

# 4

## Hermite Interpolation

The curve and surface methods of the preceding chapters are based on points. Using polynomials, it is easy to construct a parametric curve segment (or surface patch) that passes through a given one-dimensional array or two-dimensional grid of points.

The downside of these methods is that they are not interactive. If the resulting curve or surface isn't the one the designer wants, the only way to modify it is to add points. Moving the points is not an option because the curve has to pass through the original data points. Adding points provides some control over the shape of the curve, but slows down the computations.

A practical, useful curve/surface design algorithm should be interactive. It should provide user-controlled parameters that modify the shape of the curve in a predictable, intuitive way. The Hermite interpolation approach, the topic of this chapter, is such a method.

Hermite interpolation is based on two points  $\mathbf{P}_1$  and  $\mathbf{P}_2$  and two tangent vectors  $\mathbf{P}_1^t$  and  $\mathbf{P}_2^t$ . It computes a curve segment that starts at  $\mathbf{P}_1$ , going in direction  $\mathbf{P}_1^t$  and ends at  $\mathbf{P}_2$  moving in direction  $\mathbf{P}_2^t$ . Before delving into the details, the reader may find it useful to peruse Figure 4.1 where several such curves are shown, with their endpoints and extreme tangent vectors.

It is obvious that a single Hermite segment can take on many different shapes. It can even have a cusp and can develop a loop. A complete curve, however, normally requires several segments connected with  $C^0$ ,  $C^1$ , or  $C^2$  continuities, as illustrated in Section 1.4.2. Spline methods for constructing such a curve are discussed in Chapter 5.

The method is called Hermite interpolation after Charles Hermite who developed it and derived its blending functions in the 1870s, as part of his work on approximation and interpolation. He was not concerned with the computation of curves and surfaces (and was actually known to hate geometry), and developed his method as a way to interpolate any mathematical quantity from an initial value to a final value given the rates of change of the quantity at the start and at the end.

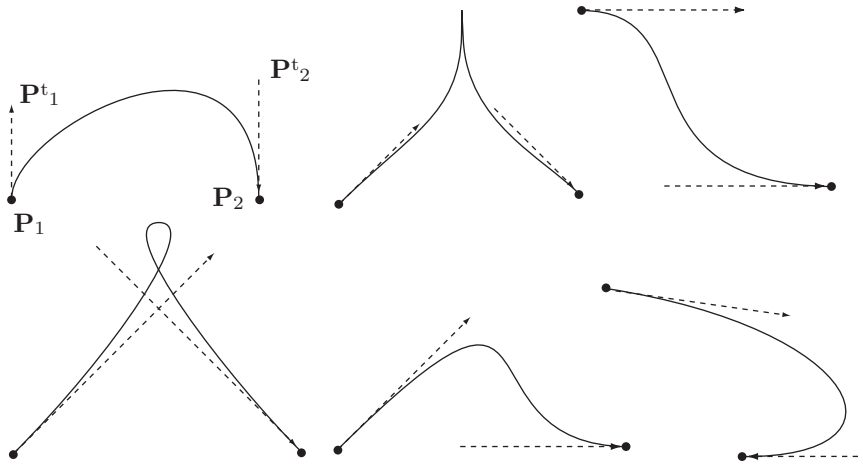


Figure 4.1: Various Hermite Curve Segments.

[Hermite] had a kind of positive hatred of geometry and once curiously reproached me with having made a geometrical memoir.

—Jacques Hadamard.

## 4.1 Interactive Control

Hermite interpolation has an important advantage; it is interactive. If a Hermite curve segment has a wrong shape, the user can edit it by modifying the tangent vectors.

- ◇ **Exercise 4.1:** In the case of a four-point PC, we can change the shape of the curve by moving the points. Why then is the four-point method considered noninteractive?

Figure 4.1 illustrates how the shape of the curve depends on the directions of the tangent vectors. Figure 4.2 shows how the curve can be edited by modifying the magnitudes of those vectors. The figure shows three curves that start in a  $45^\circ$  direction and end up going vertically down. The effect illustrated here is simple. As the magnitude of the start tangent increases, the curve continues longer in the original direction. This behavior implies that short tangents produce a curve that changes its direction early and starts moving straight toward the final point. Such a curve is close to a straight segment, so we conclude that a long tangent results in a loose curve and a short tangent produces a tight curve (see also exercise 4.7).

The reason the magnitudes, and not just the directions, of the tangents affect the shape of the curve is that the three-dimensional Hermite segment is a PC and calculating a PC involves four coefficients, each a triplet, for a total of 12 unknown numbers. The two endpoints supply six known quantities and the two tangents should supply the remaining six. However, if we consider only the direction of a vector and not its magnitude, then the vectors  $(1, 0.5, 0.3)$ ,  $(2, 1, 0.6)$ , and  $(4, 2, 1.2)$  are all equal. In such a case, only two of the three vector components are independent and two vectors supply only four independent quantities.

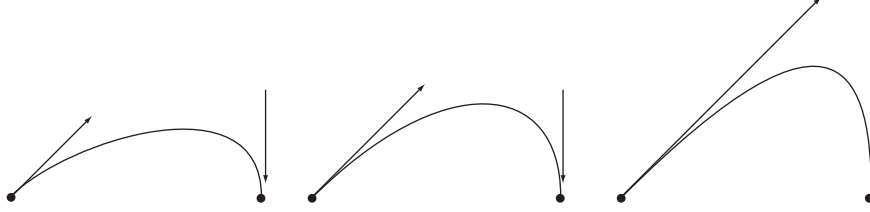


Figure 4.2: Effects of Varying the Tangent's Magnitude.

◇ **Exercise 4.2:** Discuss this claim in detail.

A sketch tells as much in a glance as a dozen pages of print.  
—Ivan Turgenev, *Fathers and Sons* (1862).

## 4.2 The Hermite Curve Segment

The Hermite curve segment is easy to derive. It is a PC curve (a degree-3 polynomial in  $t$ ) with four coefficients that depend on the two points and two tangents. The basic equation of a PC curve is Equation (3.1) duplicated here

$$\mathbf{P}(t) = \mathbf{a}t^3 + \mathbf{b}t^2 + \mathbf{c}t + \mathbf{d} = (t^3, t^2, t, 1)(\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d})^T = \mathbf{T}(t)\mathbf{A}. \quad (3.1)$$

This is the *algebraic representation* of the curve, in which the four coefficients are still unknown. Once these coefficients are expressed in terms of the known quantities, which are geometric, the curve will be expressed geometrically.

The tangent vector to a curve  $\mathbf{P}(t)$  is the derivative  $d\mathbf{P}(t)/dt$ , which we denote by  $\mathbf{P}^t(t)$ . The tangent vector of a PC curve is therefore

$$\mathbf{P}^t(t) = 3\mathbf{a}t^2 + 2\mathbf{b}t + \mathbf{c}. \quad (4.1)$$

We denote the two given points by  $\mathbf{P}_1$  and  $\mathbf{P}_2$  and the two given tangents by  $\mathbf{P}_1^t$  and  $\mathbf{P}_2^t$ . The four quantities are now used to calculate the *geometric representation* of the PC by writing equations that relate the four unknown coefficients  $\mathbf{a}$ ,  $\mathbf{b}$ ,  $\mathbf{c}$ , and  $\mathbf{d}$  to the four known ones,  $\mathbf{P}_1$ ,  $\mathbf{P}_2$ ,  $\mathbf{P}_1^t$ , and  $\mathbf{P}_2^t$ . The equations are  $\mathbf{P}(0) = \mathbf{P}_1$ ,  $\mathbf{P}(1) = \mathbf{P}_2$ ,  $\mathbf{P}^t(0) = \mathbf{P}_1^t$ , and  $\mathbf{P}^t(1) = \mathbf{P}_2^t$  [compare with Equations (3.2)]. Their explicit forms are

$$\begin{aligned} \mathbf{a} \cdot 0^3 + \mathbf{b} \cdot 0^2 + \mathbf{c} \cdot 0 + \mathbf{d} &= \mathbf{P}_1, \\ \mathbf{a} \cdot 1^3 + \mathbf{b} \cdot 1^2 + \mathbf{c} \cdot 1 + \mathbf{d} &= \mathbf{P}_2, \\ 3\mathbf{a} \cdot 0^2 + 2\mathbf{b} \cdot 0 + \mathbf{c} &= \mathbf{P}_1^t, \\ 3\mathbf{a} \cdot 1^2 + 2\mathbf{b} \cdot 1 + \mathbf{c} &= \mathbf{P}_2^t. \end{aligned} \quad (4.2)$$

They are easy to solve and the solutions are

$$\mathbf{a} = 2\mathbf{P}_1 - 2\mathbf{P}_2 + \mathbf{P}_1^t + \mathbf{P}_2^t, \quad \mathbf{b} = -3\mathbf{P}_1 + 3\mathbf{P}_2 - 2\mathbf{P}_1^t - \mathbf{P}_2^t, \quad \mathbf{c} = \mathbf{P}_1^t, \quad \mathbf{d} = \mathbf{P}_1. \quad (4.3)$$

Substituting these solutions into Equation (3.1) gives

$$\mathbf{P}(t) = (2\mathbf{P}_1 - 2\mathbf{P}_2 + \mathbf{P}_1^t + \mathbf{P}_2^t)t^3 + (-3\mathbf{P}_1 + 3\mathbf{P}_2 - 2\mathbf{P}_1^t - \mathbf{P}_2^t)t^2 + \mathbf{P}_1^t t + \mathbf{P}_1, \quad (4.4)$$

which, after rearranging, becomes

$$\begin{aligned} \mathbf{P}(t) &= (2t^3 - 3t^2 + 1)\mathbf{P}_1 + (-2t^3 + 3t^2)\mathbf{P}_2 + (t^3 - 2t^2 + t)\mathbf{P}_1^t + (t^3 - t^2)\mathbf{P}_2^t \\ &= F_1(t)\mathbf{P}_1 + F_2(t)\mathbf{P}_2 + F_3(t)\mathbf{P}_1^t + F_4(t)\mathbf{P}_2^t \\ &= (F_1(t), F_2(t), F_3(t), F_4(t))(\mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_1^t, \mathbf{P}_2^t)^T \\ &= \mathbf{F}(t)\mathbf{B}, \end{aligned} \quad (4.5)$$

where

$$\begin{aligned} F_1(t) &= (2t^3 - 3t^2 + 1), & F_2(t) &= (-2t^3 + 3t^2) = 1 - F_1(t), \\ F_3(t) &= (t^3 - 2t^2 + t), & F_4(t) &= (t^3 - t^2), \end{aligned} \quad (4.6)$$

$\mathbf{B}$  is the column  $(\mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_1^t, \mathbf{P}_2^t)^T$ , and  $\mathbf{F}(t)$  is the row  $(F_1(t), F_2(t), F_3(t), F_4(t))$ . Equations (4.4) and (4.5) are the geometric representation of the Hermite PC segment.

Functions  $F_i(t)$  are the Hermite blending functions. They create any point on the curve as a blend of the four given quantities. They are shown in Figure 4.3. Note that  $F_1(t) + F_2(t) \equiv 1$ . These two functions blend points, not tangent vectors, and should therefore be barycentric. We can also write  $F_1(t) = (t^3, t^2, t, 1)(2, -3, 0, 1)^T$  and similarly for  $F_2(t)$ ,  $F_3(t)$ , and  $F_4(t)$ . In matrix notation this becomes

$$\mathbf{F}(t) = (t^3, t^2, t, 1) \begin{pmatrix} 2 & -2 & 1 & 1 \\ -3 & 3 & -2 & -1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} = \mathbf{T}(t) \mathbf{H}.$$

The curve can now be written

$$\mathbf{P}(t) = \mathbf{F}(t)\mathbf{B} = \mathbf{T}(t) \mathbf{H} \mathbf{B} = (t^3, t^2, t, 1) \begin{pmatrix} 2 & -2 & 1 & 1 \\ -3 & 3 & -2 & -1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{P}_1 \\ \mathbf{P}_2 \\ \mathbf{P}_1^t \\ \mathbf{P}_2^t \end{pmatrix}. \quad (4.7)$$

Equation (3.1) tells us that  $\mathbf{P}(t) = \mathbf{T}(t) \mathbf{A}$ , which implies  $\mathbf{A} = \mathbf{H} \mathbf{B}$ . Matrix  $\mathbf{H}$  is called the Hermite basis matrix.

The following is *Mathematica* code to display a single Hermite curve segment.

```
Clear[T,H,B]; (* Hermite Interpolation *)
T={t^3,t^2,t,1};
H={{2,-2,1,1},{-3,3,-2,-1},{0,0,1,0},{1,0,0,0}};
B={{0,0},{2,1},{1,1},{1,0}};
ParametricPlot[T.H.B,{t,0,1},PlotRange->All]
```

- ◇ **Exercise 4.3:** Express the midpoint  $\mathbf{P}(0.5)$  of a Hermite segment in terms of the two endpoints and two tangent vectors. Draw a diagram to illustrate the geometric interpretation of the result.

### 4.2.1 Hermite Blending Functions

The four Hermite blending functions of Equation (4.6) are illustrated graphically in Figure 4.3. An analysis of these functions is essential for a thorough understanding of the Hermite interpolation method.

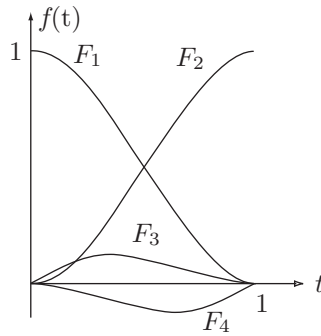


Figure 4.3: Hermite Weight Functions

Function  $F_1(t)$  is the weight assigned to the start point  $\mathbf{P}_1$ . It goes down from its maximum  $F_1(0) = 1$  to  $F_1(1) = 0$ . This shows why for small values of  $t$  the curve is close to  $\mathbf{P}_1$  and why  $\mathbf{P}_1$  has little or no influence on the curve for large values of  $t$ . The opposite is true for  $F_2(t)$ , the weight of the endpoint  $\mathbf{P}_2$ . Function  $F_3(t)$  is a bit trickier. It starts at zero, has a maximum at  $t = 1/3$ , then drops slowly back to zero. This behavior is interpreted as follows:

1. For small values of  $t$ , function  $F_3(t)$  has almost no effect. The curve stays close to  $\mathbf{P}_1$  regardless of the extreme tangents or anything else.

2. For  $t$  values around  $1/3$ , weight  $F_3(t)$  exerts some influence on the curve. For these  $t$  values, weight  $F_4(t)$  is small, and the curve is (approximately) the sum of (1) point  $F_1(t)\mathbf{P}_1$  (large contribution), (2) point  $F_2(t)\mathbf{P}_2$  (small contribution), and (3) vector  $F_3(t)\mathbf{P}_1^t$ . The sum of a point  $\mathbf{P} = (x, y)$  and a vector  $\mathbf{v} = (v_x, v_y)$  is a point located at  $(x + v_x, y + v_y)$ , which is how weight  $F_3(t)$  “pulls” the curve in the direction of tangent vector  $\mathbf{P}_1^t$ .

3. For large  $t$  values, function  $F_3(t)$  again has almost no effect. The curve moves closer to  $\mathbf{P}_2$  because weight  $F_2(t)$  becomes dominant.

Function  $F_4(t)$  is interpreted in a similar way. It has almost no effect for small and for large values of  $t$ . Its maximum (actually, minimum, because it is negative) occurs at  $t = 2/3$ , so it affects the curve only in this region. For  $t$  values close to  $2/3$ , the curve is the sum of point  $F_2(t)\mathbf{P}_2$  (large contribution), point  $F_1(t)\mathbf{P}_1$  (small contribution), and vector  $-|F_4(t)|\mathbf{P}_2^t$ . Because  $F_4(t)$  is negative, this sum is equivalent to  $(x - v_x, y - v_y)$ , which is why the curve approaches endpoint  $\mathbf{P}_2$  while moving in direction  $\mathbf{P}_2^t$ .

Another important feature of the Hermite weight functions is that  $F_1(t)$  and  $F_2(t)$  are barycentric. They have to be, since they blend two points, and a detailed look at the four equations (4.2) explains why they are. The first of these equations is simply  $\mathbf{d} = \mathbf{P}_1$ , which reduces the second one to  $\mathbf{a} + \mathbf{b} + \mathbf{c} + \mathbf{d} = \mathbf{P}_2$  or  $\mathbf{a} + \mathbf{b} + \mathbf{c} = \mathbf{P}_2 - \mathbf{P}_1$ . The third equation solves  $\mathbf{c}$ , and the fourth equation, combined with the second equation, is finally used to compute  $\mathbf{a}$  and  $\mathbf{b}$ . All this implies that  $\mathbf{a}$  and  $\mathbf{b}$  have the form  $\mathbf{a} = \alpha(\mathbf{P}_2 - \mathbf{P}_1) + \dots$ ,  $\mathbf{b} = \beta(\mathbf{P}_2 - \mathbf{P}_1) + \dots$ . The final PC therefore has the form

$$\mathbf{P}(t) = \mathbf{a}t^3 + \mathbf{b}t^2 + \mathbf{c}t + \mathbf{d} = (\alpha\mathbf{P}_2 - \alpha\mathbf{P}_1 + \dots)t^3 + (\beta\mathbf{P}_2 - \beta\mathbf{P}_1 + \dots)t^2 + (\dots)t + \mathbf{P}_1,$$

where the ellipsis represent parts that depend only on the tangent vectors, not on the endpoints. When this is rearranged, the result is

$$\mathbf{P}(t) = (-\alpha t^3 - \beta t^2 + 1)\mathbf{P}_1 + (\alpha t^3 + \beta t^2)\mathbf{P}_2 + (\dots)\mathbf{P}_1^t + (\dots)\mathbf{P}_2^t,$$

which is why the coefficients of  $\mathbf{P}_1$  and  $\mathbf{P}_2$  add up to unity.

#### 4.2.2 Hermite Derivatives

The concept of blending can be applied to the calculation of the derivatives of a curve, not just to the curve itself. One way to calculate  $\mathbf{P}^t(t)$  is to differentiate  $\mathbf{T}(t) = (t^3, t^2, t, 1)$ . The result is

$$\mathbf{P}^t(t) = \mathbf{T}^t(t)\mathbf{H}\mathbf{B} = (3t^2, 2t, 1, 0)\mathbf{H}\mathbf{B}.$$

A more general method is to use the relation  $\mathbf{P}(t) = \mathbf{F}(t)\mathbf{B}$ , which implies

$$\mathbf{P}^t(t) = \mathbf{F}^t(t)\mathbf{B} = (F_1^t(t), F_2^t(t), F_3^t(t), F_4^t(t))\mathbf{B}.$$

The individual derivatives  $F_i^t(t)$  can be obtained from Equation (4.6). The results can be expressed as

$$\mathbf{P}^t(t) = (t^3, t^2, t, 1) \begin{bmatrix} 0 & 0 & 0 & 0 \\ 6 & -6 & 3 & 3 \\ -6 & 6 & -4 & -2 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{P}_1 \\ \mathbf{P}_2 \\ \mathbf{P}_1^t \\ \mathbf{P}_2^t \end{bmatrix} = \mathbf{T}(t)\mathbf{H}_t\mathbf{B}. \quad (4.8)$$

Similarly, the second derivatives of the Hermite segment can be expressed as

$$\mathbf{P}^{tt}(t) = (t^3, t^2, t, 1) \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 12 & -12 & 6 & 6 \\ -6 & 6 & -4 & -2 \end{bmatrix} \begin{bmatrix} \mathbf{P}_1 \\ \mathbf{P}_2 \\ \mathbf{P}_1^t \\ \mathbf{P}_2^t \end{bmatrix} = \mathbf{T}(t)\mathbf{H}_{tt}\mathbf{B}. \quad (4.9)$$

These expressions make it easy to calculate the first and second derivatives at any point on a Hermite segment. Similar expressions can be derived for any other curves that are based on the blending of geometrical quantities.

◇ **Exercise 4.4:** What is  $H_{ttt}$ ?

**Example:** The two two-dimensional points  $\mathbf{P}_1 = (0, 0)$  and  $\mathbf{P}_2 = (1, 0)$  and the two tangents  $\mathbf{P}_1^t = (1, 1)$  and  $\mathbf{P}_2^t = (0, -1)$  are given. The segment should therefore start at the origin, going in a  $45^\circ$  direction, and end at point  $(1, 0)$ , going straight down. The calculation of  $\mathbf{P}(t)$  is straightforward:

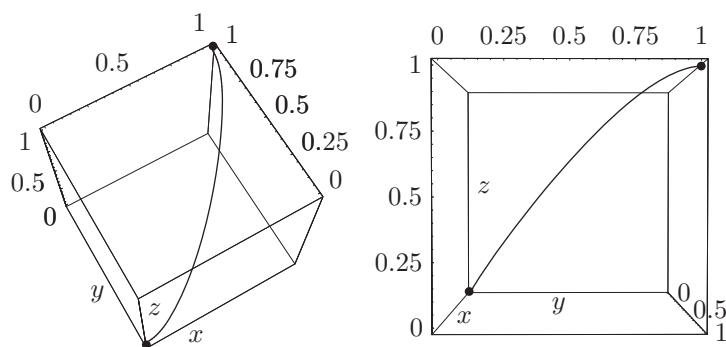
$$\begin{aligned}
 \mathbf{P}(t) &= \mathbf{T}(t) \mathbf{A} = \mathbf{T}(t) \mathbf{H} \mathbf{B} \\
 &= (t^3, t^2, t, 1) \begin{bmatrix} 2 & -2 & 1 & 1 \\ -3 & 3 & -2 & -1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} (0, 0) \\ (1, 0) \\ (1, 1) \\ (0, -1) \end{bmatrix} \\
 &= (t^3, t^2, t, 1) \begin{bmatrix} 2(0, 0) - 2(1, 0) + 1(1, 1) + 1(0, -1) \\ -3(0, 0) + 3(1, 0) - 2(1, 1) - 1(0, -1) \\ 0(0, 0) + 0(1, 0) + 1(1, 1) + 0(0, -1) \\ 1(0, 0) + 0(1, 0) + 0(1, 1) + 0(0, -1) \end{bmatrix} \\
 &= (t^3, t^2, t, 1) \begin{bmatrix} (-1, 0) \\ (1, -1) \\ (1, 1) \\ (0, 0) \end{bmatrix} \\
 &= (-1, 0)t^3 + (1, -1)t^2 + (1, 1)t. \tag{4.10}
 \end{aligned}$$

- ◇ **Exercise 4.5:** Use Equation (4.10) to show that the segment really passes through points  $(0, 0)$  and  $(1, 0)$ . Calculate the tangent vectors and use them to show that the segment really starts and ends in the right directions.
- ◇ **Exercise 4.6:** Repeat the example above with  $\mathbf{P}_1^t = (2, 2)$ . The new curve segment should go through the same points, in the same directions. However, it should continue longer in the original  $45^\circ$  direction, since the size of the new tangent is  $\sqrt{2^2 + 2^2} = 2\sqrt{2}$ , twice as long as the previous one, which is  $\sqrt{1^2 + 1^2} = \sqrt{2}$ .
- ◇ **Exercise 4.7:** Calculate the Hermite curve for two given points  $\mathbf{P}_1$  and  $\mathbf{P}_2$  assuming that the tangent vectors at the two points are zero (indeterminate). What kind of a curve is this?
- ◇ **Exercise 4.8:** Use the Hermite method to calculate PC segments for the cases where the known quantities are as follows:
1. The three tangent vectors at the start, middle, and end of the segment.
  2. The two interior points  $\mathbf{P}(1/3)$  and  $\mathbf{P}(2/3)$ , and the two extreme tangent vectors  $\mathbf{P}^t(0)$  and  $\mathbf{P}^t(1)$ .
  3. The two extreme points  $\mathbf{P}(0)$  and  $\mathbf{P}(1)$ , and the two interior tangent vectors  $\mathbf{P}^t(1/3)$  and  $\mathbf{P}^t(2/3)$  (this is similar to case 2, so it's easy).

**Example:** Given the two three-dimensional points  $\mathbf{P}_1 = (0, 0, 0)$  and  $\mathbf{P}_2 = (1, 1, 1)$  and the two tangent vectors  $\mathbf{P}_1^t = (1, 0, 0)$  and  $\mathbf{P}_2^t = (0, 1, 0)$ , the curve segment is the

simple cubic polynomial shown in Figure 4.4

$$\begin{aligned} \mathbf{P}(t) &= (t^3, t^2, t, 1) \begin{bmatrix} 2 & -2 & 1 & 1 \\ -3 & 3 & -2 & -1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} (0, 0, 0) \\ (1, 1, 1) \\ (1, 0, 0) \\ (0, 1, 0) \end{bmatrix} \\ &= (-t^3 + t^2 + t, -t^3 + 2t^2, -2t^3 + 3t^2). \end{aligned} \quad (4.11)$$



```
<<:Graphics:ParametricPlot3D.m; (* Hermite 3D example *)
Clear[T,H,B];
T={t^3,t^2,t,1};
H={{2,-2,1,1},{-3,3,-2,-1},{0,0,1,0},{1,0,0,0}};
B={{0,0,0},{1,1,1},{1,0,0},{0,1,0}};
ParametricPlot3D[T.H.B,{t,0,1}, Compiled->False,
  ViewPoint->{-0.846, -1.464, 3.997}, DefaultFont->{"cmr10", 10}];
(* ViewPoint->{3.119, -0.019, 0.054} alt view *)
```

Figure 4.4: A Hermite Curve Segment in Space.

I'm retired—goodbye tension, hello pension!  
—Anonymous.

### 4.2.3 Hermite Segments With Tension

This section shows how to create a Hermite curve segment under tension by employing a nonuniform Hermite segment. Such a segment is obtained when the parameter  $t$  varies in the interval  $[0, \Delta]$ , where  $\Delta$  can be any real positive number. The derivation of this case is similar to the uniform case. Equation (4.2) becomes

$$\begin{aligned} \mathbf{a} \cdot 0^3 + \mathbf{b} \cdot 0^2 + \mathbf{c} \cdot 0 + \mathbf{d} &= \mathbf{P}_1, \\ \mathbf{a} \Delta^3 + \mathbf{b} \Delta^2 + \mathbf{c} \Delta + \mathbf{d} &= \mathbf{P}_2, \\ 3\mathbf{a} \cdot 0^2 + 2\mathbf{b} \cdot 0 + \mathbf{c} &= \mathbf{P}_1^t, \\ 3\mathbf{a} \Delta^2 + 2\mathbf{b} \Delta + \mathbf{c} &= \mathbf{P}_2^t, \end{aligned}$$



with solutions

$$\begin{aligned} \mathbf{a} &= \frac{2(\mathbf{P}_1 - \mathbf{P}_2)}{\Delta^3} + \frac{\mathbf{P}_1^t + \mathbf{P}_2^t}{\Delta^2}, \\ \mathbf{b} &= \frac{3(\mathbf{P}_2 - \mathbf{P}_1)}{\Delta^2} - \frac{2\mathbf{P}_1^t}{\Delta} - \frac{\mathbf{P}_2^t}{\Delta}, \\ \mathbf{c} &= \mathbf{P}_1^t, \\ \mathbf{d} &= \mathbf{P}_1. \end{aligned}$$

The curve segment can now be expressed, similar to Equation (4.7), in the form

$$\mathbf{P}_{nu}(t) = (t^3, t^2, t, 1) \begin{pmatrix} \frac{2}{\Delta^3} & \frac{-2}{\Delta^3} & \frac{1}{\Delta^2} & \frac{1}{\Delta^2} \\ \frac{-3}{\Delta^2} & \frac{3}{\Delta^2} & \frac{-2}{\Delta} & \frac{-1}{\Delta} \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{P}_1 \\ \mathbf{P}_2 \\ \mathbf{P}_1^t \\ \mathbf{P}_2^t \end{pmatrix} = \mathbf{T}(t)\mathbf{H}_{nu}\mathbf{B}. \quad (4.12)$$

It is easy to verify that matrix  $\mathbf{H}_{nu}$  reduces to  $\mathbf{H}$  for  $\Delta = 1$ . Figure 4.5 shows a typical nonuniform Hermite segment drawn three times for  $\Delta = 0.5, 1$ , and  $2$ . Careful examination of the three curves shows that increasing the value of  $\Delta$  causes the curve segment to continue longer in its initial and final directions; it has the same effect as increasing the magnitudes of the tangent vectors of the uniform Hermite segment. Once this is grasped, the reader should not be surprised to learn that the nonuniform curve of Equation (4.12) can also be expressed as

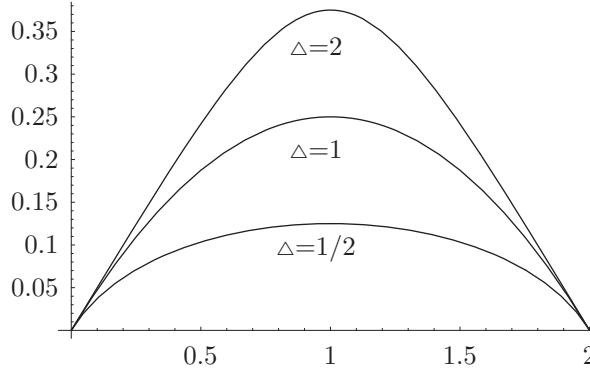
$$\mathbf{P}_{nu}(t) = (t^3, t^2, t, 1) \begin{pmatrix} 2 & -2 & 1 & 1 \\ -3 & 3 & -2 & -1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{P}_1 \\ \mathbf{P}_2 \\ \Delta\mathbf{P}_1^t \\ \Delta\mathbf{P}_2^t \end{pmatrix}. \quad (4.13)$$

This shows that the nonuniform Hermite curve segment is a special case of the uniform curve. Any nonuniform Hermite curve can also be obtained as a uniform Hermite curve by adjusting the magnitudes of the tangent vectors. However, varying the magnitudes of both tangent vectors has an important geometric interpretation, it changes the *tension* of the curve segment. Imagine that the two endpoints are nails driven into the page and the curve segment is a rubber string. When the string is pulled at both sides, its shape approaches a straight line. Figure 4.5 shows how decreasing  $\Delta$  results in a curve with higher tension, so instead of working with nonuniform Hermite segments, we can consider  $\Delta$  a tension parameter. Practical curve methods that create a spline curve out of individual Hermite segments can add a tension parameter to the spline, thereby making the method more interactive. An example is the cardinal splines method (Section 5.4).

#### 4.2.4 PC Conic Approximations

Hermite interpolation can be applied to compute (approximate) conic sections (see Appendix A for more on conics). Given three points  $\mathbf{P}_0$ ,  $\mathbf{P}_1$ , and  $\mathbf{P}_2$  and a scalar  $\alpha$ , we construct the 4-tuple

$$(\mathbf{P}_0, \mathbf{P}_2, 4\alpha(\mathbf{P}_1 - \mathbf{P}_0), 4\alpha(\mathbf{P}_2 - \mathbf{P}_1)), \quad \text{where } 0 \leq \alpha \leq 1, \quad (4.14)$$



```

Clear[T,H,B]; (* Nonuniform Hermite segments *)
T={t^3,t^2,t,1};
H={{2,-2,1,1},{-3,3,-2,-1},{0,0,1,0},{1,0,0,0}};
B[delta_]:={{0,0},{2,0},delta{2,1},delta{2,-1}};
g1=ParametricPlot[T.H.B[0.5],{t,0,1},Compiled->False,
  DisplayFunction->Identity];
g2=ParametricPlot[T.H.B[1],{t,0,1},Compiled->False,
  DisplayFunction->Identity];
g3=ParametricPlot[T.H.B[1.5],{t,0,1},Compiled->False,
  DisplayFunction->Identity];
Show[g1,g2,g3, DisplayFunction->$DisplayFunction, DefaultFont->{"cmr10", 10}]

```

Figure 4.5: Three Nonuniform Hermite Segments.

to become our two points and two extreme tangent vectors and compute a segment that approximates a conic section. We obtain an ellipse when  $0 \leq \alpha < 0.5$ , a parabola when  $\alpha = 0.5$ , and a hyperbola when  $0.5 < \alpha \leq 1$  (see below for a circle).

The tangent vectors at the two ends are  $\mathbf{P}^t(0) = 4\alpha(\mathbf{P}_1 - \mathbf{P}_0)$  and  $\mathbf{P}^t(1) = 4\alpha(\mathbf{P}_2 - \mathbf{P}_1)$  (note their directions). The tangent vector halfway is  $\mathbf{P}^t(0.5) = (1.5 - \alpha)(\mathbf{P}_2 - \mathbf{P}_0)$ . It is parallel to the vector  $\mathbf{P}_2 - \mathbf{P}_0$ .

The case of the parabola is especially useful and is explicitly shown here. Substituting  $\alpha = 0.5$  in Equation (4.14) and applying Equation (4.7) yields the Hermite segment

$$\begin{aligned}
 \mathbf{P}(t) &= (t^3, t^2, t, 1) \begin{bmatrix} 2 & -2 & 1 & 1 \\ -3 & 3 & -2 & -1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{P}_0 \\ \mathbf{P}_2 \\ 2(\mathbf{P}_1 - \mathbf{P}_0) \\ 2(\mathbf{P}_2 - \mathbf{P}_1) \end{bmatrix} \\
 &= (1-t)^2 \mathbf{P}_0 + 2t(1-t) \mathbf{P}_1 + t^2 \mathbf{P}_2.
 \end{aligned}$$

This is the parabola produced in Exercises 4.9 and 6.2.

- ◇ **Exercise 4.9:** We know that any three points  $\mathbf{P}_0$ ,  $\mathbf{P}_1$ , and  $\mathbf{P}_2$  define a unique parabola (i.e., a triangle defines a parabola). Use Hermite interpolation to calculate the parabola from  $\mathbf{P}_0$  to  $\mathbf{P}_2$  whose start and end tangents go in the directions from  $\mathbf{P}_0$  to  $\mathbf{P}_1$  and from  $\mathbf{P}_1$  to  $\mathbf{P}_2$ , respectively.

Hermite interpolation provides a simple way to construct approximate circles and circular arcs. Figure 4.6a shows how this method is employed to construct a circular arc of unit radius about the origin. We assume that an arc spanning an angle  $2\theta$  is needed and we place its two endpoints  $\mathbf{P}_1$  and  $\mathbf{P}_2$  at locations  $(\cos \theta, -\sin \theta)$  and  $(\cos \theta, \sin \theta)$ , respectively. This arc is symmetric about the  $x$  axis, but we later show how to rotate it to have an arbitrary arc. Since a circle is always perpendicular to its radius, we select as our start and end tangents two vectors that are perpendicular to  $\mathbf{P}_1$  and  $\mathbf{P}_2$ . They are  $\mathbf{P}_1^t = a(\sin \theta, \cos \theta)$  and  $\mathbf{P}_2^t = a(-\sin \theta, \cos \theta)$ , where  $a$  is a parameter to be determined. The Hermite curve segment defined by these points and vectors is, as usual,

$$\begin{aligned} \mathbf{P}(t) &= (t^3, t^2, t, 1) \begin{bmatrix} 2 & -2 & 1 & 1 \\ -3 & 3 & -2 & -1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} (\cos \theta, -\sin \theta) \\ (\cos \theta, \sin \theta) \\ a(\sin \theta, \cos \theta) \\ a(-\sin \theta, \cos \theta) \end{bmatrix} \\ &= (2t^3 - 3t^2 + 1)(\cos \theta, -\sin \theta) + (-2t^3 + 3t^2)(\cos \theta, \sin \theta) \\ &\quad + (t^3 - 2t^2 + t)a(\sin \theta, \cos \theta) + (t^3 - t^2)a(-\sin \theta, \cos \theta). \end{aligned} \quad (4.15)$$

We need an equation in order to determine  $a$  and we obtain it by requiring that the curve segment passes through the circular arc at its center, i.e.,  $\mathbf{P}(0.5) = (1, 0)$ . This produces the equation

$$\begin{aligned} (1, 0) = \mathbf{P}(0.5) &= \left(\frac{2}{8} - \frac{3}{4} + 1\right)(\cos \theta, -\sin \theta) + \left(-\frac{2}{8} + \frac{3}{4}\right)(\cos \theta, \sin \theta) \\ &\quad + \left(\frac{1}{8} - \frac{2}{4} + \frac{1}{2}\right)a(\sin \theta, \cos \theta) + \left(\frac{1}{8} - \frac{1}{4}\right)a(-\sin \theta, \cos \theta) \\ &= \frac{1}{8}(8 \cos \theta + 2a \sin \theta, 0), \end{aligned}$$

whose solution is

$$a = \frac{4(1 - \cos \theta)}{\sin \theta}.$$

The curve can now be written in the form

$$\mathbf{P}(t) = (t^3, t^2, t, 1) \begin{bmatrix} 2 & -2 & 1 & 1 \\ -3 & 3 & -2 & -1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} (\cos \theta, -\sin \theta) \\ (\cos \theta, \sin \theta) \\ \left(4(1 - \cos \theta), \frac{4(1 - \cos \theta)}{\tan \theta}\right) \\ \left(-4(1 - \cos \theta), \frac{4(1 - \cos \theta)}{\tan \theta}\right) \end{bmatrix}.$$

This curve provides an excellent approximation to a circular arc, even for angles  $\theta$  as large as  $90^\circ$ .

- ◇ **Exercise 4.10:** Write Equation (4.15) for  $\theta = 90^\circ$ ; calculate  $\mathbf{P}(0.25)$  and the deviation of the curve from a true circle at this point.

In general, an arc with a unit radius is not symmetric about the  $x$  axis but may look as in Figure 4.6b, where  $\mathbf{P}_1$  and  $\mathbf{P}_2$  are any points at a distance of one unit from

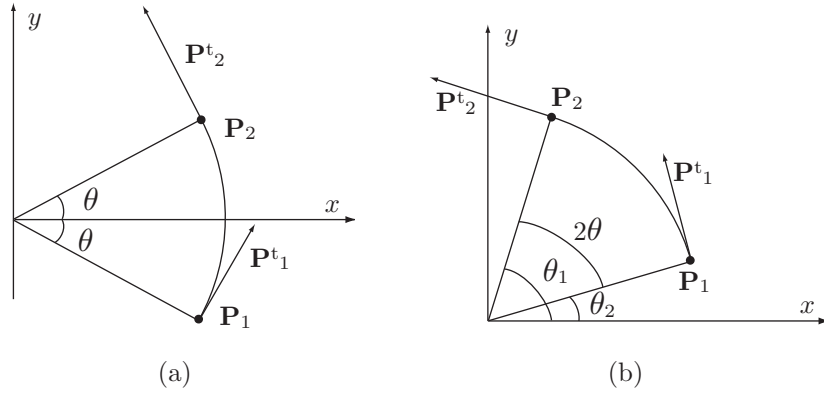


Figure 4.6: Hermite Segment and a Circular Arc.

the origin. All that's necessary to calculate the arc from Equation (4.15) is the value of  $\theta$  (where  $2\theta$  is the angle between  $\mathbf{P}_1$  and  $\mathbf{P}_2$ ) and this can be calculated numerically from the two points using the relations

$$\begin{aligned}\theta &= (\theta_1 - \theta_2)/2, \quad \cos \theta_1 = \mathbf{P}_1 \bullet (1, 0), \quad \cos \theta_2 = \mathbf{P}_2 \bullet (1, 0), \\ \cos(2\theta) &= \cos(\theta_1 - \theta_2) = \cos \theta_1 \cos \theta_2 + \sin \theta_1 \sin \theta_2, \\ \cos \theta &= \pm \sqrt{[1 + \cos(2\theta)]/2}, \quad \sin \theta = \sqrt{1 - \cos^2 \theta}.\end{aligned}$$

### 4.3 Degree-5 Hermite Interpolation

It is possible to extend the basic idea of Hermite interpolation to polynomials of higher degree. Naturally, more data is needed in order to calculate such a polynomial, and this data is provided by the user, normally in the form of higher-order derivatives of the curve. If the user specifies the two endpoints, the two extreme tangent vectors, and the two extreme second derivatives, the software can use these six data items to calculate the six coefficients of a fifth-degree polynomial that interpolates the two points. In general, if the two endpoints and the first  $k$  pairs of derivatives at the extreme points are known (a total of  $2k + 2$  items), they can be used to calculate an interpolating polynomial of degree  $2k + 1$ . These higher-degree polynomials are not as useful as the cubic, but the fifth-degree polynomial is shown here, as a demonstration of the power of Hermite interpolation (see also Section 5.3).

Given two endpoints  $\mathbf{P}_1$  and  $\mathbf{P}_2$ , the values of two tangent vectors  $\mathbf{P}_1^t$  and  $\mathbf{P}_2^t$ , and of two second derivatives  $\mathbf{P}_1^{tt}$  and  $\mathbf{P}_2^{tt}$ , we can calculate the polynomial

$$\mathbf{P}(t) = \mathbf{a}t^5 + \mathbf{b}t^4 + \mathbf{c}t^3 + \mathbf{d}t^2 + \mathbf{e}t + \mathbf{f} \quad (4.16)$$

by writing the six equations

$$\begin{aligned}
\mathbf{P}(0) &= \mathbf{a}t^5 + \mathbf{b}t^4 + \mathbf{c}t^3 + \mathbf{d}t^2 + \mathbf{e}t + \mathbf{f}|_0 = \mathbf{f} = \mathbf{P}_1, \\
\mathbf{P}(1) &= \mathbf{a}t^5 + \mathbf{b}t^4 + \mathbf{c}t^3 + \mathbf{d}t^2 + \mathbf{e}t + \mathbf{f}|_1 = \mathbf{a} + \mathbf{b} + \mathbf{c} + \mathbf{d} + \mathbf{e} + \mathbf{f} = \mathbf{P}_2, \\
\mathbf{P}^t(0) &= 5\mathbf{a}t^4 + 4\mathbf{b}t^3 + 3\mathbf{c}t^2 + 2\mathbf{d}t + \mathbf{e}|_0 = \mathbf{e} = \mathbf{P}_1^t, \\
\mathbf{P}^t(1) &= 5\mathbf{a}t^4 + 4\mathbf{b}t^3 + 3\mathbf{c}t^2 + 2\mathbf{d}t + \mathbf{e}|_1 = 5\mathbf{a} + 4\mathbf{b} + 3\mathbf{c} + 2\mathbf{d} + \mathbf{e} = \mathbf{P}_2^t, \\
\mathbf{P}^{tt}(0) &= 20\mathbf{a}t^3 + 12\mathbf{b}t^2 + 6\mathbf{c}t + 2\mathbf{d}|_0 = 2\mathbf{d} = \mathbf{P}_1^{tt}, \\
\mathbf{P}^{tt}(1) &= 20\mathbf{a}t^3 + 12\mathbf{b}t^2 + 6\mathbf{c}t + 2\mathbf{d}|_1 = 20\mathbf{a} + 12\mathbf{b} + 6\mathbf{c} + 2\mathbf{d} = \mathbf{P}_2^{tt}.
\end{aligned}$$

Solving for the six unknown coefficients yields the degree-5 Hermite interpolating polynomial

$$\begin{aligned}
\mathbf{P}(t) &= F_1(t)\mathbf{P}_1 + F_2(t)\mathbf{P}_2 + F_3(t)\mathbf{P}_1^t + F_4(t)\mathbf{P}_2^t + F_5(t)\mathbf{P}_1^{tt} + F_6(t)\mathbf{P}_2^{tt} \\
&= (-6t^5 + 15t^4 - 10t^3 + 1)\mathbf{P}_1 + (6t^5 - 15t^4 + 10t^3)\mathbf{P}_2 \\
&\quad + (-3t^5 + 8t^4 - 6t^3 + t)\mathbf{P}_1^t + (-3t^5 + 7t^4 - 4t^3)\mathbf{P}_2^t \\
&\quad + (-(1/2)t^5 + (3/2)t^4 - (3/2)t^3 + (1/2)t^2)\mathbf{P}_1^{tt} + ((1/2)t^5 - t^4 + (1/2)t^3)\mathbf{P}_2^{tt} \\
&= (t^5, t^4, t^3, t^2, t, 1) \begin{bmatrix} -6 & 6 & -3 & -3 & -1/2 & 1/2 \\ 15 & -15 & 8 & 7 & 3/2 & -1 \\ -10 & 10 & -6 & -4 & -3/2 & 1/2 \\ 0 & 0 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{P}_1 \\ \mathbf{P}_2 \\ \mathbf{P}_1^t \\ \mathbf{P}_2^t \\ \mathbf{P}_1^{tt} \\ \mathbf{P}_2^{tt} \end{bmatrix}. \tag{4.17}
\end{aligned}$$

## 4.4 Controlling the Hermite Segment

The Hermite method is interactive. In general, the points cannot be moved, but the tangent vectors can be varied. Even if their directions cannot be changed, their magnitudes normally are not fixed by the user and can be modified to edit the shape of the curve segment.

The simple experiment of this section illustrates the amount of editing and controlling that can be achieved just by varying the magnitudes of the tangents. We start with the Hermite segment defined by the two endpoints  $\mathbf{P}_1 = (0, 0)$  and  $\mathbf{P}_2 = (2, 1)$  and by the two tangent vectors  $\mathbf{P}^t(0) = (1, 1)$  and  $\mathbf{P}^t(1) = (1, 0)$ . The curve starts in the  $45^\circ$  direction and ends in a horizontal direction. The curve is easy to calculate. Its expression is

$$\mathbf{P}(t) = (t^3, t^2, t, 1) \begin{bmatrix} 2 & -2 & 1 & 1 \\ -3 & 3 & -2 & -1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} (0, 0) \\ (2, 1) \\ (1, 1) \\ (1, 0) \end{bmatrix} = -(2, 1)t^3 + (3, 1)t^2 + (1, 1)t. \tag{4.18}$$

Suppose that the user wants to raise the curve a bit, but also keep the same start and end directions and endpoints. The only way to edit the curve is to change the magnitudes of the tangents.

To keep the same directions, the new tangent vectors should have the form  $(a, a)$  and  $(b, 0)$ , where  $a$  and  $b$  are two new parameters that have to be computed. To raise the curve, we go through the following steps:

1. Calculate the midpoint of the curve. This is  $\mathbf{P}(0.5) = (1, 5/8)$ .
2. Decide by how much to raise it. Let's say we decide to raise the midpoint to  $(1, 1)$ .
3. Construct a new curve  $\mathbf{Q}(t)$ , based on the tangents  $(a, a)$  and  $(b, 0)$ .
4. Require that the new curve pass through  $(1, 1)$  as its midpoint and determine  $a$  and  $b$  from this requirement.

The general form of the new curve is

$$\begin{aligned} \mathbf{Q}(t) &= (t^3, t^2, t, 1) \begin{bmatrix} 2 & -2 & 1 & 1 \\ -3 & 3 & -2 & -1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} (0, 0) \\ (2, 1) \\ (a, a) \\ (b, 0) \end{bmatrix} \\ &= (a + b - 4, a - 2)t^3 + (-2a - b + 6, 3 - 2a)t^2 + (a, a)t. \end{aligned} \quad (4.19)$$

The requirement  $\mathbf{Q}(0.5) = (1, 1)$  can now be written

$$(a + b - 4, a - 2)/8 + (-2a - b + 6, 3 - 2a)/4 + (a, a)/2 = (1, 1),$$

which yields the two equations  $a + b - 4 + 2(-2a - b + 6) + 4a = 8$  and  $a - 2 + 2(3 - 2a) + 4a = 8$ . The solutions are  $a = b = 4$ , so the new curve has the form

$$\mathbf{Q}(t) = (4, 2)t^3 - (6, 5)t^2 + (4, 4)t. \quad (4.20)$$

A simple check verifies that this curve really starts at  $(0, 0)$ , ends at  $(2, 1)$ , has the extreme tangents  $(4, 4)$  and  $(4, 0)$ , and passes midway through  $(1, 1)$ .

Raising the midpoint from  $(1, 5/8)$  to  $(1, 1)$  has completely changed the curve (Equations (4.18) and (4.20) are different). The new curve starts going in the same  $45^\circ$  direction, then starts going up, reaches point  $(1, 1)$ , starts going down, and still has "time" to arrive at point  $(2, 1)$  moving horizontally. An interesting question is: How much can we raise the midpoint? If we raise it from  $(1, 5/8)$  to, say,  $(1, 100)$ , would the curve be able to change directions, climb up, pass through the new midpoint, dive down, and still approach  $(2, 1)$  moving horizontally?

To check this, let's assume that we raise the midpoint from  $(1, 5/8)$  to  $(1, 5/8 + \alpha)$ , where  $\alpha$  is a real number. The curve is constrained by  $\mathbf{Q}(0.5) = (1, 5/8 + \alpha)$ , which yields the equation

$$(a + b - 4, a - 2)/8 + (-2a - b + 6, 3 - 2a)/4 + (a, a)/2 = (1, 5/8 + \alpha).$$

The solutions are  $a = b = 1 + 8\alpha$ . This means that  $\alpha$  can vary without limit. When  $\alpha$  is positive, the curve is pulled up. Negative values of  $\alpha$  push the curve down. The

value  $\alpha = -1/8$  is special. It implies  $a = b = 0$  and results in the curve  $\mathbf{Q}(t) = (6t^2 - 4t^3, 3t^2 - 2t^3)$ . The parameter substitution  $u = 3t^2 - 2t^3$  yields  $\mathbf{Q}(u) = (2u, u)$ . This curve is the straight line from  $(0, 0)$  to  $(2, 1)$ . Its midpoint is  $(1, 1/2)$ .

- ◇ **Exercise 4.11:** Values  $\alpha < -1/8$  result in negative  $a$  and  $b$ . Can they still be used in Equation (4.19)?
- ◇ **Exercise 4.12:** How can we coerce the curve of Equation (4.19) to have point  $(1, 0)$  as its midpoint?

**Note:** Raising the curve is done by increasing the size of the tangent vectors. This forces the curve to continue longer in the initial and final directions. This is also the reason why too much raising causes undesirable effects. Figure 4.7 shows the original curve ( $\alpha = 0$ ) and the effects of increasing  $\alpha$ . For  $\alpha = 0.4$ , the curve is raised and still has a reasonable shape. However, for larger values of  $\alpha$ , the curve gets tight, develops a *cusp* (a kink), then starts looping on itself. It is easy to see that when  $\alpha = 5/8$ , the tangent vector becomes indefinite at the midpoint ( $t = 0.5$ ). To show this, we differentiate the curve of Equation (4.19) to obtain the tangent

$$\mathbf{Q}^t(t) = 3(a + b - 4, a - 2)t^2 + 2(-2a - b + 6, 3 - 2a)t + (a, a).$$

From  $a = b = 1 + 8\alpha$ , we get

$$\mathbf{Q}^t(t) = (48\alpha - 6, 24\alpha - 3)t^2 + (6 - 48\alpha, 2 - 32\alpha)t + (1 + 8\alpha, 1 + 8\alpha).$$

For  $\alpha = 5/8$ , this reduces to  $\mathbf{Q}^t(t) = (24, 12)t^2 - (24, 18)t + (6, 6)$ , so  $\mathbf{Q}^t(0.5) = (0, 0)$ .

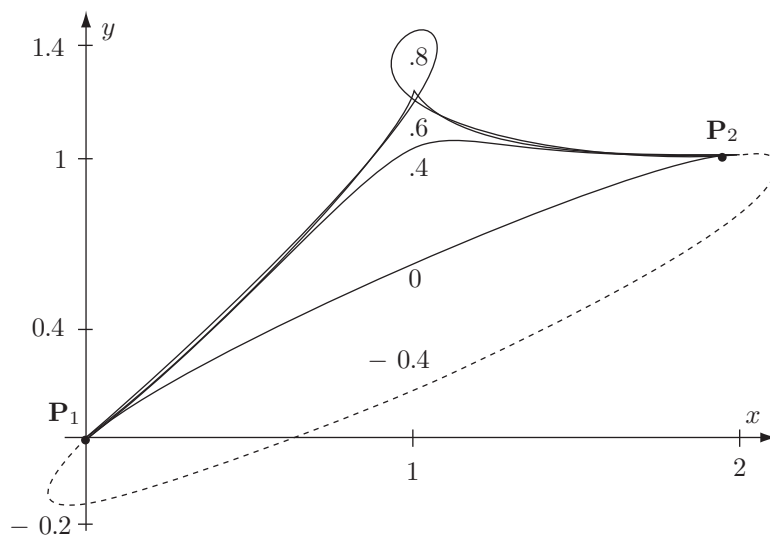


Figure 4.7: Effects of Changing  $\alpha$ .

- ◇ **Exercise 4.13:** Given the two endpoints  $\mathbf{P}_1 = (0, 0)$  and  $\mathbf{P}_2 = (1, 0)$  and the two tangent vectors  $\mathbf{P}_1^t = \alpha(\cos \theta, \sin \theta)$  and  $\mathbf{P}_2^t = \alpha(\cos \theta, -\sin \theta)$  (Figure 4.8), calculate the value of  $\alpha$  for which the Hermite segment from  $\mathbf{P}_1$  to  $\mathbf{P}_2$  has a cusp.

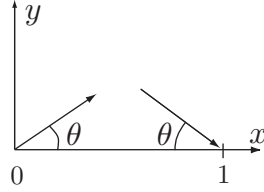


Figure 4.8: Tangents for Exercise 4.13.

The following problem may sometimes occur in practice. Given two endpoints  $\mathbf{P}_1$  and  $\mathbf{P}_2$ , two *unit* tangent vectors  $\mathbf{T}_1$  and  $\mathbf{T}_2$ , and a third point  $\mathbf{P}_3$ , find scale factors  $\alpha$  and  $\beta$  such that the Hermite segment  $\mathbf{P}(t)$  defined by points  $\mathbf{P}_1$  and  $\mathbf{P}_2$  and tangents  $\alpha\mathbf{T}_1$  and  $\beta\mathbf{T}_2$ , respectively, will pass through  $\mathbf{P}_3$ . Also find the value  $t_0$  for which  $\mathbf{P}(t_0) = \mathbf{P}_3$ .

We start with Equation (4.5), which in our case becomes

$$\mathbf{P}_3 = F_1(t_0)\mathbf{P}_1 + F_2(t_0)\mathbf{P}_2 + F_3(t_0)\alpha\mathbf{T}_1 + F_4(t_0)\beta\mathbf{T}_2,$$

where the  $F_i(t)$  are given by Equation (4.6). Since  $F_1(t) + F_2(t) \equiv 1$  we can write

$$\mathbf{P}_3 - \mathbf{P}_1 = F_2(t_0)(\mathbf{P}_2 - \mathbf{P}_1) + \alpha F_3(t_0)\mathbf{T}_1 + \beta F_4(t_0)\mathbf{T}_2.$$

This can now be written as the three scalar equations

$$\begin{aligned} x_3 - x_1 &= F_2(t_0)(x_2 - x_1) + \alpha F_3(t_0)T_{1x} + \beta F_4(t_0)T_{2x}, \\ y_3 - y_1 &= F_2(t_0)(y_2 - y_1) + \alpha F_3(t_0)T_{1y} + \beta F_4(t_0)T_{2y}, \\ z_3 - z_1 &= F_2(t_0)(z_2 - z_1) + \alpha F_3(t_0)T_{1z} + \beta F_4(t_0)T_{2z}. \end{aligned} \quad (4.21)$$

This is a system of three equations in the three unknowns  $\alpha$ ,  $\beta$ , and  $t_0$ . In principle, it should have a unique solution, but solving it is awkward since  $t_0$  is included in the  $F_i(t_0)$  functions, which are degree-3 polynomials in  $t_0$ . The first step is to isolate the two products  $\alpha F_3(t_0)$  and  $\beta F_4(t_0)$  in the first two equations. This yields

$$\begin{pmatrix} \alpha F_3(t_0) \\ \beta F_4(t_0) \end{pmatrix} = \begin{pmatrix} T_{1x} & T_{2x} \\ T_{1y} & T_{2y} \end{pmatrix}^{-1} \left[ \begin{pmatrix} x_3 - x_1 \\ y_3 - y_1 \end{pmatrix} - \begin{pmatrix} x_2 - x_1 \\ y_2 - y_1 \end{pmatrix} F_2(t_0) \right].$$

This result is used in step two to eliminate  $\alpha F_3(t_0)$  and  $\beta F_4(t_0)$  from the third equation:

$$\begin{aligned} z_3 - z_1 &= F_2(t_0)(z_2 - z_1) + (T_{1z}, T_{2z}) \begin{pmatrix} \alpha F_3(t_0) \\ \beta F_4(t_0) \end{pmatrix} \\ &= F_2(t_0)(z_2 - z_1) \\ &\quad + (T_{1z}, T_{2z}) \begin{pmatrix} T_{1x} & T_{2x} \\ T_{1y} & T_{2y} \end{pmatrix}^{-1} \left[ \begin{pmatrix} x_3 - x_1 \\ y_3 - y_1 \end{pmatrix} - \begin{pmatrix} x_2 - x_1 \\ y_2 - y_1 \end{pmatrix} F_2(t_0) \right]. \end{aligned}$$



We now have an equation with the single unknown  $t_0$ . Step three is to simplify the result above by using the value  $F_2(t_0) = -2t_0^3 + 3t_0^2$ :

$$\begin{vmatrix} x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ T1_x & T1_y & T1_z \\ T2_x & T2_y & T2_z \end{vmatrix} (-2t_0^3 + 3t_0^2) = \begin{vmatrix} x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \\ T1_x & T1_y & T1_z \\ T2_x & T2_y & T2_z \end{vmatrix}. \quad (4.22)$$

Step four is to solve Equation (4.22) for  $t_0$ . Once  $t_0$  is known,  $\alpha$  and  $\beta$  can be computed from the other equations. Equation (4.22), however, is cubic in  $t_0$ , so it may have to be solved numerically and it may have between zero and three real solutions  $t_0$ . Any acceptable solution  $t_0$  must be a real number in the range  $[0, 1]$  and must result in positive  $\alpha$  and  $\beta$ .

This, of course, is a slow, tedious approach and should only be used as a last resort, when nothing else works.

## 4.5 Truncating and Segmenting

Surfaces and solid objects are constructed of curves. When surfaces are joined, clipped, or intersected, there is sometimes a need to truncate curves. In general, the problem of truncating a curve starts with a parametric curve  $\mathbf{P}(t)$  and the two values  $t_i$  and  $t_j$ . A new curve  $\mathbf{Q}(T)$  needs be determined, that is identical to the segment  $\mathbf{P}(t_i) \rightarrow \mathbf{P}(t_j)$  (Figure 4.9a) when  $T$  varies from 0 to 1. The discussion in this section is limited to Hermite segments. The endpoints of the new curve are  $\mathbf{Q}(0) = \mathbf{P}(t_i)$  and  $\mathbf{Q}(1) = \mathbf{P}(t_j)$ . To understand how the two extreme tangent vectors of  $\mathbf{Q}(T)$  are calculated, we first need to discuss reparametrization of parametric curves.

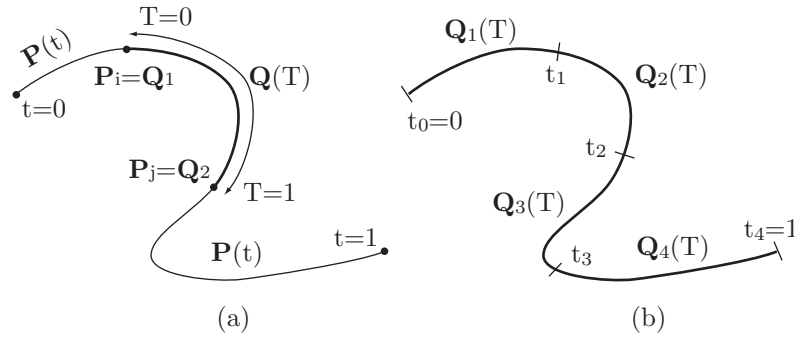


Figure 4.9: Truncating and Segmenting.

Reparametrization is the case where a new parameter  $T(t)$  is substituted for the original parameter  $t$ . Notice that  $T(t)$  is a function of  $t$ . One example of reparametrization is reversing the direction of a curve. It is easy to see that when  $t$  varies from 0 to 1, the simple function  $T = 1 - t$  varies from 1 to 0. The two curves  $\mathbf{P}(t)$  and  $\mathbf{P}(1 - t)$  have the same shape and location but move in opposite directions. Another example of

reparametrization is a curve  $\mathbf{P}(t)$  with a parameter  $0 \leq t \leq 1$  being transformed to a curve  $\mathbf{Q}(T)$  with a parameter  $a \leq T \leq b$  (Section 5.1.6 has an example). The simplest relation between  $T$  and  $t$  is linear, i.e.,  $T = at + b$ . We can make two observations about this relation as follows:

1. At two different points  $i$  and  $j$  along the curve, the parameters are related by  $T_i = at_i + b$  and  $T_j = at_j + b$ , respectively. Subtracting yields  $T_j - T_i = a(t_j - t_i)$ , so  $a = (T_j - T_i)/(t_j - t_i)$ .

2.  $T = at + b$  gives  $dT = a dt$ .

These two observations can be combined to produce the expression

$$\frac{dt}{dT} = \frac{1}{a} = \frac{t_j - t_i}{T_j - T_i}. \quad (4.23)$$

Equation (4.23) is used to calculate the extreme tangent vectors of our new curve  $\mathbf{Q}(T)$ . Since it goes from point  $\mathbf{P}(t_i)$  (where  $T = 0$ ) to point  $\mathbf{P}(t_j)$  (where  $T = 1$ ), we have  $T_j - T_i = 1$ . The tangent vectors of  $\mathbf{Q}(T)$  are therefore

$$\mathbf{Q}^T(T) = \frac{d\mathbf{Q}(T)}{dT} = \frac{d\mathbf{P}(t)}{dt} \frac{dt}{dT} = \mathbf{P}^t(t) \cdot (t_j - t_i).$$

The two extreme tangents are  $\mathbf{Q}^T(0) = (t_j - t_i)\mathbf{P}^t(t_i)$  and  $\mathbf{Q}^T(1) = (t_j - t_i)\mathbf{P}^t(t_j)$ . The new curve can now be calculated by

$$\mathbf{Q}(T) = (T^3, T^2, T, 1)\mathbf{H} \begin{bmatrix} \mathbf{P}(t_i) \\ \mathbf{P}(t_j) \\ (t_j - t_i)\mathbf{P}^t(t_i) \\ (t_j - t_i)\mathbf{P}^t(t_j) \end{bmatrix}, \quad (4.24)$$

where  $\mathbf{H}$  is the Hermite matrix, Equation (4.7).

◇ **Exercise 4.14:** Compute the PC segment  $\mathbf{Q}(T)$  that results from truncating  $\mathbf{P}(t) = (-1, 0)t^3 + (1, -1)t^2 + (1, 1)t$  [Equation (4.10)] from  $t_i = 0.25$  to  $t_j = 0.75$ .

Segmenting a curve is the problem of calculating several truncations. Assume that we are given values  $0 = t_0 < t_1 < t_2 < \dots < t_n = 1$ , and we want to break a given curve  $\mathbf{P}(t)$  into  $n$  segments such that segment  $i$  will go from point  $\mathbf{P}(t_{i-1})$  to point  $\mathbf{P}(t_i)$  (Figure 4.9b). Equation (4.24) gives segment  $i$  as

$$\mathbf{Q}_i(T) = (T^3, T^2, T, 1)\mathbf{H} \begin{bmatrix} \mathbf{P}(t_{i-1}) \\ \mathbf{P}(t_i) \\ (t_i - t_{i-1})\mathbf{P}^t(t_{i-1}) \\ (t_i - t_{i-1})\mathbf{P}^t(t_i) \end{bmatrix}.$$

#### 4.5.1 Special and Degenerate Hermite Segments

The following special cases result in Hermite curve segments that are either especially simple (degenerate) or especially interesting

■ The case  $\mathbf{P}_1 = \mathbf{P}_2$  and  $\mathbf{P}_1^t = \mathbf{P}_2^t = (0, 0)$ . Equation (4.4) yields  $\mathbf{P}(t) = \mathbf{P}_1$ ; the curve degenerates to a point.

- The case  $\mathbf{P}_1^t = \mathbf{P}_2^t = \mathbf{P}_2 - \mathbf{P}_1$ . The two tangents point in the same direction, from  $\mathbf{P}_1$  to  $\mathbf{P}_2$ . Equation (4.4) yields

$$\begin{aligned}\mathbf{P}(t) &= (2\mathbf{P}_1 - 2\mathbf{P}_2 + 2(\mathbf{P}_2 - \mathbf{P}_1))t^3 + (-3\mathbf{P}_1 + 3\mathbf{P}_2 - 3(\mathbf{P}_2 - \mathbf{P}_1))t^2 \\ &\quad + (\mathbf{P}_2 - \mathbf{P}_1)t + \mathbf{P}_1 \\ &= (\mathbf{P}_2 - \mathbf{P}_1)t + \mathbf{P}_1.\end{aligned}\tag{4.25}$$

The curve reduces to a straight segment.

- The case  $\mathbf{P}_1 = \mathbf{P}_2$ . Equation (4.4) yields  $\mathbf{P}(t) = (\mathbf{P}_1^t + \mathbf{P}_2^t)t^3 + (-2\mathbf{P}_1^t - \mathbf{P}_2^t)t^2 + \mathbf{P}_1^t t + \mathbf{P}_1$ . It is easy to see that this curve satisfies  $\mathbf{P}(0) = \mathbf{P}(1)$ . It is closed (but is not a circle).
- The case  $\mathbf{P}_1^t = \mathbf{P}_2^t = (x_2 - x_1, y_2 - y_1, 0)$ . Equation (4.4) yields

$$\begin{aligned}\mathbf{P}(t) &= (2\mathbf{P}_1 - 2\mathbf{P}_2 + 2(x_2 - x_1, y_2 - y_1, 0))t^3 \\ &\quad + (-3\mathbf{P}_1 + 3\mathbf{P}_2 - 3(x_2 - x_1, y_2 - y_1, 0))t^2 \\ &\quad + (x_2 - x_1, y_2 - y_1, 0)t + (x_1, y_1, z_1) \\ &= (x_1 + (x_2 - x_1)t, y_1 + (y_2 - y_1)t, z_1 + (z_2 - z_1)(3t^2 - 2t^3)).\end{aligned}$$

The  $x$  and  $y$  coordinates of this curve are linear functions of  $t$ , so its tangent vector has the form  $(\alpha, \beta, z(t))$ . Its  $x$  and  $y$  components are constants, so it always points in the same plane. Thus, the curve is planar.

## 4.6 Hermite Straight Segments

Equation (4.25) shows that the Hermite segment can sometimes degenerate into a straight segment. This section describes variations on Hermite straight segments. Specifically, we look in detail at the case where the two extreme tangent vectors point in the same direction, from  $\mathbf{P}_1$  to  $\mathbf{P}_2$ , but have different magnitudes. We denote them by  $\mathbf{P}_1^t = \alpha(\mathbf{P}_2 - \mathbf{P}_1)$  and  $\mathbf{P}_2^t = \beta(\mathbf{P}_2 - \mathbf{P}_1)$ , where  $\alpha$  and  $\beta$  can be any real numbers. Equation (4.25) is obtained in the special case  $\alpha = \beta = 1$ .

The Hermite segment is expressed as  $\mathbf{P}(t) = \mathbf{F}(t)\mathbf{B}$ , where the four  $F_i(t)$  functions are given by Equation (4.6), and  $\mathbf{B}$  is the geometry vector, which, in our case, has the form

$$\mathbf{B} = (\mathbf{P}_1, \mathbf{P}_2, \alpha(\mathbf{P}_2 - \mathbf{P}_1), \beta(\mathbf{P}_2 - \mathbf{P}_1))^T.$$

This can be written (since  $F_1(t) + F_2(t) \equiv 1$ ) in the form

$$\begin{aligned}\mathbf{P}(t) &= F_1(t)\mathbf{P}_1 + F_2(t)\mathbf{P}_2 + F_3(t)\alpha(\mathbf{P}_2 - \mathbf{P}_1) + F_4(t)\beta(\mathbf{P}_2 - \mathbf{P}_1) \\ &= \mathbf{P}_1 + (F_2(t) + \alpha F_3(t) + \beta F_4(t))(\mathbf{P}_2 - \mathbf{P}_1) \\ &= \mathbf{P}_1 + ((1 - 2t^3 + 3t^2) + \alpha(t^3 - 2t^2 + t) + \beta(t^3 - t^2))(\mathbf{P}_2 - \mathbf{P}_1) \\ &= \mathbf{P}_1 + ((\alpha + \beta - 2)t^3 - (2\alpha + \beta - 3)t^2 + \alpha t)(\mathbf{P}_2 - \mathbf{P}_1).\end{aligned}\tag{4.26}$$