Chap 3. Cubic Bézier Curves

3.1 Parametric Curves

- The graph of a function $y = 2x - x^2$ is a set of points $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 2x - x^2 \end{bmatrix}$.
- A parametric curve is of the form $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} f(t) \\ g(t) \end{bmatrix}$.
- A parametric curve for a straight line:
  $$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} (1-t)\, a_x + t\, b_x \\ (1-t)\, a_y + t\, b_y \end{bmatrix}$$

Ex 3.1

- The curve defined by $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} t \\ 2t - 2t^2 \end{bmatrix}$ is identical to the curve given as the graph of $y = 2x - 2x^2$

Ex 3.2

- By rotating the above curve by $90°$, we get a curve
  $$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -2t + 2t^2 \\ t \end{bmatrix}$$

An important difference between parametric curves and the graphs of functions:

① The concept of "zero slope" or "horizontal tangents" is important. It characterizes extreme points.

② But, for parametric curves, zero slopes do not signify geometric properties. A simple rotation changes horizontal tangents. The geometry of a curve does not change under rotations or other affine maps.

- A parametric curve in 3D is $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} f(t) \\ g(t) \\ h(t) \end{bmatrix}$

Ex: The helix is given by $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \cos(t) \\ \sin(t) \\ t \end{bmatrix}$

3.2 Cubic Bézier Curves

Ex 3.3

$$\mathbf{x}(t) = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -(1-t)^3 + t^3 \\ 3(1-t)^2 t - 3(1-t)t^2 \end{bmatrix}$$

$$= (1-t)^3 \begin{bmatrix} -1 \\ 0 \end{bmatrix} + 3(1-t)^2 t \begin{bmatrix} 0 \\ 1 \end{bmatrix} + 3(1-t)t^2 \begin{bmatrix} 0 \\ -1 \end{bmatrix} + t^3 \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

The polynomial curve is expressed in terms of a combination of points. We may compute $\mathbf{x}(0.5) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

A cubic Bézier curve is defined by

$$\mathbf{x}(t) = (1-t)^3 \mathbf{b}_0 + 3(1-t)^2 t\, \mathbf{b}_1 + 3(1-t)t^2 \mathbf{b}_2 + t^3 \mathbf{b}_3,$$

where $\mathbf{b}_i$, the Bézier control points, form the Bézier polygon of the curve.

$$\mathbf{x}(t) = \underbrace{B_0^3(t)}\cdot \mathbf{b}_0 + \underbrace{B_1^3(t)} \cdot \mathbf{b}_1 + \underbrace{B_2^3(t)} \cdot \mathbf{b}_2 + \underbrace{B_3^3(t)} \cdot \mathbf{b}_3.$$

$\hookrightarrow$ the cubic Bernstein polynomials.

Properties of Cubic Bézier Curves

1. Endpoint interpolation: $\mathbf{x}(0) = \mathbf{b}_0$, $\mathbf{x}(1) = \mathbf{b}_3$.
2. Symmetry: Two polygons $\mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ and $\mathbf{b}_3, \mathbf{b}_2, \mathbf{b}_1, \mathbf{b}_0$ describe the same curve; but the directions are different
3. Invariance under rotations:
4. Invariance under affine maps: If an affine map is applied to the control polygon, the curve is mapped by the same map.
5. Convex hull property: For $t \in [0,1]$, the point $\mathbf{x}(t)$ is in the convex hull of the control polygon.
6. Linear precision: If $\mathbf{b}_1 = \frac{2}{3}\mathbf{b}_0 + \frac{1}{3}\mathbf{b}_3$ and $\mathbf{b}_2 = \frac{1}{3}\mathbf{b}_0 + \frac{2}{3}\mathbf{b}_3$, then the curve $\mathbf{x}(t) = (1-t)\mathbf{b}_0 + t\,\mathbf{b}_3$: linear interpolation.

o. For $t \notin [0,1]$, $\mathbf{x}(t)$ may not stay within the convex hull of the control polygon.

3.3 Derivatives

$$\dot{\mathbf{x}}(t) = -3(1-t)^2\mathbf{b}_0 + [3(1-t)^2 - 6(1-t)t]\mathbf{b}_1 + [6(1-t)t - 3t^2]\mathbf{b}_2 + 3t^2\mathbf{b}_3$$

$$= (1-t)^2\cdot 3[\mathbf{b}_1 - \mathbf{b}_0] + 2(1-t)t\cdot 3[\mathbf{b}_2 - \mathbf{b}_1] + t^2\cdot 3[\mathbf{b}_3 - \mathbf{b}_2]$$

$$= (1-t)^2\cdot 3\Delta\mathbf{b}_0 + 2(1-t)t\cdot 3\Delta\mathbf{b}_1 + t^2\cdot 3\Delta\mathbf{b}_2$$

↳ the forward difference

$$= 3\left(B_0^2(t)\cdot\Delta\mathbf{b}_0 + B_1^2(t)\cdot\Delta\mathbf{b}_1 + B_2^2(t)\cdot\Delta\mathbf{b}_2\right)$$

↳ the quadratic Bernstein basis functions

- For parametric curves, "derivative curves" produce vectors rather than points.
- The coefficients are the difference vectors of the polygon, scaled by 3, the degree of the curve.
- $\dot{\mathbf{x}}(0) = 3\Delta\mathbf{b}_0$, $\dot{\mathbf{x}}(1) = 3\Delta\mathbf{b}_2$

3.4 The de Casteljau Algorithm

$$\begin{cases} \mathbf{b}_0^1(t) = (1-t)\mathbf{b}_0 + t\,\mathbf{b}_1 \\ \mathbf{b}_1^1(t) = (1-t)\mathbf{b}_1 + t\,\mathbf{b}_2 \\ \mathbf{b}_2^1(t) = (1-t)\mathbf{b}_2 + t\,\mathbf{b}_3 \end{cases} \Rightarrow \begin{cases} \mathbf{b}_0^2(t) = (1-t)\mathbf{b}_0^1(t) + t\,\mathbf{b}_1^1(t) \\ \mathbf{b}_1^2(t) = (1-t)\mathbf{b}_1^1(t) + t\,\mathbf{b}_2^1(t) \end{cases}$$

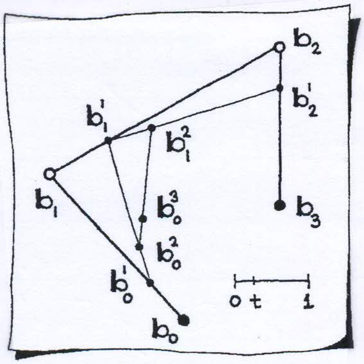
$$\Rightarrow \underbrace{\mathbf{b}_0^3(t)}_{=\,\mathbf{x}(t)} = (1-t)\mathbf{b}_0^2(t) + t\,\mathbf{b}_1^2(t).$$

a convenient schematic tool for the algorithm

$$\begin{array}{llll} \mathbf{b}_0 & & & \\ & \mathbf{b}_0^1(t) & & \\ \mathbf{b}_1 & & \mathbf{b}_0^2(t) & \\ & \mathbf{b}_1^1(t) & & \mathbf{b}_0^3(t) \\ \mathbf{b}_2 & & \mathbf{b}_1^2(t) & \\ & \mathbf{b}_2^1(t) & & \\ \mathbf{b}_3 & & & \end{array}$$

- In the implementation, $\mathbf{b}_0^1$ is calculated and loaded into $\mathbf{b}_0$ since $\mathbf{b}_0$ is never needed.
  ⇒ 1D array of control points is sufficient!

- $\dot{\mathbf{x}}(t) = 3[\mathbf{b}_1^2(t) - \mathbf{b}_0^2(t)]$:
  The derivative is essentially a byproduct of point evaluation!



Sketch 27.
The de Casteljau algorithm.

3.5 Subdivision

- Two polygons $\mathbf{b}_0, \mathbf{b}_0^1, \mathbf{b}_0^2, \mathbf{b}_0^3$ and $\mathbf{b}_0^3, \mathbf{b}_1^2, \mathbf{b}_2^1, \mathbf{b}_3$ define two segments corresponding to $[0,t]$ and $[t,1]$, respectively.
- Subdivision at $t=0.5$ splits the curve at the parameter midpoint, but the two arcs are not of equal length.
- Subdivision may be repeated:
  Each of the two new control polygons may be subdivided, ...
  The resulting sequence of control polygons converges to the curve.
- Another application is the intersection of a curve with a line.
  1. Find the AABB (axis aligned bounding box) of the polygon.
  2. If no intersection between the AABB and the line, EXIT.
  3. Else if the AABB is smaller than ε,
     report the center of the AABB.
  4. Else, subdivide the curve at $t=0.5$ into two and repeat the same procedure to each segment.
- The curve-curve intersection can be done in a similar way. In this case, the curve with a bigger AABB is subdivided.

3.6 Exploring the Properties of Bézier Curves

In this section, we will study some special Bézier curves—they will help highlight some of the important properties of this type of curve.

EXAMPLE 3.6

Let a Bézier curve be given by

$$\mathbf{b}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \mathbf{b}_1 = \begin{bmatrix} 1.5 \\ 1 \end{bmatrix}, \quad \mathbf{b}_2 = \begin{bmatrix} -0.5 \\ 1 \end{bmatrix}, \quad \mathbf{b}_3 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

It is shown in Figure 3.8. This Bézier curve has a *loop*; i.e., it self-intersects.

EXAMPLE 3.7

Let a Bézier curve be given by

$$\mathbf{b}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \mathbf{b}_1 = \begin{bmatrix} 0.7 \\ 1 \end{bmatrix}, \quad \mathbf{b}_2 = \begin{bmatrix} 0.3 \\ 1 \end{bmatrix}, \quad \mathbf{b}_3 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

It is shown in Figure 3.9. In that figure, two extra points on the curve are marked: These are *inflection points*[7]. A cubic with *two* inflection points? That does not happen for functions, but for parametric cubics, it is possible!

EXAMPLE 3.8

Next, let a Bézier curve be given by

$$\mathbf{b}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \mathbf{b}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{b}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \mathbf{b}_3 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

It is shown in Figure 3.10. At $t = 0.5$, the curve has a *cusp*. These are points where the first derivative vector vanishes.

The next example continues the investigation of our cusp curve.

EXAMPLE 3.9

Referring to the previous example, let us subdivide that curve at $t = 0.5$. The corresponding de Casteljau algorithm is given by:

$$\begin{array}{cccc}
\begin{bmatrix} 0 \\ 0 \end{bmatrix} & & & \\
\begin{bmatrix} 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} & & \\
\begin{bmatrix} 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0.5 \\ 1 \end{bmatrix} & \begin{bmatrix} 0.5 \\ 0.75 \end{bmatrix} & \\
\begin{bmatrix} 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} & \begin{bmatrix} 0.5 \\ 0.75 \end{bmatrix} & \begin{bmatrix} 0.5 \\ 0.75 \end{bmatrix}
\end{array}$$

The Bézier points $\hat{\mathbf{b}}_i$ of the segment corresponding to $t \in [0, 0.5]$ of the original curve are thus given by

$$\hat{\mathbf{b}}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \hat{\mathbf{b}}_1 = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, \quad \hat{\mathbf{b}}_2 = \begin{bmatrix} 0.5 \\ 0.75 \end{bmatrix}, \quad \hat{\mathbf{b}}_3 = \begin{bmatrix} 0.5 \\ 0.75 \end{bmatrix}.$$

The last two of these Bézier points are identical, hence the cusp.

3.7 The Matrix Form and Monomials

As a preparation for what is to follow, let us rewrite (3.3) using the formalism of dot products. It then becomes

$$\mathbf{b}(t) = \begin{bmatrix} \mathbf{b}_0 & \mathbf{b}_1 & \mathbf{b}_2 & \mathbf{b}_3 \end{bmatrix} \begin{bmatrix} (1-t)^3 \\ 3(1-t)^2 t \\ 3(1-t)t^2 \\ t^3 \end{bmatrix},$$

or taking advantage of the shorthand basis function notation in (3.4)

$$\mathbf{b}(t) = \begin{bmatrix} \mathbf{b}_0 & \mathbf{b}_1 & \mathbf{b}_2 & \mathbf{b}_3 \end{bmatrix} \begin{bmatrix} B_0^3(t) \\ B_1^3(t) \\ B_2^3(t) \\ B_3^3(t) \end{bmatrix}. \tag{3.11}$$

This is the matrix form of a Bézier curve.

Polynomials were traditionally thought of as combinations of the *monomial polynomials* or *monomials*; they are $1, t, t^2, t^3$ for the cubic case. Equation (3.3) may be rewritten in this form:

$$\mathbf{b}(t) = \mathbf{b}_0 + 3t(\mathbf{b}_1 - \mathbf{b}_0) + 3t^2(\mathbf{b}_2 - 2\mathbf{b}_1 + \mathbf{b}_0) + t^3(\mathbf{b}_3 - 3\mathbf{b}_2 + 3\mathbf{b}_1 - \mathbf{b}_0). \tag{3.12}$$

This allows a more concise formulation using matrices:

$$\mathbf{b}(t) = \begin{bmatrix} \mathbf{b}_0 & \mathbf{b}_1 & \mathbf{b}_2 & \mathbf{b}_3 \end{bmatrix} \begin{bmatrix} 1 & -3 & 3 & -1 \\ 0 & 3 & -6 & 3 \\ 0 & 0 & 3 & -3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ t \\ t^2 \\ t^3 \end{bmatrix}. \tag{3.13}$$

Equation (3.13) shows how to write a Bézier curve in monomial form. A curve in monomial form looks like this:

$$\mathbf{b}(t) = \mathbf{a}_0 + \mathbf{a}_1 t + \mathbf{a}_2 t^2 + \mathbf{a}_3 t^3.$$

Rewritten using the dot product form, this becomes

$$\mathbf{b}(t) = \begin{bmatrix} \mathbf{a}_0 & \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{bmatrix} \begin{bmatrix} 1 \\ t \\ t^2 \\ t^3 \end{bmatrix}.$$

Thus the monomial coefficients $\mathbf{a}_i$ are defined as

$$\begin{bmatrix} \mathbf{a}_0 & \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{b}_0 & \mathbf{b}_1 & \mathbf{b}_2 & \mathbf{b}_3 \end{bmatrix} \begin{bmatrix} 1 & -3 & 3 & -1 \\ 0 & 3 & -6 & 3 \\ 0 & 0 & 3 & -3 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{3.14}$$

Reviewing (3.12), it becomes apparent that the monomial form $\mathbf{a}_i$ have a different geometric interpretation than the Bézier form's $\mathbf{b}_i$. Sketch 28 illustrates that $\mathbf{a}_0$ is a point, however $\mathbf{a}_1, \mathbf{a}_2$, and $\mathbf{a}_3$ are vectors defining the derivatives of the cubic curve at $\mathbf{a}_0$.

The inverse process is not hard either: Given a curve in monomial form, how can we write it as a Bézier curve? Simply rearrange (3.14) to solve for the $\mathbf{b}_i$:

$$\begin{bmatrix} \mathbf{b}_0 & \mathbf{b}_1 & \mathbf{b}_2 & \mathbf{b}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{a}_0 & \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{bmatrix} \begin{bmatrix} 1 & -3 & 3 & -1 \\ 0 & 3 & -6 & 3 \\ 0 & 0 & 3 & -3 \\ 0 & 0 & 0 & 1 \end{bmatrix}^{-1}.$$

A matrix inversion is all that is needed here!

Notice that the square matrix in this equation is nonsingular. Because of its nonsingularity, we can conclude that any cubic curve can be written in either the Bézier or the monomial form.