and the addendum circle of radius r_a . Thus we have

$$r_a \cos \phi = r_b(\cos \theta + \theta \sin \theta) \tag{8}$$

$$r_a \sin \phi = r_b (\sin \theta - \theta \cos \theta) . \tag{9}$$

By squaring the both hand sides of equations (8) and (9), and adding them we obtain

$$r_a^2 = r_b^2 (1 + \theta^2)$$
, (10)

from which the parametric value θ_a of the intersection is derived to be

$$\theta_a = \frac{\sqrt{r_a^2 - r_b^2}}{r_b} \ . \tag{11}$$

Therefore the tooth profile is obtained for a range of the parameter θ such that $0 \le \theta \le \theta_a$ using equation (2). Further, the arc length along the tooth profile between the base circle and the addendum circle is evaluated from (5) as follows:

$$s_a = \frac{r_b \theta_a^2}{2} = \frac{r_a^2 - r_b^2}{2r_b} \ . \tag{12}$$

3 Chebyshev Approximation Formula

The Chebyshev polynomial of degree k, $T_k(x)$ is defined in [4, 11] as:

$$T_k(x) = \cos(k\arccos x) \ . \tag{13}$$

Using the trignometric identities, we have explicit expressions for $T_k(x)$,

$$T_0(x) = 1, (14)$$

$$T_1(x) = x \,, \tag{15}$$

$$T_2(x) = 2x^2 - 1, (16)$$

$$T_3(x) = 4x^3 - 3x \,, \tag{17}$$

$$T_4(x) = 8x^4 - 8x^2 + 1 (18)$$

. . .

from which we obtain the three-term recurrence relation

$$T_{k+1}(x) = 2xT_k(x) - T_{k-1}(x), \quad k \ge 1.$$
 (19)

The Chebyshev polynomials satisfy a discrete orthogonality relation as follows:

$$\sum_{k=1}^{m} T_i(x_k) T_j(x_k) = \begin{cases} 0 & i \neq j \\ m/2 & i = j \neq 0 \\ m & i = j = 0 \end{cases},$$
 (20)

where x_k $(k=1,\ldots,m)$ are the m zeros of $T_m(x)$ given by

$$x_k = \cos\left(\frac{\pi(k - \frac{1}{2})}{m}\right) . (21)$$

Suppose an arbitrary function f(x) in the interval [-1,1] is approximated by a finite sum of Chebyshev polynomials as follows:

$$f(x) \approx \sum_{k=0}^{N-1} c_k T_k(x) - \frac{1}{2} c_0$$
 (22)

If we substitute $x=x_k$ into (22), we obtain

$$f(x_k) \approx c_0 T_0(x_k) + c_1 T_1(x_k) + \dots + c_{N-1} T_{N-1}(x_k) - \frac{1}{2} c_0$$
 (23)

Multiplying $T_i(x_k)$ to the both hand sides of (23) and taking the summation from k=1 to N yields,

$$\sum_{k=1}^{N} f(x_k) T_j(x_k) \approx \sum_{k=1}^{N} c_0 T_0(x_k) T_j(x_k) + \dots + c_j T_j(x_k) T_j(x_k) + \dots + c_{N-1} T_{N-1}(x_k) T_j(x_k) - \frac{1}{2} c_0 T_j(x_k)$$

$$= c_j \frac{N}{2}, \qquad (24)$$

where the discrete orthogonality condition (20) was used. Thus the coefficients c_i become

$$c_j = \frac{2}{N} \sum_{k=1}^N f(x_k) T_j(x_k)$$

$$= \frac{2}{N} \sum_{k=1}^N f\left(\cos\left(\frac{\pi(k-\frac{1}{2})}{N}\right)\right) \cos\left(\frac{\pi j(k-\frac{1}{2})}{N}\right). \tag{25}$$

It is well known that (22) is exact for x equal to the N zeros of $T_N(x)$ [11]. By definition (13), the Chebyshev polynomials are bounded between ± 1 , and therefore for $n \ll N$, the truncated approximation

$$f(x) \approx \sum_{k=0}^{n-1} c_k T_k(x) - \frac{1}{2} c_0 ,$$
 (26)

is no larger than (22) by the sum of c_n, \ldots, c_{N-1} . Since c_k decreases sharply as k increases (see Table 1), the error is dominated by the term $c_n T_n(x)$. As the polynomial $T_n(x)$ has n+1 extrema at

$$x = \cos\left(\frac{\pi k}{n}\right) , \quad k = 0, 1, \dots, n , \qquad (27)$$

the error $c_n T_n(x)$ is distributed smoothly over the interval [-1,1]. If we exclude the extrema at the two boundaries, the number of extrema is n-1. Unlike Taylor-series expansion based approximation, where the error grows rapidly as we evaluate away from the point of expansion, the error of the Chebyshev approximation is spread out over the interval [-1,1]. The Chebyshev approximation formula (26) is very close to the minimax polynomial, which has the smallest maximum deviation from the true function among all polynomials of the same degree [11].

4 Bézier Approximation of Circle Involute Curves

In this section we explain how to obtain the Bézier approximation of a circle involute curve. Let us say, we need to obtain the profile of a tooth. The circle involute curve of the gear tooth under consideration is defined by selecting the module m, the number of gear teeth z, and the pressure angle α [3, 8]. The pitch circle radius r, the base circle radius r_b , and the addendum circle radius r_a are related by:

$$r = \frac{mz}{2} \,, \tag{28}$$

$$r_b = r \cos \alpha \,, \tag{29}$$

$$r_a = r + m. (30)$$

To apply the Chebyshev approximation, the range of the parameter of the involute curve $0 \le \theta \le \theta_a$ (see (11)) must be transformed to the range [-1,1] by using the transformation

$$x = \frac{\theta - \frac{1}{2}\theta_a}{\frac{1}{2}\theta_a} \,. \tag{31}$$

The transition from the interval $0 \le \theta \le \theta_a$ to the interval $-1 \le x \le 1$ is an affine map, and the polynomials are invariant under such transformation [5]. After the affine parameter transformation, we compute the N coefficients of the Chebyshev sum (22) using (25), and truncate the expansion at k = n - 1. Conventionally the Bézier curves are represented in the range $0 \le t \le 1$, therefore we transform the degree p = (n - 1) polynomial equation over the range $-1 \le x \le 1$ to the interval $0 \le t \le 1$ by another affine parameter transformation $t = \frac{x+1}{2}$ yielding

$$f(t) = \sum_{i=0}^{p} c_i^M t^i, \quad 0 \le t \le 1.$$
 (32)

Equation (32) can be converted to Bernstein polynomials of the form

$$f(t) = \sum_{i=0}^{p} c_i^B B_{i,p}(t) , \quad 0 \le t \le 1 ,$$
 (33)

using the following formula [6]:

$$c_i^B = \sum_{j=0}^i \frac{\binom{i}{j}}{\binom{p}{j}} c_j^M , \qquad (34)$$

where $B_{i,p}(t)$ is the i^{th} Bernstein polynomial of degree p. After this conversion, the approximation to the involute curves becomes

$$x(t) = \sum_{i=0}^{p} x_i B_{i,p}(t) , \qquad (35)$$

$$y(t) = \sum_{i=0}^{p} y_i B_{i,p}(t) . (36)$$

or in a vector form we rewrite equations (35) and (36) as:

$$\mathbf{r}(t) = \sum_{i=0}^{p} \mathbf{b}_i B_{i,p}(t) , \qquad (37)$$

where $\mathbf{r}(t) = (x(t), y(t))^T$ and $\mathbf{b}_i = (x_i, y_i)^T$.

As we pointed out in Section 2, the involute curve has a singularity and an infinite curvature at the base circle. Consequently, the approximated Bézier curve also has a singularity at t = 0 and results in a double control point at the beginning of the curve as shown in Figure 2(a). To avoid this high curvature span of the curve, one can truncate a small arc of the curve near the base curve. We can do this by setting a short arc length s_s , say one hundredth of s_a ,

and changing the parameter of the involute curve to $\theta_s \leq \theta \leq \theta_a$, where $\theta_s = \sqrt{\frac{2s_s}{r_b}}$. Figure 2(b) shows the same Bézier curve whose first 1 % of the arc length is truncated to avoid the singular point.

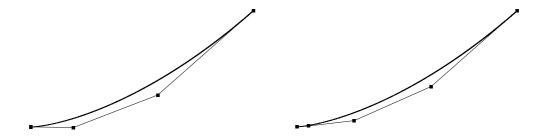


Figure 2: Circle involute curve approximated by a degree 4 Bézier curve: (a) double control point at t=0, (b) the first 1% of the curve is truncated.

5 Illustration

As an illustration we used an example from Barone [2]. The example involves a pinion with tooth number z = 17, pitch diameter d = 51mm, pressure angle $\alpha = 25^{\circ}$, and the parameter value at the addendum circle $\theta_a = 0.735$. Figure 3 shows the 3D CAD generated solid model of this model. We list the first ten coefficients of the Chebyshev approximation formula c_j (Equation (25)) with N = 50 in Table 1.

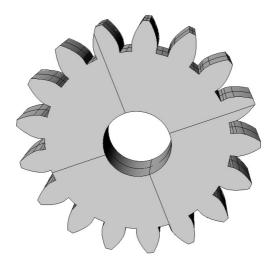


Figure 3: Solid model of the spur gear used as an illustrative example.

Deviations of the approximated gear profile from the true circle involute curve are evaluated by computing the minimum distance at equally distributed points along the profile curve. Figures 4(a), (b), (c) show the deviations normalized by the pitch diameter for approximating Bézier curves of degree 4, 6, and 8, respectively.

We also evaluated the angle difference (measured in degree) between the tangent vectors of the true circle involute curves and the approximating Bézier curves of degree 4, 6, and they are shown in Figures 5 (a), (b), respectively. For

Table 1: The first ten coefficients of the Chebyshev approximation formula (N = 50 was used).

Coefficients	x-component	y-component
c_0	50.31650288217119300000	1.73630075493736520000
c_1	2.67466952231797790000	1.29618763157396270000
c_2	0.58622346095403799000	0.50814630896878366000
c_3	-0.04656704193647897400	0.07755570080796368700
c_4	-0.00540616886595515250	-0.00282158624452299050
c_5	0.00012782053121384251	-0.00026392159909978239
c_6	0.00001000612010670920	0.00000462524384676655
c_7	-0.00000013935321220515	0.00000031124759300513
c_8	-0.00000000822238192200	-0.00000000359687314274
<i>c</i> ₉	0.000000000008119997119	-0.00000000018894385989

the degree 8 Bézier curve, the difference was nearly zero, so we did not include this in the Figure 5. Figures 6 (a), (b), (c) depict the curvature distributions of the true circle involute curve and the approximating Bézier curves of degree 4, 6, and 8 at equally distributed points along the profile curve.

In Table 2, we tabulate and compare the maximum as well as the average deviations from the true curve normalized by the pitch circle diameter using a variety of methods. We reproduce results by Barone [2] for B-spline curves with $(p=3, n_c=9), (p=3, n_c=18), (p=4, n_c=9),$ and $(p=4, n_c=18)$. Here p is degree of the B-spline curve and n_c is the number of control points. Pro/ENGINEER [13] allows the user to input a curve in a parametric representation, even as a transcendental equation. Different accuracies of the involute curves (resulting in cubic B-spline curves with $n_c=6,7,9$ and 12) are achieved by modifying the tolerance of the geometric entities within Pro/ENGINEER (Pro/E) and tabulated in Table 2. Another geometric modeling system, Praxiteles [1, 14, 12] uses an approximation algorithm to generate a cubic B-spline (p=3) curve with $(n_c=9)$ and $(n_c=18)$. All the curves mentioned in the Table 2 have full length, in other words are not truncated.

One can observe that the increase of the degree of the Bézier curve contributes significantly in the reduction of the deviations of the positions, tangent directions, and curvature, whereas for the B-spline approximation method the reduction is not as significant with an increase in the number of control points. By comparing the number of control points and the number of knots, it is clear that our method has a more compact representation. Finally, we would like to point out that the number of extrema of the oscillations, excluding the boundaries, is at most n, and is in general less than the B-spline fitting method.

This can be easily explained by the following reasoning. In the Chebyshev approximation, we truncated the approximation at k = n - 1, and therefore the approximation error is dominated by the term $c_n T_n(x)$, which has n - 1 (= p) extrema excluding the boundaries. By contrast, the B-spline fitting method has at most $n_c - 1$ stationary points, where n_c is the number of control points.

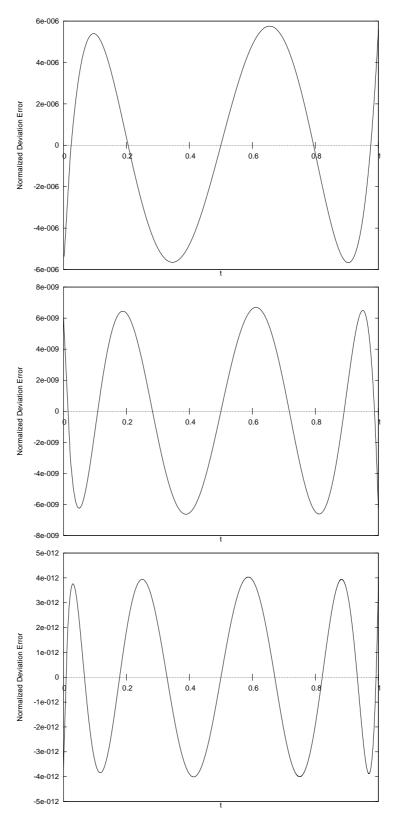


Figure 4: Position deviations of the approximated gear profile from the true circle involute curve at equally distributed points along the span of the profile curve: (a) degree 4 Bézier curve; (b) degree 6 Bézier curve; (c) degree 8 Bézier curve.

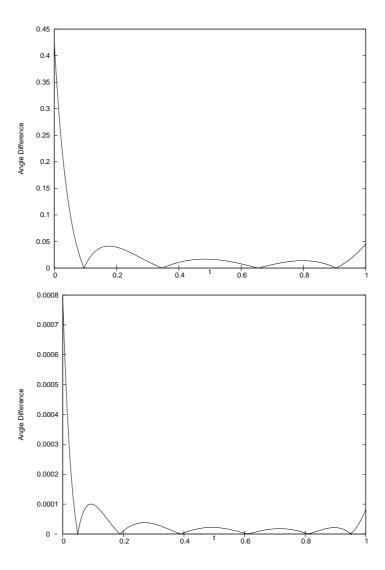


Figure 5: Angle difference (measured in degree) between the tangent vectors of the approximated gear profile and the true circle involute curve at equally distributed points along the span of the profile curve: (a) degree 4 Bézier curve; (b) degree 6 Bézier curve.

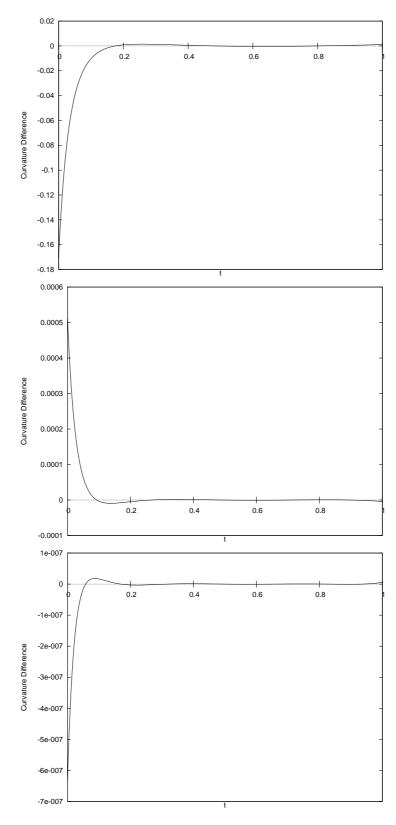


Figure 6: Curvature deviations of the approximated gear profile from the true circle involute curve at equally distributed points along the span of the profile curve: (a) degree 4 Bézier curve; (b) degree 6 Bézier curve; (c) degree 8 Bézier curve.

Table 2: Summary of the normalized deviations computed as the difference between the approximating curve and the true involute curve.

Method	Approximating	Deviations	
	Curve	Maximum	Average
Barone	B-spline ($p=3, n_c=9$)	2.157×10^{-4}	4.118×10^{-5}
Barone	B-spline ($p=3, n_c=18$)	1.745×10^{-5}	3.137×10^{-6}
Barone	B-spline ($p=4, n_c=9$)	1.353×10^{-4}	3.922×10^{-5}
Barone	B-spline ($p=4$, $n_c=18$)	1.176×10^{-5}	2.353×10^{-6}
Pro/E	B-spline ($p=3, n_c=6$)	6.318×10^{-4}	3.087×10^{-4}
Pro/E	B-spline ($p=3, n_c=7$)	3.971×10^{-4}	1.141×10^{-4}
Pro/E	B-spline ($p=3, n_c=9$)	4.443×10^{-5}	1.303×10^{-5}
Pro/E	B-spline ($p=3, n_c=12$)	1.359×10^{-5}	1.654×10^{-6}
Pro/E	B-spline ($p=3, n_c=13$)	4.640×10^{-6}	7.270×10^{-7}
Praxiteles	B-spline ($p=3, n_c=9$)	5.413×10^{-7}	9.762×10^{-8}
Praxiteles	B-spline ($p=3, n_c=18$)	5.663×10^{-7}	1.112×10^{-7}
This method	Bézier $(p=4)$	5.757×10^{-6}	3.598×10^{-6}
This method	Bézier $(p=6)$	6.690×10^{-9}	4.187×10^{-9}
This method	Bézier $(p=8)$	4.034×10^{-12}	2.524×10^{-12}

6 Conclusions

A novel approximation algorithm to generate and represent a circle involute curve is presented in this paper. This method employs the Chebyshev approximation formula to convert a transcendental equation of the circle involute to a polynomial in a monomial basis. This is followed by converting the polynomial in monomial basis to the Bernstein basis which allows for better numerical stability, less oscillations and a compact representation.

Experiments performed show that in comparison with the B-spline curve fitting method, significant improvements in accuracy is achieved with the proposed method. Another advantage of the proposed method is the decrease in oscillations (both in amplitude and frequency) of the approximated involute curve.

The proposed algorithm is expected to be of significant use in commercial CAD systems, particularly in the area of gear design and manufacturing. Major reduction in the oscillations and increased accuracy during curve approximation allows decreased machining tolerance, friction, and wear-and-tear.

One topic for further research is the introduction of features like fillets in the circle involute curve. It is our firm belief that introducing and further manipulating such features will be easier with the Bernstein representation of the circle involute curve.

References

- S. L. Abrams, L. Bardis, C. Chryssostomidis, N. M. Patrikalakis, S. T. Tuohy, F.-E. Wolter, and J. Zhou. The geometric modeling and interrogation system Praxiteles. *Journal of Ship Production*, 11(2):116–131, May 1995.
- [2] S. Barone. Gear Geometric Design by B-Spline Curve Fitting and Sweep Surface Modelling *Engineering* with Computers, 17(1):66-74, 2001.
- [3] E. Buckingham. Spur Gears. McGraw-Hill Book Company, Inc., New York, 1928.
- [4] G. Dahlquist and Å. Björck. Numerical Methods. Prentice-Hall, Inc., Englewood Cliffs, NJ, 1974.
- [5] G. Farin. Curves and Surfaces for Computer Aided Geometric Design: A Practical Guide. Academic Press, Boston, MA, 3rd edition, 1993.
- [6] H. Prautzsch W. Boehm and M. Paluszny Bézier and B-Spline Techniques. Heidelberg, Germany: Springer-Verlag, 2002
- [7] J. D. Lawrence. A Catalogue of Special Plane Curves. Dover Publications, Inc., New York, 1972.
- [8] R. L. Norton. Machine Design: An Integrated Approach Second Edition. Prentice Hall, 2000.
- [9] N. M. Patrikalakis and T. Maekawa. Shape Interrogation for Computer Aided Design and Manufacturing Heidelberg, Germany: Springer-Verlag, 2002.
- [10] L. A. Piegl and W. Tiller. The NURBS Book. Springer, New York, 1995.
- [11] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. Numerical Recipes in C. Cambridge University Press, 1988.
- [12] N. M. Patrikalakis. Approximate conversion of rational splines. Computer Aided Geometric Design, 6(2):155–165, 1989.
- [13] Pro/ENGINEER. Parametric Technology Corporation, MA, 2005. http://www.ptc.com/.
- [14] F.- E. Wolter and S. T. Tuohy. Approximation of high degree and procedural curves. *Engineering with Computers*, 8:61–80, 1992.