A Better Way to Design Railroad Transition Spirals

Louis T. Klauder Jr., PhD, PE.
Work Soft
833 Galer Dr.
Newtown Square, PA 19073
lklauder@wsof.com

May 25, 2001

ABSTRACT

The traditional rules for geometrical design of railroad track stipulate that when the curvature of the track needs to change (as, for example, when a section of tangent track is followed by a curve), the curvature should not change abruptly but rather should change linearly with distance. The geometrical curve shape in which the curvature changes linearly with distance along the curve is a form of spiral that is generally referred to by railroad track designers as a clothoid spiral. The present paper observes that the clothoid spiral is not a good form of spiral from the point of view of rail vehicle dynamics and illustrates earlier suggestions for improvement of spirals. The paper then explains a better way to approach the design of railroad track transition spirals and presents numerous examples showing how more logically designed spirals compare with traditional spirals.

1   Introduction

The author became interested in the question of spiral shape as a result of a study of the possibility of substantial reduction of travel times between Washington, D.C. and New York City through the combination of high train speeds and higher than normal superelevations of curves. It was found (ref. 1) that extra large superelevations would have great value in that corridor if they could be made practical. Leaving aside the various operational and safety questions that extra high superelevations would raise, there was also the question of whether track alignments with highly elevated curves could stay within the railroad property lines, given the long spirals that would be required. This led the author to think about the physical logic of spiral design and to wonder whether a spiral design starting from basic physics might have more degrees of freedom than the traditional clothoid spiral. Consideration of this question led to a new spiral design concept. In 1984, the FRA provided some support for testing whether or not the new design concept would allow design of very highly elevated spirals that could stay within railroad property lines in the North East Corridor. The results were positive in all respects. However, there did not seem to be very much interest at the time in putting the new design method into practice. A modified version of the 1985 report to the FRA was contributed to the AREMA Annual Meeting in Dallas, TX in September of 2000 and included in the proceedings of the Meeting (ref. 2). In comparison to reference 2, the present paper offers additional insight, new and better spiral shapes, and simpler presentation of the mathematics of the design method.

2   The traditional Spiral and its deficiency

Railroad alignments normally consist mostly of sections of tangent track and sections of curved track in each of which the curvature is constant. Tangent track can be thought of as track with constant curvature having the particular curvature value of zero.

A section of curved track over which vehicles travel with substantial speed is generally banked by having the rail on the outside of the curve raised relative to the inner rail. It is obvious that tangent track which is not banked cannot be connected directly to curved track that is banked. The simplest track shape that can provide a reasonably smooth transition between two adjacent sections of constant curvature track is a shape with curvature and bank angle at the each end that match those of the neighboring track section and in which the curvature and bank angle both vary linearly with distance along the transition. In plan view, this shape is a spiral commonly referred to by track designers as a clothoid spiral. Cartesian coordinates of points along a clothoid spiral are given by the Fresnel integrals, which are well known mathematical functions, and for which numerical tables and computer algorithms are readily available. See for example Abramowitz & Stegun (ref. 3). Use of this form of transition spiral has been a part of standard practice in North American railroad track design for many years and continues to be the standard practice today.

However, around 1960 some United States railroad engineers began to give a little attention to the question of whether an alternative to the clothoid spiral might provide a better transition between railroad track curves. A discussion of the situation may be found in the Association of American Railroads Research Department Report No. ER-37 published in 1963 (ref. 4). The problem with the clothoid spiral is that, conceptually, at the beginning of the spiral the vertical profile of the outer rail changes abruptly from the prevailing profile to a linear ramp with a different slope, and that at the end of the spiral the profile changes abruptly back to the prevailing profile. Thus, again conceptually, when a vehicle truck enters a clothoid spiral its angular velocity about the longitudinal or "roll" axis is forced

to go abruptly from zero to a positive value, and the same thing happens in reverse as the truck leaves the clothoid spiral. Conceptually, an instantaneous change in angular velocity about the roll axis would imply an infinite angular acceleration and a corresponding infinite torque. In practice, rails have some stiffness so that the change in outer rail profile is not instantaneous. In addition, the trucks at the two ends of a vehicle enter and leave a spiral at different times, and the vehicle suspension provides some cushioning of roll accelerations. As a result, the linear ramp profile of the outer rail of a clothoid spiral does not usually cause vehicles to lurch enough for passengers to object. However, from the point of view of vehicle dynamics, the traditional clothoid spiral is conceptually defective.

In Europe there is a tradition of attention to the deficiency of the clothoid spiral that goes back into the ninteenth century. Moreover, some railroads have made widespread use of alternative spiral geometries. The survey by Bjorn Kufver (ref. 5) provides a number of references.

A 1998 paper by the Austrian engineers Gerard Presle and Herbert Hasslinger (ref. 6) reports measurements of rail vehicle ride motion disturbance and of damaging systematic lateral force on the track structure caused by the clothoid spiral even when the linear ramp profile is deliberately "rounded" at each end.

3   Some previous ideas for improving spiral design

In chapter 8 of his very helpful survey published in 1997 (ref. 5), Bjorn Kufver surveys the history of proposals for improved spiral geometries and gives functions describing seven noteworthy candidates in a consistent notation. The following table lists the seven main spiral shapes described in detail by Kufver and provides for each one an illustrative plot and a description of the shape of the second derivative of the track curvature with respect to distance along the spiral.

Table 3-1: Alternative forms of curvature for track spirals as listed by Kufver (ref. 5).

Name used for spiral shape in survey by Kufver (publication year is from Kufver)	Description of the shape of the 2nd derivative of curvature with respect to distance	Plot illustrating shape of curvature and of its first two derivatives; The 2nd derivative is similar to the roll acceleration and is the shape that is most informative. Of the seven shapes listed below, the first five have 2nd derivatives that are discontinuous at the start and end of the spiral, and the first two have one or two additional discontinuities.
Helmert (1872), Schramm (1934), or biquadratic parabola	A step function with two steps: a constant positive value in the first half of the spiral and the opposite negative value in the second half.	

Ruch (1903)	A step function with three steps: a constant positive value at the beginning of the spiral, the opposite negative value at the end of the spiral, and in the center of the spiral between those two a step with value zero.	
Bloss (1936)	A straight line from a positive value at the beginning of the spiral, passing through zero in the center of the spiral, and reaching the opposite negative value at the end of the spiral.	
Cosine (Vojacec, 1868)	A half period of a cosine with the maximum of the cosine at the start of the spiral, the value zero in the middle of the spiral, and the negative minimum of the cosine at the end of the spiral.	
Gubar (1990)	The same as the preceding cosine shape except that the period of the cosine is shortened so that it falls to zero before the middle of the spiral and the positive and negative parts of the cosine curve are separated by a central zone with zero value.	

Watorek (1907)	A simple cubic function that rises from zero at the beginning of the spiral, bends down to pass through zero in the center of the spiral, and rises back up to zero at the end of the spiral.	
Sinusoidal (Klein, 1937)	A full period of a sine function whose shape is similar to that of the preceding simple cubic. Compared to the 2nd derivative of the Watorek curvature, the maximum value is a little higher, the initial slope is a little lower, and each half of the curve is symmetric about its midpoint.	

In his survey, Kufver notes that the central zone of zero acceleration in the Ruch and Gubar shapes has the beneficial effect of reducing the maximum value of the warp of the track in the spiral. He notes that additional candidate shapes could be obtained by inserting zero acceleration zones in the middle of some of the other listed shapes.

As far as the present author is aware, all prior proposals of improved transition spirals have arisen from the perspective in which a spiral is regarded as a track alignment shape. The present paper is an extension of work arising from a different perspective, namely the perspective that the basic purpose of a spiral is to change the bank angle of vehicles from one value to another. The connection between the two perspectives comes from the equation for balance between the centrifugal force due to track curvature and the perceived transverse component of gravitational force due to banking of the track. Leaving aside roll deflection in vehicle suspensions, the well-known equation that expresses this balance is

$$\text{track\_curvature} = (g/v_b^2) \tan (\text{roll\_angle}),$$

where g is the acceleration of gravity and v_b is the balance speed of the curve. Assuming for the purpose of superelevation measurement that the track gage is 60 inches, a superelevation of 6 inches corresponds to a bank angle close to 0.1 radians. When roll angles are that small the function tan (roll_angle) in the above relation can be replaced by the roll_angle (in radians) itself without introducing any more than about 0.5 % error. Therefore, for purposes of physical understanding one can write

$$\text{track\_curvature} = (g/v_b^2) \text{ roll\_angle (in radians)}.$$

Then solving for the roll angle, taking the second derivative of both sides with respect to distance along the track, and multiplying by the square of the vehicle velocity, v_v, we obtain

$$\text{roll\_acceleration (in rad/sec}^2) = v_v^2 \; (v_b^2 / g) \; d^2(\text{track curvature})/ds^2,$$

This means that in each plot of Table 3-1, the curve for the second derivative of the track curvature shows the shape of the roll acceleration (as a function of time) to which the associated spiral would subject vehicles. Looked at from this point of view, the Watorek and sinusoidal alternatives are the only ones in which the roll acceleration is continuous and thus free from points at which the first derivative of the angular acceleration (the angular jerk) is conceptually infinite. Thus, they are the only two that are not objectionable from a dynamic point of view. This paper will propose competing shapes that are more attractive dynamically and that are more efficient from the point of view of reducing maximum track warp in spirals.

The paper by Presle and Hasslinger (ref. 6) presents test results showing substantial improvement in ride comfort and substantial reduction in damage to track through improved contouring of spirals. The alignment shape described by Presle and Hasslinger is based in part on the sinusoidal shape shown in the preceding table. However, the spiral design approach that they have advanced includes a very significant addition relative to previously tested designs. They recognize that as a vehicle traverses a spiral the vehicle rotates about some longitudinal rotation axis (i.e., roll axis). They observe that the forces required to accomplish this rotation will be minimized if the roll axis passes through the center of gravity of the vehicle. They therefore have the center of gravity of a vehicle follow a path corresponding to the sinusoidal spiral shape and infer what shape the track should have from the combination of that assumption plus some unspecified assumption about how the track bank angle should vary with distance.

The greatly improved performance that Presle and Hasslinger have reported results to a significant degree from the raising of the height of the vehicle roll rotation axis above the plane of the track. That feature is included in the spiral design approach advocated by the present author. Presle and Hasslinger credit an article published in 1968 by vonDonges for the proposal that spiral performance could be improved by having the nominal spiral shape apply to the path of the vehicle center of gravity, which is above the plane of the track, and by devising a shape for the track that would secure that behavior.

As noted above, prior thinking about spiral design has begun with search for a "good" shape for the alignment of the spiral. The present paper argues that that is not the best way to begin.

4   The proper objective of spiral design

It is generally accepted that a spiral should provide a connection between two adjacent sections of track with differing constant values of curvature and superelevation and that at each end of the spiral its curvature and superelevation should match those of the neighboring track. It is also generally accepted that, for a vehicle traversing a spiral at design speed, the bank or roll angle of the track should vary along with the curvature of the track alignment so that the centripetal acceleration resulting from the curvature stays in balance with the perceived transverse component of gravity throughout the length of the spiral.

It is argued here that subsidiary to the forgoing two constraints the primary objective of spiral design should be to manage the roll motion that a vehicle executes as it traverses the spiral. That is to say, the

design process should begin with choice of an advantageous form of roll motion. From that roll motion the corresponding spiral shape should then be derived.

Before looking at roll motion per se, it may be helpful to review simple linear motion. The simplest form of linear motion that can take an object from one location to another "gracefully" is illustrated in Figure 4-1 below.

In that figure, the horizontal axis represents time. Think of the motion as being vertical.

The shallow "S" shaped curve represents the height of the object as a function of time.

The "bell" shaped curve is the first derivative of position with respect to time and is thus the velocity. The velocity should be zero at the beginning and end of the motion and can be seen to be so.

The curve consisting of three sloped straight-line segments represents the 2nd derivative of position with respect to time and is thus the acceleration. In this example, the acceleration is also zero at the beginning and end of the motion. When there is a desire to make a motion "smooth", attention will be directed particularly to the manner in which the acceleration varies with time.

When motion applies to people, an abrupt change in acceleration is perceived as a "jerk". Accordingly, the derivative of acceleration with respect to time is commonly referred to as the jerk. The example in this plot was created by selecting a functional form for the jerk that keeps its value limited but that is otherwise as simple as possible. It is a three-step function that begins with a constant positive value, then drops to the negative of that value, and then returns to the original value.

Figure 4-1: Simple form of motion with finite jerk.
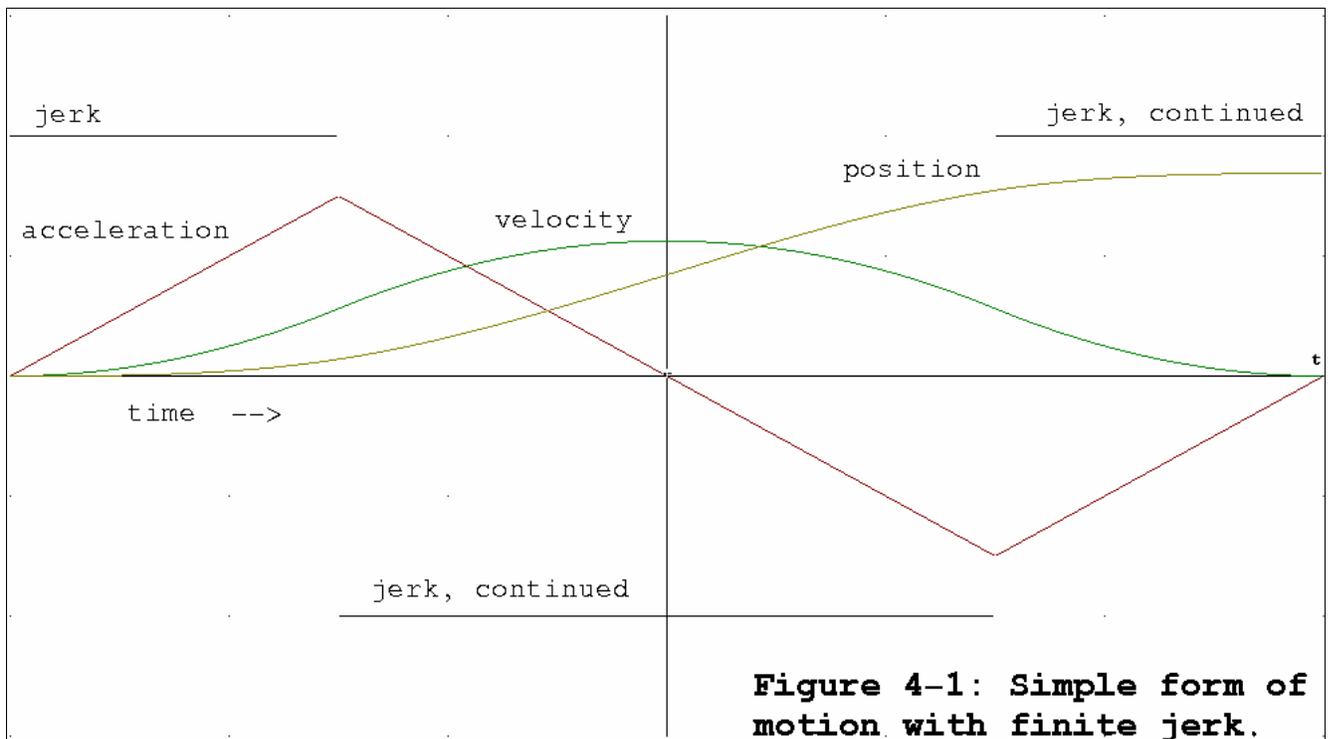
The purpose of this illustration is to get across the concept that in order for a motion to be considered "smooth" its jerk function needs to be limited so that its acceleration is not discontinuous and does not change too rapidly with time.

We apply this same perspective directly to the roll motion of a vehicle about its longitudinal axis. In the above plot, the curve for position versus time could equally well represent vehicle roll angle as a function of time.

5   Candidate models for the Roll Acceleration of a vehicle traversing a spiral

This section presents fifteen candidate forms for roll angle as a function of location for a vehicle traversing a spiral. Each candidate is illustrated by a plot that shows the shape of the roll angle, the roll velocity, and the roll acceleration. To facilitate comparison among the models, each plot has its distance axis scaled to extend from –2.0 to +2.0 and takes the roll angle from 0.0 to 0.2.

Each candidate is defined by a mathematical formula that is as simple as possible consistent with possession of its characteristic features. The formulae given below embody the following conventions.
a) Distance along the spiral is called 's', and s = 0.0 at the midpoint of the spiral.
b) The spiral extends from   s = – a   to   s = + a so that the spiral has length   2 a.
c) If a model has a central zone in which roll acceleration is identically zero, then the central zone extends from   s = - f a   to   s = + f a, so that   'f'   is the ratio of the length of the central zone to the length of the spiral.
d) The final roll angle minus the initial roll angle is called   'roll_change'.

The formula for the roll acceleration is given for each candidate. The formulae for the roll velocity and roll angle can be obtained in closed form (i.e., in terms of standard mathematical functions) for each candidate by successive integrations. The integration constant for the roll velocity is always zero. The integration constant for the roll angle is the roll angle at the beginning of the spiral.

5.1   Linear up – down at ends

The first roll model was discussed in the author's preceding paper (ref. 2). It is referred to in this paper as the "linear up-down" model and is illustrated in the following plot.

Figure 5-1:
Linear "up-down" roll model
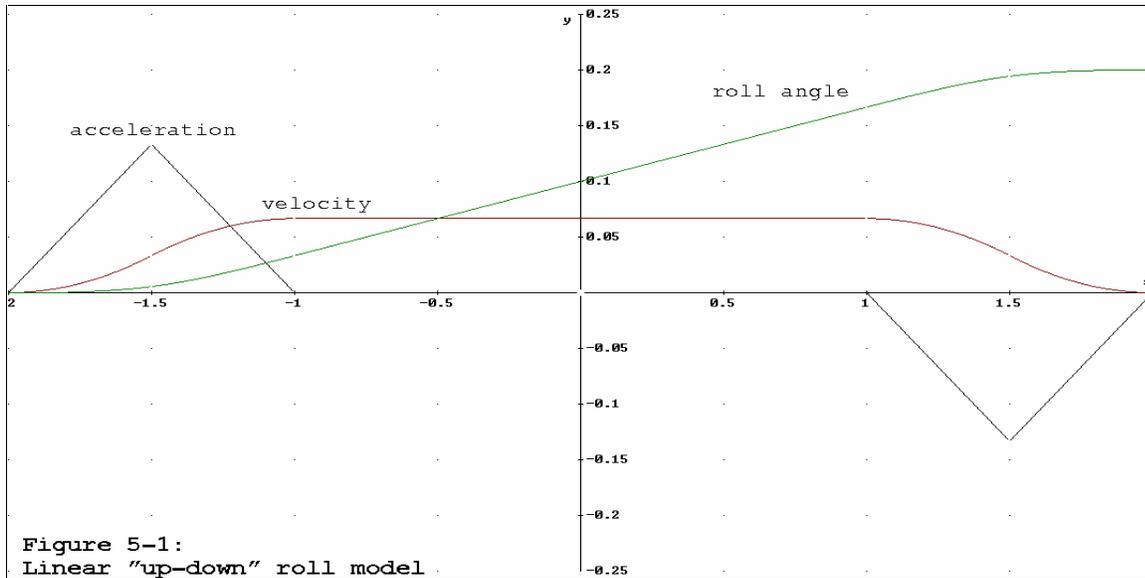
The characteristics of this model are that the roll acceleration is continuous and piecewise linear with a zone in the center of the spiral in which the roll acceleration is zero. The central zone in which the roll acceleration is identically zero is shown as going from -1.0 to $+1.0$ for illustration. However, the width of this zone is an adjustable parameter of the model. The roll acceleration in the first zone where

$$- a <= s <= - f a$$

is given by the formula

$$4 \ roll\_change \cdot (a + s) / (a^3 \cdot (1 + f) \cdot (1 - f)^2).$$

Similar expressions apply for the other three zones in which the acceleration is not zero.

For computational purposes is is convenient to use a single more general expression that gives the value of the roll acceleration throughout the length of the spiral. The general expression that applies in this case is (expressed in C program language notation):

```
   -2*roll_change
     *(  SIGN(2*fabs(s)-a*(f+1)) * (a*(f+1)*SIGN(s)-2*s)
       + SIGN(fabs(s)-a*f)          * (s-a*f*SIGN(s))
       + SIGN(fabs(s)-a)            * (s-a*SIGN(s))
      ) / (pow(a,3)*(f+1)*(pow(f,2)-2*f+1))
```

where SIGN(s) $=$ $+1$ for positive s and $= -1$ for negative s and fabs(x) is the absolute value function. Since the general expressions are lengthy and not intuitively illuminating they will not be written out for the other piecewise roll models.

5.2   _Linear up - flat - down at ends_

This model is like the linear up-down model except that each zone of non-zero acceleration is divided into three sub-zones with the roll acceleration held constant in the central sub-zone. Considering the distance of length   (a – f a)   at each end in which the roll acceleration is not identically zero, we take a central fraction   c   of   that distance, make the acceleration constant in that fraction, and on either side fit a linear up or down ramp that meets the constant acceleration to either side. This fraction   c

becomes an additional variable parameter of the model. For illustration we continue with the value   f = ½   and   adopt   c = 1/3. The plot is:

Figure 5-2:
Linear "up-flat-down" roll model
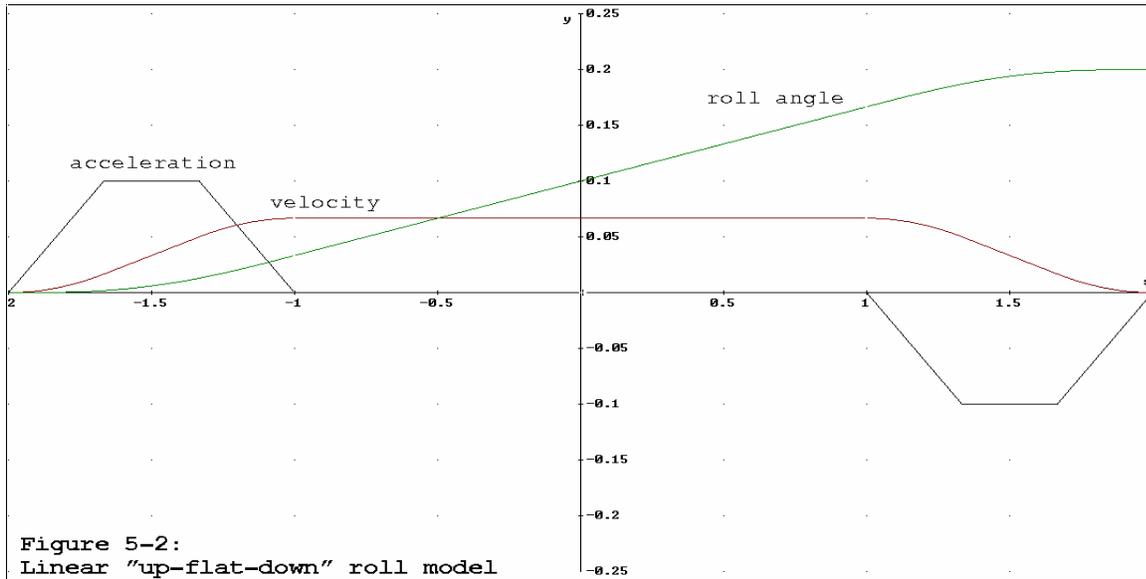
The roll acceleration in the first zone where $-a <= s <= -(1+f+c-cf) a/2$ is given by the formula

$$- 4 \text{ roll\_change} \cdot (a \text{ SIGN}(s) - s) / (a^3 \cdot (1 - c^2) \cdot (1 + f) \cdot (1 - f)^2).$$

Similar expressions can be written for the other five zones where the acceleration is not zero.

5.3   *quartic at ends*

After some preliminary exploration the author judged that it would be desirable to work primarily with roll models in which, in addition to the roll acceleration being a continuous function with zero value at each end of the spiral, its derivative, the angular jerk, would also be a continuous function with zero value at each end of the spiral. This and the following roll models all embody that feature.

This roll model is referred to here as quartic because the roll acceleration is given by a 4th order polynomial except in the central zone where it is identically zero. The plot for this model is:

Figure 5-3:
Quartic Roll Model
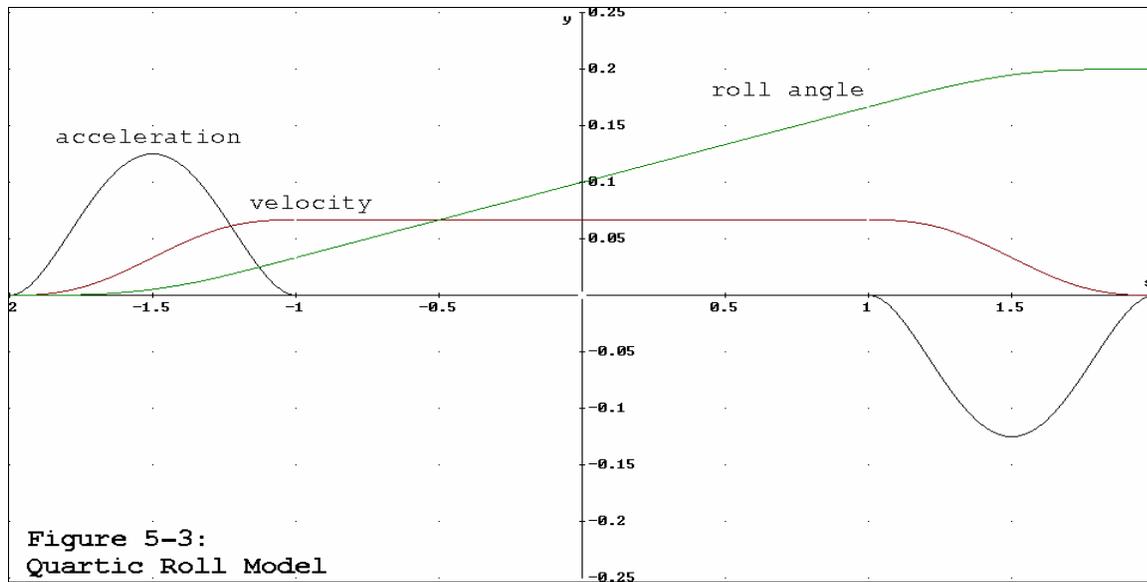
Outside of the central zone the function for the roll acceleration may be written as

$$\frac{30 \ \text{roll\_change} \cdot \text{SIGN}(s) \cdot (|s| - a)^2 \cdot (|s| - f\,a)^2}{a^6 \cdot (f + 1) \cdot (f - 1)^5}$$

where $\text{SIGN}(s) = +1$ for positive s and $= -1$ for negative s.

5.4   *quartic & flat at ends*

The preceding quartic example can be modified by analogy with the earlier linear up-flat-down example. Considering in the quartic & flat model the distance of length $(a - f\,a)$ at each end in which the roll acceleration is not identically zero, we take a central fraction c of that distance, make the acceleration constant in that fraction, and on either side fit a "compressed" half of the former quartic that connects smoothly with the constant acceleration to either side. This fraction c becomes an additional variable parameter of the model. Continuing with the values $f = ½$ and $c = 1/3$ being used for illustration, the plot is:

Figure 5-4:
Quartic & Flat Roll Model
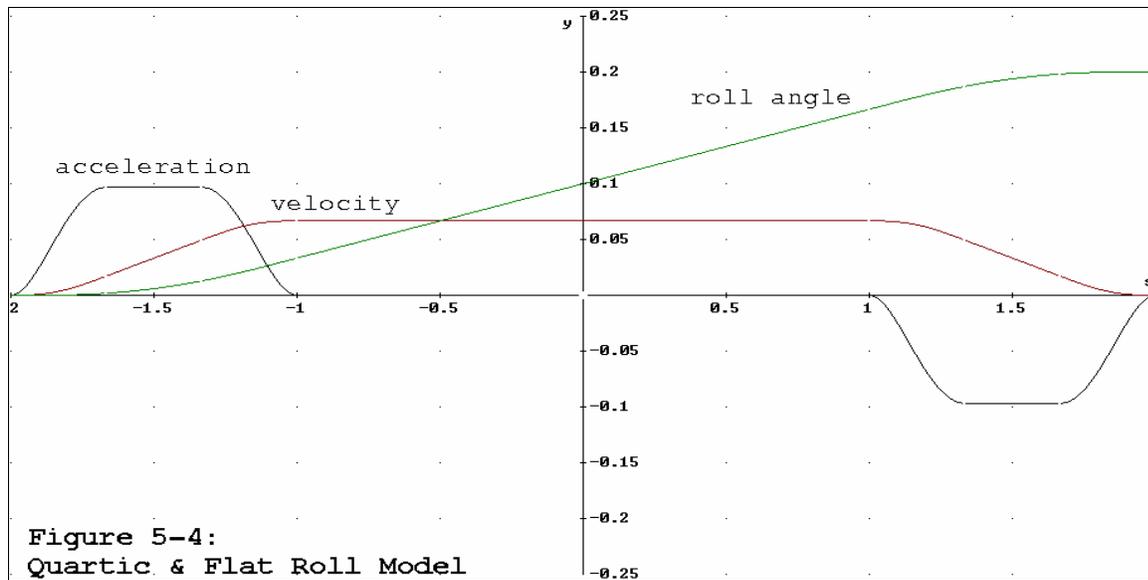
The formula for the outer 4th order segment at the left is

$$\frac{120 \ \ roll\_change \cdot (a + s)^2 \cdot (a \cdot (c \cdot (f - 1) + f + 1) + 2 \cdot s)^2}{a^6 \cdot (c - 1)^2 \cdot (f + 1) \cdot (1 - f)^5 \cdot (89 \ c^3 + 23 \ c^2 + 7 \ c + 1)}$$

and similar expressions apply in the other zones where the roll acceleration is not constant. In a previous paper (ref. 2), the author judged based on qualitative considerations that spiral roll motion is more likely to be constrained by limits on angular jerk and track warp than by a limit on magnitude of roll acceleration. If that judgment is born out, then models with zones of constant non-zero roll acceleration are not likely to be needed in practice. If it turns out that they are needed, then practitioners are likely to prefer the "elevated sine & flat" model that is given below. Its results would be close to those of this "quartic & flat" model and its formulae are more attractive.

5.5   *Hexic at ends*

This model increases the order of zero at |s| = a   and |s| = f a   from 2 to 3. The plot is

Figure 5-5:
Hexic Roll Model
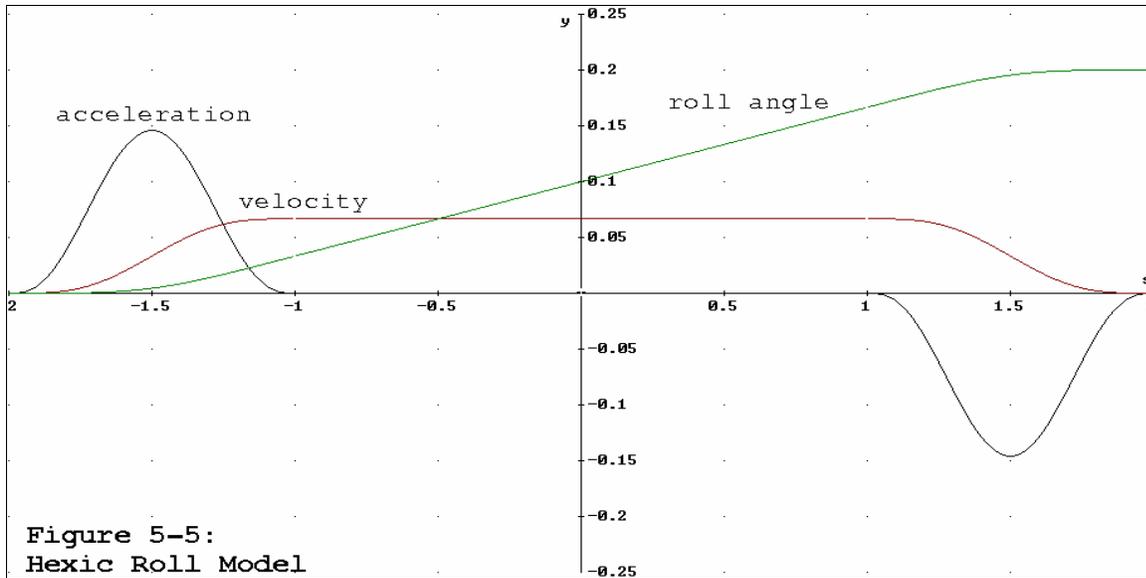
The formula for the roll acceleration outside of the middle zone is

$$\frac{-\ 140\ \text{SIGN}(s) \cdot \text{roll\_change} \cdot (\ |s| - a\ )^{3} \cdot (\ |s| - a \cdot f\ )^{3}}{a^{8} \cdot (\ 1 + f\ ) \cdot (\ 1 - f\ )^{7}}$$

5.6   *Elevated Sine at ends*

This next roll model looks and behaves very much like the preceding quartic model. However, where
its acceleration is non-zero at each end it is formed by elevating a full cycle of a sine curve. The plot is

Figure 5-6:
Elevated Sine Roll Model
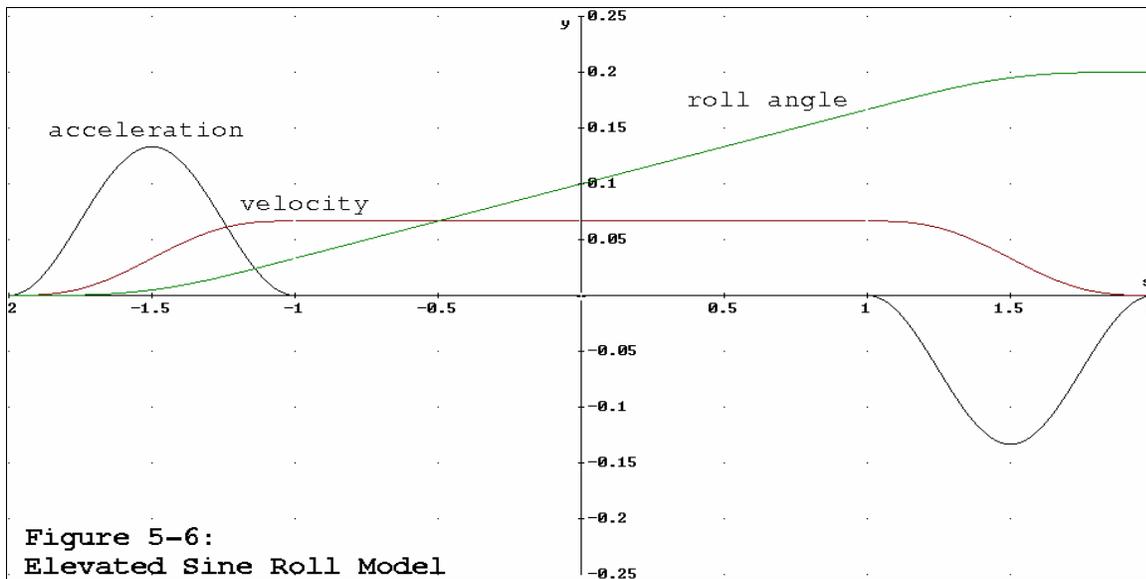
Outside of the central zone the acceleration is given by the expression

$$\frac{-\ \text{roll\_change} \cdot \text{SIGN(s)} \cdot (\text{SIN}(\ (\ 4\ \text{pi} \cdot |s| \ -\ \text{pi} \cdot a \cdot (3f + 1)\ )/(2\ a \cdot (1 - f\ ))\ ) + 1)}{a^{2} \cdot (1 - f^{2})}$$

5.7 *Elevated Sine & flat   at ends*

This is a variant of the previous model. It is analogous to the quartic & flat model. Its appearance and behavior are close to those of the quartic & flat model. The illustrative plot is

Figure 5-7:
Raised Sine & Flat Roll Model
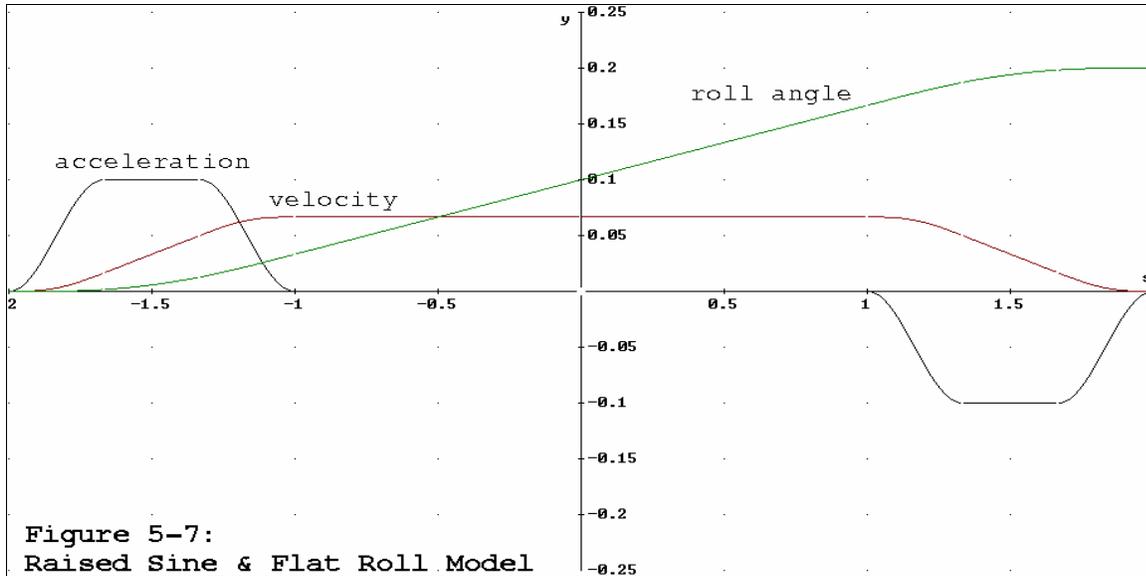
With the additional parameter   c   defined as stated for the quartic & flat model, the formula for the outer half cycle of the sine on the left where -a $<=$ s $<=$ - $(1 + f + c - c\ f\ )\ a\ /\ 2$ is:

$$\frac{-\ \text{roll\_change} \cdot (\ \text{COS}(\ 2\ \text{pi} \cdot (a - |s|\ )\ /\ (a \cdot (1 - f\ ) \cdot (1 - c))\ )\ ) - 1\ )}{a^{2} \cdot (\ 1 + c\ ) \cdot (\ 1 - f^{2})}$$

Similar expressions apply in the other three zones in which roll acceleration is not constant.

5.8 *Order (2,1)*

Each of the preceding roll models is derived from a roll acceleration function constructed with multiple zones and with the mathematical form changing from zone to zone. By way of contrast, this model and those that follow are based on respective single polynomial expressions that apply over the whole length of the spiral. This roll model is referred to as order (2,1) to indicate that the roll acceleration curve has a 2nd order zero at each end of the spiral and a 1st order zero in the center of the spiral. The models that follow are labeled analogously by the order of the zero at each end of the spiral and the order of the zero in the center of the spiral.
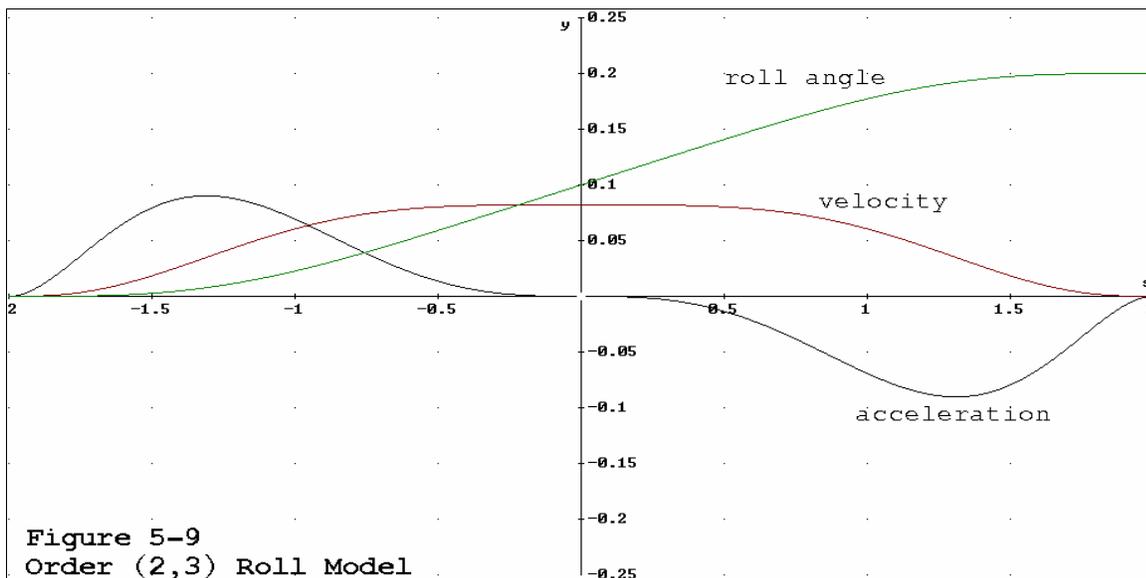
The plot for this model is

Figure 5-8:
Order (2,1) Roll Model
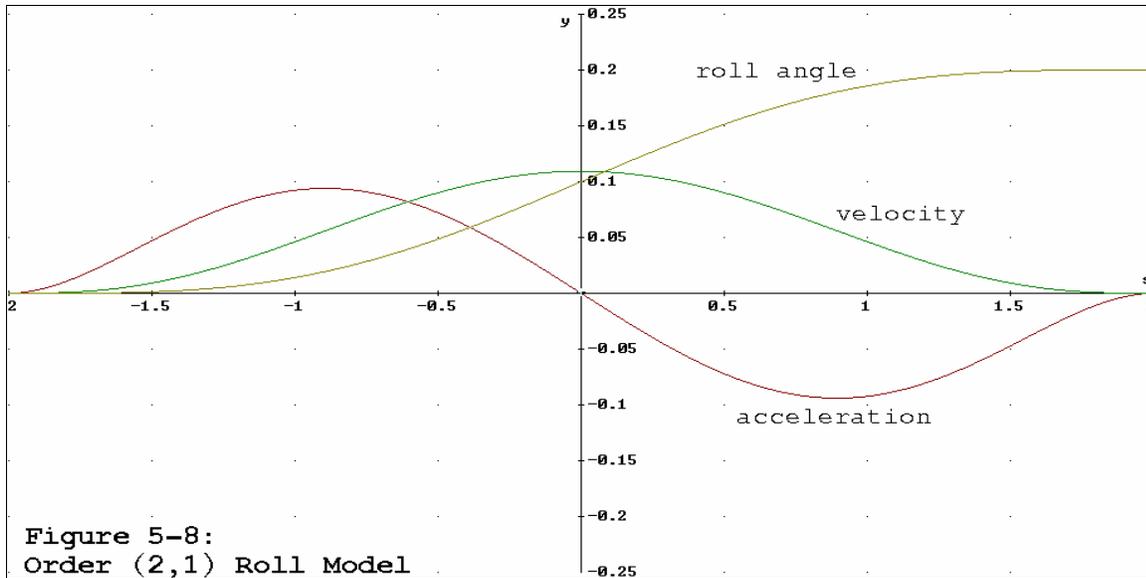
The spiral yielded by this model is similar to the Watorek spiral reviewed by Kufver. The differences are the change of the zero at each end from a 1st order zero to a 2nd order zero (to make the jerk zero at each end) and the mathematically small difference that this model is applied to the roll motion whereas the Watorek model was applied to the track curvature. The formula for the roll acceleration is:

$$\frac{-\ 105\ \ \text{roll\_change} \cdot (a + s)^2 \cdot (a - s)^2 \cdot s}{16\ a^7}$$

5.9   Order (2,3)

Compared to the preceding order (2,1) model, this model has the zero at the center flattened out so that the roll acceleration is shifted somewhat from the center toward the ends. The plot for this model is
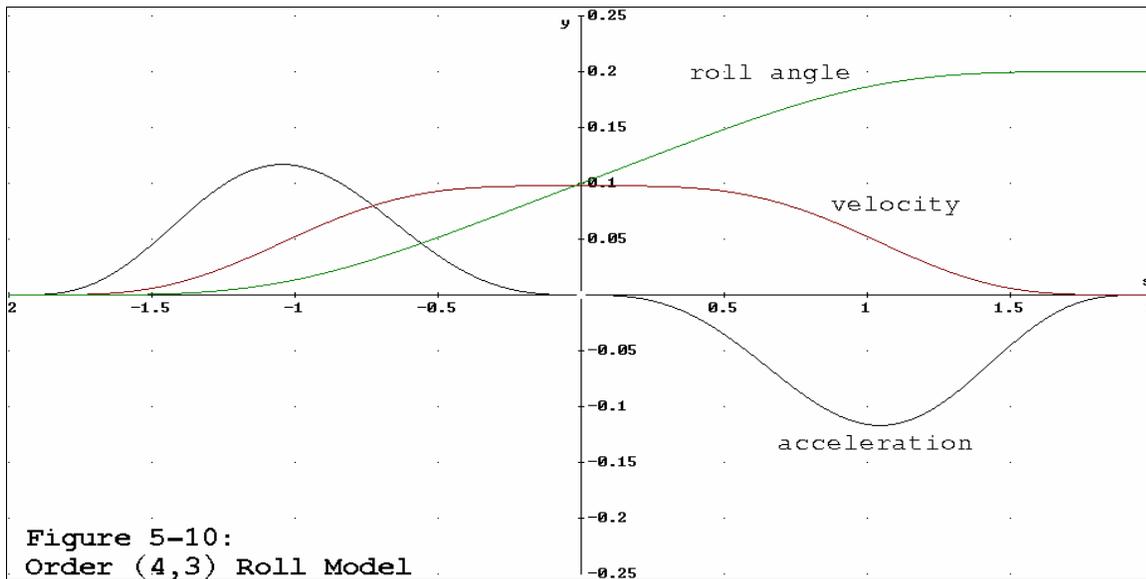

Figure 5-9
Order (2,3) Roll Model

The expression for the roll acceleration is

$$\frac{-\ 315\ \text{roll\_change}\cdot(a+s)^2\cdot(a-s)^2\cdot s^3}{16\ a^9}$$

5.10 *Order (4,3)*

This model has roll acceleration with a 4th order zero at each end and a 3rd order zero in the center. The plot is:


Figure 5-10:
Order (4,3) Roll Model

Compared to order (2,3), this model builds and drops roll acceleration more gently and symmetrically but will require larger curve offsets and spiral lengths for given limits on track warp.

The formula for the roll acceleration in this model is

$$\frac{-\ 15015\ \text{roll\_change}\cdot(a+s)^4\cdot(a-s)^4\cdot s^3}{256\ a^{13}}$$

5.11 *Order (2,5)*

This roll model is like the order (2,3) model except that the acceleration is further flattened near the center by raising the order of the zero there from 3rd order to 5th order. The motive for pushing the acceleration away from the center and toward the ends is that, other things being equal, this reduces the track warp in the middle of the spiral. The plot is
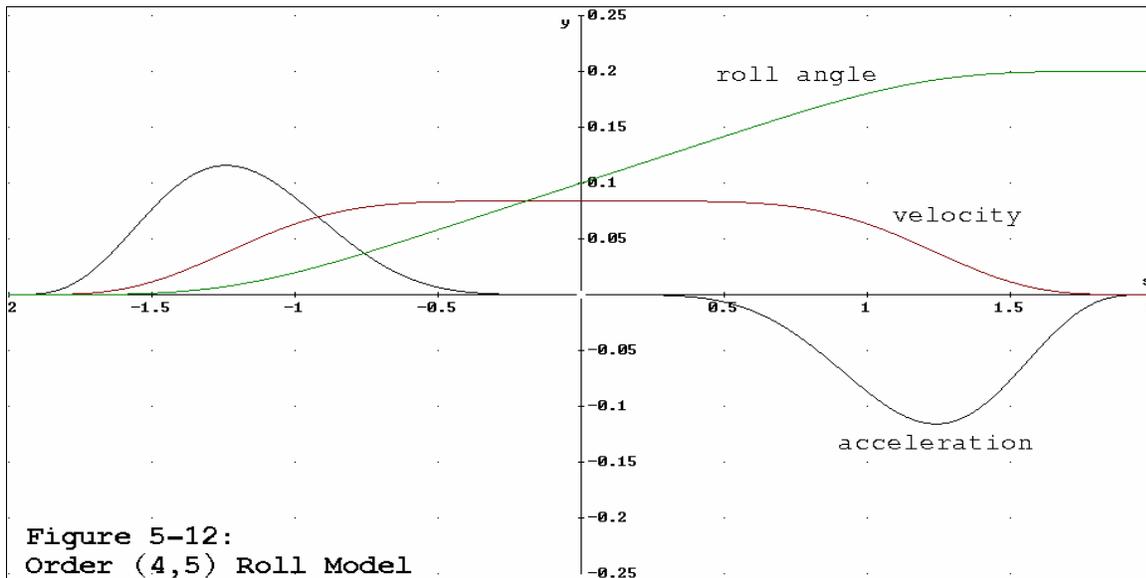
Figure 5-11:
Order (2,5) Roll Model
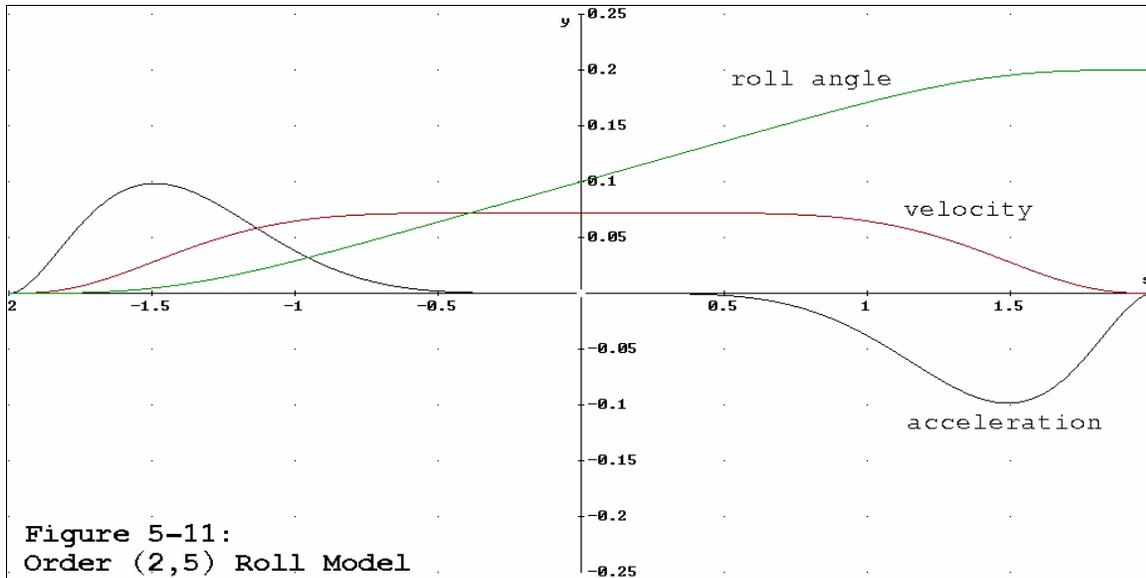
and the expression for the roll acceleration is

$$\frac{-\ 693\ \text{roll\_change} \cdot (a+s)^2 \cdot (a-s)^2 \cdot s^5}{16\ a^{11}}$$

5.12 Order (4,5)

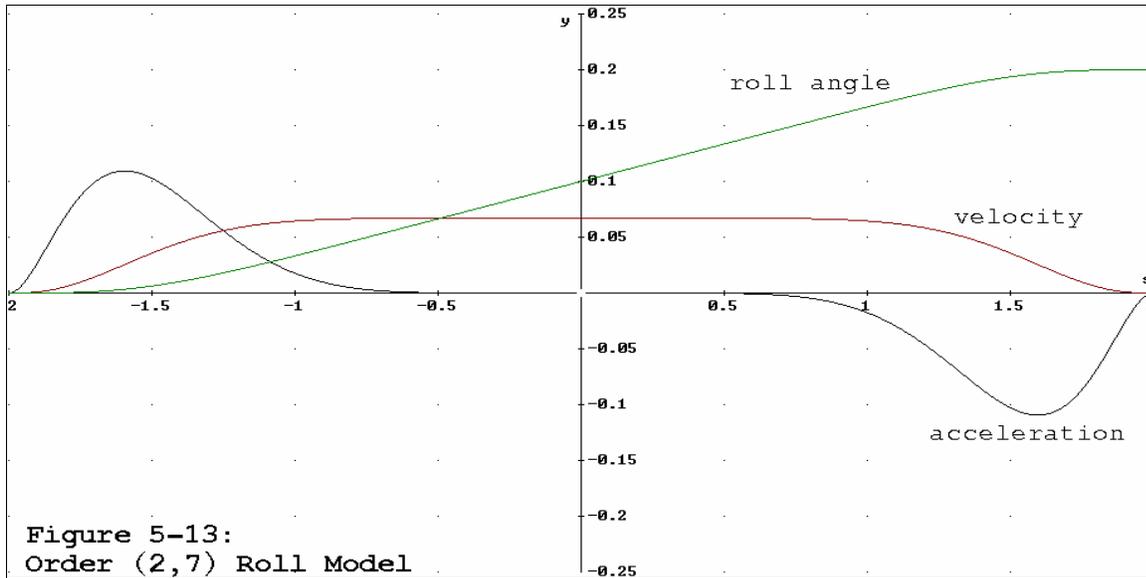The plot for this model is

Figure 5-12:
Order (4,5) Roll Model

The formula for roll acceleration in this model is

$$\frac{-\ 45045\ \text{roll\_change} \cdot (a+s)^4 \cdot (a-s)^4 \cdot s^5}{256\ a^{15}}$$

5.13 *Order (2,7)*

The plot for this model is

Figure 5-13:
Order (2,7) Roll Model

The formula for the roll acceleration for this model is

$$-\ \frac{1287\ \text{roll\_change}\ \cdot (a+s)^2 \cdot (a-s)^2 \cdot s^7}{16\ a^{13}}$$

5.14 *Order (3,7)*

The continuous polynomial roll acceleration models with angular jerk zero at each end of the spiral and continuous throughout that have been displayed so far have a 2nd or 4th order zero at each end. The model presented in this section has a 3rd order zero at each end. The plot (showing just the left half in light of the symmetries) is:
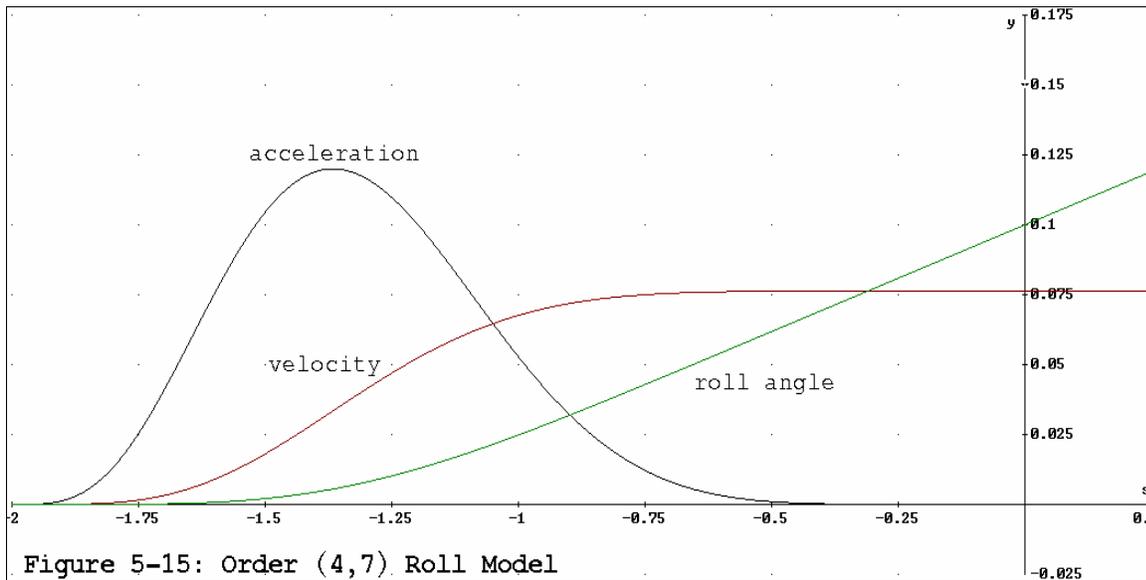
Figure 5-14: Order (3,7) Roll Model
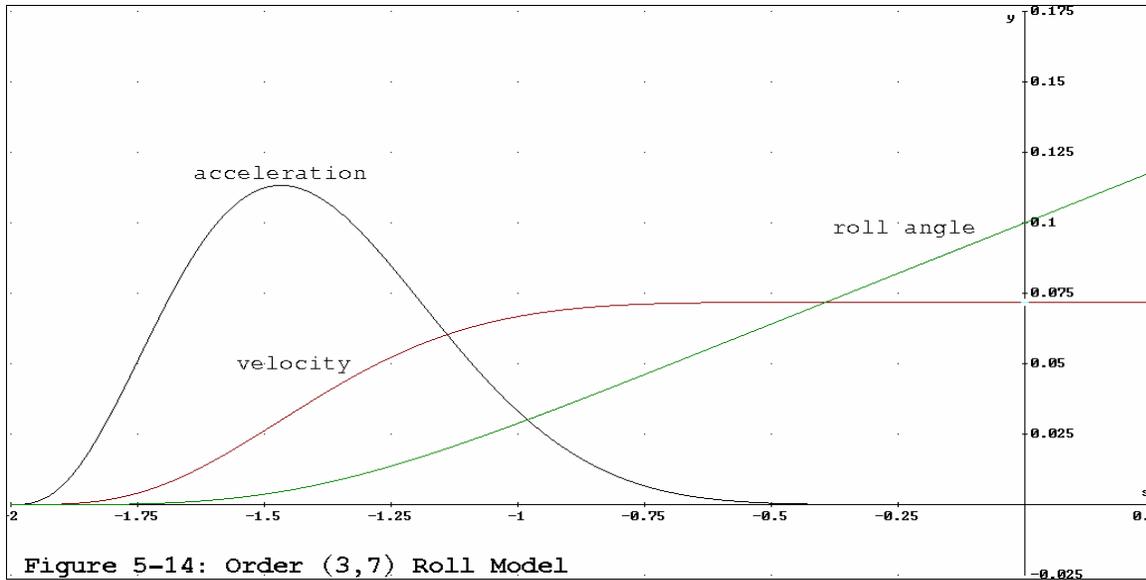
and the formula for the roll acceleration is:

$$\frac{6435 \ \text{roll\_change} \cdot (a + s)^3 \cdot (s - a)^3 \cdot s^7}{32 \ a^{15}}$$

5.15 *Order (4,7)*

The plot is

Figure 5-15: Order (4,7) Roll Model

and the formula for the roll acceleration is

$$\frac{-\ 109395 \ \text{roll\_change} \cdot (a + s)^4 \cdot (a - s)^4 \cdot s^7}{256 \ a^{17}}$$

5.16   *Comparison of the roll model shapes*

The next four plots show features of selected roll models together for comparison.

Figure 5-16 shows how selected models compare with respect to rate of rise of roll acceleration in the first quarter of the spiral's length.

Figure 5-17 shows the full shapes of the roll accelerations for the continuous polynomial models and includes the piecewise elevated sine model for comparison.

Figure 5-18 shows how the shape of the roll acceleration curve for one of the piecewise models varies with variation of the parameter   f   that specifies the width of the central zone in which roll acceleration is identically zero.

In these figures different curves have different maximum acceleration values because all the curves are scaled to correspond to the same value for the change in roll angle between the ends of the spiral, namely 0.2 radians.
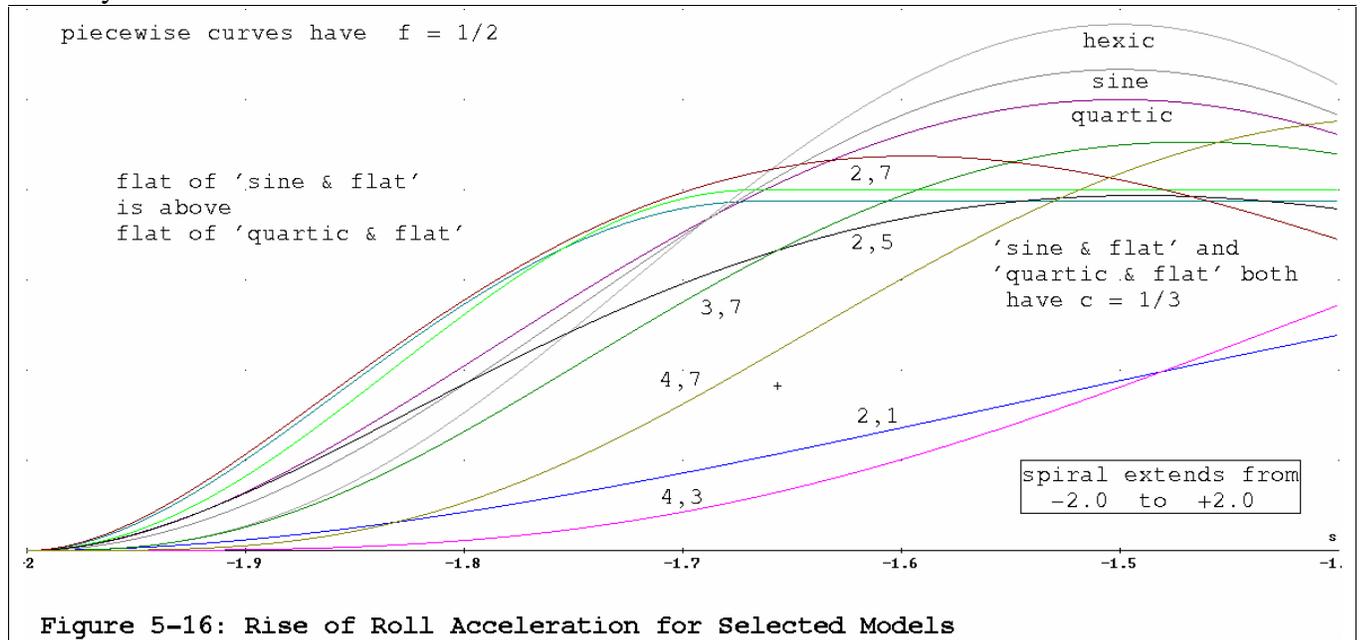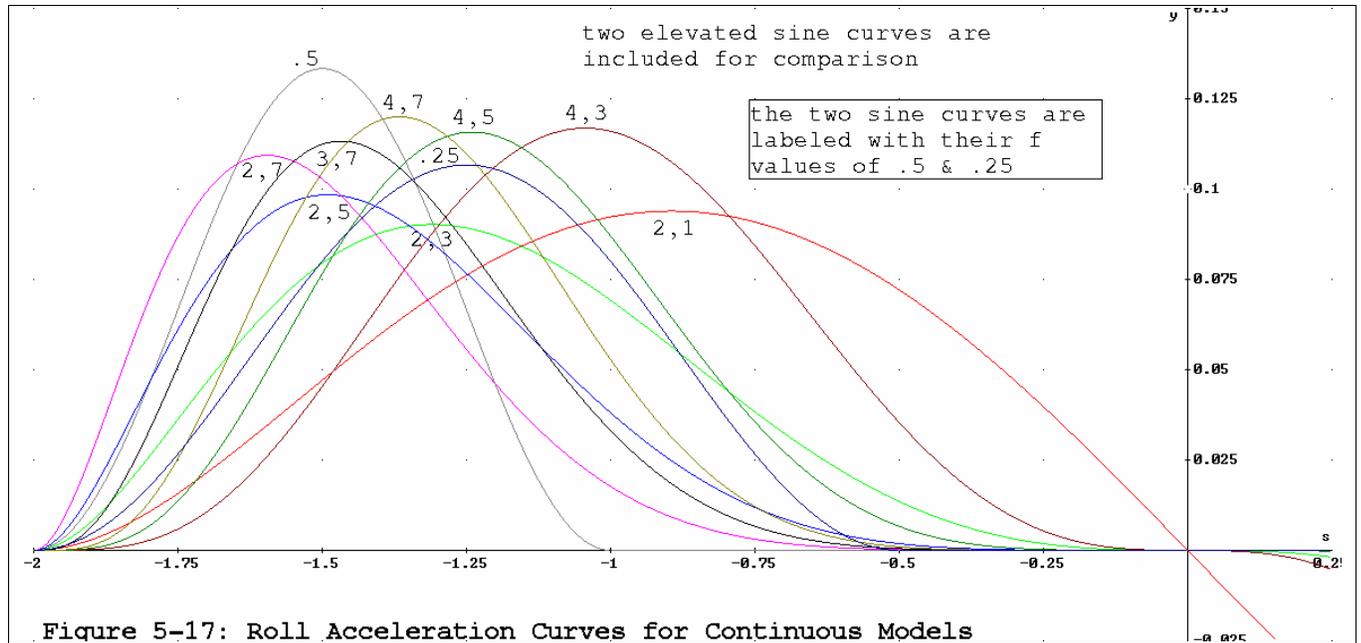

Figure 5-16: Rise of Roll Acceleration for Selected Models

two elevated sine curves are
included for comparison

the two sine curves are
labeled with their f
values of .5 & .25

Figure 5-17: Roll Acceleration Curves for Continuous Models

Consider two or more continuous polynomial roll models that have the same spiral length and accomplish the same roll_change. If the roll acceleration functions of those models are added together after being multiplied by respective weighting factors whose sum = 1.0, the sum will be the roll acceleration function of a new roll model that will have the same spiral length and roll_change. The same is true for the roll velocity and roll angle functions. As an example, denoting the roll angle functions for the order (4,7) and order (4,3) models by ra47(s) and ra43(s) respectively and letting p be an adjustable parameter that varies between 0 and 1, a new roll model with a single adjustable parameter may be defined based on the roll angle function ra_new(s) = p ra47(s) + (1 – p) ra43(s). Varying parameter p would shift the character of the combination spiral between the greater efficiency of the order (4,7) model and the greater gentleness of the order (4,3) spiral in much the same way that varying the parameter f varies the character of the piecewise raised sine spiral. This idea is not explored any further in the present paper, but it could turn out to be useful when standard practices are developed based on the methods of this paper.

It should be kept in mind that through variation of parameter f, the quartic, elevated sine, and hexic models can be made somewhat similar to any one of the continuous polynomial models. This flexibility of the piecewise models is illustrated for the case of the elevated sine model in Figure 5-18.

Figure 5-18: Sine Model Roll Acceleration curves for sample values of f.
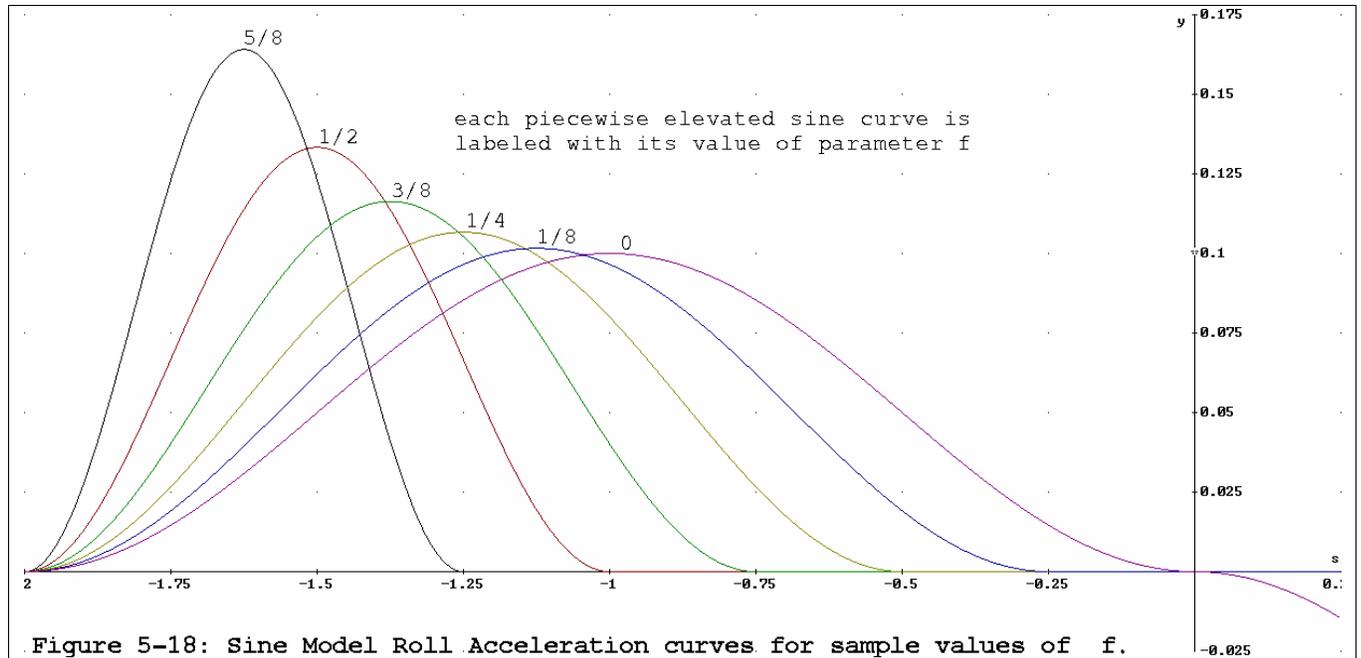
After reviewing the shapes of the roll acceleration curves for these candidate roll models one can look back at the previously proposed alternatives to the clothoid spiral as listed by Kufver and illustrated in table 3-1 above. Recall that the curve shown in table 3-1 for the 2nd derivative of curvature of each prior proposal is approximately proportional to the roll acceleration in the corresponding spiral. As noted in the comments following that table, in each of the first five prior proposals as listed by Kufver, the roll acceleration has a large discontinuity at the beginning and end of the spiral and so incorporates conceptually infinite angular jerks. In the present author's opinion, if prior investigators had had the perspective that a primary objective of a spiral is to rotate vehicles intelligently, then of the 7 prior proposals that are illustrated, only the last two, the Watorek and the "sinusoidal" would ever have been proposed. Those two could have been included among the candidates considered in this paper. The reason they are not is that the present author has found a motive, which will be explained shortly, to prefer forms of roll acceleration in which the angular jerk is zero at the ends of the spiral and continuous throughout.

In the next plot seven continuous roll functions are shown together along with the piecewise quartic and elevated sine roll functions. (The roll functions are all antisymmetric about the point (0.0, 0.1) so that showing the left half is sufficient.)

Figure 5-19: Roll Angle vs. Distance for selected Models
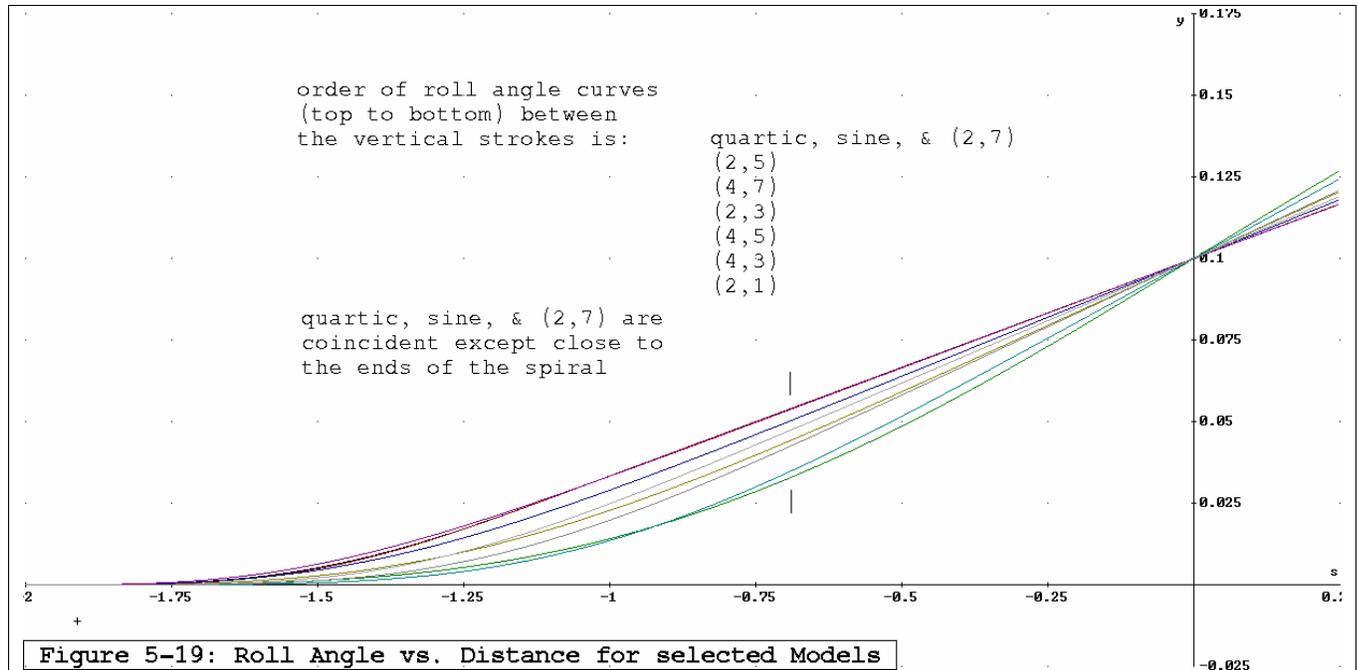
6   Raising the roll axis height above the rails

As has already been noted and as emphasized by Presle and Hasslinger (ref. 6), the forces required to change a vehicle's roll angle will be minimized if the longitudinal axis of rotation passes through the center of gravity of the vehicle.

Thinking in terms of passenger vehicles and passenger comfort, one would like to have the vehicle roll while traversing a spiral take place about an axis at an elevation somewhere near the shoulders of the average seated passenger. Such an axis would probably be somewhat higher than the vehicle center of gravity.

When we look later in this paper at the way that the proposed spirals change with change of roll axis height we will find a third benefit of having the roll axis height above the plane of the rails. That is that for given curve offset, raising the roll axis allows the spiral to become a good deal longer and thereby substantially reduces the maximum track warp and maximum roll acceleration in the spiral. This is expected to be of practical significance when existing tracks need to be upgraded for higher speeds.

7   Computing the spiral geometry that embodies a roll motion model

The procedure for calculating the spiral geometry yielded by a given roll model is as follows:

a) Choose an initial length for the spiral. If the roll model being used has a central section of zero roll acceleration whose width is adjustable, select a value for that width. Together with the known values of roll angle at the two ends of the spiral, these choices establish the roll angle function, r (s), as a definite function of the path length, s, along the spiral.

b) Integrate the balance equation
$$db/ds = (g / v_b^2) \tan (r (s)) \qquad (7\text{-}1)$$

where g is the acceleration of gravity, v_b is the balancing speed of the spiral, and the curvature db/ds is the derivative of the bearing angle of the track with respect to path length. This will yield the bearing angle b(s) as a function of path length.

c) Integrate the pair of equations

$$dx/ds = \cos (b(s)) \qquad (7\text{-}2)$$

and

$$dy/ds = \sin (b(s)) \qquad (7\text{-}3)$$

to obtain the x and y coordinates of points along the spiral.

As mentioned previously, the difference between tan(r(s)) and r(s) in equation (7-1) above could be ignored without very much error. If that approximation were introduced, then the integration of step b) could be done in closed form for each of the roll models listed above. However, in the roll models that appear most attractive the expressions for r(s) are either transcendental or else polynomial of degree higher than 2. For these models the integrations called for in this step to obtain the x and y coordinates of points on the spiral cannot be done in closed form even if the foregoing approximation were introduced. It is therefore necessary to employ numerical integration in this step, and, since it must be used here, it may as well be used in step b) as well. With personal computing resources what they are today, the use of numerical integration does not seem to be a handicap.

d) Based on the vehicle roll angle and the height selected for the vehicle roll axis locate the track centerline relative to the roll axis at each end of the spiral.

e) Compute the offset that would exist between the two curves (or tangent and curve) being joined based on the computed shape of the spiral. If the spiral is being designed to fit an existing pair of curves whose offset is already established, repeat steps a) through e) adjusting the length the length of the spiral to find the length that gives the required offset.

f) If the roll model being used has a central zone with zero roll acceleration whose length is adjustable, the length of that zone can be varied to find a good balance between the competing characteristics of spiral length and maximum track warp on the one hand and maximum roll acceleration on the other.

8   <u>Improved spirals derived from the candidate roll models</u>

This section provides comparative spiral characteristics and plots illustrating the fitting of spirals designed as proposed herein to reverse curves 308 & 309 on the NE Corridor south of Philadelphia. Attention is limited here to this one location to avoid excessive proliferation of variables and because reverse curves call for large roll angle changes and yield nice looking plots.

The fifth tabulated case, labeled W308_0_0_6 and illustrated in the first plot, has the vehicle roll axis at top-of-rail height so that some geometrical consequences of roll axis height can be observed. All the other cases have the vehicle roll axis raised 7.0 ft. above top-of-rail. The 7.0 ft. height was chosen partly with passenger comfort in mind and partly to emphasize the effect of roll axis height. Future simulations and testing may to show that dynamic behavior for some mix of vehicle types and loadings is better with some different height.

The following table shows how some basic spiral features vary from model to model and as the width of the central zero-acceleration zone is varied in the case of models that include that degree of freedom. It also indicates which cases are illustrated by figures. In the labeling, W308 indicates reverse curves 308 & 309 with offset = 2.031 ft. The three cases labeled W308os3 are for the same two reverse curves but with the offset between them increased hypothetically to 3.0 ft. to illustrate the improved dynamic performance that can be obtained when curve offset is more adequate.

Table 8-1: Spiral Characteristics for selected Roll Models and parameter values

Model	Width of center zone as % of whole	Designation	Maximum value of		Figure	Spiral length (ft.)
			Warp (inch rise in 62 ft)	Roll accel (rad/sec^2)		
Up-down	0	W308_7_0_0	1.86	0.045		775
	20	W308_7_0_2	1.68	0.055	8-3	716
	40	W308_7_0_4	1.59	0.076		649
	60	W308_7_0_6	1.54	0.124	8-2	585
	60	W308_0_0_6, axis height = 0.0	1.99	0.206	8-1	453
quartic	0	W308_7_2_0	1.82	0.040		795
	20	W308_7_2_2	1.66	0.050	8-4	725
	20	W308os3_7_2_2	1.46	0.039	8-15	823
	40	W308_7_2_4	1.58	0.071	8-5	653
	60	W308_7_2_6	1.54	0.115	8-6	586
Raised sine	0	W308_7_3_0	1.79	0.042		805
	20	W308_7_3_2	1.65	0.053		730
	40	W308_7_3_4	1.57	0.075		655
	60	W308_7_3_6	1.54	0.123		587
order (2,7)		W308_7_27	1.58	0.079	8-7	613
order (3,7)		W308_7_37	1.59	0.072	8-8	653
order (4,7)		W308_7_47	1.60	0.068	8-9	690
order (2,5)		W308_7_25	1.61	0.064	8-10	647
order (3,5)		W308_7_35	1.63	0.059	8-11	695
order (4,5)		W308_7_45	1.64	0.057	8-12	740
		W308os3_7_45	1.44	0.044	8-16	839
order (2,3)		W308_7_23	1.68	0.049	8-13	704
		W308os3_7_23	1.48	0.038	8-17	799
order (3,3)		W308_7_33	1.70	0.048	8-14	766
order (4,3)		W308_7_43	1.72	0.047		822
order (2,1)		W308_7_21	1.92	0.038		822

The author began exploration of the spiral design approach advocated herein by looking at spirals obtained from the "linear up-down" roll model. Figure 8-1 illustrates that with the vehicle roll axis located in the plane of the rails (i.e., roll axis height = 0.0), the resulting spiral has track curvature that varies smoothly with distance along the spiral. However, as seen in Figure 8-2, with the roll axis height raised to 7.0 ft the track curvature acquires a triangular "jog" at each end of the spiral and the derivative of the track curvature with respect to distance becomes discontinuous. Each "jog" mirrors the shape of the roll acceleration of the "linear up-down" model and arises as a matter of geometry from the combination of the roll acceleration and the elevation of the roll axis. (In reference 2 the author failed to include this effect, and the plots labeled as track curvature in that paper are actually showing the curvature of the path followed by the vehicle roll axis.) Discontinuity in the slope of track curvature like that illustrated in Figure 8-2 can be considered unaesthetic.

One would like to know whether discontinuity in the slope of the track curvature has an adverse effect on vehicle motion. There is a discontinuity in the slope of the track curvature at each end of a clothoid spiral. However, the deficiency of the clothoid spiral is considered to come not from its track curvature per se but rather from the manner in which it mishandles vehicle roll motion. Some information that appears to bear on this question has been provided to the author by Steven Chrismer of LTK Engineering Services (private communication). Chrismer used the NUCARS (ref. 7) rail vehicle motion simulation program to simulate the movement of a freight locomotive over a spiral with geometry like that of Figure 8-2. The program predicted that the "jogs" in the track curvature would excite substantial "hunting" oscillation on the part of the locomotive trucks. The author's first reaction to this problem was to reduce the width of the model's central zone of zero roll acceleration so as to spread the roll acceleration out and reduce the angular jerk of the model. The resulting spiral was like that illustrated in Figure 8-3. In that figure the discontinuities in the slope of the track curvature are barely noticeable, and for that spiral the predicted amplitude of truck "hunting" was small enough to have no visible effect on vehicle ride motion. However, spreading roll acceleration away from the ends of the spiral and toward the center is not an efficient way to solve the problem because doing so causes an increase in the track warp.

Chrismer's information suggested that rail vehicle truck rolling motion could be adversely affected by any abrupt change in the slope of the track curvature. In light of that information the author decided to look for roll models that would yield spirals whose track curvatures would have continuous first derivatives even with the vehicle roll axis raised above the plain of the track. Mathematically that meant looking for roll motion models with angular jerk equal to zero at each end and continuous throughout. This was the motivation for developing the "quartic", "raised sine", and "hexic" piecewise roll models and the order (m,n) continuous roll models set forth in section 5 above.

Note in figures 8-2 through 8-15 that at the beginning of the spiral, the initial change of curvature is in the direction opposite to the change accomplished by the spiral as a whole. This is a characteristic feature of spirals with raised roll axis and a feature to which Presle and Hasslinger (ref. 6) have called attention. It corresponds to the manner in which the rider of a motorcycle steers when approaching a turn. The beginning of the steering maneuver serves to set up the bank angle for the turn as a whole.

The "quartic" and "raised sine" roll models yield very similar spirals. Examples are plotted for "quartic" model spirals, but "raised sine" model spirals would be equally useful in practice and might be preferred.

The rest of this section consists of the sample plots. Some further comments are offered in the next section.
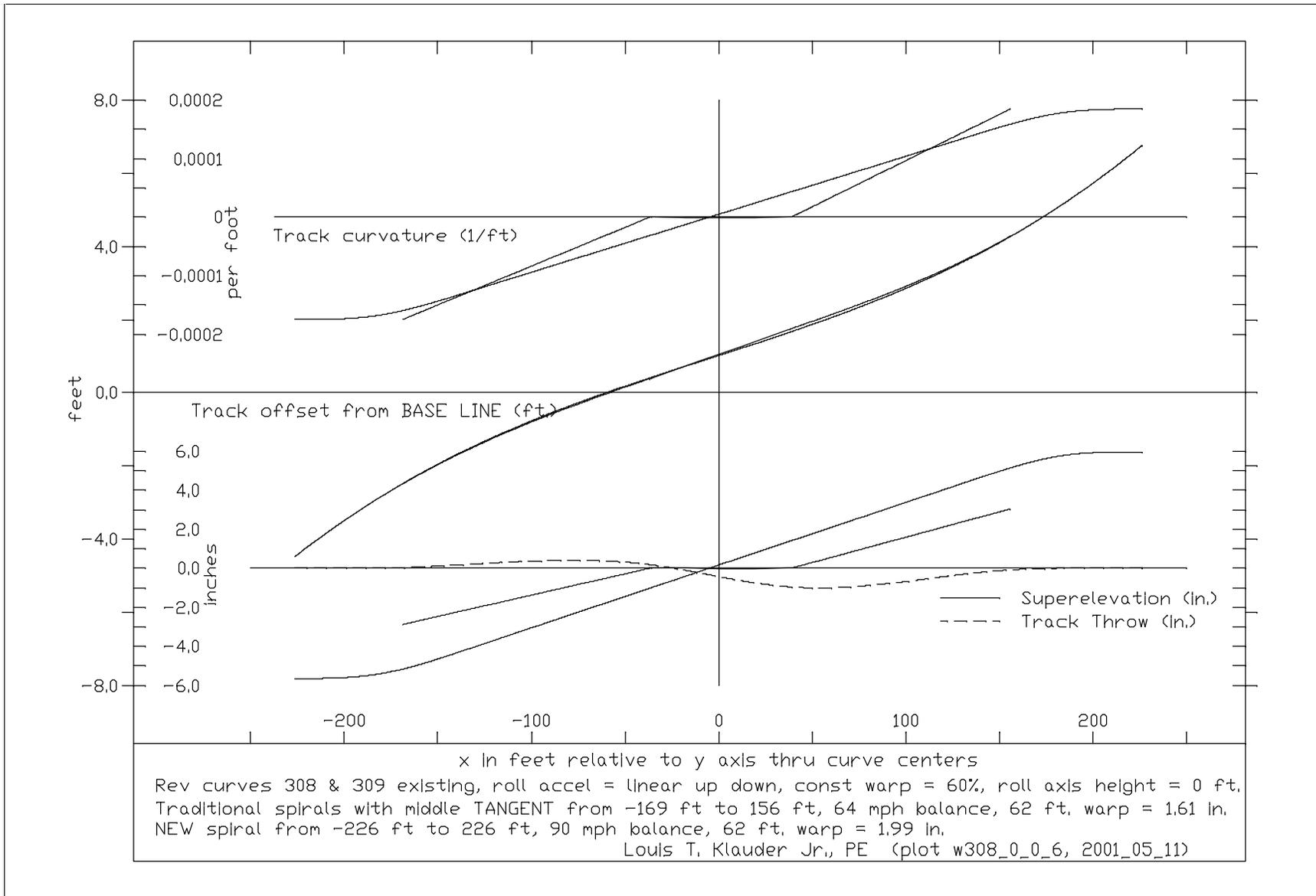
Figure 8-1: Plot W308_0_0_6: Roll axis height = 0 ft; Roll model = "Linear Up-Down"; Width of central 0.0 acceleration zone = 60%
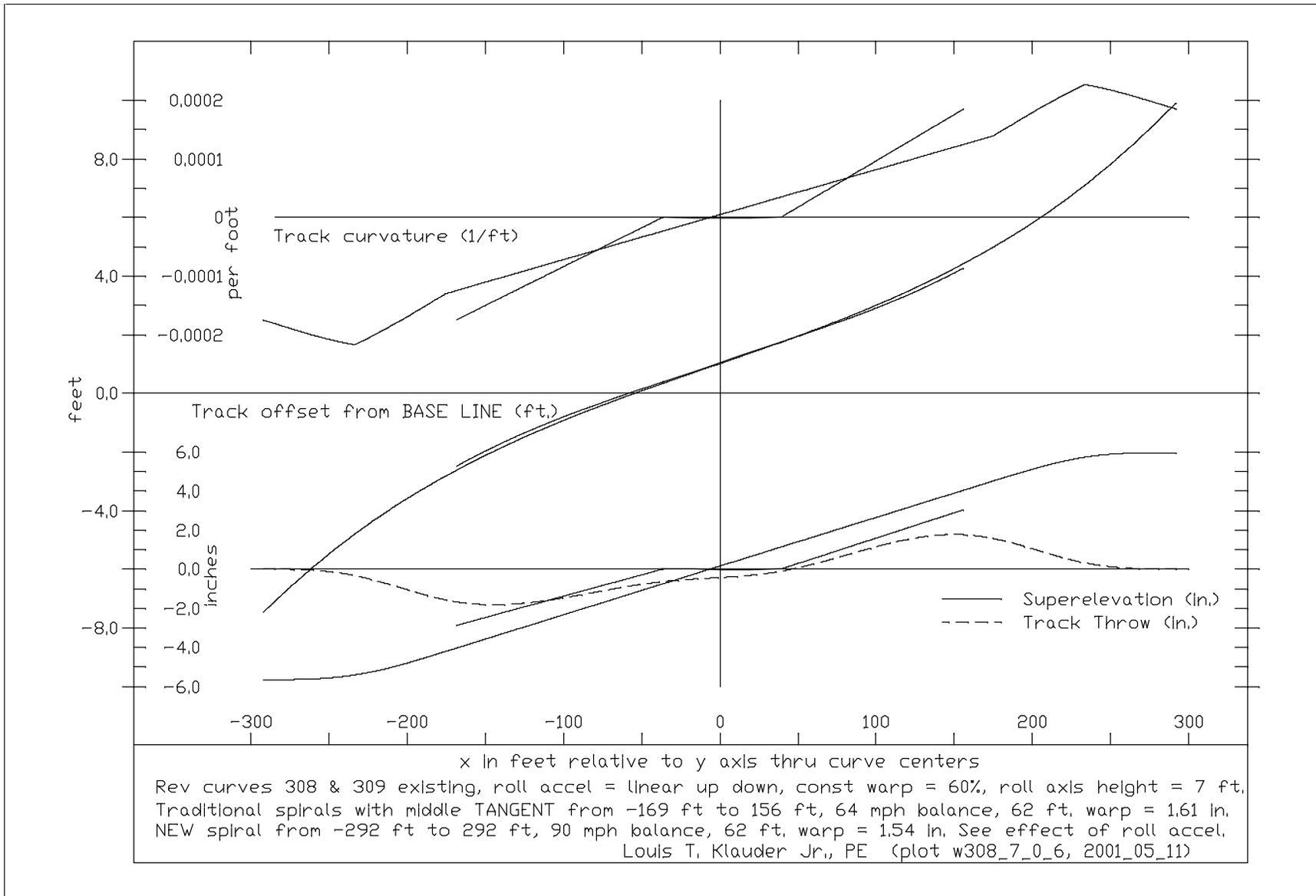
Figure 8-2: Plot W308_7_0_6: Roll axis height = 7 ft; Roll model = "Linear Up-Down"; Width of central 0.0 acceleration zone = 60%
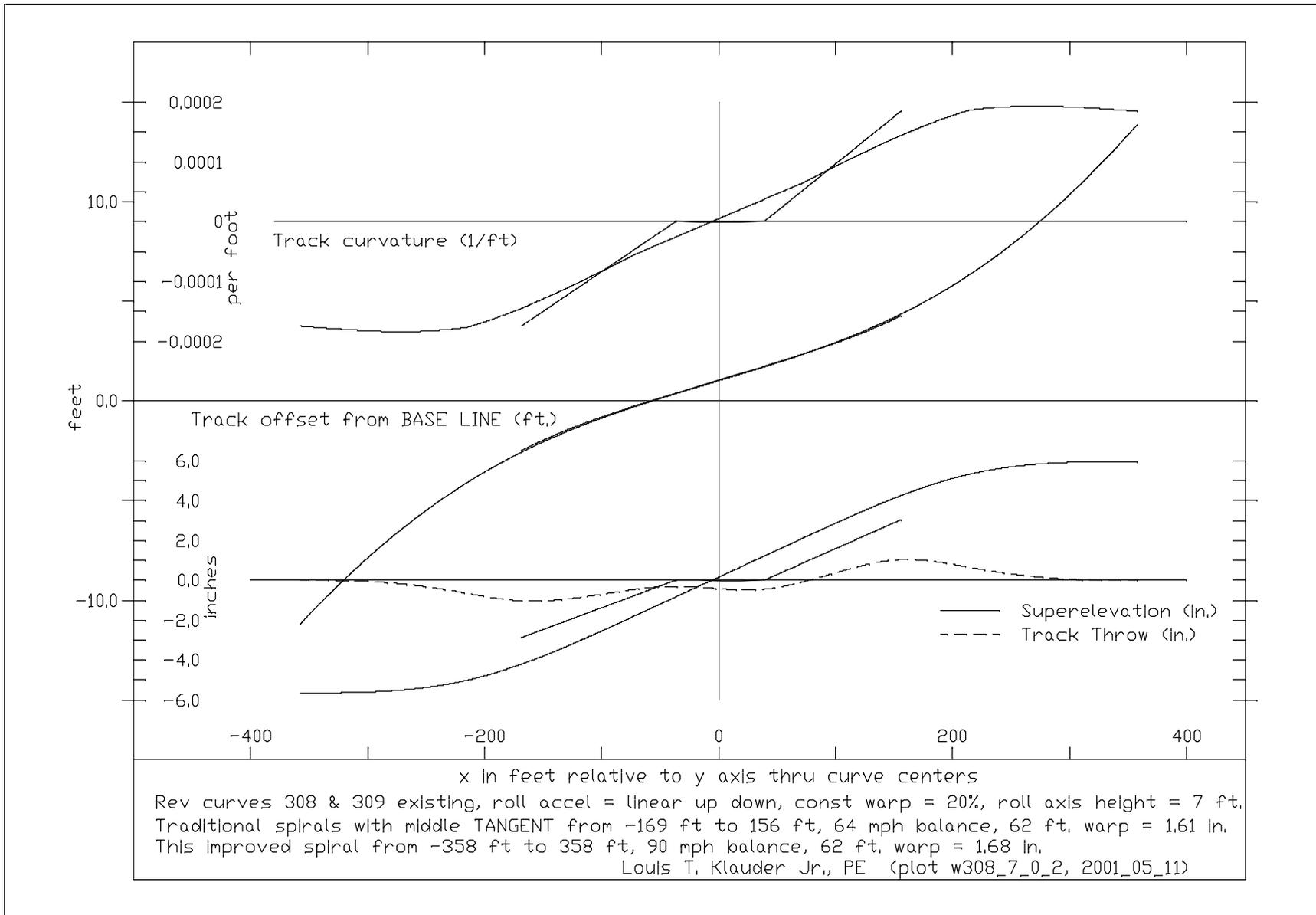
Figure 8-3: Plot W308_7_0_2: Roll axis height = 7 ft; Roll model = "Linear Up-Down"; Width of central 0.0 acceleration zone = 20%
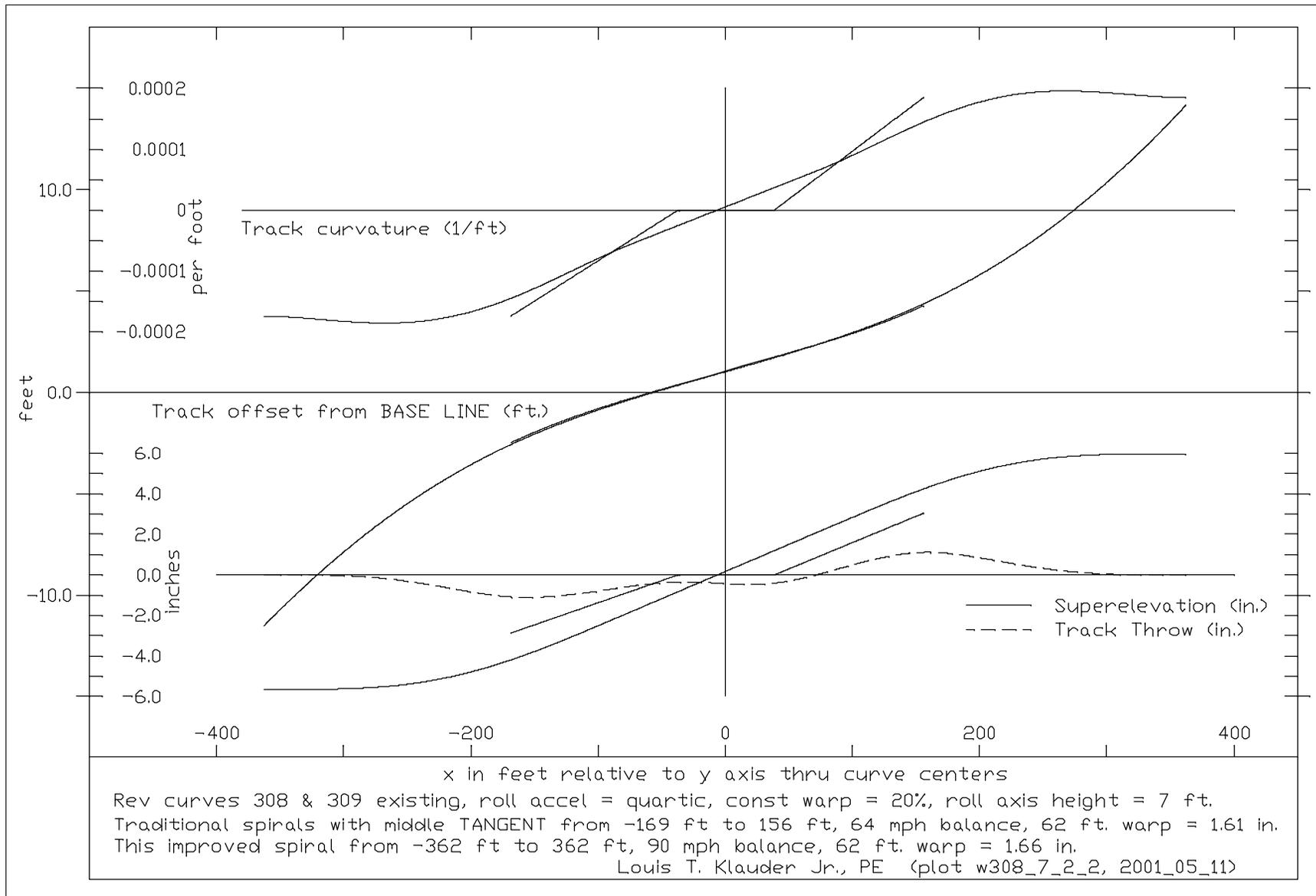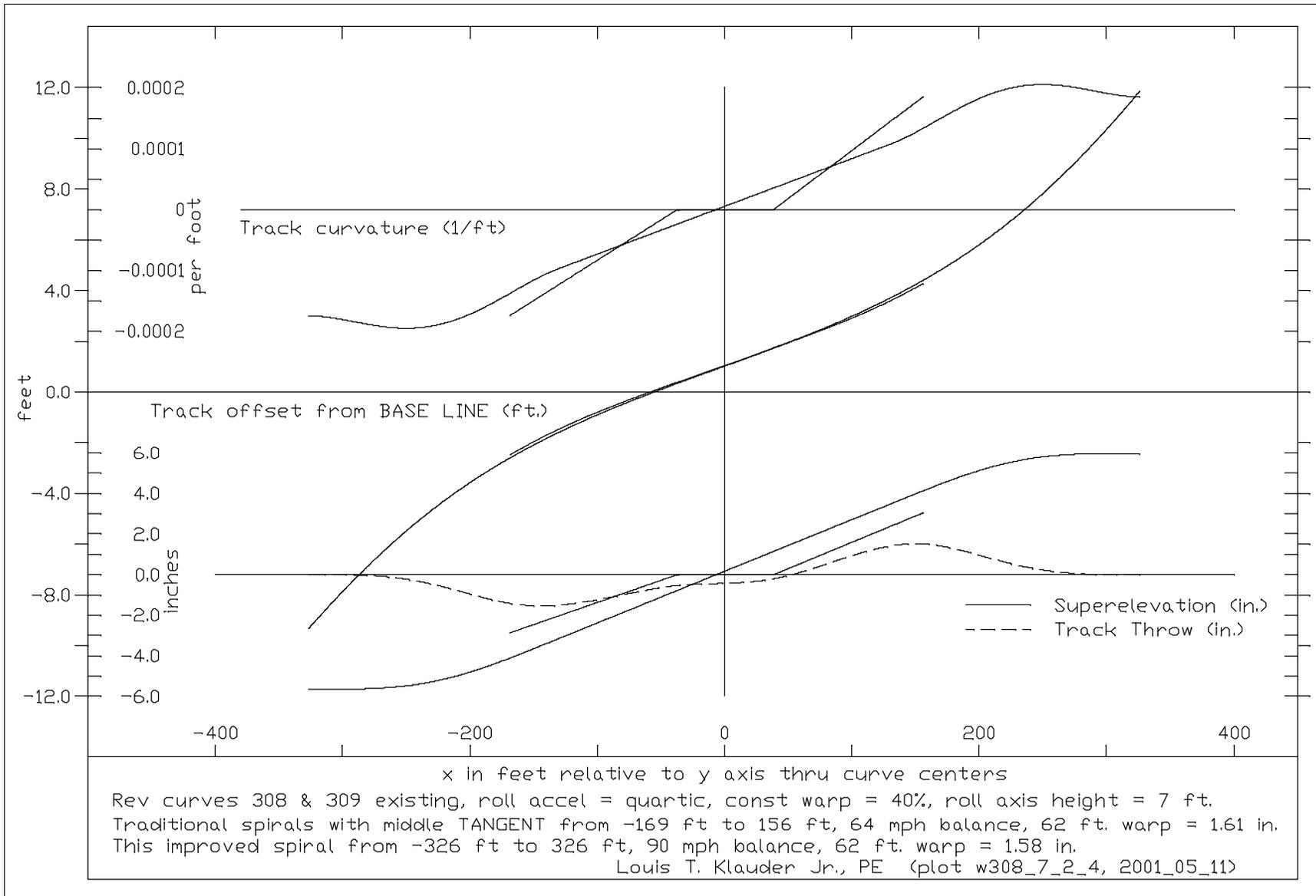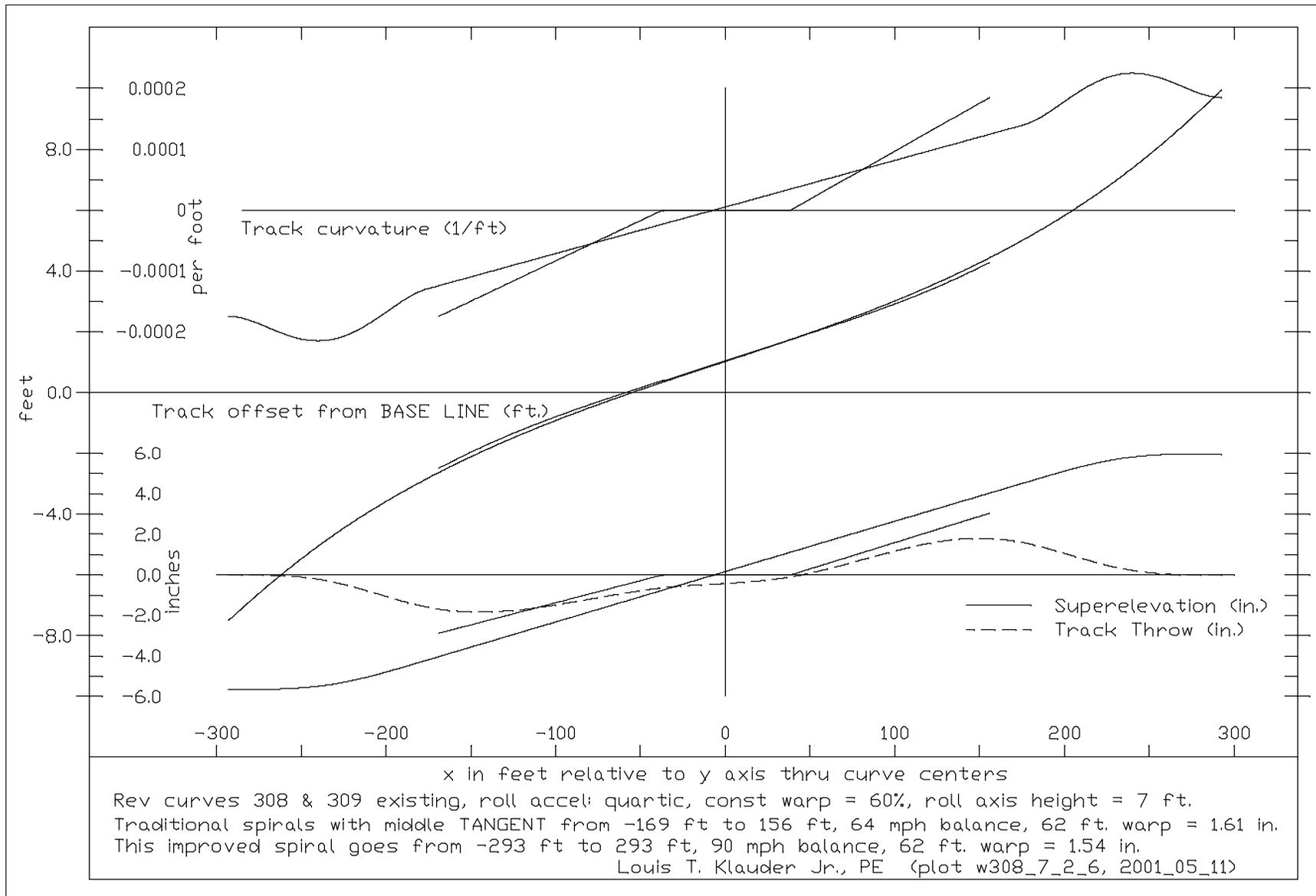
Figure 8-4: Plot W308_7_2_2: Roll axis height = 7 ft; Roll model = "Piecewise Quartic"; Width of central 0.0 acceleration zone = 20%
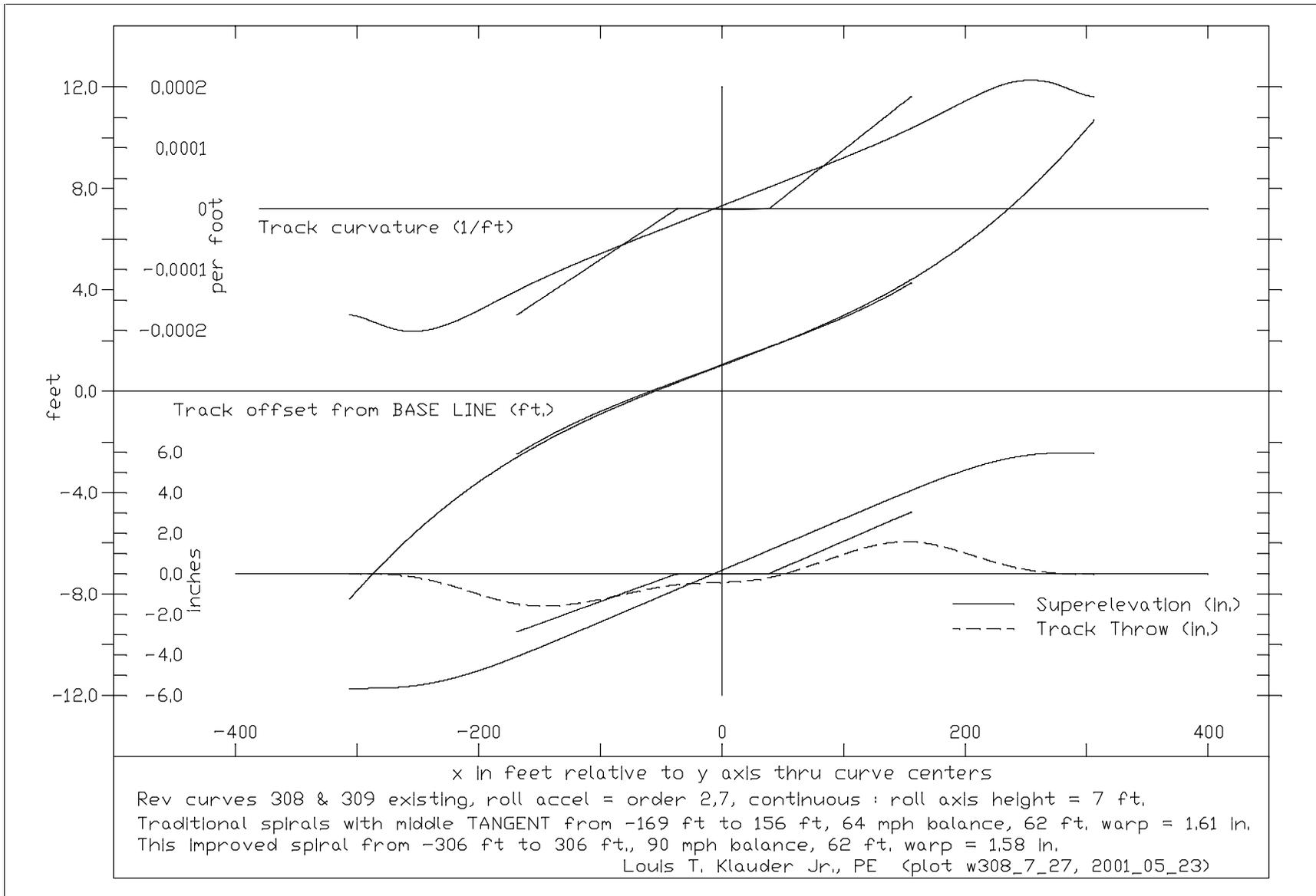
Figure 8-5: Plot W308_7_2_4: Roll axis height = 7 ft; Roll model = "Piecewise Quartic"; Width of central 0.0 acceleration zone = 40%

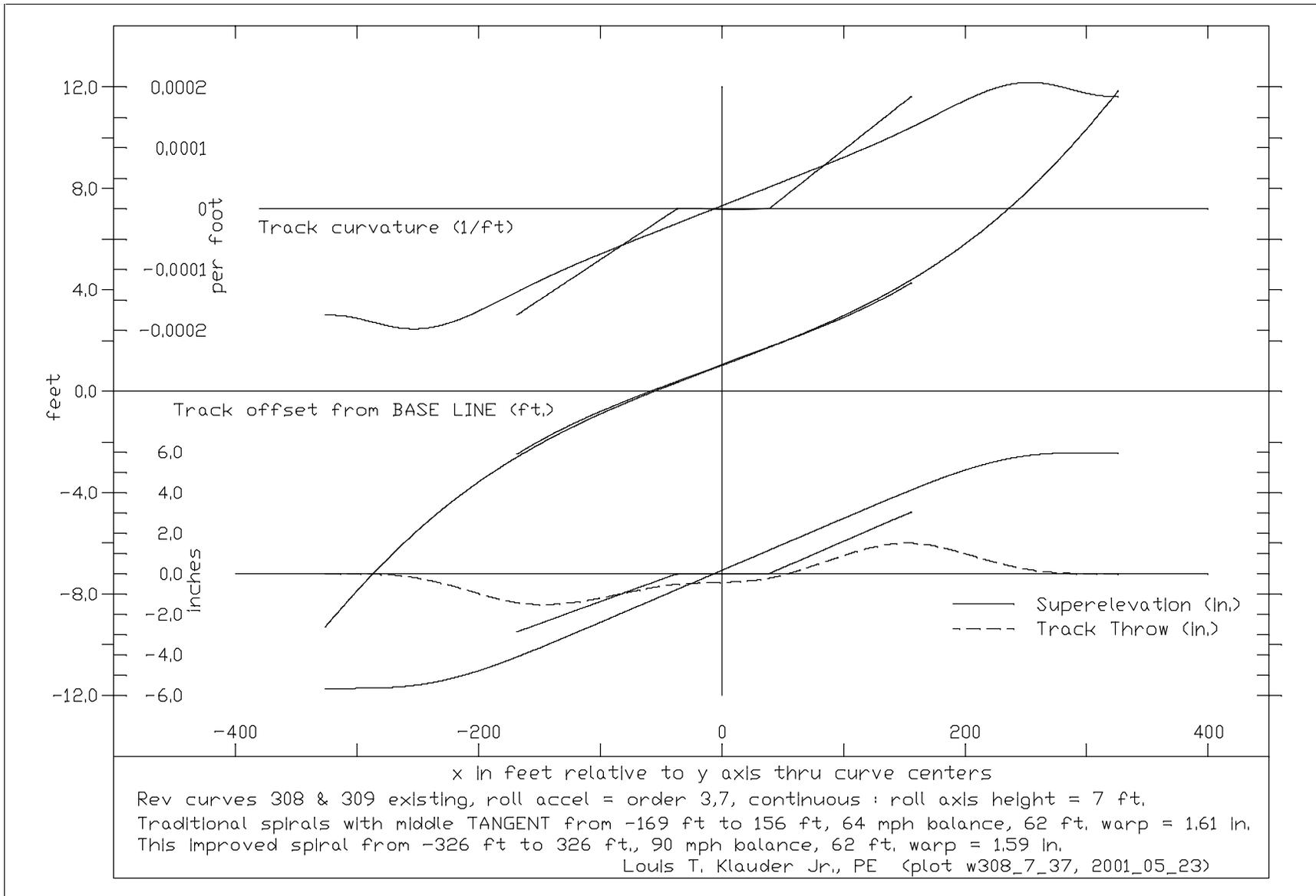Figure 8-6: Plot W308_7_2_6: Roll axis height = 7 ft; Roll model = "Piecewise Quartic"; Width of central 0.0 acceleration zone = 60%

Figure 8-7: Plot W308_7_27: Roll axis height = 7 ft; Roll model = Order (2,7)

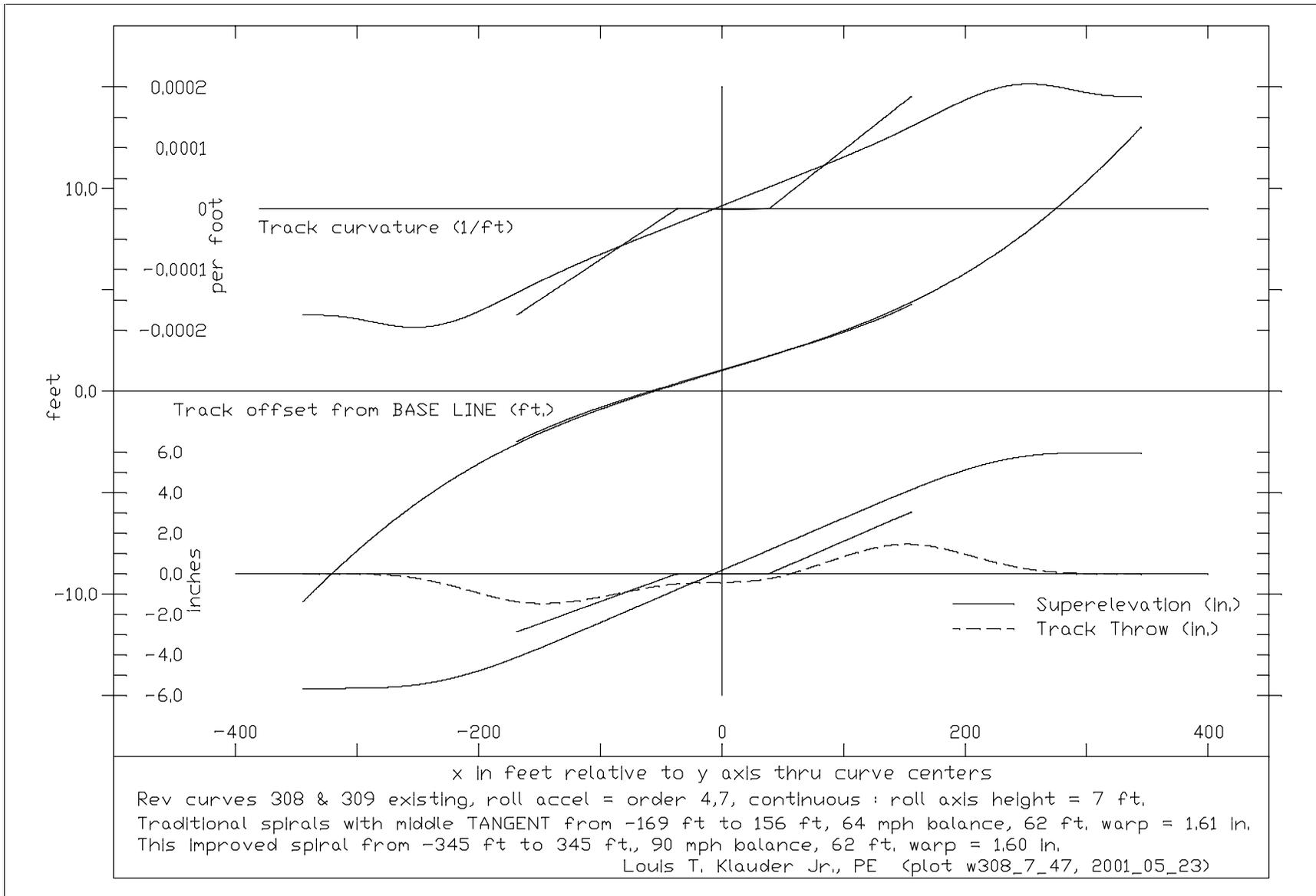Figure 8-8: Plot W308_7_37: Roll axis height = 7 ft; Roll model = Order (3,7)

Figure 8-9: Plot W308_7_47: Roll axis height = 7 ft; Roll model = Order (4,7)
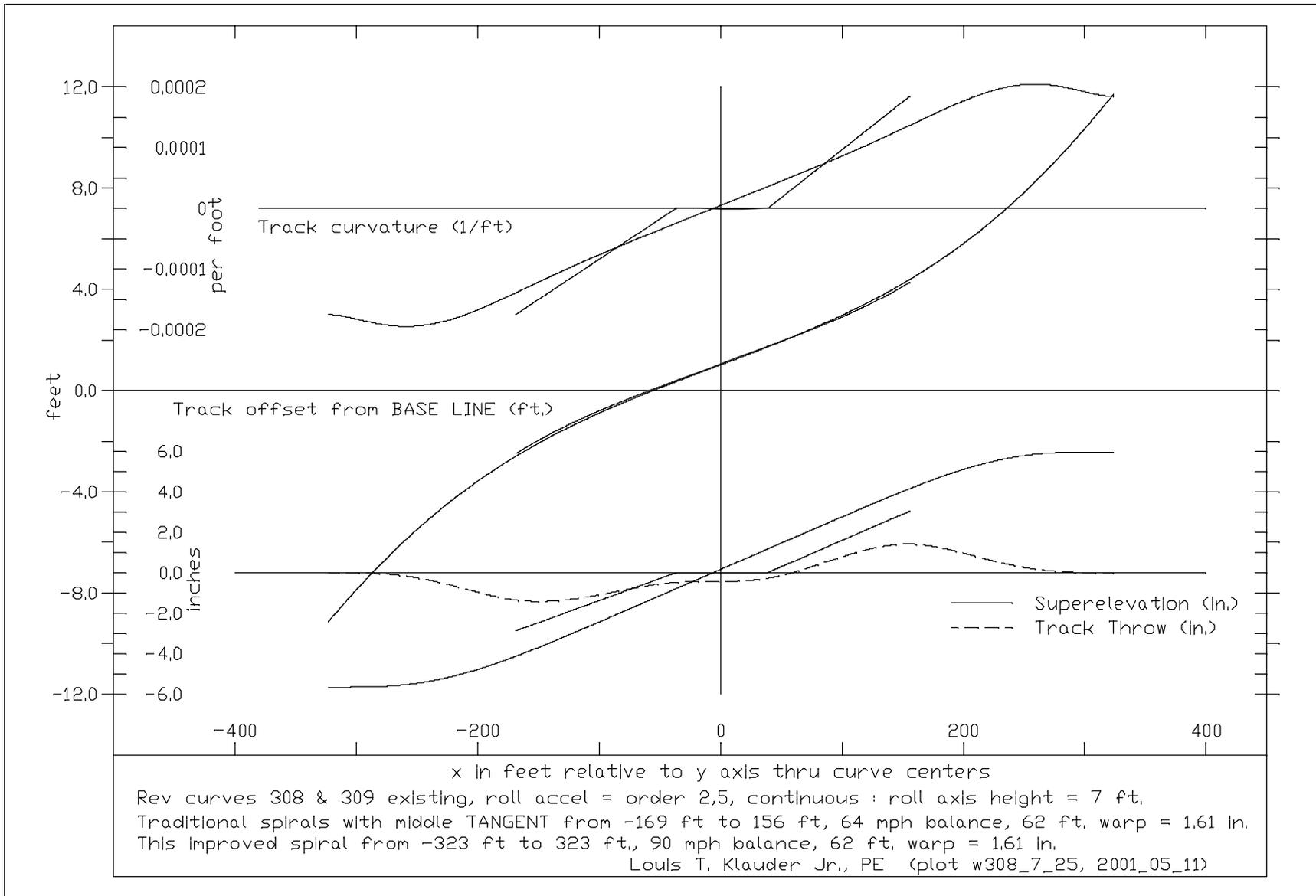
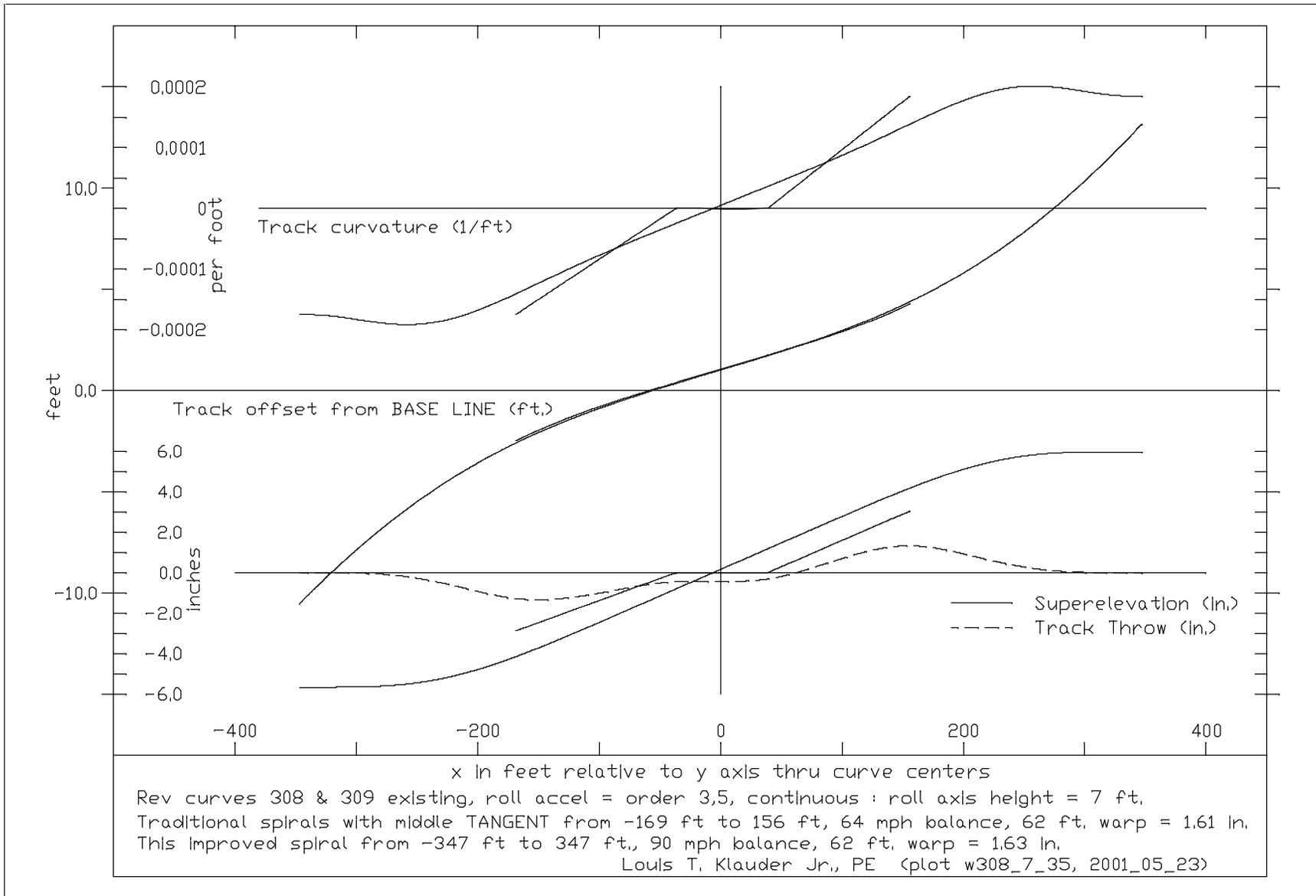Figure 8-10: Plot W308_7_25: Roll axis height = 7 ft; Roll model = Order (2,5)

Figure 8-11: Plot W308_7_35: Roll axis height = 7 ft; Roll model = Order (3,5)
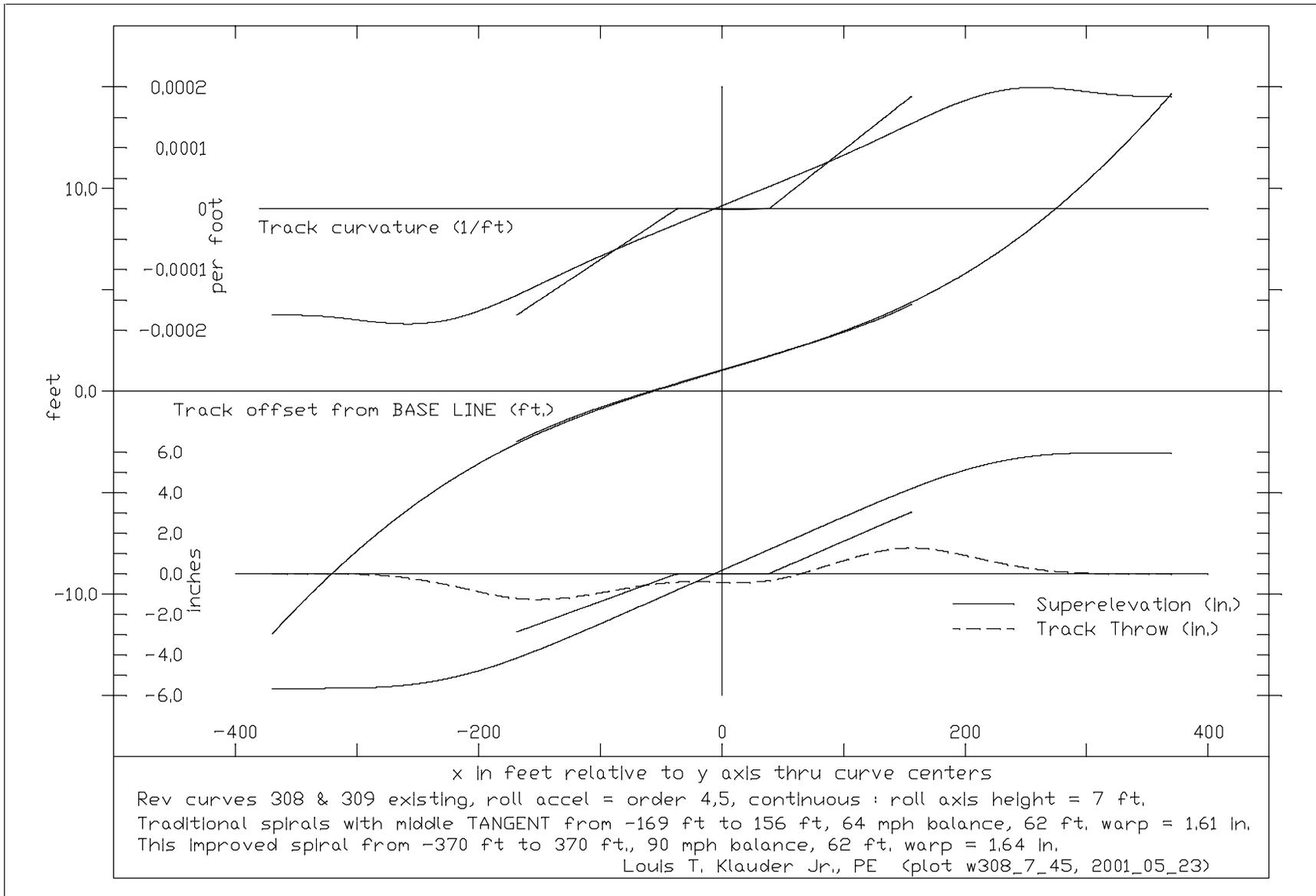
Figure 8-12: Plot W308_7_45: Roll axis height = 7 ft; Roll model = Order (4,5)
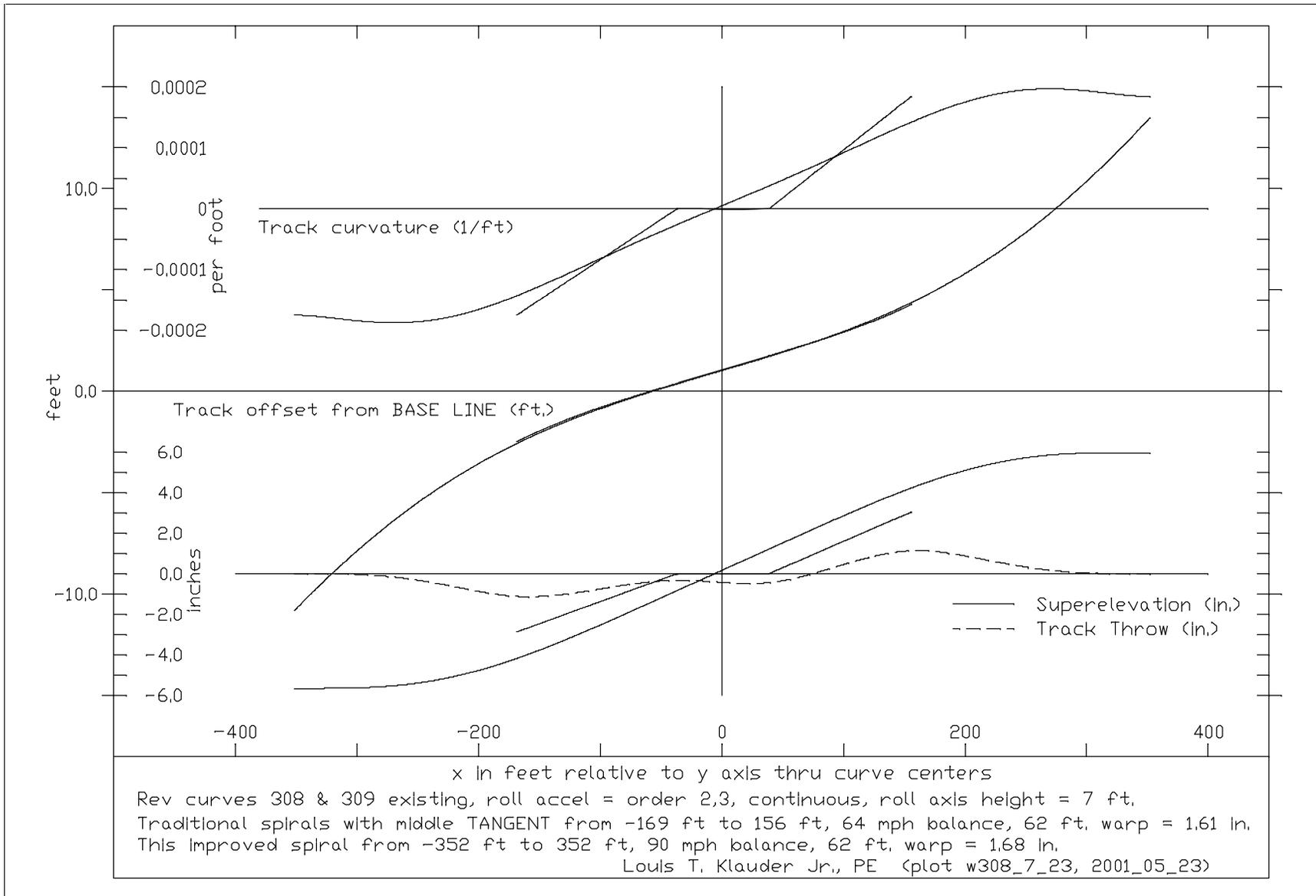
Figure 8-13: Plot W308_7_23: Roll axis height = 7 ft; Roll model = Order (2,3)
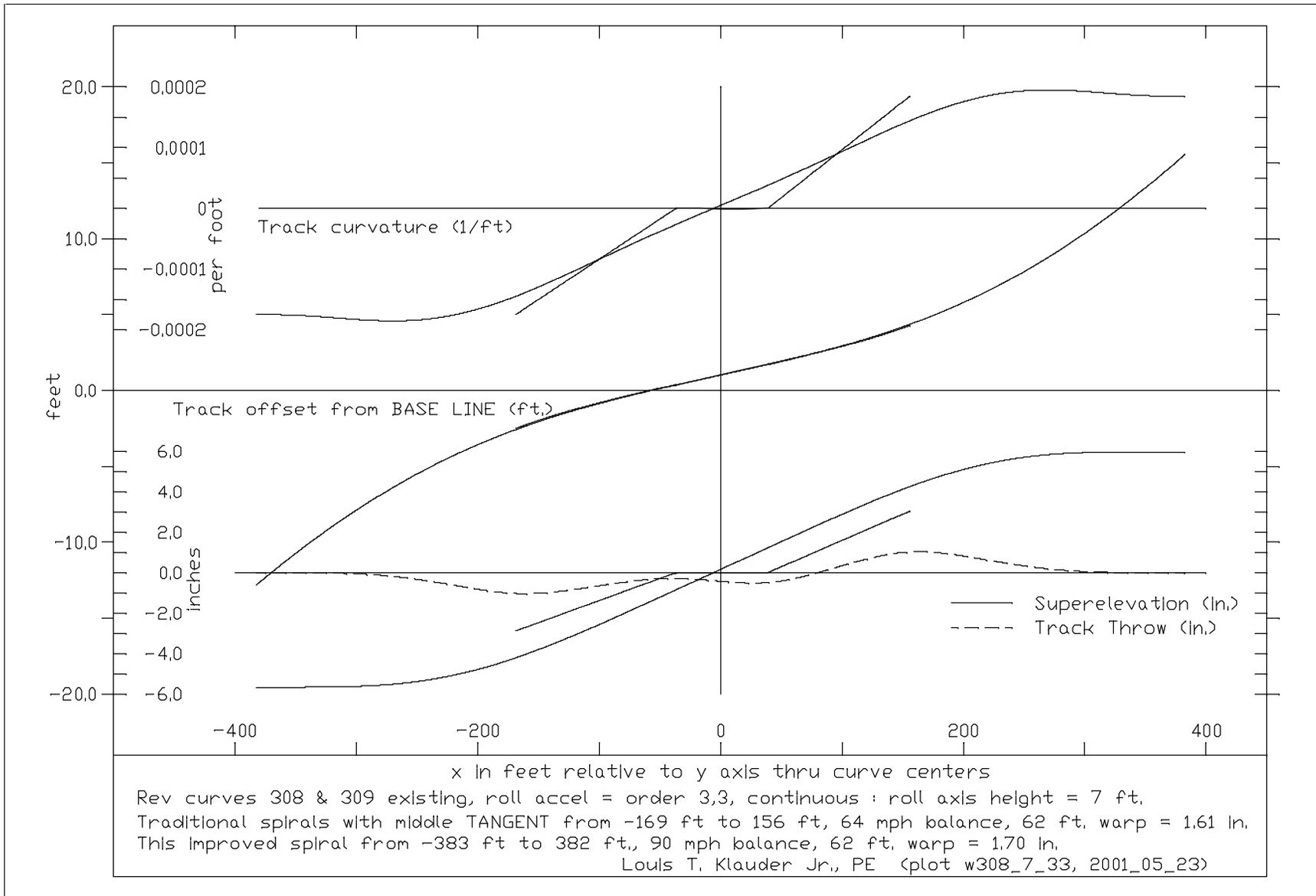
Figure 8-14: Plot W308_7_33: Roll axis height = 7 ft; Roll model = Order (3,3)
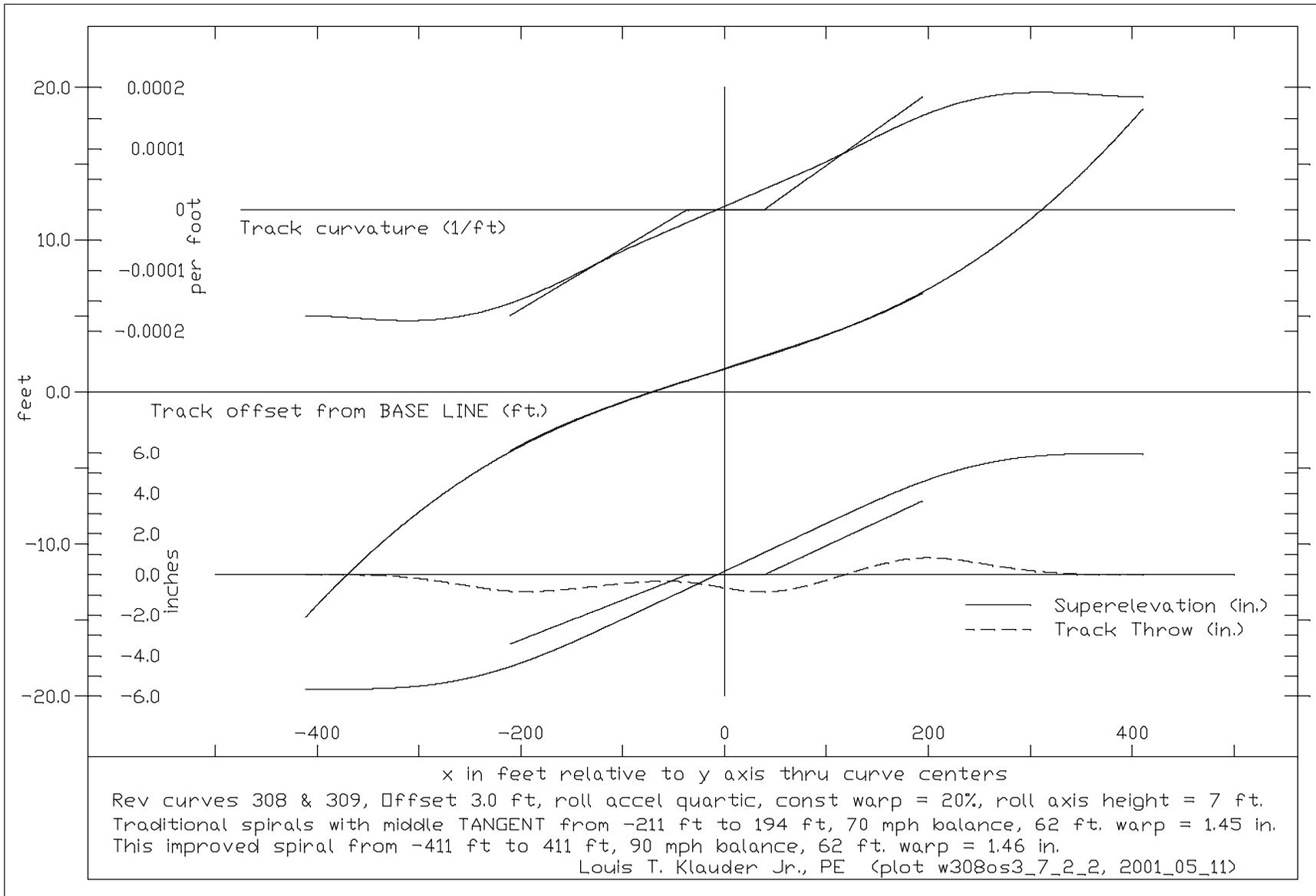
Figure 8-15: Plot W308os3_7_2_2: Curve Offset raised to 3 ft; Roll axis height = 7 ft; Roll model = "Quartic"; Center Zone width 20%
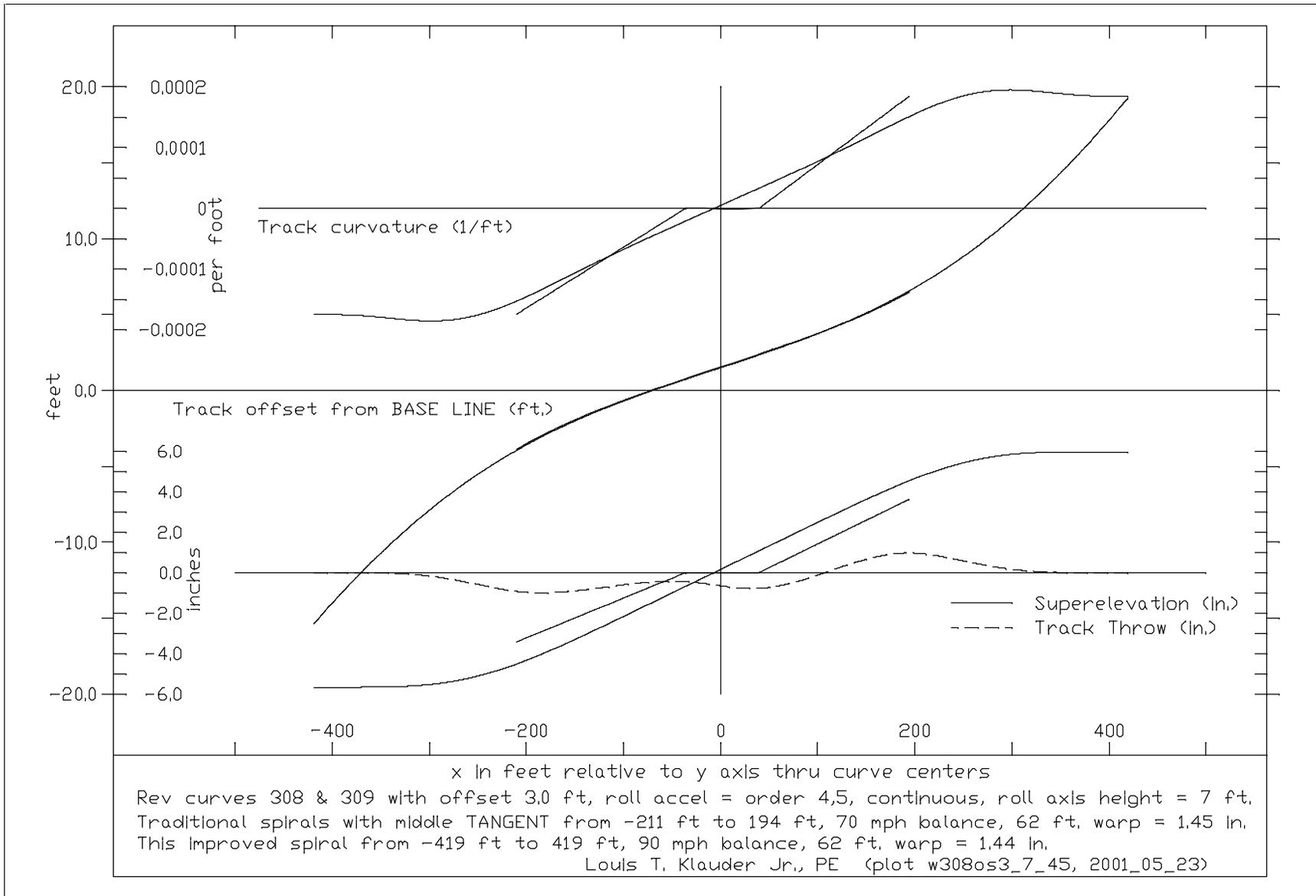
Figure 8-16: Plot W308os3_7_45: Curve Offset raised to 3 ft; Roll axis height = 7 ft; Roll model = Order (4,5)
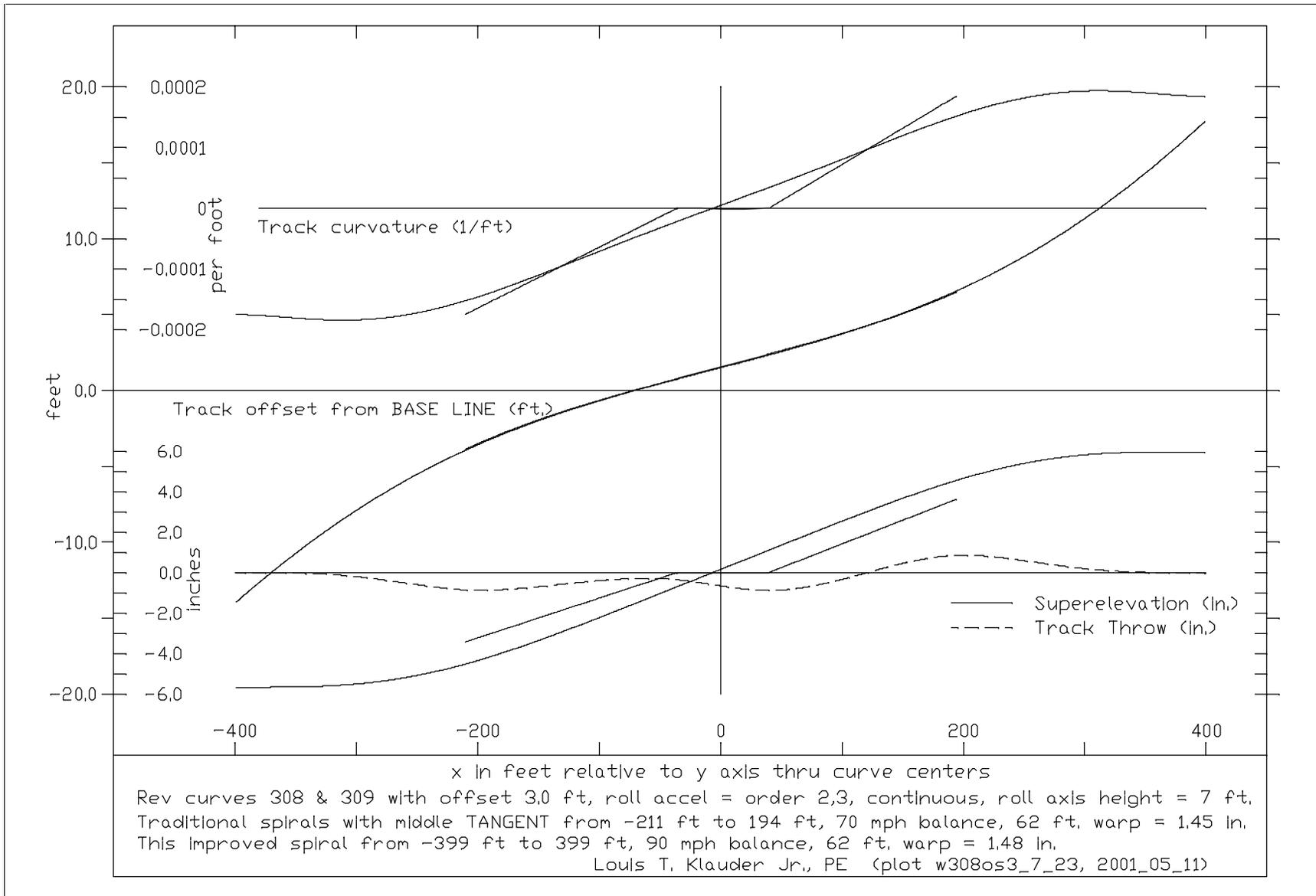
Figure 8-17: Plot W308os3_7_23: Curve Offset raised to 3 ft; Roll axis height = 7 ft; Roll model = Order (2,3)

9   Trade-off between track warp and roll acceleration

Here we observe some of the interplay between the various spiral parameters. This interplay is also discussed by Kufver (ref. 5).

When a new route is being designed in open country, the choice of spirals should generally be easy. Offsets can be made large and spirals can be made very long, inefficient spiral shapes based on roll models such as order (2,1), order (5,3), or raised sine with c = 0.0 can used to construct very gentle spirals. The benefit of gentler spirals that can be used when offsets are more generous may be seen in the previous examples by comparing spirals whose labels begin W3080s3_ (offset = 3 ft.) with counterparts whose labels begin W308_ (offset = about 2 ft.). If the offset of curves 308 & 309 were raised to 3 ft., then the spirals labeled W308os3_7_2_2, W308os3_7_23, and W308os3_7_45 would give results that are similar to one another and quite attractive.

A challenge arises when an existing route with fixed and inadequate curve offsets needs to be upgraded. Looking at the characteristic values in the table of the preceding section, one can observe that reducing the roll acceleration and jerk requires spreading the acceleration toward the middle of the spiral. When roll acceleration is spread out, then, in order to achieve the required change in roll angle, the track warp near the middle of the spiral has to be larger. One reason that use of the clothoid spiral has continued for so long is that by moving all of the acceleration to the very ends of the spiral it does the best job of minimizing the track warp when the length of the spiral is restricted. Of the solutions for the existing 2 ft. offset of curves 308 & 309, none have track warp quite as low as 1.5 in. in 62 ft. The solutions that come closest to that low a warp have roll accelerations that may be a little higher than desired. Of the examples shown, the ones labeled W308_7_2_4 and W308_7_25 might be considered the most attractive.

Note the similarity between spirals W308os3_7_2_2 and W308os3_7_23 on the on hand and on the other hand between W308_7_2_4 and W308_7_25. This similarity illustrates the way that variation of the width of the central zero acceleration zone in a model that has such a zone gives it some flexibility that can be used to help accommodate to differing situations.

10   The question of allowable track warp

The AREMA Manual for Railway Engineering (ref. 8), Volume 1 Track, Section 3.1.1 recommends that the rate of change of superelevation in spirals in main line tracks should not exceed 1.0 inch per 62 feet in order to avoid applying excessive torsion to the suspensions of 85 foot long cars. That limit may be more conservative than necessary in relation to its stated rationale. (The Manual has another criterion for maximum rate of change of superelevation in main line track spirals for avoiding discomfort to passengers.)

The Federal Railroad Administration publishes various rules among which is rule Track Safety Standards, Part 213, Subpart G, Class of Track 6 and Higher (ref. 9). Section 213.331(a) of these rules states among other things: "The difference in crosslevel between any two points less than 62 feet apart may not be more than 1.5 inches." In Subpart A to F, Class of Track 1-5 of the same standard, section 213.63 states the corresponding limits in spirals where the local superelevation is less than 6.0 inches as 3 in., 2.25 in., 2 in., and 1.75 in. in 62 ft. for track classes 1, 2, 3, and 4 respectively.

The FRA and AREMA rules are assumed to be based mainly on experience over many years relating to comfort, to safety, and perhaps also to track maintenance cost. That experience has been based on operation with clothoid spirals, and the resulting practice has had to cope both with the defective dynamic behavior that is experienced at each end of the clothoid spiral and with the twisting of vehicle suspensions associated with the warp of the track in the interior of the spiral. It would be a coincidence if the track warp limit needed because of defective spiral dynamics turned out to be the same as the track warp limit needed to ensure safe suspension performance in the middle of the spiral.

If a clothoid spiral is to be replaced by an improved spiral, then, depending on vehicle suspension characteristics, it may be possible to allow a more generous value of track warp in the interior of the spiral.

11   Conclusions

11.1   Improved spiral geometry is practical

In earlier times when most North American railroad track traditions were developed, there were two obstacles to consideration of alternatives to the clothoid spiral. It is argued that neither obstacle applies any longer.

First, the mathematical functions (the Fresnel integrals) used for layout of a clothoid spiral were available in tabular form so that any civil engineer could lay spirals out. In contrast, any other form of spiral is more complex, and layout for any other spirals would have required tedious manual calculations. Because personal computers and their software are now so well developed and economical, there is no longer any computational obstacle to using alternative spiral shapes.

Second, when track lining was done by hand, lining of curved track was more expensive than lining of tangent track. As spirals are a form of curved track, there was an economic incentive to keep spirals as short as the needs of ride comfort and safety would allow. With the advent of computer based surveying and tamping, the cost for lining curved track should not any more than the cost for lining straight track. There should therefore no longer be any prejudice against allowing spirals to be long. Kufver (ref. 5) and Presle & Hasslinger (ref. 6) make this same point.

It should also be noted from the sample plots of improved spirals that the lateral shifting of the track required to move from a traditional spiral design to an improved design tends to be of the order of 1 or 2 inches. While clearances are always looked at with great care, shifts of this magnitude are expected to manageable without wayside structure changes and without compromise of clearance standards.

11.2   Improved spirals will eliminate a source of ride discomfort
Spreading and reducing the roll acceleration in comparison to that experienced upon entrance to and exit from a clothoid spiral and raising the axis about which vehicle roll takes place will work together to make travel over spirals as comfortable as travel over the neighboring sections of track.

11.3   Improved spirals will eliminate a cause of track degradation
Spreading and reducing the roll acceleration in comparison to that experienced upon entrance to and exit from a clothoid spiral and raising the roll axis about which the acceleration occurs to a height at or somewhat above the average vehicle center of gravity height will eliminate corresponding lateral reaction forces applied locally and systematically by vehicles against the track. As demonstrated by

Presle and Hasslinger (ref. 6), this eliminates systematic degradation of track alignment at those points. It is expected that the geometry developed for some spiral situations using the concepts presented here could be similar to geometry that Presle and Hasslinger would develop for those situations and that the performance improvements relative to traditional spirals could also be similar. However, the author believes that designs based on the models presented herein can sometimes be a little better and, more importantly, that the approach to spiral design described herein represents an improvement from a conceptual point of view.

11.4 *Reverse curves can be fully elevated*

In traditional North American track design practice there is a requirement that adjacent reverse spirals be separated by a section of tangent track (ref. 10). This requirement is a mistake from the point of view of vehicle dynamics. If the initial reason for that requirement was that it made the tasks of line location and surveying easier when those tasks had to be carried out manually, then there may no longer be any reason to perpetuate that requirement. This observation is independent of the question of how spirals may best be shaped.

However, the combination of dropping that requirement and using the spiral design method advocated here offers a distinctive benefit for reverse curves. Namely, as illustrated in the plots presented herein, that combination will allow the desired superelevations for some reverse curves that cannot be adequately superelevated based on traditional practice.

12 **References**

1. Louis T. Klauder, Jr., "Engineering Options for the Northeast Corridor", Transportation Research Record 1028, Transportation Research Board, National Research Council, Washington, D.C., 1985

2. Proceedings of the 2000 Annual Conference, American Railroad Engineering and Maintenance of Way Association (AREMA), Landover, MD, 2000 (available from AREMA printed or as a CDROM.)

3. M. Abramowitz and I. Stegun, Handbook of Mathematical Functions, National Bureau of Standards, U.S. Government Printing Office, Washington, D.C. 1964

4. Report No. ER-37, "Length of Railway Transition Spiral: Analysis and Running Tests", Engineering Research Division, A.A.R. Research Center, Chicago, IL, 1963.

5. Bjorn Kufver, VTI rapport 420A, "Mathematical description of railway alignments and some preliminary comparative studies", Swedish National Road and Transport Research Institute, 1997

6. Gerard Presle & Herbert L Hasslinger, "Entwicklung und Grundlagen neuer Gleisgeometrie", ZEV + DET Glas. Ann. 122, 1998, 9/10, September/Oktober, page 579

7. For information about the NUCARS rail vehicle motion simulation program, contact Transportation Technology Center, Inc. 55500 DOT Rd, P.O.Box 11130, Pueblo, CO 81001.

8. The 1999 Manual for Railway Engineering, AREMA, 8201 Corporate Dr., #1125, Landover, MD 20785

9. Track Safety Standards, PART 213, subpart G, Class of Track 6 and Higher, dated October 1998, Published by DOT, FRA, Office of Safety, Washington, D.C., Distributed by the Railway Educational Bureau, Omaha, NE

10. Northeast Corridor Improvement Project Task 104: Design Manual: Engineering, Section 3.3.3, "Tangent Lengths", De Leuw, Cather/Parsons, Final Report to U.S. Department of Transportation, Federal Railroad Administration, November, 1977. Copies may be available from National Technical Information Service, Springfield, VA 22161, www.ntis.gov, (703) 605-6000. This section reads: "The minimum tangent length between two curves shall not be less than 100 ft. nor be less than: L = 2.2 V, where L = length of tangent in feet" and "V = scheduled velocity of high speed trains for the curve in miles per hour."

The 1999 Manual for Railway Engineering (AREMA, 8201 Corporate Dr., #1125, Landover, MD 20785), Section 3.5 deals with minimum tangent lengths required between reverse curves for yard operations. Section 3.5.2 pertains to reverse curves with spirals and superelevation and states that such reverse spirals should be separated by a section of tangent track at least as long as the longest car expected to traverse the curves. This section appears to be dealing with the static geometry of adjacent coupled cars traversing sharp reverse curves and the associated coupler deflections and adjacent car end displacements rather than with dynamic conditions on main line tracks.