Kochanek-Bartels Cubic Splines (TCB Splines)

David Eberly
Geometric Tools, LLC
http://www.geometrictools.com/

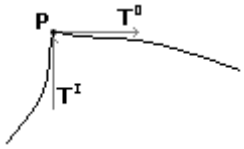
Created: July 16, 1999
Last Modified: February 14, 2008

Contents

1 Introduction

A sequence of positional key frames is $\{(t_i, \mathbf{P}_i, \mathbf{T}_i^I, \mathbf{T}_i^O)\}_{i=0}^{n-1}$ where t_i is the *sample time*, $\mathbf{P}_i$ is the *sample position*, $\mathbf{T}_i^I$ is the *incoming tangent vector*, and $\mathbf{T}_i^O$ is the *outgoing tangent vector*. The Kochanek-Bartels spline [1] specifies a cubic polynomial interpolation between each pair of key frames by choosing the incoming and outgoing tangents in a special way. The incoming and outgoing tangent vectors at $\mathbf{P}_i$ need not be the same vector, in which case the spline has a derivative discontinuity at that position. The tangents at $\mathbf{P}_i$ are chosen based on neighboring positions and on three parameters that have some visual appeal. The parameters are *tension* τ_i which controls how sharply the curve bends at a control point, *continuity* γ_i which controls the continuity (or discontinuity) at a position, and *bias* β_i which controls the direction of the path at $\mathbf{P}_i$ by taking weighted combination of one-sided derivatives at that position. Figure 1.1 shows a typical position, tangent vectors, and the curve segments passing through the position.

Figure 1.1 A cubic spline curve passing through point $\mathbf{P}$ with incoming tangent $\mathbf{T}^I$ and outgoing tangent $\mathbf{T}^O$.



The Kochanek-Bartels splines are sometimes called *TCB splines*, the acronoym referring to tension, continuity, and bias.

2 Cubic Polynomial Curves

In general, a cubic polynomial curve can be constructed to represent the curve connecting a pair of key frames $(t_i, \mathbf{P}_i, \mathbf{T}_i^I, \mathbf{T}_i^O)$ and $(t_{i+1}, \mathbf{P}_{i+1}, \mathbf{T}_{i+1}^I, \mathbf{T}_{i+1}^O)$. Such a curve is of the form

$$\mathbf{X}_i(t) = \mathbf{A}_i + \left(\frac{t-t_i}{\Delta_i}\right)\mathbf{B}_i + \left(\frac{t-t_i}{\Delta_i}\right)^2\mathbf{C}_i + \left(\frac{t-t_i}{\Delta_i}\right)^3\mathbf{D}_i$$

where $\Delta_i = t_{i+1} - t_i$ and where $t \in [t_i, t_{i+1}]$. The vector-valued coefficients of the polynomial are determined by requiring

$$\mathbf{X}_i(t_i) = \mathbf{P}_i, \;\; \mathbf{X}_i(t_{i+1}) = \mathbf{P}_{i+1}, \;\; \mathbf{X}_i'(t_i) = \mathbf{T}_i^O, \;\; \mathbf{X}_i'(t_{i+1}) = \mathbf{T}_{i+1}^I$$

That is, the curve must pass through the end points and the curve derivatives at the end points must match the specified tangent vectors. The four linear equations implied by these conditions are

$$\begin{aligned}
\mathbf{A}_i &= \mathbf{P}_i \\
\mathbf{A}_i + \mathbf{B}_i + \mathbf{C}_i + \mathbf{D}_i &= \mathbf{P}_{i+1} \\
\mathbf{B}_i &= \Delta_i \mathbf{T}_i^O \\
\mathbf{B}_i + 2\mathbf{C}_i + 3\mathbf{D}_i &= \Delta_i \mathbf{T}_{i+1}^I
\end{aligned}$$

The solution is

$$\begin{aligned}\mathbf{A}_i &= \mathbf{P}_i \\ \mathbf{B}_i &= \Delta_i \mathbf{T}_i^O \\ \mathbf{C}_i &= 3(\mathbf{P}_{i+1} - \mathbf{P}_i) - \Delta_i(2\mathbf{T}_i^O + \mathbf{T}_{i+1}^I) \\ \mathbf{D}_i &= -2(\mathbf{P}_{i+1} - \mathbf{P}_i) + \Delta_i(\mathbf{T}_i^O + \mathbf{T}_{i+1}^I)\end{aligned}$$

The curve may be rewritten to use a Hermite interpolation basis $H_0(s) = 2s^3 - 3s^2 + 1$, $H_1(s) = -2s^3 + 3s^2$, $H_2(s) = s^3 - 2s^2 + s$, and $H_3(s) = s^3 - s^2$, namely

$$\mathbf{X}_i(t) = H_0\left(\frac{t-t_i}{\Delta_i}\right)\mathbf{P}_i + H_1\left(\frac{t-t_i}{\Delta_i}\right)\mathbf{P}_{i+1} + H_2\left(\frac{t-t_i}{\Delta_i}\right)\Delta_i\mathbf{T}_i^O + H_3\left(\frac{t-t_i}{\Delta_i}\right)\Delta_i\mathbf{T}_{i+1}^I \tag{1}$$

A spline is formed by the sequence of curves $\mathbf{X}_i(t)$, $0 \le i \le n-2$. Each segment is continuous and differentiable. The spline is continuous over all segments since it passes through all the sample positions. Whether or not the spline is differentiable at the sample positions depends on the choice of incoming and outgoing tangent vectors. The spline is said to have G^1 *continuity* at position $\mathbf{P}_i$ if the incoming and outgoing tangents $\mathbf{T}_i^I$ and $\mathbf{T}_i^O$ are in the same direction (not necessarily of equal lengths). The spline is said to have C^1 *continuity* at the sample position if the incoming and outgoing tangents are equal (in the same direction *and* of equal length).

3 Choosing Tangent Vectors ($t_i = i$)

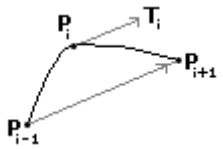
We now look at the choices for tangent vectors based on the tension, continuity, and bias parameters. You will notice that the section heading mentions that the sample times are $t_i = i$. The formulation in [1] implicitly assumes this, but mention is made later about "adjustments" to the tangents when nonuniform sample times are used in order to obtain continuity of speed, that is, continuity of the length of the tangent vectors to the curve. This issue needs closer attention than what is provided in [1]; I will do so in the next section.

The default values for tension τ_i, continuity γ_i, and bias β_i are all zero. In this case the incoming and outgoing tangent vectors are chosen to be the same vector,

$$\mathbf{T}_i^I = \mathbf{T}_i^O = \frac{1}{2}\left((\mathbf{P}_{i+1} - \mathbf{P}_i) + (\mathbf{P}_i - \mathbf{P}_{i-1})\right) = \frac{\mathbf{P}_{i+1} - \mathbf{P}_{i-1}}{2} \tag{2}$$

It is important to notice that the units of the tangent vectors are position divided by time. The right-hand side of equation (2) appears to have units of position. To make the left-hand sides and the right-hand side are commensurate, you must think of the denominator 2 of the right-hand side as 2 units of time, thus making the units match. This choice of tangents leads to the *Catmull-Rom spline*. The tangent vector at $\mathbf{P}_i$ is chosen to be parallel to the secant vector from $\mathbf{P}_{i-1}$ to $\mathbf{P}_{i+1}$ as shown in Figure 3.1.

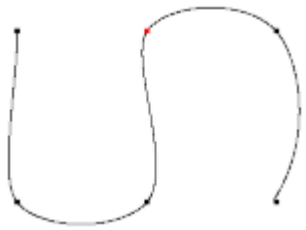
Figure 3.1 The Catmull-Rom spline. The tangent vector at $\mathbf{P}_i$ is chosen to be parallel to the secant vector connecting the neighboring positions.



The tangent is also viewed as the average change in position between the sample position and its neighboring sample positions.

Figure 3.2 shows a Catmull-Rom spline for six sample positions. The spline has C^1 continuity everywhere.

Figure 3.2 The Catmull-Rom spline for six sample positions with uniform sampling in time (Δt_i is constant).



We are going to modify each of the tension, continuity, and bias parameters at the sample position shown in red to see how the shape of the curve varies at that point.

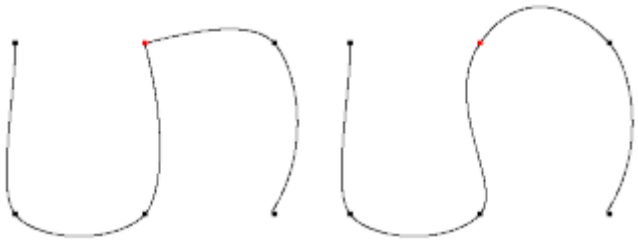
3.1 Tension

Tension $\tau_i \in [-1, 1]$ is introduced as a parameter in the equations (2) as shown below,

$$\mathbf{T}_i^I = \mathbf{T}_i^O = \frac{(1-\tau_i)}{2}((\mathbf{P}_{i+1} - \mathbf{P}_i) + (\mathbf{P}_i - \mathbf{P}_{i-1})). \tag{3}$$

As noted earlier, the units must match between the left-hand and right-hand sides. The tension is dimensionless, so the 2 in the denominator must be thought of as having units of time. If $\tau_i = 0$, then we have the Catmull-Rom spline. For τ_i near 1 the curve is "tightened" at the control point while τ_i near -1 produces "slack" at the control point. Varying τ_i changes the length of the tangent at the control point, a smaller tangent leading to a tightening and a larger tangent leading to a slackening. Figure 3.3 shows nonzero tension values applied to the red sample position of Figure 3.2.

Figure 3.3 Left: Tension $\tau_i = 1$ at the sample position. Right: Tension $\tau_i = -1$ at the sample position.



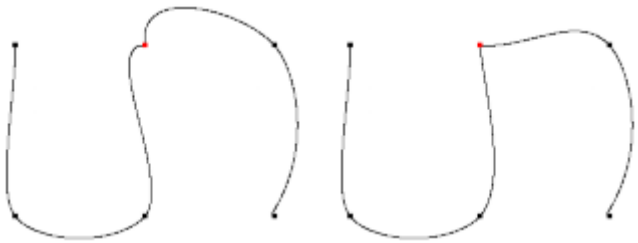
3.2 Continuity

Continuity $\gamma_i \in [-1, 1]$ is introduced as a parameter in the equations (2) as shown below. Notice that the incoming and outgoing tangents are generally different.

$$\mathbf{T}_i^I = \frac{1+\gamma_i}{2}(\mathbf{P}_{i+1} - \mathbf{P}_i) + \frac{1-\gamma_i}{2}(\mathbf{P}_i - \mathbf{P}_{i-1}), \quad \mathbf{T}_i^O = \frac{1-\gamma_i}{2}(\mathbf{P}_{i+1} - \mathbf{P}_i) + \frac{1+\gamma_i}{2}(\mathbf{P}_i - \mathbf{P}_{i-1}) \tag{4}$$

Once again the 2 in the denominator must be viewed as having units of time so that the units match between the left-hand and right-hand sides. When $\gamma_i = 0$, the curve has a continuous tangent vector at the control point. As $|\gamma_i|$ increases, the resulting curve has a "corner" at the control point, the direction of the corner depending on the sign of γ_i. Figure 3.4 shows nonzero continuity values applied to the red sample position of Figure 3.2.

Figure 3.4 Left: Continuity $\gamma_i = 1$ at the sample position. Right: Continuity $\gamma_i = -1$ at the sample position.



In both images of the figure it is apparent that the curve has neither C^1 nor G^1 continuity at the red sample position.

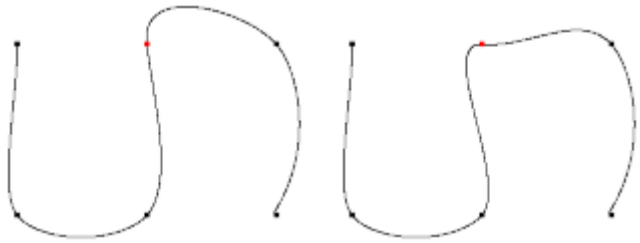
3.3 Bias

Bias $\beta_i \in [-1, 1]$ is introduced as a parameter in the equations (2) as shown below. As with tension, the incoming and outgoing tangents are chosen to be equal.

$$\mathbf{T}_i^I = \mathbf{T}_i^O = \frac{1-\beta_i}{2}(\mathbf{P}_{i+1} - \mathbf{P}_i) + \frac{1+\beta_i}{2}(\mathbf{P}_i - \mathbf{P}_{i-1}) \tag{5}$$

As in the other cases, the 2 in the denominator has units of time in order that the units of the left-hand and right-hand sides match. When $\beta_i = 0$ the left and right one-sided tangents are equally weighted, producing the Catmull-Rom spline. For β_i near -1, the outgoing tangent dominates the direction of the path of the curve through the control point (undershooting). For β_i near 1, the incoming tangent dominates (overshooting). Figure 3.5 shows nonzero continuity values applied to the red sample position of Figure 3.2.

Figure 3.5 Left: Bias $\beta_i = 1$ at the sample position. Right: Bias $\beta_i = -1$ at the sample position.



The three effects may be combined into a single set of equations

$$\mathbf{T}_i^I = \frac{(1-\tau_i)(1+\gamma_i)(1-\beta_i)}{2}(\mathbf{P}_{i+1} - \mathbf{P}_i) + \frac{(1-\tau_i)(1-\gamma_i)(1+\beta_i)}{2}(\mathbf{P}_i - \mathbf{P}_{i-1}) \tag{6}$$

and

$$\mathbf{T}_i^O = \frac{(1-\tau_i)(1-\gamma_i)(1-\beta_i)}{2}(\mathbf{P}_{i+1} - \mathbf{P}_i) + \frac{(1-\tau_i)(1+\gamma_i)(1+\beta_i)}{2}(\mathbf{P}_i - \mathbf{P}_{i-1}) \tag{7}$$

4 Adjustments for Nonuniform Sample Times

The discussion in [1] includes mention of adjustments of the tangent vectors in equations (6) and (7) in order to account for nonuniform sample times. In particular, the authors assume default continuity, $\gamma_i = 0$, the implication being that $\mathbf{T}_i^I = \mathbf{T}_i^O$ in equations (6) and (7). Although the incoming and outgoing tangents are the same (implying G^1 continuity), the authors mention that the speeds of the two segments meeting at $\mathbf{P}_i$ may be different due to nonuniform sample times. The proposed solution is to adjust the lengths of the tangent vectors to force the speed at $\mathbf{P}_i$ to be continuous (implying C^1 continuity). The adjustments are stated as

$$\mathbf{T}_i^I = \frac{2N_i}{N_{i-1} + N_i}\left(\frac{(1-\tau_i)(1+\gamma_i)(1-\beta_i)}{2}(\mathbf{P}_{i+1} - \mathbf{P}_i) + \frac{(1-\tau_i)(1-\gamma_i)(1+\beta_i)}{2}(\mathbf{P}_i - \mathbf{P}_{i-1})\right) \tag{8}$$

and

$$\mathbf{T}_i^O = \frac{2N_{i-1}}{N_{i-1}+N_i}\left(\frac{(1-\tau_i)(1-\gamma_i)(1-\beta_i)}{2}(\mathbf{P}_{i+1}-\mathbf{P}_i)+\frac{(1-\tau_i)(1+\gamma_i)(1+\beta_i)}{2}(\mathbf{P}_i-\mathbf{P}_{i-1})\right) \tag{9}$$

where N_{i-1} is the number of inbetweens that will be used for the time interval $[t_{i-1}, t_i]$ and N_i is the number of inbetweens that will be used for the time interval $[t_i, t_{i+1}]$.

Although this approach has some intuitive appeal, it is not the correct way to deal with continuity of speed at $\mathbf{P}_i$. Continuity at a single point is *not* related to the quantity and selection of a discrete set of inbetweens. Continuity of speed is a calculus concept that requires the one-sided limits of the speed function $\sigma(t)$ at t_i to be finite and equal. That is

$$\lim_{t\to t_i^-}\sigma(t) = \lim_{t\to t_i^+}\sigma(t) = \sigma(t_i).$$

In our current setting, this is just a statement that the lengths of the incoming and outgoing tangents at $\mathbf{P}_i$ must be the same. Clearly the incoming and outgoing tangents defined by equations (8) and (9) do not have the same length when $N_i \neq N_{i-1}$. So the question remains: How do we include the nonuniform sample times in the incoming and outgoing tangent formulas to force continuity of speed at t_i? My opinion is that if you choose to introduce tension, continuity, and bias differently for each of $\mathbf{P}_{i+1}-\mathbf{P}_i$ and $\mathbf{P}_i-\mathbf{P}_{i-1}$, then the inclusion of the time differences should be those relevant to the positional differences. The incoming and outgoing tangents should be

$$\mathbf{T}_i^I = \frac{(1-\tau_i)(1+\gamma_i)(1-\beta_i)}{2}\,\frac{\mathbf{P}_{i+1}-\mathbf{P}_i}{\Delta_i}+\frac{(1-\tau_i)(1-\gamma_i)(1+\beta_i)}{2}\,\frac{\mathbf{P}_i-\mathbf{P}_{i-1}}{\Delta_{i-1}} \tag{10}$$

and

$$\mathbf{T}_i^O = \frac{(1-\tau_i)(1-\gamma_i)(1-\beta_i)}{2}\,\frac{\mathbf{P}_{i+1}-\mathbf{P}_i}{\Delta_i}+\frac{(1-\tau_i)(1+\gamma_i)(1+\beta_i)}{2}\,\frac{\mathbf{P}_i-\mathbf{P}_{i-1}}{\Delta_{i-1}} \tag{11}$$

When $\gamma_i = 0$ the incoming and outgoing tangents are equal, so the curve has C^1 continuity at $\mathbf{P}_i$ regardless of the choice of tension and bias. When $\gamma_i \neq 0$, if you decide you want continuity of speed even though you do not have continuity of direction, an adjustment is made to obtain incoming and outgoing tangents that do have the same length. One such choice is

$$\mathbf{V}_i^I = \frac{2|\mathbf{T}_i^O|}{|\mathbf{T}_i^I|+|\mathbf{T}_i^O|}\,\mathbf{T}_i^I \tag{12}$$

and

$$\mathbf{V}_i^O = \frac{2|\mathbf{T}_i^I|}{|\mathbf{T}_i^I|+|\mathbf{T}_i^O|}\,\mathbf{T}_i^O \tag{13}$$

where $\mathbf{T}_i^I$ and $\mathbf{T}_i^O$ are defined by equations (10) and (11). It is simple to verify that $|\mathbf{V}_i^I| = |\mathbf{V}_i^O|$, so continuity in speed occurs at $\mathbf{P}_i$. Observe that both adjustment coefficients are 1 in the case of $\gamma_i = 0$.

References

[1] Doris H. U. Kochanek and Richard H. Bartels, *Interpolating splines with local tension, continuity, and bias control*, ACM SIGGRAPH 1984, vol. 18, no. 3, pp. 33-41.