This notes were taken from :

"Introduction to Applied Mathematics" book by Gilbert Strang

Interpolation: Displacements and Slopes

After writing down the four types of end conditions, I realized that the list might be misleading. It suggests that the conditions must force u or its derivatives to be *zero* at each end. That is not at all true. The boundary conditions can equally well be inhomogeneous, as in

$$u(0) = a, \frac{du}{dx}(0) = b, u(1) = c, \frac{du}{dx}(1) = d. \tag{23}$$

The ends of this beam are still fixed, but they are not necessarily lined up. In fact we can have zero force, $f = 0$, and the beam will bend in order to satisfy these boundary conditions. It is curved by the forces at its ends.

The deflection is still governed by the differential equation

$$\frac{d^2}{dx^2}\left(c\,\frac{d^2u}{dx^2}\right) = f, \quad \text{or} \quad \frac{d^4u}{dx^4} = 0,$$

assuming that c is constant. Any third degree polynomial like $x^3 + 1$ is a solution to this equation; the fourth derivative of a cubic is automatically zero. But there is only one cubic polynomial whose four coefficients match all four conditions in (23). It is the ***Hermite cubic***, and we can write it out in full:

$$u = a(x-1)^2(2x+1) + b(x-1)^2 x + cx^2(3-2x) - dx^2(x-1). \tag{24}$$

This is certainly a third-degree polynomial, so $d^4u/dx^4 = 0$. Each term is chosen to match one of the boundary conditions in (23). The coefficient of d is $x^3 - x^2$, and it is zero at both ends. Its slope $3x^2 - 2x$ equals zero at $x = 0$ and one at $x = 1$. The graph is drawn upside down in Fig. 3.5b, showing the shape of a beam when one end is held horizontal and the other is fixed at $-45°$. The left figure also shows the coefficient of a—the shape of a beam when one end is raised but still cantilevered.

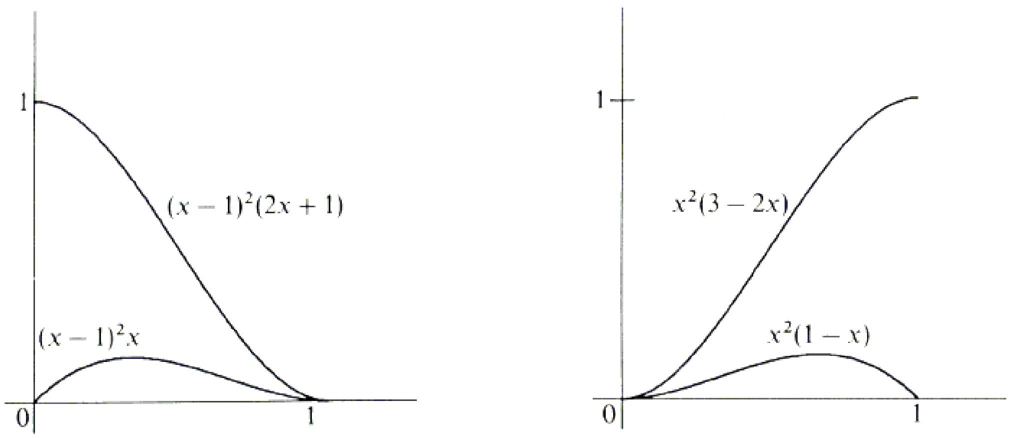


Fig. 3.5. Cubics for the four boundary conditions: coefficients of a, b, c, and $-d$.

These cubics also appear in problems that have nothing to do with beams. They are used for ***interpolation***—constructing a curve that goes through a given set of points. It is a different process from least squares fitting, which only approximates the given values. The least squares fit comes as close as possible with a small number of degrees of freedom, whereas ***interpolation is exact***. The curve will go right through the points—it must have enough degrees of freedom to do so—and the question is what kind of curve to use.

Suppose we are given the height and also the slope of the unknown curve at the points $x = 0, 1, \ldots, n$. That is $2n + 2$ pieces of information, and they can be matched with a single polynomial of high degree. The degree would be $2n + 1$, producing a polynomial with $2n + 2$ coefficients. However the result would be absolutely terrible. Even if the values and slopes are taken from a smooth curve, the interpolating polynomial is almost certain to oscillate. Away from the interpolation points it is far from the right curve. A much better alternative is to fit the data by a ***piecewise polynomial***, which has low degree in each interval between interpolation points.

When heights and slopes are both given, the natural choice is a piecewise cubic. It is exactly the one constructed above in (24). The four coefficients of the cubic satisfy four conditions—two at each end of every interval. These "Hermite cubic" pieces fit together with matching height and slope at the interpolation points. Of course the second and third derivatives at those points can be expected to jump, as in the following example:

$$\text{Heights} \quad u_0 = 0, u_1 = 1, u_2 = 0 \quad \text{at} \quad x = 0, 1, 2$$

$$\text{Slopes} \quad s_0 = 1, s_1 = 0, s_2 = 0 \quad \text{at} \quad x = 0, 1, 2.$$

The cubic between 0 and 1 comes directly from (24), with $a = d = 0$ and $b = c = 1$ (since $s_0 = u_1 = 1$):

$$u(x) = (x-1)^2 x + x^2(3 - 2x) = -x^3 + x^2 + x.$$

The second piece goes from $x = 1$ to $x = 2$. It also comes from (24), but now the height $u_1 = 1$ is at the left end of the interval. Thus $a = 1$ and $b = c = d = 0$. At the same time x in (24) is changed to $x - 1$, to shift the interval to $1 \le x \le 2$, so that

$$u(x) = (x-2)^2(2x-1) = 2x^3 - 9x^2 + 12x - 4.$$

At the junction $x = 1$, these cubics $u(x)$ have the same height $u = 1$ and the same slope $u' = 0$. The second derivative jumps from -4 to -6, and the third derivative jumps from -6 to 12. It is hardly possible to see that effect, and we have exchanged the total smoothness of a sensitive high degree polynomial for the stability of a low degree piecewise polynomial.

Cubic Splines

Now comes the more usual interpolation, when only the heights are given. We know the values $u_0, u_1, \ldots, u_n$ at the points $x = 0, 1, \ldots, n$, and we want to fit them with a curve. Again a high degree polynomial is unstable, and a low degree piecewise polynomial is better. The easiest method of all is to connect the points by straight lines—which is piecewise linear interpolation. But then the slopes change suddenly and are unreliable. A much better choice is to use ***cubic splines***.†

A cubic spline is a piecewise cubic in which not only the function u and the slope du/dx but also ***the second derivative*** d^2u/dx^2 ***is continuous***. Physically, it comes from bending a long thin beam to give it the correct heights $u_0, \ldots, u_n$ at the interpolation points. I think of rings at those points, and the beam going through the rings. At all other points the force is zero, the beam is free to choose its own shape, and the solution to the beam equation $d^4u/dx^4 = 0$ is an ordinary cubic. But something must happen at an interpolation point like $x = 0$, where the ring imparts a concentrated load of unknown magnitude f_0 and

$$\frac{d^4u}{dx^4} = f_0 \delta(x). \tag{25}$$

† The name and the idea come from naval architects.

Remember that the delta function is a "spike" of infinite height and infinitesimal width, and it has unit area:

$$\boxed{\int \delta(x)dx = 1 \quad \text{and} \quad \int f(x)\delta(x)dx = f(0).} \tag{26}$$

It is not a genuine function, but it has been legalized as a *distribution*—which means that $\delta(x)$ is known by its effect on smooth functions $f(x)$.

Integrating both sides of (25), we discover that the third derivative of u jumps by f_0:

$$\int \frac{d^4u}{dx^4}\,dx = \int f_0\delta(x)dx \quad \text{or} \quad \left[\frac{d^3u}{dx^3}\right]_{x=0-}^{x=0+} = f_0.$$

In other words d^3u/dx^3 is a *step function*. It is constant on each side of $x = 0$, but those constants differ by f_0. At $x = 1$ there is another step of height f_1. When we integrate the step function d^3u/dx^3, we conclude that the second derivative d^2u/dx^2 is continuous.

Now the exact shape of the spline can be determined. If slopes $s_0, \ldots, s_n$ were known at the interpolation points, the problem would already be solved. Between those points it is a cubic, and earlier in the section we constructed the only cubic—the Hermite cubic—with the required heights $u_0, \ldots, u_n$ and the required slopes. At present the slopes are unknown, but we have $n+1$ new conditions; d^2u/dx^2 is continuous at each point $x = 0, \ldots, n$. This will give $n+1$ equations for the slopes, as follows.

In the infinite interval to the left of $x = 0$, the beam is straight (but not necessarily horizontal). There is nothing to produce curvature, so that $d^2u/dx^2 = 0$ at $x = 0$. To make the second derivative also zero coming from the right of $x = 0$, (24) requires

$$2s_0 + s_1 = 3u_1 - 3u_0. \tag{27}$$

This is the first of the $n+1$ equations. At the next node $x = 1$, the continuity of d^2u/dx^2 connects the cubic on the left to the cubic on the right. With some patience this calculation gives

$$(2s_1 + s_2) + (2s_1 + s_0) = (3u_2 - 3u_1) + (3u_1 - 3u_0). \tag{28}$$

There is a similar equation at each of the points $x = 1, \ldots, n-1$ and an equation like (27) at $x = n$. Altogether these $n+1$ equations for the unknown slopes go into matrix form as

$$\begin{bmatrix} 2 & 1 & & & \\ 1 & 4 & 1 & & \\ & \cdot & \cdot & \cdot & \\ & & 1 & 4 & 1 \\ & & & 1 & 2 \end{bmatrix} \begin{bmatrix} s_0 \\ s_1 \\ \cdot \\ s_{n-1} \\ s_n \end{bmatrix} = 3 \begin{bmatrix} u_1 - u_0 \\ u_2 - u_0 \\ \cdot \\ u_n - u_{n-2} \\ u_n - u_{n-1} \end{bmatrix}. \tag{29}$$

The matrix is symmetric and positive definite, and the slopes are found from this equation. The piecewise cubic which has these slopes, and which takes on the given values $u_0, \ldots, u_n$, is the ***interpolating spline***.

3D With heights and slopes both given, there is one possible piecewise cubic. It is determined in every interval by two conditions at each end, as in (24), and its first derivative is continuous. If only the heights $u_0, \ldots, u_n$ are given, the slopes can be chosen according to (29) and then the piecewise cubic becomes a spline; its second derivative is also continuous.

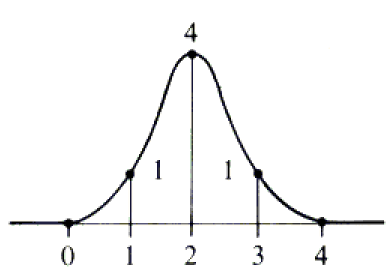
EXAMPLE $\quad u_0 = 0, u_1 = 1, u_2 = 4, u_3 = 1, u_4 = 0$

The equation (29) for the slopes is

$$\begin{bmatrix} 2 & 1 & & & \\ 1 & 4 & 1 & & \\ & 1 & 4 & 1 & \\ & & 1 & 4 & 1 \\ & & & 1 & 2 \end{bmatrix} \begin{bmatrix} s_0 \\ s_1 \\ s_2 \\ s_3 \\ s_4 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 4 \\ 0 \\ -4 \\ -1 \end{bmatrix}. \qquad (30)$$

The solution has $s_1 = 3$ and $s_3 = -3$, with the other slopes $s_0 = s_2 = s_4 = 0$. This is a very special spline, drawn in Fig. 3.6. It starts with everything zero except the third derivative; $u = x^3$ over the first interval. After four intervals it returns to zero—again with $u = du/dx = d^2u/dx^2 = 0$. If it is extended to the additional values $u_5 = 0$, $u_6 = 0, \ldots,$ the spline will be identically zero on these extra intervals.

This example is the spline that "dies" the fastest. It begins with $u = 0$ to the left of $x = 0$, and returns to $u = 0$ in only four intervals. It is called a basic spline, or ***B-spline***—and all others are combinations of these special four-interval splines.



$u = x^3$, then $u = 4 - 6(2-x)^2 \pm 3(2-x)^3$, then $u = (4-x)^3$

Fig. 3.6. The B-spline: a cubic with continuous second derivative.

There is also a minimum principle. The spline has the smallest bending energy

$$P(u) = \int_0^n \left(\frac{d^2 u}{dx^2} \right)^2 dx \tag{31}$$

among all functions with the correct heights $u_0, \ldots, u_n$ at the points $x = 0, \ldots, n$ (where the rings are). That gives the Euler equation $d^4u/dx^4 = 0$ away from the rings. In every interval u is a cubic, and the spline bends as little as possible. It is an excellent way to pass a curve through the prescribed points.