

Calculation of Reference Frames along a Space Curve

Jules Bloomenthal

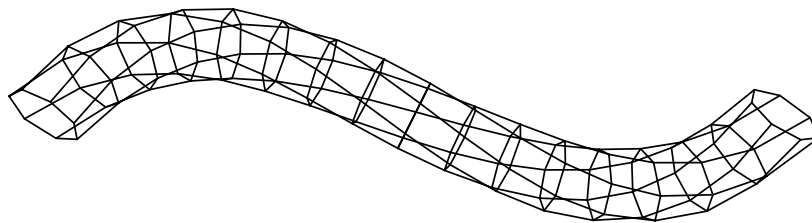
(published in *Graphics Gems*, vol. 1)

Three-dimensional space curves can represent the path of an object or the boundary of a surface patch. They can also participate in various free-form geometric constructions. For example, the *generalized cylinder* (a cylinder with arbitrary cross-sections along a central, space curve axis) is used in Computer Graphics to good effect. Establishing reference frames for the cross-sections of a generalized cylinder, or for any other geometric use, is the subject of this article.

We restrict the central axis to the familiar three-dimensional cubic curve, which we represent by its polynomial coefficients, the three-dimensional vectors **A**, **B**, **C**, and **D**. A point *P* on the curve is computed according to its parametric position, *t*:

$$P = \mathbf{A}t^3 + \mathbf{B}t^2 + \mathbf{C}t + \mathbf{D}.$$

When constructing a polygonal generalized cylinder, each cross-section must be properly aligned with its neighbors so that the structure does not twist. This alignment is usually provided by a *reference frame*, a point and three orthogonal vectors that define position and orientation along the central axis of the cylinder.



Polygons resulting from twisting reference frames.

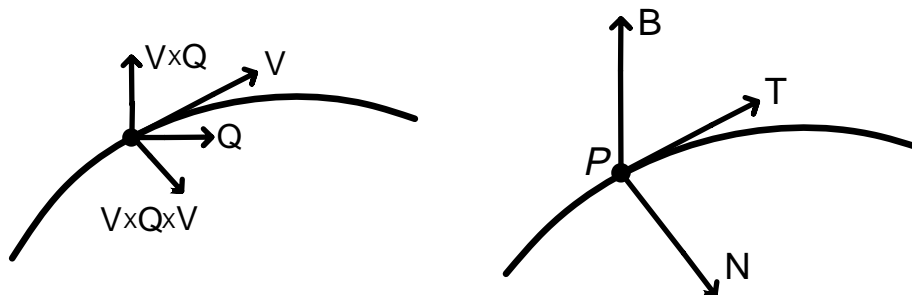
One of the more intuitive reference frames is due to Frenet; the frame consists of a unit length tangent, **T**, to the central axis, a principal normal, **N**, and a binormal, **B**. **T** is simply computed as the unit length velocity vector, **V**; **V** is the derivative of the curve:

$$\mathbf{V} = 3\mathbf{A}t^2 + 2\mathbf{B}t + \mathbf{C}.$$

The principal normal is often defined to be in the direction of curvature, $\mathbf{K} = \mathbf{V} \times \mathbf{Q} \times \mathbf{V} / |\mathbf{V}|^4$.

Q is the acceleration of the curve, *i.e.*, the derivative of velocity, $6\mathbf{A}t + 2\mathbf{B}$. Thus,

$$\mathbf{T} = \mathbf{V} / |\mathbf{V}|, \mathbf{N} = \mathbf{V} \times \mathbf{Q} \times \mathbf{V} / |\mathbf{V} \times \mathbf{Q} \times \mathbf{V}|, \text{ and } \mathbf{B} = \mathbf{T} \times \mathbf{N}.$$



Curvature (left) and a Frenet frame (right).

The Frenet frame is convenient because it can be analytically computed at arbitrary points along the curve. Unfortunately, it is undefined wherever the curvature vanishes, such as at points of inflection or along straight sections of curve. Worse, the curvature vector can reverse direction on either side of an inflection point, inflicting a violent twist in a progression of Frenet frames.

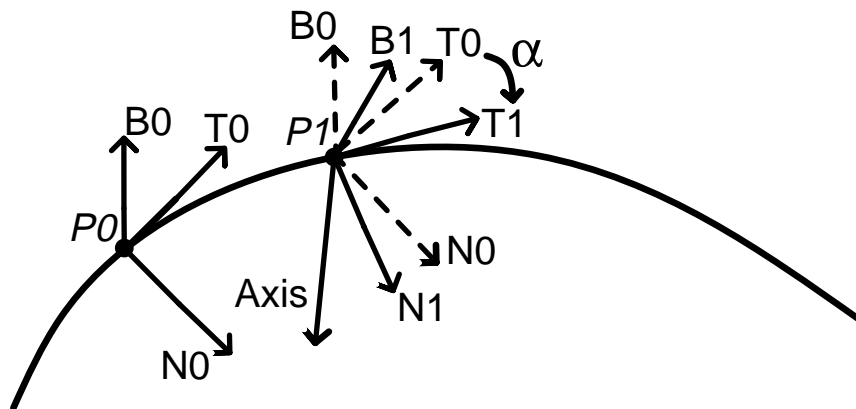
This problem was discussed by U. Shani and D. H. Ballard (*Splines as Embeddings for Generalized Cylinders*, Computer Vision, Graphics, and Image Processing, Vol. 27, August 1984), who proposed an iterative solution to minimize *torsion*, the rotation around the tangent to the curve. This technique was used to compute reference frames for the tree branches in my *Modeling the Mighty Maple* (Proceedings of SIGGRAPH 1985).

Papers by F. Klok (*Two Moving Coordinate Frames for Sweeping along a 3D Trajectory*, Computer Aided Geometric Design, Vol. 3, 1986) and M. Bloomenthal (*Approximation of Sweep Surfaces by Tensor Product B-splines*, University of Utah Technical Report UUCS-88-008, August, 1988) discuss rotation minimizing frames in some detail. They both observe that a torsion minimizing frame does not necessarily produce the intuitively desired result; in the case of a helical curve, for example, the Frenet frame appears more desirable.

The mechanism to minimize torsion defines an initial reference frame at the beginning of the curve; this frame is propagated along the curve using small, local rotations. This method is immune to degeneracies in the curvature vector; it does not, unfortunately, permit analytical computation of a reference frame.

The first frame usually can be computed using curvature, as illustrated above. If the curvature is zero, then \mathbf{N} can be any unit length normal perpendicular to \mathbf{T} . Given the initial frame, subsequent frames are generated, *in order*, by computing P and \mathbf{T} at the new location on the curve. The old reference frame is then rotated such that the old \mathbf{T} aligns itself with the new \mathbf{T} . The rotation creates a new \mathbf{N} and \mathbf{B} , which, with the new P and \mathbf{T} , define a new reference frame.

The axis of rotation, \mathbf{A} , is given by $(\mathbf{T}_0 \times \mathbf{T}_1) / (|\mathbf{T}_0 \times \mathbf{T}_1|)$ and $\alpha = \cos^{-1}((\mathbf{T}_0 \cdot \mathbf{T}_1) / (|\mathbf{T}_0| |\mathbf{T}_1|))$. In the illustration below, $\{P_0, \mathbf{T}_0, \mathbf{N}_0, \mathbf{B}_0\}$ becomes $\{P_1, \mathbf{T}_1, \mathbf{N}_1, \mathbf{B}_1\}$. As the curve becomes relatively straight the difference between \mathbf{T}_0 and \mathbf{T}_1 becomes small. If $\mathbf{T}_0 = \mathbf{T}_1$ their cross-product is undefined and no axis is available to perform the rotation; this is not a problem, because the rotation is zero.



Computing a reference frame from the previous frame.

\mathbf{N}_1 may be computed as a product of \mathbf{N}_0 with a rotation matrix, \mathbf{R} , that represents the rotation from the old to the new reference frame. Assuming \mathbf{T}_0 , \mathbf{T}_1 , and \mathbf{A} are of unit length,

$$\mathbf{R} = \begin{bmatrix} sq_x + (1 - sq_x) \cos & xycos1 + zsin & zxcos1 - ysin \\ xycos1 - zsin & sq_y + (1 - sq_y) \cos & yzcos1 + xsin \\ zxcos1 + ysin & yzcos1 - xsin & sq_z + (1 - sq_z) \cos \end{bmatrix},$$

where

$$sq_x = \mathbf{A}_x \mathbf{A}_x,$$

$$sq_y = \mathbf{A}_y \mathbf{A}_y,$$

$$sq_z = \mathbf{A}_z \mathbf{A}_z,$$

$$\cos = \mathbf{T0} \cdot \mathbf{T1},$$

$$\cos1 = 1 - \cos,$$

$$xycos1 = \mathbf{A}_x \mathbf{A}_y \cos1,$$

$$yzcos1 = \mathbf{A}_y \mathbf{A}_z \cos1,$$

$$zxcos1 = \mathbf{A}_x \mathbf{A}_z \cos1,$$

$$\sin = \text{sqrt}(1 - \cos^2),$$

$$xsin = \mathbf{A}_x \sin,$$

$$ysin = \mathbf{A}_y \sin, \text{ and}$$

$$zsin = \mathbf{A}_z \sin.$$

This formula requires one square root, 19 multiplications, and 15 additions.

Pat Hanrahan mentions an alternate computation of \mathbf{R} that utilizes two intermediate reference frames: $\{\mathbf{T0}, \mathbf{A}, \mathbf{A0}\}$ and $\{\mathbf{T1}, \mathbf{A}, \mathbf{A1}\}$, where $\mathbf{A0} = \mathbf{T0} \times \mathbf{A}$ and $\mathbf{A1} = \mathbf{T1} \times \mathbf{A}$; \mathbf{R} rotates $\{\mathbf{T0}, \mathbf{A}, \mathbf{A0}\}$ into $\{\mathbf{T1}, \mathbf{A}, \mathbf{A1}\}$. Assuming $\mathbf{T0}$, $\mathbf{T1}$, \mathbf{A} , $\mathbf{A0}$, and $\mathbf{A1}$ are of unit length (if $\mathbf{T0}$, $\mathbf{T1}$, and \mathbf{A} are of unit length, so will be $\mathbf{T0} \times \mathbf{A}$ and $\mathbf{T1} \times \mathbf{A}$), the intermediate reference frames can be represented as pure rotation matrices:

$$\mathbf{M0} = \begin{bmatrix} \mathbf{T0}_x & \mathbf{A}_x & \mathbf{A0}_x \\ \mathbf{T0}_y & \mathbf{A}_y & \mathbf{A0}_y \\ \mathbf{T0}_z & \mathbf{A}_z & \mathbf{A0}_z \end{bmatrix}; \mathbf{M1} = \begin{bmatrix} \mathbf{T1}_x & \mathbf{A}_x & \mathbf{A1}_x \\ \mathbf{T1}_y & \mathbf{A}_y & \mathbf{A1}_y \\ \mathbf{T1}_z & \mathbf{A}_z & \mathbf{A1}_z \end{bmatrix}.$$

\mathbf{R} is given by $\mathbf{M0}^{-1} \mathbf{M1}$; Since $\mathbf{M0}$ is a pure rotation, $\mathbf{M0}^{-1} = \mathbf{M0}^T$, and

$$\mathbf{R} = \begin{bmatrix} \mathbf{T0}_x & \mathbf{T0}_y & \mathbf{T0}_z \\ \mathbf{A}_x & \mathbf{A}_y & \mathbf{A}_z \\ \mathbf{A0}_x & \mathbf{A0}_y & \mathbf{A0}_z \end{bmatrix} \begin{bmatrix} \mathbf{T1}_x & \mathbf{A}_x & \mathbf{A1}_x \\ \mathbf{T1}_y & \mathbf{A}_y & \mathbf{A1}_y \\ \mathbf{T1}_z & \mathbf{A}_z & \mathbf{A1}_z \end{bmatrix}.$$

This formulation, including the computation of $\mathbf{A0}$ and $\mathbf{A1}$, requires one square root, 45 multiplications, and 26 additions.

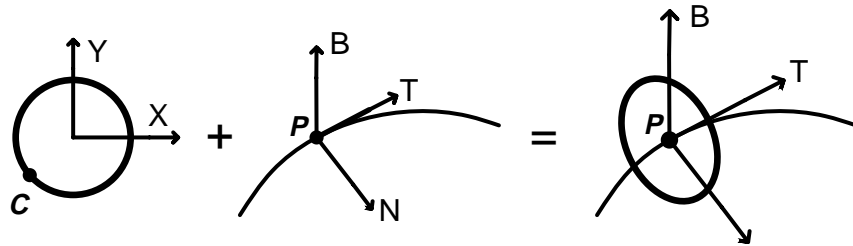
Ken Sloan suggests a simpler means to propagate reference frames:

$$\mathbf{N1} = \mathbf{B0} \times \mathbf{T1}$$

$$\mathbf{B1} = \mathbf{T1} \times \mathbf{N1}$$

Although this method produces correct results for $\mathbf{T1} = \mathbf{T0}$, it is subject to other degeneracies; for example, the results are undefined for $\mathbf{T1} = \mathbf{B0}$. Note that $\mathbf{B1} = \mathbf{T1} \times \mathbf{N1} = \mathbf{T1} \times (\mathbf{B0} \times \mathbf{T1}) = (\mathbf{T1} \cdot \mathbf{T1})\mathbf{B0} - (\mathbf{T1} \cdot \mathbf{B0})\mathbf{T1}$; that is, this method restricts $\mathbf{B1}$ to the plane defined by $\mathbf{B0}$ and $\mathbf{T1}$. The torsion minimizing method does not; thus, the two methods produce different results.

Although the tangent is needed to compute the reference frame, only the point P , normal \mathbf{N} and binormal \mathbf{B} are needed to transform the cross-section into the plane defined by \mathbf{N} and \mathbf{B} .



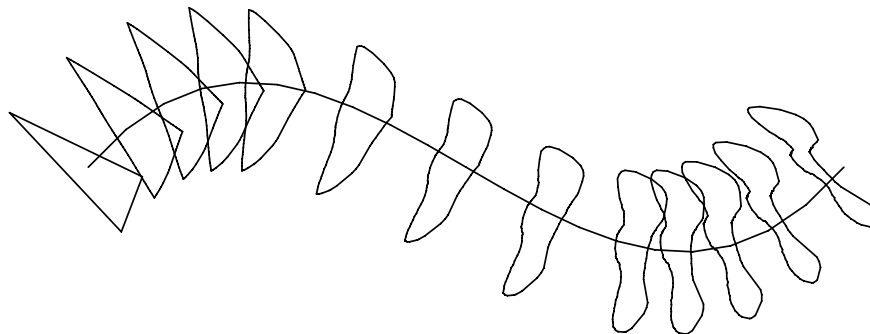
Positioning and orienting a cross-section.

If $\mathbf{C} = (C_x, C_y)$ is on the two-dimensional cross-section, $(P_x + C_x \mathbf{N}_x + C_y \mathbf{B}_x, P_y + C_x \mathbf{N}_y + C_y \mathbf{B}_y, P_z + C_x \mathbf{N}_z + C_y \mathbf{B}_z)$ is a three-dimensional point positioned on the surface of the generalized cylinder. This is conveniently expressed in matrix form:

$$\mathbf{P}_{surface} = [C_x \ C_y \ 1] [\mathbf{M}], \text{ where } \mathbf{M} = \begin{bmatrix} \mathbf{N}_x & \mathbf{N}_y & \mathbf{N}_z \\ \mathbf{B}_x & \mathbf{B}_y & \mathbf{B}_z \\ P_x & P_y & P_z \end{bmatrix}$$

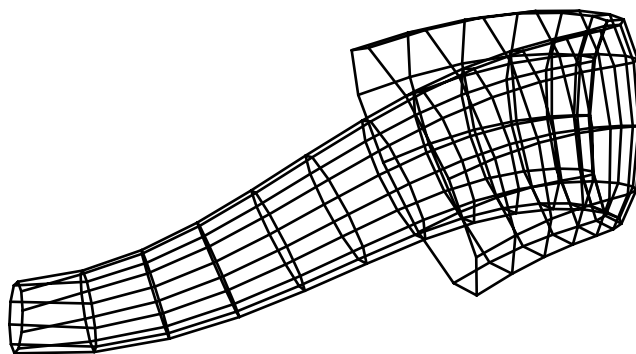
A cross-section at the beginning of a closed curve will not necessarily align with the cross-section at the end of the curve. Paul Heckbert observes that in such cases a corrective twist can be added to each frame.

Note that the results will depend on the distance between successive reference frames. Reference frames a small distance apart follow the path of the curve more closely. With large distances it is possible to miss turns of the curve; such an error is then propagated along the curve. Implementors may find it advantageous to create several intermediate reference frames in order to establish one at a desired location. Also, as shown in the following example, the number of reference frames can be increased at points of relatively high curvature.



A generalized cylinder with changing cross-sections.

In the example above, the cross-sections can change as they progress along the curve. The number of cross-sections can also depend on the change in cross-sections (whether radius or shape). Lastly, cross-sections may “double back” upon the curve; for example, a cross-section may be located according to its parametric position, t . t normally ranges monotonically from 0 through 1, but it might be a function of another parameter, $t = f(u)$, $u \in [0, 1]$. This enables the production of cylinders similar to the one below.



A generalized cylinder with $t = f(u)$.