CHAPTER 4

Knot insertion

We have already seen that the control polygon of a spline provides a rough sketch of the spline itself. In this chapter we will study another aspect of the relationship between a spline and its control polygon. Namely, we shall see that as the distance between the knots of a spline is reduced, the control polygon approaches the spline it represents.

To reduce the knot spacing we perform *knot insertion*. Knot insertion amounts to what the name suggests, namely insertion of knots into an existing knot vector. This results in a new spline space with more B-splines and therefore more flexibility than the original spline space. This can be useful in many situations, for example in interactive design of spline curves. It turns out that the new spline space contains the original spline space as a subspace, so any spline in the original space can also be represented in terms of the B-splines in the refined space. As mentioned above, an important property of this new representation is that the control polygon will have moved closer to the spline itself. In fact, by inserting sufficiently many knots, we can make the distance between the spline and its control polygon as small as we wish. This has obvious advantages for practical computations since we can represent a function consisting of infinitely many points by a polygon with a finite number of vertexes. By combining this with other properties of splines like the convex hull property, we obtain a very powerful toolbox for algorithmic manipulation of spline functions.

We start, in Section 4.1, by showing that the control polygon of a spline converges to the spline as the knot spacing goes to zero. To prove this we make use of one property of splines that is proved later, in Chapter 8. The obvious way to reduce the knot spacing is via knot insertion, and in Section 4.2 we develop algorithms for expressing the B-spline coefficients relative to the new refined knot vector in terms of the B-spline coefficients relative to the original knot vector. In Section 4.3 we give a characterisation the B-spline coefficients as functions of the knots. This characterisation is often useful for developing the theory of splines, and in Section 4.4 this characterisation is used to obtain formulas for inserting one new knot into a spline function.

4.1 Convergence of the control polygon for spline functions

Recall that for a spline function $f(x) = \sum_i c_i B_{i,d,\boldsymbol{t}}$, the control polygon is the piecewise linear interpolant to the points (t_i^*, c_i), where $t_i^* = (t_{i+1} + \cdots + t_{i+d})/d$ is the ith knot average. Lemma 4.1 below shows that this is indeed a 'good' definition of the control polygon since c_i is close to $f(t_i^*)$, at least when the spacing in the knot vector is small. The proof of the lemma makes use of the fact that the size of a B-spline coefficient c_i can be bounded in terms of the size of the spline on the interval $[t_{i+1}, t_{i+d+1}]$. More specifically, we define the size of f on the interval $[a, b]$ in terms of the max-norm

$$\|f\|_{[a,b]} = \max_{x \in [a,b]} \big|f(x)\big|,$$

where we take the limit from the right at a and the limit from the left at b. We will prove in Lemma 9.16 that there exists a constant K_d that depends on d, but *not* on $\boldsymbol{t}$, such that the inequality

$$|c_i| \leq K_d \|f\|_{[t_{i+1}, t_{i+d}]} \tag{4.1}$$

holds.

Lemma 4.1. *Let f be a spline in $\mathbb{S}_{d,\boldsymbol{t}}$ with coefficients (c_i). Then*

$$|c_i - f(t_i^*)| \leq K(t_{i+d} - t_{i+1})^2 \|D^2 f\|_{[t_{i+1}, t_{i+d}]}, \tag{4.2}$$

where $t_i^* = (t_{i+1} + \cdots + t_{i+d})/d$, the operator D^2 denotes (one-sided) differentiation (from the right), and the constant K only depends on d.

Proof. Let i be fixed. If $t_{i+1} = t_{i+d}$ then we know from property 5 in Lemma 2.6 that $B_{i,d}(t_i^*) = 1$ so $c_i = f(t_i^*)$ and there is nothing to prove. Assume for the rest of the proof that the interval $J = (t_{i+1}, t_{i+d})$ is nonempty. Since J contains at most $d - 2$ knots, it follows from the continuity property of B-splines that f has at least two continuous derivatives on J. Let x_0 be a number in the interval J and consider the spline

$$g(x) = f(x) - f(x_0) - (x - x_0) Df(x_0)$$

which is the error in a first order Taylor expansion of f at x_0. The ith B-spline coefficient of g in $\mathbb{S}_{d,\boldsymbol{t}}$ is given by

$$b_i = c_i - f(x_0) - (t_i^* - x_0) Df(x_0).$$

Choosing $x_0 = t_i^*$ we have $b_i = c_i - f(t_i^*)$ and according to the inequality (4.1) and the error term in first order Taylor expansion we find

$$|c_i - f(t_i^*)| = |b_i| \leq K_d \|g\|_J \leq \frac{K_d (t_{i+d} - t_{i+1})^2}{2} \|D^2 f\|_J.$$

The inequality (4.2) therefore holds with $K = K_d/2$ and the proof is complete. ∎

Lemma 4.1 shows that the corners of the control polygon converge to the spline as the knot spacing goes to zero, but what does this really mean? So far we have considered the knots of a spline to be given, fixed numbers, but it is in fact possible to represent a spline on many different knot vectors. Suppose for example that the given spline f is a polynomial of degree d on the interval $[a, b] = [t_{d+1}, t_{m+1}]$, and that the knot vector $\boldsymbol{t} = (t_i)_{i=1}^{m+d+1}$ is $d+1$-regular. From Section 3.1.2 we know that f lies in $\mathbb{S}_{d,\boldsymbol{t}}$ regardless of how the interior knots in $[a, b]$ are chosen. We can therefore think of the B-spline coefficients as functions of the knots, and the difference $c_i - f(t_i^*)$ is then also a function of the knots. Lemma 4.1

tells us that this difference is bounded by $(t_{i+d}-t_{i+1})^2$ and therefore tends to zero as t_{i+1} tends to t_{i+d}. It is important to realise that this argument is in general only valid if f is kept fixed. Otherwise it may happen that $\|D^2 f\|_{[t_{i+1},t_{i+d}]}$ increases when $(t_{i+d}-t_{i+1})^2$ decreases with the result that the product remains fixed.

If f is a true piecewise polynomial with jumps in some derivatives at the knots in (a,b), we can introduce some auxiliary knots in (a,b) where the jumps in the derivatives are zero. These knots we can then move around in such a way that the control polygon approaches f.

Since the corners of the control polygon converges to the spline it is not surprising that the control polygon as a whole also converges to the spline.

Theorem 4.2. *Let $f=\sum_{i=1}^m c_i B_{i,d}$ be a spline in $\mathbb{S}_{d,\boldsymbol{t}}$, and let $\Gamma_{d,\boldsymbol{t}}(f)$ be its control polygon. Then*

$$\big\|\Gamma_{d,\boldsymbol{t}}(f)-f\big\|_{[t_1^*,t_m^*]} \le Kh^2\|D^2 f\|_{[t_1,t_{m+d+1}]}, \tag{4.3}$$

where $h=\max_i\{t_{i+1}-t_i\}$ and the constant K only depends on d.

Proof. As usual, we assume that $\boldsymbol{t}$ is $d+1$-regular (if not we extend it with $d+1$-tuple knots at either ends and add zero coefficients). Suppose that x is in $[t_1^*,t_m^*]$ and let j be such that $t_j^*\le x<t_{j+1}^*$. Observe that since the interval $J^*=(t_j^*,t_{j+1}^*)$ is nonempty we have $t_{j+1}<t_{j+d+1}$ and J^* contains at most $d-1$ knots. From the continuity property of B-splines we conclude that f has a continuous derivative and the second derivative of f is at least piecewise continuous on J^*. Let

$$g(x)=\frac{(t_{j+1}^*-x)f(t_j^*)+(x-t_j^*)f(t_{j+1}^*)}{t_{j+1}^*-t_j^*}$$

be the linear interpolant to f on this interval. We will show that both $\Gamma=\Gamma_{d,\boldsymbol{t}}(f)$ and f are close to g on J^* and then deduce that Γ is close to f because of the triangle inequality

$$\big|\Gamma(x)-f(x)\big|\le\big|\Gamma(x)-g(x)\big|+\big|g(x)-f(x)\big|. \tag{4.4}$$

Let us first consider the difference $\Gamma-g$. Note that

$$\Gamma(x)-g(x)=\frac{(t_{j+1}^*-x)(b_j-f(t_j^*))+(x-t_j^*)(b_{j+1}-f(t_{j+1}^*))}{t_{j+1}^*-t_j^*}$$

for any x in J^*. We therefore have

$$\big|\Gamma(x)-g(x)\big|\le\max\Big\{\big|b_j-f(t_j^*)\big|,\big|b_{j+1}-f(t_{j+1}^*)\big|\Big\},$$

for $x\in J^*$. From Lemma 4.1 we then conclude that

$$|\Gamma(x)-g(x)|\le K_1h^2\|D^2 f\|_J,\quad x\in J^*, \tag{4.5}$$

where $J=[t_1,t_{m+d+1}]$ and K_1 depending only on d.

The second difference $f(x)-g(x)$ in (4.4) is the error in linear interpolation to f at the endpoints of J^*. For this process we have the standard error estimate

$$\big|f(x)-g(x)\big|\le\frac{1}{8}(t_{j+1}^*-t_j^*)^2\|D^2 f\|_{J^*}\le\frac{1}{8}h^2\|D^2 f\|_J,\quad x\in J^*. \tag{4.6}$$

If we now combine (4.5) and (4.6) as indicated in (4.4), we obtain the Theorem with constant $K=K_1+1/8$. ∎

Because of the factor h^2 in Theorem 4.2 we say (somewhat loosely) that the control polygon converges quadratically to the spline.

4.2 Knot insertion

In Section 4.1 we showed that the control polygon converges to the spline it represents when the knot spacing tends to zero. In this section we shall develop two algorithms for reducing the knot spacing by inserting new (artificial) knots into a spline. The two algorithms for knot insertion are closely related to Algorithms 2.20 and 2.21; in fact these two algorithms are special cases of the algorithms we develop here.

Knot insertion is exactly what the name suggests: extension of a given knot vector by adding new knots. Let us first define precisely what we mean by knot insertion, or knot refinement as it is also called.

Definition 4.3. *A knot vector $\boldsymbol{t}$ is said to be a refinement of a knot vector $\boldsymbol{\tau}$ if any real number occurs at least as many times in $\boldsymbol{t}$ as in $\boldsymbol{\tau}$.*

A simple example of a knot vector and a refinement is given by

$$\boldsymbol{\tau} = (0,0,0,3,4,5,5,6,6,6) \qquad \text{and} \qquad \boldsymbol{t} = (0,0,0,2,2,3,3,4,5,5,5,6,6,6).$$

Here two knots have been inserted at 2, one at 3 and one at 5.

Note that if $\boldsymbol{t}$ is a refinement of $\boldsymbol{\tau}$ then $\boldsymbol{\tau}$ is a subsequence of $\boldsymbol{t}$, and this we will write $\boldsymbol{\tau} \subseteq \boldsymbol{t}$. The term knot insertion is used because in most situations the knot vector $\boldsymbol{\tau}$ is given and $\boldsymbol{t}$ is obtained by 'inserting' knots into $\boldsymbol{\tau}$.

With some polynomial degree d given, we can associate the spline spaces $\mathbb{S}_{d,\boldsymbol{\tau}}$ and $\mathbb{S}_{d,\boldsymbol{t}}$ with the two knot vectors $\boldsymbol{\tau}$ and $\boldsymbol{t}$. When $\boldsymbol{\tau}$ is a subsequence of $\boldsymbol{t}$, the two spline spaces are also related.

Lemma 4.4. *Let d be a positive integer and let $\boldsymbol{\tau}$ be a knot vector with at least $d+2$ knots. If $\boldsymbol{t}$ is a knot vector which contains $\boldsymbol{\tau}$ as a subsequence then $\mathbb{S}_{d,\boldsymbol{\tau}} \subseteq \mathbb{S}_{d,\boldsymbol{t}}$.*

Proof. Suppose first that both $\boldsymbol{\tau}$ and $\boldsymbol{t}$ are $d+1$-regular knot vectors with common knots at the ends. By the Curry-Schoenberg theorem (Theorem 3.25) we know that $\mathbb{S}_{d,\boldsymbol{t}}$ contains all splines with smoothness prescribed by the knot vector $\boldsymbol{t}$. Since all knots occur at least as many times in $\boldsymbol{t}$ as in $\boldsymbol{\tau}$, we see that at any knot, a spline in $\mathbb{S}_{d,\boldsymbol{\tau}}$ is at least as smooth as required for a spline in $\mathbb{S}_{d,\boldsymbol{t}}$. We therefore conclude that $\mathbb{S}_{d,\boldsymbol{\tau}} \subseteq \mathbb{S}_{d,\boldsymbol{t}}$.

A proof in the general case, where $\boldsymbol{\tau}$ and $\boldsymbol{t}$ are not $d+1$-regular with common knots at the ends, is outlined in exercise 5. ∎

Suppose that f is a spline in $\mathbb{S}_{d,\boldsymbol{\tau}}$ with B-spline coefficients $\boldsymbol{c} = (c_j)$ so that $f = \sum_j c_j B_{j,d,\boldsymbol{\tau}}$. If $\boldsymbol{\tau}$ is a subsequence of $\boldsymbol{t}$, we know from Lemma 4.4 that $\mathbb{S}_{d,\boldsymbol{\tau}}$ is a subspace of $\mathbb{S}_{d,\boldsymbol{t}}$ so f must also lie in $\mathbb{S}_{d,\boldsymbol{t}}$. Hence there exist real numbers $\boldsymbol{b} = (b_i)$ with the property that $f = \sum_i b_i B_{i,d,\boldsymbol{t}}$, i.e., the vector $\boldsymbol{b}$ contains the B-spline coefficients of f in $\mathbb{S}_{d,\boldsymbol{t}}$. Knot insertion is therefore nothing but a change of basis from the B-spline basis in $\mathbb{S}_{d,\boldsymbol{\tau}}$ to the B-spline basis in $\mathbb{S}_{d,\boldsymbol{t}}$.

Since $\mathbb{S}_{d,\boldsymbol{\tau}} \subseteq \mathbb{S}_{d,\boldsymbol{t}}$ all the B-splines in $\mathbb{S}_{d,\boldsymbol{\tau}}$ are also in $\mathbb{S}_{d,\boldsymbol{t}}$ so that

$$B_{j,d,\boldsymbol{\tau}} = \sum_{i=1}^{m} \alpha_{j,d}(i) B_{i,d,\boldsymbol{t}}, \quad j = 1,2,\ldots,n, \tag{4.7}$$

for certain numbers $\alpha_{j,d}(i)$. In matrix form this can be written

$$\boldsymbol{B}_{\boldsymbol{\tau}}^T = \boldsymbol{B}_{\boldsymbol{t}}^T \boldsymbol{A}, \tag{4.8}$$

where $\boldsymbol{B}_{\boldsymbol{\tau}}^T = (B_{1,d,\boldsymbol{\tau}}, \ldots, B_{n,d,\boldsymbol{\tau}})$ and $\boldsymbol{B}_{\boldsymbol{t}}^T = (B_{1,d,\boldsymbol{t}}, \ldots, B_{m,d,\boldsymbol{t}})$ are row vectors, and the $m \times n$-matrix $\boldsymbol{A} = \big(\alpha_{j,d}(i)\big)$ is the basis transformation matrix. Using this notation we can write f in the form

$$f = \boldsymbol{B}_{\boldsymbol{\tau}}^T \boldsymbol{c} = \boldsymbol{B}_{\boldsymbol{t}}^T \boldsymbol{b},$$

where $\boldsymbol{b}$ and $\boldsymbol{c}$ are related by

$$\boldsymbol{b} = \boldsymbol{A}\boldsymbol{c}, \qquad \text{or} \qquad b_i = \sum_{j=1}^{n} a_{i,j} c_j \quad \text{for } i = 1,\, 2,\, \ldots,\, m. \tag{4.9}$$

The basis transformation $\boldsymbol{A}$ is called *the knot insertion matrix of degree d from $\boldsymbol{\tau}$ to $\boldsymbol{t}$* and we will use the notation $\alpha_{j,d}(i) = \alpha_{j,d,\boldsymbol{\tau},\boldsymbol{t}}(i)$ for its entries. The discrete function $\alpha_{j,d}$ has many properties similar to those of $B_{j,d}$, and it is therefore called a *discrete B-spline on $\boldsymbol{t}$ with knots $\boldsymbol{\tau}$.*

To illustrate these ideas, let us consider a couple of simple examples of knot insertion for splines.

Example 4.5. Let us determine the transformation matrix $\boldsymbol{A}$ for splines with $d = 0$, when the coarse knot vector is given by $\boldsymbol{\tau} = (0, 1, 2)$, and the refined knot vector is $\boldsymbol{t} = (0, 1/2, 1, 3/2, 2) = (t_i)_{i=1}^5$. In this case

$$\mathbb{S}_{d,\boldsymbol{\tau}} = \operatorname{span}\{B_{1,0,\boldsymbol{\tau}}, B_{2,0,\boldsymbol{\tau}}\} \quad \text{and} \quad \mathbb{S}_{d,\boldsymbol{t}} = \operatorname{span}\{B_{1,0,\boldsymbol{t}}, B_{2,0,\boldsymbol{t}}, B_{3,0,\boldsymbol{t}}, B_{4,0,\boldsymbol{t}}\}.$$

We clearly have

$$B_{1,0,\boldsymbol{\tau}} = B_{1,0,\boldsymbol{t}} + B_{2,0,\boldsymbol{t}}, \quad B_{2,0,\boldsymbol{\tau}} = B_{3,0,\boldsymbol{t}} + B_{4,0,\boldsymbol{t}}.$$

This means that the knot insertion matrix in this case is given by

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}.$$

Example 4.6. Let us also consider an example with linear splines. Let $d = 1$, and let $\boldsymbol{\tau}$ and $\boldsymbol{t}$ be as in the preceding example. In this case $\dim \mathbb{S}_{d,\boldsymbol{\tau}} = 1$ and we find that

$$B(x \,|\, 0, 1, 2) = \frac{1}{2} B(x \,|\, 0, 1/2, 1) + B(x \,|\, 1/2, 1, 3/2) + \frac{1}{2} B(x \,|\, 1, 3/2, 2).$$

The situation is shown in Figure 4.1. The linear B-spline on $\boldsymbol{\tau}$ is a weighted sum of the three B-splines (dashed) on $\boldsymbol{t}$. The knot insertion matrix $\boldsymbol{A}$ is therefore the 3×1-matrix, or row vector, given by

$$\boldsymbol{A} = \begin{pmatrix} 1/2 \\ 1 \\ 1/2 \end{pmatrix}.$$

4.2.1 Formulas and algorithms for knot insertion

To develop algorithms for performing knot insertion we need to study the matrix $\boldsymbol{A}$ in some more detail. Suppose as before that we have two knot vectors $\boldsymbol{\tau}$ and $\boldsymbol{t}$ with $\boldsymbol{\tau} \subseteq \boldsymbol{t}$

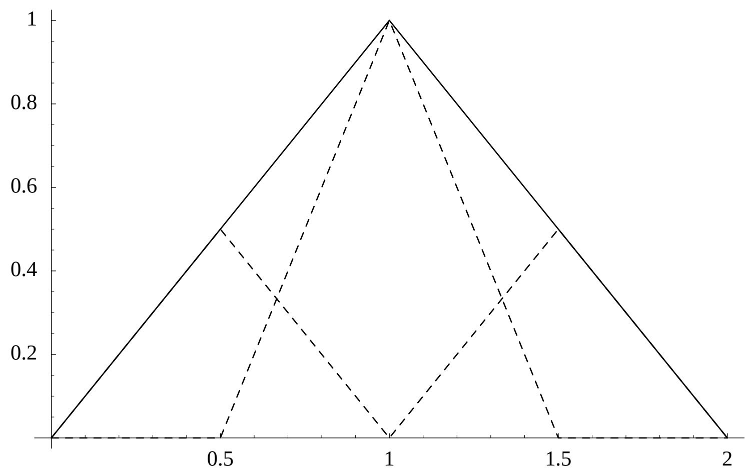


Figure 4.1. Refining a linear B-spline.

and a spline function $f = \sum_j c_j B_{j,d,\boldsymbol{\tau}}$ in $\mathbb{S}_{d,\boldsymbol{\tau}}$. Since the (i,j)-entry of $\boldsymbol{A}$ is $\alpha_{j,d}(i)$, the B-spline coefficients of f relative to $\mathbb{S}_{d,\boldsymbol{t}}$ are given by

$$b_i = \sum_{j=1}^{n} \alpha_{j,d}(i) c_j \tag{4.10}$$

for $i = 1, \ldots, m$, see (4.9). Similarly, equation (4.8) can now be written

$$B_{j,d,\boldsymbol{\tau}} = \sum_{i=1}^{m} \alpha_{j,d}(i) B_{i,d,\boldsymbol{t}} \tag{4.11}$$

for $j = 1, \ldots, n$.

In the following we make the assumption that $\boldsymbol{\tau} = (\tau_j)_{j=1}^{n+d+1}$ and $\boldsymbol{t} = (t_i)_{i=1}^{m+d+1}$ are both $d+1$-regular knot vectors with common knots at the ends so that $\tau_1 = t_1$ and $t_{n+1} = t_{m+1}$. Exercise 6 shows that this causes no loss of generality. The following theorem gives an explicit formula for the knot insertion matrix $\boldsymbol{A}$.

Recall from Theorem 2.18 the the B-spline matrix $\boldsymbol{R}_k(x) = \boldsymbol{R}^{\mu}_{k,\boldsymbol{\tau}}(x)$ is given by

$$\boldsymbol{R}^{\mu}_{k,\boldsymbol{\tau}}(x) = \begin{pmatrix} \frac{\tau_{\mu+1}-x}{\tau_{\mu+1}-\tau_{\mu+1-k}} & \frac{x-\tau_{\mu+1-k}}{\tau_{\mu+1}-\tau_{\mu+1-k}} & 0 & \cdots & 0 \\ 0 & \frac{\tau_{\mu+2}-x}{\tau_{\mu+2}-\tau_{\mu+2-k}} & \frac{x-\tau_{\mu+2-k}}{\tau_{\mu+2}-\tau_{\mu+2-k}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{\tau_{\mu+k}-x}{\tau_{\mu+k}-\tau_{\mu}} & \frac{x-\tau_{\mu}}{\tau_{\mu+k}-\tau_{\mu}} \end{pmatrix}.$$

Theorem 4.7. *Let the polynomial degree d be given, and let $\boldsymbol{\tau} = (\tau_j)_{j=1}^{n+d+1}$ and $\boldsymbol{t} = (t_i)_{i=1}^{m+d+1}$ be two $d+1$-regular knot vectors with common knots at the ends and $\boldsymbol{\tau} \subseteq \boldsymbol{t}$. In row i of the knot insertion matrix $\boldsymbol{A}$ the entries are given by $\alpha_{j,d}(i) = 0$ for $j < \mu - d$*

and $j > \mu$, where μ is determined by $\tau_\mu \le t_i < \tau_{\mu+1}$ and

$$\boldsymbol{\alpha}_d(i)^T = \big(\alpha_{\mu-d,d}(i), \ldots, \alpha_{\mu,d}(i)\big) = \begin{cases} 1, & \text{if } d = 0, \\ \boldsymbol{R}^\mu_{1,\boldsymbol{\tau}}(t_{i+1})\boldsymbol{R}^\mu_{2,\boldsymbol{\tau}}(t_{i+2}) \cdots \boldsymbol{R}^\mu_{d,\boldsymbol{\tau}}(t_{i+d}), & \text{if } d > 0. \end{cases} \tag{4.12}$$

If $f = \sum_j c_j B_{j,d,\boldsymbol{\tau}}$ is a spline in $\mathbb{S}_{d,\boldsymbol{\tau}}$, with B-spline coefficients $\boldsymbol{b}$ in $\mathbb{S}_{d,\boldsymbol{t}}$, then b_i is given by

$$b_i = \sum_{j=\mu-d}^{\mu} \alpha_{j,d}(i) c_j = \boldsymbol{R}^\mu_{1,\boldsymbol{\tau}}(t_{i+1}) \cdots \boldsymbol{R}^\mu_{d,\boldsymbol{\tau}}(t_{i+d}) \boldsymbol{c}_d, \tag{4.13}$$

where $\boldsymbol{c}_d = (c_{\mu-d}, \ldots, c_\mu)$.

Proof. We already know that the two B-spline bases are related through the relation (4.11); we will obtain the result by analysing this equation for different values of j.

Let ν be the largest integer such that $t_\nu = t_i$. This means that $\nu - d \le i \le \nu$ (note that $\nu = i$ only if $t_i < t_{i+1}$). Since $\boldsymbol{\tau}$ is a subsequence of $\boldsymbol{t}$ we have $[t_\nu, t_{\nu+1}) \subseteq [\tau_\mu, \tau_{\mu+1})$. Restrict x to the interval $[t_\nu, t_{\nu+1})$ so that $B_{\ell,d,\boldsymbol{t}}(x) = 0$ for $\ell < \nu - d$ and $\ell > \nu$. Equation (4.11) can then be written

$$B_{j,d,\boldsymbol{\tau}}(x) = \sum_{\ell=\nu-d}^{\nu} \alpha_{j,d}(\ell) B_{\ell,d,\boldsymbol{t}}(x), \quad \text{for } j = 1, \ldots, n. \tag{4.14}$$

Since $x \in [\tau_\mu, \tau_{\mu+1})$ we have $B_{j,d,\boldsymbol{\tau}}(x) = 0$ for $j < \mu - d$ or $j > \mu$. For these values of j the left-hand side of (4.14) is therefore zero, and since the B-splines $\{B_{\ell,d,\boldsymbol{t}}\}_{\ell=\nu-d}^{\nu}$ are linearly independent we must have $\alpha_{j,d}(\ell) = 0$ for $\nu - d \le \ell \le \nu$, and in particular $\alpha_{j,d}(i) = 0$, for $j < \mu - d$ and $j > \mu$.

To establish (4.12) we consider the remaining values of j, namely $j = \mu - d, \ldots, \mu$. On the interval $[t_\nu, t_{\nu+1})$, the nonzero part of the two B-spline bases can be represented by the two vectors $\boldsymbol{B}_{d,\boldsymbol{\tau}} = (B_{k,d,\boldsymbol{\tau}})_{k=\mu-d}^{\mu}$ and $\boldsymbol{B}_{d,\boldsymbol{t}} = (B_{\ell,d,\boldsymbol{t}})_{\ell=\nu-d}^{\nu}$. We also have the vectors of dual polynomials $\boldsymbol{\sigma}_d(y) = \big(\sigma_{k,d}(y)\big)_{k=\mu-d}^{\mu}$ and $\boldsymbol{\rho}_d(y) = \big(\rho_{\ell,d}(y)\big)_{\ell=\nu-d}^{\nu}$ given by

$$\begin{aligned} \sigma_{k,d}(y) &= (y - \tau_{k+1}) \cdots (y - \tau_{k+d}), \\ \rho_{\ell,d}(y) &= (y - t_{\ell+1}) \cdots (y - t_{\ell+d}). \end{aligned}$$

From Corollary 3.2 we have that the two sets of dual polynomials are related by

$$\rho_{\ell,d}(y) = \boldsymbol{R}_1(t_{\ell+1}) \cdots \boldsymbol{R}_d(t_{\ell+d}) \boldsymbol{\sigma}_d(y) \tag{4.15}$$

(in this proof we omit the second subscript and the superscript to the $\boldsymbol{R}$ matrices). Combining this with two versions of Marsden's identity (3.10) then yields

$$\begin{aligned} \boldsymbol{B}_{d,\boldsymbol{\tau}}(x)^T \boldsymbol{\sigma}_d(y) = (y - x)^d &= \sum_{\ell=\nu-d}^{\nu} \rho_{\ell,d}(y) B_{\ell,d,\boldsymbol{t}}(x) \\ &= \sum_{\ell=\nu-d}^{\nu} B_{\ell,d,\boldsymbol{t}}(x) \boldsymbol{R}_1(t_{\ell+1}) \cdots \boldsymbol{R}_d(t_{\ell+d}) \boldsymbol{\sigma}_d(y). \end{aligned}$$

The linear independence of the $d+1$ dual polynomials $\boldsymbol{\sigma}_d(y)$ allows us to conclude that

$$\boldsymbol{B}_{d,\boldsymbol{\tau}}(x)^T = \sum_{\ell=\nu-d}^{\nu} B_{\ell,d,\boldsymbol{t}}(x)\boldsymbol{R}_1(t_{\ell+1})\cdots\boldsymbol{R}_d(t_{\ell+d}) = \boldsymbol{B}_{d,\boldsymbol{t}}(x)^T \begin{pmatrix} \boldsymbol{R}_1(t_{\nu-d+1})\cdots\boldsymbol{R}_d(t_\nu) \\ \boldsymbol{R}_1(t_{\nu-d+2})\cdots\boldsymbol{R}_d(t_{\nu+1}) \\ \vdots \\ \boldsymbol{R}_1(t_{\nu+1})\cdots\boldsymbol{R}_d(t_{\nu+d}) \end{pmatrix}$$

for any x in the interval $[t_\nu, t_{\nu+1})$. Comparing this equation with (4.14) and making use of the linear independence of B-splines shows that the matrix on the right is the submatrix of $\boldsymbol{A}$ given by $(\alpha_{j,d}(i))_{i=\nu-d,j=\mu-d}^{\nu,\mu}$. Since $\nu-d \le i \le \nu$ equation (4.12) follows. Equation (4.13) now follows from (4.10). ∎

Note that if no new knots are inserted ($\boldsymbol{\tau}=\boldsymbol{t}$) then the two sets of B-spline coefficients $\boldsymbol{c}$ and $\boldsymbol{b}$ are obviously the same. Equation (4.13) then shows that

$$c_i = \boldsymbol{R}^{\mu}_{1,\boldsymbol{\tau}}(\tau_{i+1})\cdots\boldsymbol{R}^{\mu}_{d,\boldsymbol{\tau}}(\tau_{i+d})\boldsymbol{c}_d. \tag{4.16}$$

This simple observation will be useful later.

An example will illustrate the use of Theorem 4.7.

Example 4.8. We consider quadratic splines ($d=2$) on the knot vector $\boldsymbol{\tau}=(-1,-1,-1,0,1,1,1)$, and insert two new knots, at $-1/2$ and $1/2$ so $\boldsymbol{t}=(-1,-1,-1,-1/2,0,1/2,1,1,1)$. We note that $\tau_3 \le t_i < \tau_4$ for $1 \le i \le 4$ so the first three entries of the first four rows of the 6×4 knot insertion matrix $\boldsymbol{A}$ are given by

$$\boldsymbol{\alpha}_2(i) = \boldsymbol{R}^3_{1,\boldsymbol{\tau}}(t_{i+1})\boldsymbol{R}^3_{2,\boldsymbol{\tau}}(t_{i+2})$$

for $i=1,\ldots,4$. Since

$$\boldsymbol{R}^3_{1,\boldsymbol{\tau}}(x) = \begin{pmatrix} -x & 1+x \end{pmatrix}, \qquad \boldsymbol{R}^3_{2,\boldsymbol{\tau}}(x) = \begin{pmatrix} -x & 1+x & 0 \\ 0 & (1-x)/2 & (1+x)/2 \end{pmatrix},$$

we have from (4.12)

$$\boldsymbol{\alpha}_2(i) = \frac{1}{2}\begin{pmatrix} 2t_{i+1}t_{i+2}, & 1-t_{i+1}-t_{i+2}-3t_{i+1}t_{i+2}, & (1+t_{i+1})(1+t_{i+2}) \end{pmatrix}.$$

Inserting the correct values for t_{i+1} and t_{i+2} and adding one zero at the end of each row, we find that the first four rows of $\boldsymbol{A}$ are given by

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 0 & 3/4 & 1/4 & 0 \\ 0 & 1/4 & 3/4 & 0 \end{pmatrix}.$$

To determine the remaining two rows of $\boldsymbol{A}$ we have to move to the interval $[\tau_4,\tau_5)=[0,1)$. Here we have

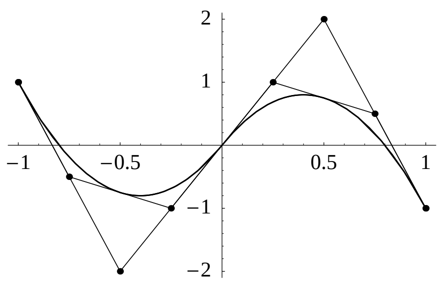
$$\boldsymbol{R}^4_{1,\boldsymbol{\tau}}(x) = \begin{pmatrix} 1-x & x \end{pmatrix} \qquad \boldsymbol{R}^4_{2,\boldsymbol{\tau}}(x) = \begin{pmatrix} (1-x)/2 & (1+x)/2 & 0 \\ 0 & 1-x & x \end{pmatrix},$$

so

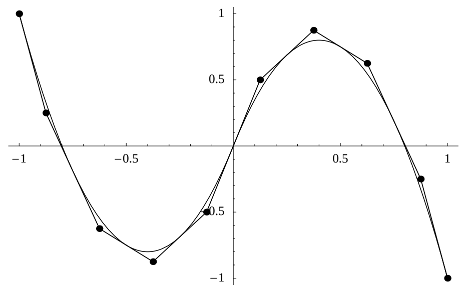
$$\boldsymbol{a}_2(i) = \boldsymbol{R}^4_{1,\boldsymbol{\tau}}(t_{i+1})\boldsymbol{R}^4_{2,\boldsymbol{\tau}}(t_{i+2}) = \frac{1}{2}\begin{pmatrix} (1-t_{i+1})(1-t_{i+2}), & 1+t_{i+1}+t_{i+2}-3t_{i+1}t_{i+2}, & 2t_{i+1}t_{i+2} \end{pmatrix}.$$

Evaluating this for $i=5$, 6 and inserting one zero as the first entry, we obtain the last two rows as

$$\begin{pmatrix} 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$



(a)



(b)

Figure 4.2. A quadratic spline together with its control polygon relative to a coarse and a finer knot vector (a), and the same spline as in (a) with its control polygon relative to an even more refined knot vector (b).

To see visually the effect of knot insertion, let $f = B_{1,2,\boldsymbol{\tau}} - 2B_{2,2,\boldsymbol{\tau}} + 2B_{3,2,\boldsymbol{\tau}} - B_{4,2,\boldsymbol{\tau}}$ be a spline in $\mathbb{S}_{d,\boldsymbol{\tau}}$ with B-spline coefficients $\boldsymbol{c} = (1, -2, 2, -1)^T$. Its coefficients $\boldsymbol{b} = (b_i)_{i=1}^6$ are then given by

$$\boldsymbol{b} = \boldsymbol{A}\boldsymbol{c} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 0 & 3/4 & 1/4 & 0 \\ 0 & 1/4 & 3/4 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -2 \\ 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1/2 \\ -1 \\ 1 \\ 1/2 \\ -1 \end{pmatrix}.$$

Figure 4.2 (a) shows a plot of f together with its control polygons relative to $\boldsymbol{\tau}$ and $\boldsymbol{t}$. We note that the control polygon relative to $\boldsymbol{t}$ is much closer to f and that both control polygons give a rough estimate of f.

The knot insertion process can be continued. If we insert one new knot halfway between each old knot in $\boldsymbol{t}$, we obtain the new knot vector

$$\boldsymbol{t}^1 = (-1, -1, -1, -3/4, -1/2, -1/4, 0, 1/4, 1/2, 3/4, 1, 1, 1).$$

A plot of f and its control polygon relative to this knot vector is shown in Figure 4.2 (b).

Example 4.9. Let us again consider quadratic splines on a uniform knot vector with multiple knots at the ends,

$$\boldsymbol{\tau} = (\tau_j)_{j=1}^{n+3} = (3, 3, 3, 4, 5, 6, \dots, n, n+1, n+1, n+1),$$

and form $\boldsymbol{t}$ by inserting one knot half way between each pair of old knots,

$$\boldsymbol{t} = (t_i)_{i=1}^{2n+1} = (3, 3, 3, 7/2, 4, 9/2, 5, \dots, n, (2n+1)/2, n+1, n+1, n+1).$$

Since $\dim \mathbb{S}_{d,\boldsymbol{\tau}} = n$ and $\dim \mathbb{S}_{d,\boldsymbol{t}} = 2n-2$, the knot insertion matrix $\boldsymbol{A}$ is now a $(2n-2) \times n$ matrix. As in Example 4.8 we find that the first three columns of the first four rows of $\boldsymbol{A}$ are

$$\begin{pmatrix} 1 & 0 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 3/4 & 1/4 \\ 0 & 1/4 & 3/4 \end{pmatrix}.$$

To determine rows $2\mu - 3$ and $2\mu - 2$ with $4 \le \mu \le n-1$, we need the matrices $\boldsymbol{R}_{1,\boldsymbol{\tau}}^{\mu}$ and $\boldsymbol{R}_{2,\boldsymbol{\tau}}^{\mu}$ which are given by

$$\boldsymbol{R}_{1,\boldsymbol{\tau}}^{\mu}(x) = \begin{pmatrix} \mu + 1 - x & x - \mu \end{pmatrix}, \qquad \boldsymbol{R}_{2,\boldsymbol{\tau}}^{\mu}(x) = \begin{pmatrix} (\mu+1-x)/2 & (x+1-\mu)/2 & 0 \\ 0 & (\mu+2-x)/2 & (x-\mu)/2 \end{pmatrix}.$$

Observe that $\tau_i = i$ for $i = 3, \dots, n+1$ and $t_i = (i+3)/2$ for $i = 3, \dots, 2n-1$. Entries $\mu - 2$, $\mu - 1$ and μ of row $2\mu - 3$ are therefore given by

$$\boldsymbol{R}_{1,\boldsymbol{\tau}}^{\mu}(t_{2\mu-2})\boldsymbol{R}_{2,\boldsymbol{\tau}}^{\mu}(t_{2\mu-1}) = \boldsymbol{R}_{1,\boldsymbol{\tau}}^{\mu}(\mu + 1/2)\boldsymbol{R}_{2,\boldsymbol{\tau}}^{\mu}(\mu+1) = \begin{pmatrix} 1/2 & 1/2 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1/2 & 1/2 \end{pmatrix} = \begin{pmatrix} 0 & 3/4 & 1/4 \end{pmatrix}.$$

Similarly, entries $\mu-3$, $\mu-2$ and μ of row $2\mu-2$ are given by

$$\boldsymbol{R}^{\mu}_{1,\boldsymbol{\tau}}(t_{2\mu-1})\boldsymbol{R}^{\mu}_{2,\boldsymbol{\tau}}(t_{2\mu}) = \boldsymbol{R}^{\mu}_{1,\boldsymbol{\tau}}(\mu+1)\boldsymbol{R}^{\mu}_{2,\boldsymbol{\tau}}(\mu+3/2) = \begin{pmatrix}0 & 1\end{pmatrix}\begin{pmatrix}-1/4 & 5/4 & 0\\ 0 & 1/4 & 3/4\end{pmatrix} = \begin{pmatrix}0 & 1/4 & 3/4\end{pmatrix}.$$

Finally, we find as in Example 4.8 that the last three entries of the last two rows are

$$\begin{pmatrix}0 & 1/2 & 1/2\\ 0 & 0 & 1\end{pmatrix}.$$

The complete knot insertion matrix is therefore

$$\boldsymbol{A} = \begin{pmatrix}
1 & 0 & 0 & 0 & \dots & 0 & 0 & 0\\
1/2 & 1/2 & 0 & 0 & \dots & 0 & 0 & 0\\
0 & 3/4 & 1/4 & 0 & \dots & 0 & 0 & 0\\
0 & 1/4 & 3/4 & 0 & \dots & 0 & 0 & 0\\
0 & 0 & 3/4 & 1/4 & \dots & 0 & 0 & 0\\
0 & 0 & 1/4 & 3/4 & \dots & 0 & 0 & 0\\
\vdots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots\\
0 & 0 & 0 & 0 & \dots & 3/4 & 1/4 & 0\\
0 & 0 & 0 & 0 & \dots & 1/4 & 3/4 & 0\\
0 & 0 & 0 & 0 & \dots & 0 & 1/2 & 1/2\\
0 & 0 & 0 & 0 & \dots & 0 & 0 & 1
\end{pmatrix}.$$

The formula for $\boldsymbol{\alpha}_d(i)$ shows very clearly the close relationship between B-splines and discrete B-splines, and it will come as no surprise that $\alpha_{j,d}(i)$ satisfies a recurrence relation similar to that of B-splines, see Definition 2.1. The recurrence for $\alpha_{j,d}(i)$ is obtained by setting $x = t_{i+d}$ in the recurrence (2.1) for $B_{j,d}(x)$,

$$\alpha_{j,d}(i) = \frac{t_{i+d}-\tau_j}{\tau_{j+d}-\tau_j}\alpha_{j,d-1}(i) + \frac{\tau_{j+1+d}-t_{i+d}}{\tau_{j+1+d}-\tau_{j+1}}\alpha_{j+1,d-1}(i), \tag{4.17}$$

starting with $\alpha_{j,0}(i) = B_{j,0}(t_i)$.

The two evaluation algorithms for splines, Algorithms 3.17 and 3.18, can be adapted to knot insertion quite easily. For historical reasons these algorithms are usually referred to as *the Oslo algorithms*.

Algorithm 4.10 (Oslo-Algorithm 1)**.** *Let the polynomial degree d, and the two $d+1$-regular knot vectors $\boldsymbol{\tau} = (\tau_j)_{j=1}^{n+d+1}$ and $\boldsymbol{t} = (t_i)_{i=1}^{m+d+1}$ with common knots at the ends be given. To compute the $m\times n$ knot insertion matrix $\boldsymbol{A} = \big(\alpha_{j,d}(i)\big)_{i,j=1}^{m,n}$ from $\boldsymbol{\tau}$ to $\boldsymbol{t}$ perform the following steps:*

1. *For $i = 1, \dots, m$.*

  1.1 *Determine μ such that $\tau_\mu \le t_i < \tau_{\mu+1}$.*

  1.2 *Compute entries $\mu-d, \dots, \mu$ of row i by evaluating*

$$\boldsymbol{\alpha}_d(i)^T = \big(\alpha_{\mu-d,d}(i), \dots, \alpha_{\mu,d}(i)\big)^T = \begin{cases}1, & \text{if } d = 0.\\ \boldsymbol{R}_1(t_{i+1})\cdots\boldsymbol{R}_d(t_{i+d}), & \text{if } d > 0.\end{cases}$$

  All other entries in row i are zero.

An algorithm for converting a spline from a B-spline representation in $\mathbb{S}_{d,\boldsymbol{\tau}}$ to $\mathbb{S}_{d,\boldsymbol{t}}$ is as follows.

Algorithm 4.11 (Oslo-Algorithm 2). *Let the polynomial degree d, and the two $d+1$-regular knot vectors $\boldsymbol{\tau} = (\tau_j)_{j=1}^{n+d+1}$ and $\boldsymbol{t} = (t_i)_{i=1}^{m+d+1}$ with common knots at the ends be given together with the spline f in $\mathbb{S}_{d,\boldsymbol{\tau}}$ with B-spline coefficients $\boldsymbol{c} = (c_j)_{j=1}^{n}$. To compute the B-spline coefficients $\boldsymbol{b} = (b_i)_{i=1}^{m}$ of f in $\mathbb{S}_{d,\boldsymbol{t}}$ perform the following steps:*

1. *For $i = 1, \ldots, m$.*

  1.1 *Determine μ such that $\tau_\mu \le t_i < \tau_{\mu+1}$.*

  1.2 *Set $\boldsymbol{c}_d = (c_j)_{j=\mu-d}^{\mu}$ and compute b_i by evaluating*

$$b_i = \begin{cases} c_\mu, & \text{if } d = 0. \\ \boldsymbol{R}_1(t_{i+1}) \cdots \boldsymbol{R}_d(t_{i+d})\boldsymbol{c}_d, & \text{if } d > 0. \end{cases}$$

4.3 B-spline coefficients as functions of the knots

Knot insertion allows us to represent the same spline function on different knot vectors. In fact, any spline function can be given any real numbers as knots, as long as we also include the original knots. It therefore makes sense to consider the B-spline coefficients as functions of the knots, and we shall see that this point of view allows us to characterise the B-spline coefficients completely by three simple properties.

Initially, we assume that the spline $f = \sum_{j=1}^{n} c_j B_{j,d,\boldsymbol{\tau}}$ is a polynomial represented on a $d+1$-extended knot vector $\boldsymbol{\tau}$. On the knot interval $[\tau_\mu, \tau_{\mu+1})$ we know that f can be written as

$$f(x) = \boldsymbol{R}_1(x) \cdots \boldsymbol{R}_d(x)\boldsymbol{c}_d, \tag{4.18}$$

where $\boldsymbol{c}_d = (c_{\mu-d}, \ldots, c_\mu)^T$, see Section 2.3. Since f is assumed to be a polynomial this representation is valid for all real numbers x, although when x is outside $[\tau_\mu, \tau_{\mu+1})$ it is no longer a true B-spline representation.

Consider the function

$$F(x_1, \ldots, x_d) = \boldsymbol{R}_1(x_1) \cdots \boldsymbol{R}_d(x_d)\boldsymbol{c}_d. \tag{4.19}$$

We recognise the right-hand side of this expression from equation (4.13) in Theorem 4.7: If we have a knot vector that includes the knots $(x_0, x_1, \ldots, x_d, x_{d+1})$, then $F(x_1, \ldots, x_d)$ gives the B-spline coefficient that multiplies the B-spline $B(x \mid x_0, \ldots, x_{d+1})$ in the representation of the polynomial f on the knot vector $\boldsymbol{x}$. When f is a polynomial, it turns out that the function F is completely independent of the knot vector $\boldsymbol{\tau}$ that underlie the definition of the $\boldsymbol{R}$-matrices in (4.19). The function F is referred to as the *blossom* of f, and the whole theory of splines can be built from properties of this function.

4.3.1 The blossom

In this subsection we develop some of the properties of the blossom. We will do this in an abstract fashion, by starting with a formal definition of the blossom. In the next subsection we will then show that the function F in (4.19) satisfies this definition.

Definition 4.12. *A function on the form $f(x) = ax$, where a is a real number, is called a linear function. A function on the form $f(x) = ax + b$ with a and b real constants is called an affine function. A function of d variables $f(x_1, \ldots, x_d)$ is said to be affine if it is affine viewed as a function of each x_i for $i = 1, \ldots, d$, with the other variables fixed.*

A symmetric affine function is an affine function that is not altered when the order of the variables is changed.

It is common to say that a polynomial $p(x) = a+bx$ of degree one is a linear polynomial, even when a is nonzero. According to Definition 4.12 such a polynomial is an affine polynomial, and this (algebraic) terminology will be used in the present section. Outside this section however, we will use the term linear polynomial.

For a linear function of one variable we have

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y), \quad x, y \in \mathbb{R} \tag{4.20}$$

for all real numbers α and β, while for an affine function f with $b \neq 0$ equation (4.20) only holds if $\alpha + \beta = 1$. This is in fact a complete characterisation of affine functions: If (4.20) holds with $\alpha + \beta = 1$, then f is affine, see exercise 9.

A general affine function of 2 variables is given by

$$\begin{aligned} f(x_1, x_2) &= ax_2 + b = (a_2 x_1 + b_2)x_2 + a_1 x_1 + b_1 \\ &= c_0 + c_1 x_1 + c_2 x_2 + c_{1,2} x_1 x_2. \end{aligned} \tag{4.21}$$

Similarly, an affine function of three variables is a function on the form

$$f(x_1, x_2, x_3) = c_0 + c_1 x_1 + c_2 x_2 + c_3 x_3 + c_{1,2} x_1 x_2 + c_{1,3} x_1 x_3 + c_{2,3} x_2 x_3 + c_{1,2,3} x_1 x_2 x_3.$$

In general, an affine function can be written as a linear combination of 2^d terms. This follows by induction as in (4.21) where we passed from one argument to two.

A symmetric and affine function satisfies the equation

$$f(x_1, x_2, \ldots, x_d) = f(x_{\pi_1}, x_{\pi_2}, \ldots, x_{\pi_d}),$$

for any permutation $(\pi_1, \pi_2, \ldots, \pi_d)$ of the numbers 1, 2, ..., d. We leave it as an exercise to show that symmetric, affine functions of two and three variables can be written in the form

$$\begin{aligned} f(x_1, x_2) &= a_0 + a_1(x_1 + x_2) + a_2 x_1 x_2, \\ f(x_1, x_2, x_3) &= a_0 + a_1(x_1 + x_2 + x_3) + a_2(x_1 x_2 + x_1 x_3 + x_2 x_3) + a_3 x_1 x_2 x_3. \end{aligned}$$

We are now ready to give the definition of the blossom of a polynomial.

Definition 4.13. *Let p be a polynomial of degree at most d. The blossom $\mathcal{B}[p](x_1, \ldots, x_d)$ of p is a function of d variables with the properties:*

1. *Symmetry. The blossom is a symmetric function of its arguments,*

   $$\mathcal{B}[p](x_1, \ldots, x_d) = \mathcal{B}[p](x_{\pi_1}, \ldots, x_{\pi_d})$$

   for any permutation π_1, ..., π_d of the integers 1, ..., d.

2. *Affine. The blossom is affine in each of its variables,*

   $$\mathcal{B}[p](\ldots, \alpha x + \beta y, \ldots) = \alpha \mathcal{B}[p](\ldots, x, \ldots) + \beta \mathcal{B}[p](\ldots, y, \ldots)$$

   whenever $\alpha + \beta = 1$.

3. *Diagonal property. The blossom agrees with p on the diagonal,*

$$\mathcal{B}[p](x,\dots,x) = p(x)$$

for all real numbers x.

The blossom of a polynomial exists and is unique.

Theorem 4.14. *Each polynomial p of degree d has a unique blossom $\mathcal{B}[p](x_1,\dots,x_d)$. The blossom acts linearly on p, i.e., if p_1 and p_2 are two polynomials and c_1 and c_2 are two real constants then*

$$\mathcal{B}[c_1p_1 + c_2p_2](x_1,\dots,x_d) = c_1\mathcal{B}[p_1](x_1,\dots,x_d) + c_2\mathcal{B}[p_2](x_1,\dots,x_d). \tag{4.22}$$

Proof. The proof of uniqueness follows along the lines sketched at the beginning of this section for small d. Start with a general affine function F of d variables

$$F(x_1,\dots,x_d) = c_0 + \sum_{j=1}^{d} \sum_{1\le i_1<\dots<i_j\le d} c_{i_1,\dots,i_j} x_{i_1}\cdots x_{i_j}.$$

Symmetry forces all the coefficients multiplying terms of the same degree to be identical. To see this we note first that

$$F(1,0,\dots,0) = c_0 + c_1 = F(0,\dots,1,\dots,0) = c_0 + c_i$$

for all i with $1 \le i \le d$. Hence we have $c_1 = \dots = c_d$. To prove that the terms of degree j all have the same coefficients we use induction and set j of the variables to 1 and the rest to 0. By the induction hypothesis we know that all the terms of degree less than j are symmetric; denote the contribution from these terms by p_{j-1}. Symmetry then gives

$$p_{j-1} + c_{1,2,\dots,j} = p_{j-1} + c_{1,2,\dots,j-1,j+1} = \dots = p_{j-1} + c_{d-j+1,\dots,d}.$$

From this we conclude that all the coefficients multiplying terms of degree j must be equal. We can therefore write F as

$$F(x_1,\dots,x_d) = a_0 + \sum_{j=1}^{d} a_j \sum_{1\le i_1<\dots<i_j\le d} x_{i_1}\cdots x_{i_j}, \tag{4.23}$$

for suitable constants $(a_j)_{j=0}^d$. From the diagonal property $F(x,\dots,x) = f(x)$ the coefficients $(a_j)_{j=0}^d$ are all uniquely determined (since 1, x, $\dots$, x^d is basis for π_d).

The linearity of the blossom with regards to p follows from its uniqueness: The right-hand side of (4.22) is affine in each of the x_i, it is symmetric, and it reduces to $c_1p_1(x) + c_2p_2(x)$ on the diagonal $x_1 = \dots = x_d = x$. ∎

Recall that the elementary symmetric polynomials

$$s_j(x_1,\dots,x_d) = \Big(\sum_{1\le i_1<\dots<i_j\le d} x_{i_1}x_{i_2}\cdots x_{i_j}\Big)\Big/\binom{d}{j}$$

that appear in (4.23) (apart from the binomial coefficient) agree with the B-spline coefficients of the polynomial powers,

$$\sigma_{k,d}^{j} = s_j(\tau_{k+1}, \dots, \tau_{k+d}),$$

see Corollary 3.5. In fact, the elementary symmetric polynomials are the blossoms of the powers,

$$\mathcal{B}[x^j](x_1, \dots, x_d) = s_j(x_1, \dots, x_d) \qquad \text{for } j = 0, \dots, d.$$

They can also be defined by the relation

$$(x - x_1) \cdots (x - x_d) = \sum_{k=0}^{d} (-1)^{d-k} \binom{d}{k} s_{d-k}(x_1, \dots, x_d) x^k.$$

Note that the blossom depends on the degree of the polynomial in a nontrivial way. If we consider the polynomial $p(x) = x$ to be of degree one, then $\mathcal{B}[p](x_1) = x_1$. But we can also think of p as a polynomial of degree three (the cubic and quadratic terms are zero); then we obviously have $\mathcal{B}[p](x_1, x_2, x_3) = (x_1 + x_2 + x_3)/3$.

4.3.2 B-spline coefficients as blossoms

Earlier in this chapter we have come across a function that is both affine and symmetric. Suppose we have a knot vector $\boldsymbol{\tau}$ for B-splines of degree d. On the interval $[\tau_\mu, \tau_{\mu+1})$ the only nonzero B-splines are $\boldsymbol{B}_d = (B_{\mu-d,d}, \dots, B_{\mu,d})^T$ which can be expressed in terms of matrices as

$$\boldsymbol{B}_d(x)^T = \boldsymbol{R}_1(x) \cdots \boldsymbol{R}_d(x).$$

If we consider the polynomial piece $f = \boldsymbol{B}_d^T \boldsymbol{c}_d$ with coefficients $\boldsymbol{c}_d = (c_{\mu-d}, \dots, c_\mu)^T$ we can define a function F of d variables by

$$F(x_1, \dots, x_d) = \boldsymbol{R}_1(x_1) \cdots \boldsymbol{R}_d(x_d) \boldsymbol{c}_d. \tag{4.24}$$

From equation(4.13) we recognise $F(x_1, \dots, x_d)$ as the coefficient multiplying a B-spline with knots $x_0, x_1, \dots, x_{d+1}$ in the representation of the polynomial f.

Equation (3.7) in Lemma 3.3 shows that F is a symmetric function. It is also affine in each of its variables. To verify this, we note that because of the symmetry it is sufficient to check that it is affine with respect to the first variable. Recall from Theorem 2.18 that $\boldsymbol{R}_1 = \boldsymbol{R}_{1,\boldsymbol{\tau}}$ is given by

$$\boldsymbol{R}_1(x) = \left(\frac{\tau_{\mu+1} - x}{\tau_{\mu+1} - \tau_\mu}, \quad \frac{x - \tau_\mu}{\tau_{\mu+1} - \tau_\mu} \right)$$

which is obviously an affine function of x.

The function F is also related to the polynomial f in that $F(x, \dots, x) = f(x)$. We have proved the following lemma.

Lemma 4.15. *Let $f = \sum_{j=\mu-d}^{\mu} c_j B_{j,d}$ be a polynomial represented in terms of the B-splines in $\mathbb{S}_{d,\boldsymbol{\tau}}$ on the interval $[\tau_\mu, \tau_{\mu+1})$, with coefficients $\boldsymbol{c}_d = (c_{\mu-d}, \dots, c_\mu)^T$. Then the function*

$$F(x_1, \dots, x_d) = \boldsymbol{R}_1(x_1) \cdots \boldsymbol{R}_d(x_d) \boldsymbol{c}_d$$

is symmetric and affine, and agrees with f on the diagonal,

$$F(x, \dots, x) = f(x).$$

Lemma 4.15 and Theorem 4.14 show that the blossom of f is given by

$$\mathcal{B}[f](x_1,\dots,x_d) = \boldsymbol{R}_1(x_1)\cdots\boldsymbol{R}_1(x_d)\boldsymbol{c}_d.$$

Blossoming can be used to give explicit formulas for the B-spline coefficients of a spline.

Theorem 4.16. *Let $f = \sum_{j=1}^{n} c_j B_{j,d,\boldsymbol{\tau}}$ be a spline on a $d+1$-regular knot vector $\boldsymbol{\tau} = (\tau_j)_{j=1}^{n+d+1}$. It's B-spline coefficients are then given by*

$$c_j = \mathcal{B}[f_k](\tau_{j+1},\dots,\tau_{j+d}), \quad \text{for } k = j,\, j+1,\, \dots,\, j+d, \tag{4.25}$$

provided $\tau_k < \tau_{k+1}$. Here $f_k = f|_{(\tau_k,\tau_{k+1})}$ is the restriction of f to the interval (τ_k, τ_{k+1}).

Proof. Let us first restrict x to the interval $[\tau_\mu, \tau_{\mu+1})$ and only consider one polynomial piece f_μ of f. From Lemma 4.15 we know that $\mathcal{B}[f_\mu](x_1,\dots,x_d) = \boldsymbol{R}_1(x_1)\cdots\boldsymbol{R}_d(x_d)\boldsymbol{c}_d$, where $\boldsymbol{c}_d = (c_j)_{j=\mu-d}^{\mu}$ are the B-spline coefficients of f active on the interval $[\tau_\mu, \tau_{\mu+1})$. From (4.16) we then obtain

$$c_j = \mathcal{B}[f_\mu](\tau_{j+1},\dots,\tau_{j+d}) \tag{4.26}$$

which is (4.25) in this special situation.

To prove (4.25) in general, fix j and choose the integer k in the range $j \le k \le j+d$. We then have

$$f_k(x) = \sum_{i=k-d}^{k} c_i B_{i,d}(x), \tag{4.27}$$

By the choice of k we see that the sum in (4.27) includes the term $c_j B_{j,d}$. Equation (4.25) therefore follows by applying (4.26) to f_k. ∎

The affine property allows us to perform one important operation with the blossom; we can change the arguments.

Lemma 4.17. *The blossom of p satisfies the relation*

$$\mathcal{B}[p](\dots,x,\dots) = \frac{b-x}{b-a}\mathcal{B}[p](\dots,a\dots) + \frac{x-a}{b-a}\mathcal{B}[p](\dots,b,\dots) \tag{4.28}$$

for all real numbers a, b and x with $a \ne b$.

Proof. Observe that x can be written as an affine combination of a and b,

$$x = \frac{b-x}{b-a}a + \frac{x-a}{b-a}b.$$

Equation (4.28) then follows from the affine property of the blossom. ∎

The next result will be useful later.

Lemma 4.18. *Let $\mathcal{B}_x\big[p(x,y)\big]$ denote the blossom of p with respect to the variable x. Then*

$$\mathcal{B}_x\big[(y-x)^k\big](x_1,\dots,x_d) = \frac{k!}{d!}D^{d-k}\big((y-x_1)\cdots(y-x_d)\big), \tag{4.29}$$

for $k = 0, 1, \ldots, d$, and

$$\mathcal{B}_x\big[(y_1 - x)\cdots(y_\ell - x)\big](x_1, \ldots, x_d) = \frac{(d-\ell)!}{d!} \sum_{1\le i_1,\ldots,i_\ell \le d} (y_1 - x_{i_1})\cdots(y_\ell - x_{i_\ell}), \quad (4.30)$$

where the sum is over all distinct choices $i_1, \ldots, i_\ell$ of ℓ integers from the d integers $1, \ldots, d$.

Proof. For $k = d$ equation (4.29) follows since the right-hand side is symmetric and affine in each of the variables x_i and it agrees with $(y - x)^d$ on the diagonal $x_1 = \cdots = x_d = x$. The general result is then obtained by differentiating both sides k times.

Equation (4.30) follows since the right-hand side is affine, symmetric and reduces to $(y_1 - x)\cdots(y_\ell - x)$ when $x = x_1 = \cdots = x_d$, i.e., it must be the blossom of $(y - x)^d$. ∎

4.4 Inserting one knot at a time

With blossoming we have a simple but powerful tool for determining the B-spline coefficients of splines. Here we will apply blossoming to develop an alternative knot insertion strategy. Instead of inserting all new knots simultaneously we can insert them sequentially. We insert one knot at a time and update the B-spline coefficients between each insertion. This leads to simple, explicit formulas.

Lemma 4.19 (Böhm's method). *Let $\boldsymbol{\tau} = (\tau_j)_{j=1}^{n+d+1}$ be a given knot vector and let $\boldsymbol{t} = (t_i)_{i=1}^{n+d+2}$ be the knot vector obtained by inserting a knot z in $\boldsymbol{\tau}$ in the interval $[\tau_\mu, \tau_{\mu+1})$. If*

$$f = \sum_{j=1}^{n} c_j B_{j,d,\boldsymbol{\tau}} = \sum_{i=1}^{n+1} b_i B_{i,d,\boldsymbol{t}},$$

then $(b_i)_{i=1}^{n+1}$ can be expressed in terms of $(c_j)_{j=1}^{n}$ through the formulas

$$b_i = \begin{cases} c_i, & \text{if } 1 \le i \le \mu - d; \\ \dfrac{z - \tau_i}{\tau_{i+d} - \tau_i} c_i + \dfrac{\tau_{i+d} - z}{\tau_{i+d} - \tau_i} c_{i-1}, & \text{if } \mu - d + 1 \le i \le \mu; \\ c_{i-1}, & \text{if } \mu + 1 \le i \le n + 1. \end{cases} \quad (4.31)$$

Proof. Observe that for $j \le \mu$ we have $\tau_j = t_j$. For $i \le \mu - d$ and with k an integer such that $i \le k \le i + d$ it therefore follows from (4.25) that

$$b_i = \mathcal{B}[f^k](t_{i+1}, \ldots, t_{i+d}) = \mathcal{B}[f^k](\tau_{i+1}, \ldots, \tau_{i+d}) = c_i.$$

Similarly, we have $t_i = \tau_{i-1}$ for $i \ge \mu + 1$ so

$$b_i = \mathcal{B}[f^k](t_{i+1}, \ldots, t_{i+d}) = \mathcal{B}[f^k](\tau_i, \ldots, \tau_{i+d-1}) = c_{i-1}$$

for such values of i.

When i satisfies $\mu - d + 1 \le i \le \mu$ we note that z will appear in the sequence $(t_{i+1}, \ldots, t_{i+d})$. From (4.25) we therefore obtain

$$b_i = \mathcal{B}[f^\mu](t_{i+1}, \ldots, z, \ldots, t_{i+d}) = \mathcal{B}[f^\mu](\tau_{i+1}, \ldots, z, \ldots, \tau_{i+d-1})$$

since we now may choose $k = \mu$. Applying Lemma 4.17 with $x = z$, $a = \tau_i$ and $b = \tau_{i+d}$ yields

$$b_i = \frac{\tau_{i+d} - z}{\tau_{i+d} - \tau_i} \mathcal{B}[f^\mu](\tau_{i+1}, \ldots, \tau_i, \ldots, \tau_{i+d}) + \frac{z - \tau_i}{\tau_{i+d} - \tau_i} \mathcal{B}[f^\mu](\tau_i, \ldots, \tau_{i+d}, \ldots, \tau_{i+d-1}).$$

Exploiting the symmetry of the blossom and again applying (4.25) leads to the middle formula in (4.31). ∎

It is sometimes required to insert the same knot several times; this can of course be accomplished by applying the formulas in (4.31) several times. Since blossoms have the property $\mathcal{B}[f](z, \ldots, z) = f(z)$, we see that inserting a knot d times in a spline of degree d gives as a by-product the function value of f at z. This can be conveniently illustrated by listing old and new coefficients in a triangular scheme. Consider the following triangle ($d = 3$),

$$\begin{array}{ccccccccccccccc}
\cdots & & c^0_{\mu-4} & & c^0_{\mu-3} & & c^0_{\mu-2} & & c^0_{\mu-1} & & c^0_{\mu} & & c^0_{\mu+1} & & \cdots \\
 & & & & & c^1_{\mu-2} & & c^1_{\mu-1} & & c^1_{\mu} & & & & & \\
 & & & & & & c^2_{\mu-1} & & c^2_{\mu} & & & & & & \\
 & & & & & & & c^3_{\mu} & & & & & & &
\end{array}$$

In the first row we have the coefficients of f on the original knot vector $\boldsymbol{\tau}$. After inserting z in $(\tau_\mu, \tau_{\mu+1})$ once, the coefficients relative to the knot vector $\boldsymbol{\tau}^1 = \boldsymbol{\tau} \cup \{z\}$ are

$$(\ldots, c^0_{\mu-4}, c^0_{\mu-3}, c^1_{\mu-2}, c^1_{\mu-1}, c^1_{\mu}, c^0_{\mu}, c^0_{\mu+1}, \ldots),$$

i.e., we move down one row in the triangle. Suppose that z is inserted once more. The new B-spline coefficients on $\boldsymbol{\tau}^2 = \boldsymbol{\tau}^1 \cup \{z\}$ are now found by moving down to the second row, across this row, and up the right hand side,

$$(\ldots, c^0_{\mu-4}, c^0_{\mu-3}, c^1_{\mu-2}, c^2_{\mu-1}, c^2_{\mu}, c^1_{\mu}, c^0_{\mu}, c^0_{\mu+1}, \ldots).$$

Similarly, if z is inserted 3 times, we move around the whole triangle. We can also insert z a full $d = 4$ times. We then simply repeat c^3_μ two times in the last row.

Lemma 4.19 shows that Oslo Algorithm 2 (Algorithm 4.11) is not always efficient. To compute a new coefficient in the case where only one new knot is inserted requires at most one convex combination according to Lemma 4.19 while Algorithm 4.11 requires the computation of a full triangle (two nested loops). More efficient versions of the Oslo algorithms can be developed, but this will not be considered here.

The simplicity of the formulas (4.31) indicates that the knot insertion matrix $\boldsymbol{A}$ must have a simple structure when only one knot is inserted. Setting $\boldsymbol{c} = (c_i)_{i=1}^n$ and $\boldsymbol{b} = (b_i)_{i=1}^{n+1}$

and remembering that $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{c}$, we see that $\boldsymbol{A}$ is given by the $(n+1) \times n$ matrix

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0 & & & & & & & \\ & \ddots & \ddots & & & & & & \\ & & 1 & 0 & & & & & \\ & & 1-\lambda_{\mu-d+1} & \lambda_{\mu-d+1} & & & & & \\ & & & \ddots & \ddots & & & & \\ & & & & 1-\lambda_{\mu} & \lambda_{\mu} & & & \\ & & & & 0 & 1 & & & \\ & & & & & \ddots & \ddots & & \\ & & & & & & 0 & 1 & \end{pmatrix}, \tag{4.32}$$

where $\lambda_i = (z - \tau_i)/(\tau_{i+d} - \tau_i)$ for $\mu - d + 1 \leq i \leq \mu$. All the entries off the two diagonals are zero and such matrices are said to be bi-diagonal. Since z lies in the interval $[\tau_\mu, \tau_{\mu+1})$ all the entries in $\boldsymbol{A}$ are nonnegative. This property generalises to arbitrary knot insertion matrices.

Lemma 4.20. *Let $\boldsymbol{\tau} = (\tau_j)_{j=1}^{n+d+1}$ and $\boldsymbol{t} = (t_i)_{i=1}^{m+d+1}$ be two knot vectors for splines of degree d with $\boldsymbol{\tau} \subseteq \boldsymbol{t}$. All the entries of the knot insertion matrix $\boldsymbol{A}$ from $\mathbb{S}_{d,\boldsymbol{\tau}}$ to $\mathbb{S}_{d,\boldsymbol{t}}$ are nonnegative and $\boldsymbol{A}$ can be factored as*

$$\boldsymbol{A} = \boldsymbol{A}_{m-n}\boldsymbol{A}_{m-n-1}\cdots\boldsymbol{A}_1, \tag{4.33}$$

where $\boldsymbol{A}_i$ is a bi-diagonal $(n+i) \times (n+i-1)$-matrix with nonnegative entries.

Proof. Let us denote the $m - n$ knots that are in $\boldsymbol{t}$ but not in $\boldsymbol{\tau}$ by $(z_i)_{i=1}^{m-n}$. Set $\boldsymbol{t}^0 = \boldsymbol{\tau}$ and $\boldsymbol{t}^i = \boldsymbol{t}^{i-1} \cup (z_i)$ for $i = 1, \ldots, m-n$. Denote by $\boldsymbol{A}_i$ the knot insertion matrix from $\boldsymbol{t}^{i-1}$ to $\boldsymbol{t}^i$. By applying Böhm's method $m - n$ times we obtain (4.33). Since all the entries in each of the matrices $\boldsymbol{A}_i$ are nonnegative the same must be true of $\boldsymbol{A}$. ∎

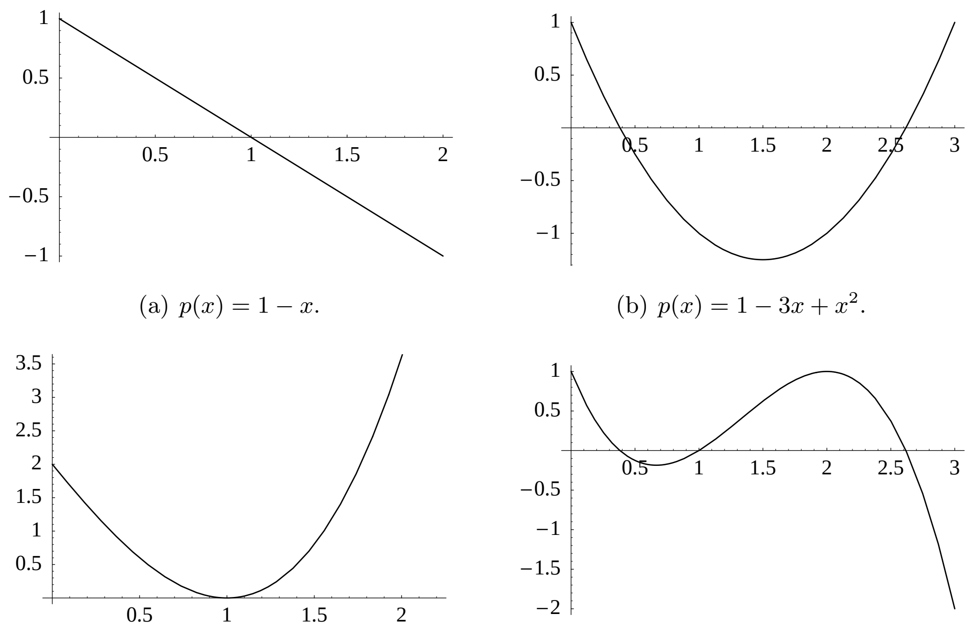
4.5 Bounding the number of sign changes in a spline

In this section we will make use of Böhm's method for knot insertion to prove that the number of spline changes in a spline function is bounded by the number of sign changes in its B-spline coefficient vector. This provides a generalisation of an interesting property of polynomials known as Descartes' rule of signs. Bearing the name of Descartes, this result is of course classical, but it is seldom mentioned in elementary mathematics textbooks. Before stating Descartes' rule of signs let us record what we mean by sign changes in a definition.

Definition 4.21. *Let $\boldsymbol{c} = (c_i)_{i=1}^n$ be a vector of real numbers. The number of sign changes in $\boldsymbol{c}$ (zeros are ignored) is denoted $S^-(\boldsymbol{c})$. The number of sign changes in a function f in an interval (a,b) is denoted $S^-_{(a,b)}(f) = S^-(f)$, provided this number is finite. It is given by the largest possible integer r such that an increasing sequence of $r+1$ real numbers $x_1 < \cdots < x_{r+1}$ in (a,b) can be found with the property that $S^-\big(f(x_1), \ldots, f(x_{r+1})\big) = r$.*

Example 4.22. Let us consider some simple examples of counting sign changes. It is easily checked that

$$\begin{aligned} S^-(1,-2) &= 1, & S^-(1,0,-1,3) &= 2, \\ S^-(1,0,2) &= 0, & S^-(2,0,0,0,-1) &= 1, \\ S^-(1,-1,2) &= 2, & S^-(2,0,0,0,1) &= 0. \end{aligned}$$



(a) $p(x) = 1 - x$.

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

(c) $p(x) = 2 - 3x + x^2$.

(d) $p(x) = 1 - 4x + 4x^2 - x^3$.

Figure 4.3. Illustrations of Descartes' rule of signs: the number of zeros in $(0, \infty)$ is no greater than the number of strong sign changes in the coefficients.

As stated in the definition, we simply count sign changes by counting the number of jumps from positive to negative values and from negative to positive, ignoring all components that are zero.

Descartes' rule of signs bounds the number of zeros in a polynomial by the number of sign changes in its coefficients. Recall that z is a zero of f of multiplicity $r \geq 1$ if $f(z) = Df(z) = \cdots = D^{r-1}f(z) = 0$ but $D^r f(z) \neq 0$.

Theorem 4.23 (Descartes' rule of signs)**.** *Let $p = \sum_{i=0}^{d} c_i x^i$ be a polynomial of degree d with coefficients $\boldsymbol{c} = (c_0, \ldots, c_d)^T$, and let $Z(p)$ denote the total number of zeros of p in the interval $(0, \infty)$, counted with multiplicities. Then*

$$Z(p) \leq S^-(\boldsymbol{c}),$$

i.e., the number of zeros of p is bounded by the number of sign changes in its coefficients.

Figures 4.3 (a)–(d) show some polynomials and their zeros in $(0, \infty)$.

Our aim is to generalise this result to spline functions, written in terms of B-splines. This is not so simple because it is difficult to count zeros for splines. In contrast to polynomials, a spline may for instance be zero on an interval without being identically zero. In this section we will therefore only consider zeros that are also sign changes. In the next section we will then generalise and allow multiple zeros.

To bound the number of sign changes of a spline we will investigate how knot insertion influences the number of sign changes in the B-spline coefficients. Let $\mathbb{S}_{d,\boldsymbol{\tau}}$ and $\mathbb{S}_{d,\boldsymbol{t}}$ be two spline spaces of degree d, with $\mathbb{S}_{d,\boldsymbol{\tau}} \subseteq \mathbb{S}_{d,\boldsymbol{t}}$. Recall from Section 4.4 that to get from the knot vector $\boldsymbol{\tau}$ to the refined knot vector $\boldsymbol{t}$, we can insert one knot at a time. If there are ℓ

more knots in $\boldsymbol{\tau}$ than in $\boldsymbol{t}$, this leads to a factorisation of the knot insertion matrix $\boldsymbol{A}$ as

$$\boldsymbol{A} = \boldsymbol{A}_\ell \boldsymbol{A}_{\ell-1} \cdots \boldsymbol{A}_1, \tag{4.34}$$

where $\boldsymbol{A}_k$ is a $(n+k) \times (n+k-1)$ matrix for $k = 1, \ldots, \ell$, if $\dim \mathbb{S}_{d,\boldsymbol{\tau}} = n$. Each of the matrices $\boldsymbol{A}_k$ corresponds to insertion of only one knot, and all the nonzero entries of the bi-diagonal matrix $\boldsymbol{A}_k$ are found in positions (i,i) and $(i+1,i)$ for $i = 1, \ldots, n+k-1$, and these entries are all nonnegative (in general many of them will be zero).

We start by showing that the number of sign changes in the B-spline coefficients is reduced when the knot vector is refined.

Lemma 4.24. *Let $\mathbb{S}_{d,\boldsymbol{\tau}}$ and $\mathbb{S}_{d,\boldsymbol{t}}$ be two spline spaces such that $\boldsymbol{t}$ is a refinement of $\boldsymbol{\tau}$. Let $f = \sum_{j=1}^n c_j B_{j,d,\boldsymbol{\tau}} = \sum_{i=1}^m b_i B_{i,d,\boldsymbol{t}}$ be a spline in $\mathbb{S}_{d,\boldsymbol{\tau}}$ with B-spline coefficients $\boldsymbol{c}$ in $\mathbb{S}_{d,\boldsymbol{\tau}}$ and $\boldsymbol{b}$ in $\mathbb{S}_{d,\boldsymbol{t}}$. Then $\boldsymbol{b}$ has no more sign changes than $\boldsymbol{c}$, i.e.,*

$$S^-(\boldsymbol{A}\boldsymbol{c}) = S^-(\boldsymbol{b}) \leq S^-(\boldsymbol{c}), \tag{4.35}$$

where $\boldsymbol{A}$ is the knot insertion matrix from $\boldsymbol{\tau}$ to $\boldsymbol{t}$.

Proof. Since we can insert the knots one at a time, it clearly suffices to show that (4.35) holds in the case where there is only one more knot in $\boldsymbol{t}$ than in $\boldsymbol{\tau}$. In this case we know from Lemma 4.19 that $\boldsymbol{A}$ is bidiagonal so

$$b_i = \alpha_{i-1}(i)c_{i-1} + \alpha_i(i)c_i, \qquad \text{for } i = 1, \ldots n+1,$$

where $\big(\alpha_j(i)\big)_{i,j=1}^{n+1,n}$ are the entries of $\boldsymbol{A}$ (for convenience of notation we have introduced two extra entries that are zero, $\alpha_0(1) = \alpha_{n+1}(n+1) = 0$). Since $\alpha_{i-1}(i)$ and $\alpha_i(i)$ both are nonnegative, the sign of b_i must be the same as either c_{i-1} or c_i (or be zero). Since the number of sign changes in a vector is not altered by inserting zeros or a number with the same sign as one of its neighbours we have

$$S^-(\boldsymbol{c}) = S^-(b_1, c_1, b_2, c_2, \ldots, b_{n-1}, c_{n-1}, b_n, c_n, b_{n+1}) \geq S^-(\boldsymbol{b}).$$

The last inequality follows since the number of sign changes in a vector is always reduced when entries are removed. ∎

From Lemma 4.24 we can quite easily bound the number of sign changes in a spline in terms of the number of sign changes in its B-spline coefficients.

Theorem 4.25. *Let $f = \sum_{j=1}^n c_j B_{j,d}$ be a spline in $\mathbb{S}_{d,\boldsymbol{\tau}}$. Then*

$$S^-(f) \leq S^-(\boldsymbol{c}) \leq n-1. \tag{4.36}$$

Proof. Suppose that $S^-(f) = \ell$, and let $(x_i)_{i=1}^{\ell+1}$ be $\ell+1$ points chosen so that $S^-(f) = S^-\big(f(x_1), \ldots, f(x_{\ell+1})\big)$. We form a new knot vector $\boldsymbol{t}$ that includes $\boldsymbol{\tau}$ as a subsequence, but in addition each of the x_i occurs exactly $d+1$ times in $\boldsymbol{t}$. From our study of knot insertion we know that f may be written $f = \sum_j b_j B_{j,d,\boldsymbol{t}}$ for suitable coefficients (b_j), and from Lemma 2.6 we know that each of the function values $f(x_i)$ will appear as a B-spline coefficient in $\boldsymbol{b}$. We therefore have

$$S^-(f) \leq S^-(\boldsymbol{b}) \leq S^-(\boldsymbol{c}),$$

the last inequality following from Lemma 4.24. The last inequality in (4.36) follows since an n-vector can only have $n-1$ sign changes. ∎

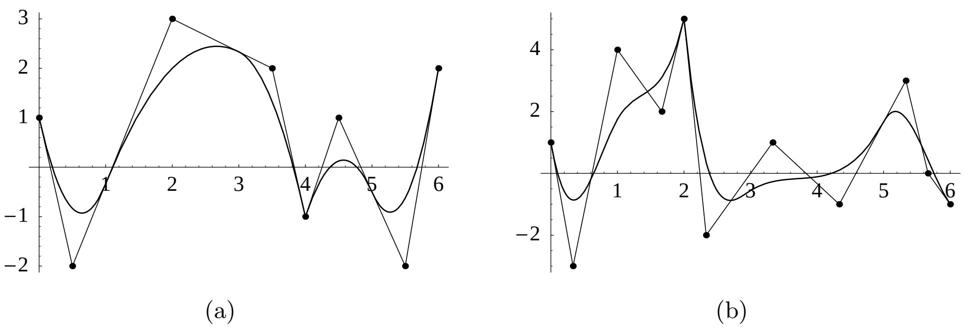


Figure 4.4. A quadratic spline (a) and a cubic spline (b) with their control polygons.

The validity of Theorem 4.25 can be checked with the two plots in Figure 4.4 as well as all other figures which include both a spline function and its control polygon.

Exercises for Chapter 4

4.1 In this exercise we are going to study a change of polynomial basis from the Bernstein basis to the Monomial basis. Recall that the Bernstein basis of degree d is defined by

$$B_j^d(x) = \binom{d}{j} x^j (1-x)^{d-j}, \qquad \text{for } j = 0, 1, \ldots, d. \tag{4.37}$$

A polynomial p of degree d is said to be written in Monomial form if $p(x) = \sum_{j=0}^{d} b_j x^j$ and in Bernstein form if $p(x) = \sum_{j=0}^{d} c_j B_j^d(x)$. In this exercise the binomial formula

$$(a+b)^d = \sum_{k=0}^{d} \binom{d}{k} a^k b^{d-k} \tag{4.38}$$

will be useful.

a) By applying (4.38), show that

$$B_j^d(x) = \sum_{i=j}^{d} (-1)^{i-j} \binom{d}{j} \binom{d-j}{i-j} x^i, \qquad \text{for } j = 0, 1, \ldots, d.$$

Also show that $\binom{d}{j}\binom{d-j}{i-j} = \binom{d}{i}\binom{i}{j}$ for $i = j, \ldots, d$ and $j = 0, \ldots, d$.

b) The two basis vectors $\boldsymbol{B}_d = \big(B_0^d(x), \ldots, B_d^d(x)\big)^T$ and $\boldsymbol{P}_d = (1, x, \ldots, x^d)^T$ are related by $\boldsymbol{B}_d^T = \boldsymbol{P}_d^T \boldsymbol{A}_d$ where $\boldsymbol{A}_d$ is a $(d+1) \times (d+1)$-matrix $\boldsymbol{A}_d$. Show that the entries of $\boldsymbol{A}_d = (a_{i,j})_{i,j=0}^{d}$ are given by

$$a_{i,j} = \begin{cases} 0, & \text{if } i < j, \\ (-1)^{i-j}\binom{d}{i}\binom{i}{j}, & \text{otherwise.} \end{cases}$$

c) Show that the entries of $\boldsymbol{A}_d$ satisfy the recurrence relation

$$a_{i,j} = \beta_i \left(a_{i-1,j-1} - a_{i-1,j}\right), \qquad \text{where } \beta_i = (d-i+1)/i.$$

Give a detailed algorithm for computing $\boldsymbol{A}_d$ based on this formula.

d) Explain how we can find the coefficients of a polynomial relative to the Monomial basis if $\boldsymbol{A}_d$ is known and the coefficients relative to the Bernstein basis are known.

4.2 In this exercise we are going to study the opposite conversion of that in Exercise 1, namely from the Monomial basis to the Bernstein basis.

a) With the aid of (4.38), show that for all x and t in $\mathbb{R}$ we have

$$\big(tx + (1-x)\big)^d = \sum_{k=0}^{d} B_k^d(x) t^k. \tag{4.39}$$

The function $G(t) = \big(tx + (1-x)\big)^d$ is called a generating function for the Bernstein polynomials.

b) Show that $\sum_{k=0}^{d} B_k^d(x) = 1$ for all x by choosing a suitable value for t in (4.39).

c) Find two different expressions for $G^{(j)}(1)/j!$ and show that this leads to the formulas

$$\binom{d}{j} x^j = \sum_{i=j}^{d} \binom{i}{j} B_k^d(x), \qquad \text{for } j = 0, \ldots, d. \tag{4.40}$$

d) Show that the entries of the matrix $\boldsymbol{B}_d = (b_{i,j})_{i,j=0}^d$ such that $\boldsymbol{P}_d^T = \boldsymbol{B}_d^T \boldsymbol{B}_d$ are given by

$$b_{i,j} = \begin{cases} 0, & \text{if } i < j, \\ \binom{i}{j}/\binom{d}{j}, & \text{otherwise.} \end{cases}$$

4.3 Let $\boldsymbol{P}$ denote the cubic Bernstein basis on the interval $[0,1]$ and let $\boldsymbol{Q}$ denote the cubic Bernstein basis on the interval $[2,3]$. Determine the matrix $\boldsymbol{A}_3$ such that $\boldsymbol{P}(x)^T = \boldsymbol{Q}(x)^T \boldsymbol{A}_3$ for all real numbers x.

4.4 Let $\boldsymbol{A}$ denote the knot insertion matrix for the linear ($d=1$) B-splines on $\boldsymbol{\tau} = (\tau_j)_{j=1}^{n+2}$ to the linear B-splines in $\boldsymbol{t} = (t_i)_{i=1}^{m+2}$. We assume that $\boldsymbol{\tau}$ and $\boldsymbol{t}$ are 2-extended with $\tau_1 = t_1$ and $\tau_{n+2} = t_{m+2}$ and $\boldsymbol{\tau} \subseteq \boldsymbol{t}$.

a) Determine $\boldsymbol{A}$ when $\boldsymbol{\tau} = (0, 0, 1/2, 1, 1)$ and $\boldsymbol{t} = (0, 0, 1/4, 1/2, 3/4, 1, 1)$.

b) Device a detailed algorithm that computes $\boldsymbol{A}$ for general $\boldsymbol{\tau}$ and $\boldsymbol{t}$ and requires $O(m)$ operations.

c) Show that the matrix $\boldsymbol{A}^T\boldsymbol{A}$ is tridiagonal.

4.5 Prove Lemma 4.4 in the general case where $\boldsymbol{\tau}$ and $\boldsymbol{t}$ are not $d+1$-regular. Hint: Augment both $\boldsymbol{\tau}$ and $\boldsymbol{t}$ by inserting $d+1$ identical knots at the beginning and end.

4.6 Prove Theorem 4.7 in the general case where the knot vectors are not $d+1$-regular with common knots at the ends. Hint: Use the standard trick of augmenting $\boldsymbol{\tau}$ and $\boldsymbol{t}$ with $d+1$ identical knots at both ends to obtain new knot vectors $\hat{\boldsymbol{\tau}}$ and $\hat{\boldsymbol{t}}$. The knot insertion matrix from $\boldsymbol{\tau}$ to $\boldsymbol{t}$ can then be identified as a sub-matrix of the knot insertion matrix from $\hat{\boldsymbol{\tau}}$ to $\hat{\boldsymbol{t}}$.

4.7 Show that if $\boldsymbol{\tau}$ and $\boldsymbol{t}$ are $d+1$-regular knot vectors with $\boldsymbol{\tau} \subseteq \boldsymbol{t}$ whose knots agree at the ends then $\sum_j \alpha_{j,d}(i) = 1$.

4.8 Implement Algorithm 4.11 and test it on two examples. Verify graphically that the control polygon converges to the spline as more and more knots are inserted.

4.9 Let f be a function that satisfies the identity

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y) \tag{4.41}$$

for all real numbers x and y and all real numbers α and β such that $\alpha + \beta = 1$. Show that then f must be an affine function. Hint: Use the alternative form of equation (4.41) found in Lemma 4.17.

4.10 Find the cubic blossom $\mathcal{B}[p](x_1, x_2, x_3)$ when p is given by:

a) $p(x) = x^3$.

b) $p(x) = 1$.

c) $p(x) = 2x + x^2 - 4x^3$.

d) $p(x) = 0$.

e) $p(x) = (x-a)^2$ where a is some real number.