in $[a, b]$. (A trivial example of this is given when $x = x_j = a + jh$, but x_j is not a Chebyshev zero. For then, $0 = |e_h(x_j)| \leq |e_T(x_j)|$.) Usually, however, if we are free to select the interpolating points at our own discretion, we prefer to use Chebyshev interpolation since we can determine an error bound *a priori* for all values of x.

There are other ways of using (5.27) to estimate the error of interpolation. For example, let $w(x)$ be a nonnegative weighting function as given in (5.1b). Multiplying the squares of both sides of (5.27) by $w(x)$ and integrating from a to b yield

$$\begin{aligned} \|e(x)\|_2 &= \left(\int_a^b (f(x) - p(x))^2 w(x)\,dx\right)^{1/2} \\ &= \left(\int_a^b \left(\frac{f^{(n+1)}(\xi)}{(n+1)!} W(x)\right)^2 w(x)\,dx\right)^{1/2} \\ &\leq \frac{\max_{a \leq x \leq b} |f^{(n+1)}(x)|}{(n+1)!} \left(\int_a^b (W(x))^2 w(x)\,dx\right)^{1/2} \\ &= \frac{K_n}{(n+1)!}\|W\|_2. \end{aligned} \tag{5.31}$$

After we investigate orthogonal polynomials in a later section, we will show how to choose the interpolating points, $\{x_j\}_{j=0}^n$, in order to minimize $\|W\|_2$. We merely state here that they are not, in most cases, the zeros of a Chebyshev polynomial.

5.2.4. Translating the Interval

Often it is necessary for practical purposes to take a problem stated on an interval $[c, d]$ and reformulate the same problem on a different interval $[a, b]$. This necessity is especially true in problems that require the numerical approximation of integrals or in problems concerning orthogonal polynomials. We have already come across this problem in Chebyshev interpolation, in which Theorem 5.6 is applicable to the specific interval $[-1, 1]$ and says nothing about the optimal choice of interpolation points in an arbitrary interval $[a, b]$. The change of variable that usually achieves our purposes of reformulating a problem on $[c, d]$ to one on $[a, b]$ is the "straight line" transformation $x = mt + \beta$ for $t \in [c, d]$ and $x \in [a, b]$. Here m and β are chosen such that $x = a$ when $t = c$ and $x = b$ when $t = d$. A simple algebraic computation will show that this transformation is given by

$$x = \left(\frac{b-a}{d-c}\right)t + \left(\frac{ad-bc}{d-c}\right). \tag{5.32}$$

To illustrate this procedure, let us suppose that we wish to perform Chebyshev interpolation on an arbitrary interval $[a, b]$ instead of the interval $[-1, 1]$. Then $t \in [-1, 1] \equiv [c, d]$ and $x \in [a, b]$; so by (5.32)

$$x = \left(\frac{b-a}{2}\right)t + \left(\frac{a+b}{2}\right)$$

or (5.33)

$$t = \frac{2}{b-a}\left(x - \frac{a+b}{2}\right) = \frac{2(x-a)}{(b-a)} - 1.$$

We now define the "shifted" Chebyshev polynomial of degree k for $x \in [a, b]$ as

$$\tilde{T}_k(x) \equiv T_k(t) = T_k\left(\frac{2(x-a)}{(b-a)} - 1\right) = \cos\left[k \cos^{-1}\left(\frac{2(x-a)}{(b-a)} - 1\right)\right]. \tag{5.34}$$

Since $T_k(t_i) = 0$ for $t_i = \cos((2i+1)\pi)/2k$, $0 \le i \le k-1$, then $x_i = ((b-a)/2)t_i + (b+a)/2$, $0 \le i \le k-1$, are the zeros of $\tilde{T}_k(x)$. Furthermore, $\tilde{T}_0(x) = 1$, $\tilde{T}_1(x) = 2(x-a)/(b-a) - 1$; and so from (5.29)

$$\begin{aligned}\tilde{T}_{k+1}(x) &= T_{k+1}(t) = 2tT_k(t) - T_{k-1}(t)\\ &= 2\left(\frac{2(x-a)}{(b-a)} - 1\right)\tilde{T}_k(x) - \tilde{T}_{k-1}(x), \qquad k \ge 1.\end{aligned} \tag{5.35}$$

We leave it to the reader to verify that the leading coefficient of $\tilde{T}_k(x)$ is $2^{k-1}(2/(b-a))^k$, $k \ge 1$. Therefore, if we interpolate in $[a, b]$ at the zeros $\{x_i\}_{i=0}^n$ of $\tilde{T}_{n+1}(x)$, we have $W(x) = 2^{-n}(2/(b-a))^{-n-1}\,\tilde{T}_{n+1}(x)$. Then

$$\max_{a \le x \le b} |W(x)| = 2^{-n}\left(\frac{2}{b-a}\right)^{-n-1}$$

since

$$\max_{a \le x \le b} |\tilde{T}_{n+1}(x)| = \max_{-1 \le t \le 1} |T_{n+1}(t)| = 1.$$

Thus the error bound, (5.30), becomes

$$|e(x)| = |f(x) - p(x)| \le \frac{K_n}{(n+1)!} \max_{a \le x \le b} |W(x)| = \frac{K_n}{2^n(n+1)!}\left(\frac{b-a}{2}\right)^{n+1}. \tag{5.36}$$

It is easily seen that Theorem 5.6 is valid on $[a, b]$ with respect to this $W(x)$, and so $W(x)$ is the smallest monic polynomial of degree $(n+1)$ with respect to

$$\|W\|_\infty \equiv \max_{a \le x \le b} |W(x)|.$$

EXAMPLE 5.9. Suppose $f(x) = \cos(x)$ and $[a, b] = [0, \pi/2]$, and suppose we wish to find a value of n and interpolating points $\{x_j\}_{j=0}^n$ such that the error of the nth-degree interpolation is less than 10^{-6} for all x in $[0, \pi/2]$. Now $K_n = 1$ for all n; so with Eq. (5.36), the smallest value of n for which

$$\frac{1}{2^n(n+1)!}\left(\frac{b-a}{2}\right)^{n+1} = \frac{1}{2^n(n+1)!}(\pi/4)^{n+1} < 10^{-6}$$

is $n = 6$. Thus the interpolating points are the zeros of $\tilde{T}_7(x)$, which are the points

$$x_j = \frac{\pi}{4}(\cos[(2j+1)\pi/14] + 1), \qquad 0 \le j \le 6.$$

From (5.36)

$$|e(x)| \le \frac{1}{2^6 7!}(\pi/4)^7 = 5.71 \times 10^{-7}, \qquad x \in [0, \pi/2].$$

The transformation technique of this section is quite simple but is used often and will be referred to in later sections.

PROBLEMS, SECTION 5.2.4

1. As in Example 5.7, use (5.28) to bound the interpolation error at $x = .15$ and $x = .35$ when $p(x)$ interpolates $f(x)$ at $x_i = .1 + i/10$, $0 \le i \le 5$, for

a) $f(x) = \cos(x)$ b) $f(x) = \ell n(x)$ c) $f(x) = \dfrac{1}{1+x}$.

Compare the bounds with the actual errors; use the program in Problem 8, Section 5.2.1.

2. Use (5.36) to bound the interpolation error at any point in $[.1, .6]$ for the functions in Problem 1 when the interpolation points are the zeros of $T_k(x)$ shifted to $[.1, .6]$. Compare your results with the results of Problem 1.

3. How large should k be if we wish to obtain an interpolation error of 10^{-6} or less throughout $[a, b]$ by interpolating $f(x)$ at the zeros of $\tilde{T}_k(x)$ for

a) $f(x) = \cos(x)$, $0.3 \le x \le 0.6$

b) $f(x) = 1/(2 + x)$, $-1 \le x \le 1$

c) $f(x) = \ln(x)$, $0.1 \le x \le 1$

d) $f(x) = e^{3x}$, $-1 \le x \le 1$.

4. Let $x_0, x_1, \ldots, x_n$ be equally spaced points in $[-2, 2]$, $x_0 = -2$, and $x_n = 2$. Call a plotting routine to sketch $W(x)$ in increments of .05 for $n = 10, 15, 20$. Note the characteristic shapes and that $|W(x)|$ is largest near the endpoints ± 2.

5. To use the error estimate (5.28), we need some way of estimating $|W(x)|$. For equally spaced interpolation points, there are a number of ways of estimating $|W(x)|$. For simplicity, suppose we are interpolating in $[-1, 1]$ with an odd number of equally spaced points. Thus suppose that n is even, $W(x) = (x - x_0)(x - x_1)\ldots$

$(x - x_n)$ where $h = 2/n$, $x_j = x_0 + jh$ for $j = 0, 1, \ldots, n$ and where $x_0 = -1$, $x_n = 1$. Let $N = n/2$ and show that

$$W(x) = (x + Nh)(x + (N - 1)h) \cdots (x + h)x(x - h) \cdots (x - (N - 1)h)(x - Nh).$$

For $x \in [-1, 1]$, write x as $x = rh$ where r is a number such that $-N \le r \le N$. Suppose next that $x_{n-1} < x < x_n$ so that $N - 1 < r < N$. Show that

$$(n - 1)!h^{n-1}|(x - x_{n-1})(x - x_n)| \le |W(x)| \le n!h^{n-1}|(x - x_{n-1})(x - x_n)|.$$

6. Show that the maximum value of the factor $|(x - x_{n-1})(x - x_n)|$ in Problem 5, for $x_{n-1} \le x \le x_n$, is $h^2/4$. Thus conclude that $|W(x)| \le n!2^{n-1}/n^{n+1}$. Also conclude, for $x = (x_n + x_{n-1})/2$, that $(n - 1)!2^{n-1}/n^{n+1} \le |W(x)|$.

7. In Problem 6, for $n = 5$, 10, and 20, evaluate the upper and lower bounds for $|W(x)|$. Contrast the upper bounds with $1/2^n$, which is the bound on $|(x - t_0)(x - t_1). . .(x - t_n)|$ where $t_0, t_1, \ldots, t_n$ are the zeros of $T_{n+1}(x)$.

8. Given $(n + 1)$ equally spaced interpolation points in $[-1, 1]$ where n is even as in Problem 5, show that $W(x)$ is an odd function [i.e., $W(x) = -W(-x)$]. Thus if we can find the maximum of $|W(x)|$ for $0 \le x \le 1$, we have a maximum for $|W(x)|$ for $-1 \le x \le 1$. Show that the maximum of $|W(x)|$ occurs somewhere in (x_{n-1}, x_n). [Hint: Show that

$$\left|\frac{W(x + h)}{W(x)}\right| > 1$$

for $0 < x < x_{n-1}$ and for x not an interpolation point.] Use the representation of $W(x)$ in Problem 5 to establish this inequality. This result means that the upper and lower bounds established in Problem 6 are upper and lower bounds for $\|W\|_\infty$. The inequality in the hint also shows why interpolation at equally spaced points is usually more reliable near the center of the entire interval.

9. Let $f(x) = 1/(x + 2)$. Using (5.28) and the results of Problems 6 and 8, give an error estimate for interpolating $f(x)$ at n equally spaced points in $[-1, 1]$ where n is even. What is the corresponding error estimate if $f(x)$ is interpolated at the zeros of $T_{n+1}(x)$?

10. Let n be odd and let $x_0, x_1, \ldots, x_n$ be equally spaced in $[-1, 1]$, $x_0 = -1$, $x_n = 1$. The length of each subinterval is $2h$ where $h = 1/n$. Show that

$$W(x) = (x + nh)(x + (n - 2)h). . .(x + h)(x - h). . .(x - (n - 2)h)(x - nh).$$

Show that the maximum value of $|W(x)|$ for $-h \le x \le h$ occurs at $x = 0$. [Hint: Group the factors of $W(x)$ as $(x^2 - n^2h^2). . .(x^2 - h^2)$ and show that $W'(x)$ and $W'(-x)$ have opposite signs for any x in $(-h, h)$.] Evaluate $W(0)$ and determine the maximum interpolation error for $f(x) = 1/(x + 2)$ with $x \in (-h, h)$.

11. Verify (5.29) by showing that

$$\cos((k + 1)\theta) = 2\cos(\theta)\cos(k\theta) - \cos((k - 1)\theta).$$

12. Using Eq. (5.29), show by induction that

a) $T_k(x)$ is a kth-degree polynomial with leading coefficient 2^{k-1};

b) $T_k(x)$ is an even (odd) function if k is even (odd) [recall that $f(x)$ is an even function if $f(x) = f(-x)$ and that $f(x)$ is an odd function if $f(x) = -f(-x)$].

13. Show that $\tilde{T}_k(x)$ defined by (5.35) has leading coefficients equal to

$$2^{k-1}\left(\frac{2}{b-a}\right)^k, \qquad k \geq 1.$$

What are the polynomials $\tilde{T}_k(x)$ in the special case in which $[a, b] = [0, 1]$?

*5.2.5. Polynomial Interpolation with Derivative Data

Examination of (5.5) and (5.6) reveals that the type of nth-degree polynomial interpolation that we have studied up to this point is based on the fact that there exists a unique solution to the $(n+1)$ linear equations, $p(x_j) = f(x_j)$, $0 \leq j \leq n$, in the $(n+1)$ unknowns, a_j [the coefficients of $p(x)$]. Thus $p(x)$ "matches" $f(x)$ at the interpolating points. It is natural to ask if we might not get a better polynomial approximation to $f(x)$ if we forced some derivatives of $p(x)$ at various points to match the respective derivatives of $f(x)$ at these points [for then we would be using more information about $f(x)$ and would expect a better approximation]. The extreme example of this procedure is to take only one point, x_0, and choose the coefficients of $p(x)$ such that $p^{(j)}(x_0) = f^{(j)}(x_0)$ for $0 \leq j \leq n$. This method yields the following $(n+1)$ equations in $\{a_j\}_{j=0}^n$:

$$\begin{aligned}
a_0 + a_1 x_0 + a_2 x_0^2 + \cdots + a_n x_0^n &= f(x_0)\\
a_1 + 2a_2 x_0 + \cdots + n a_n x_0^{n-1} &= f'(x_0)\\
2a_2 + \cdots + n(n-1) a_n x_0^{n-2} &= f''(x_0)\\
\ddots \qquad \vdots \qquad &\quad \vdots\\
n! a_n &= f^{(n)}(x_0).
\end{aligned}$$

We note that this triangular system is nonsingular (the coefficient matrix has positive diagonal elements), and so $p(x)$ exists and is unique. The reader will recall from calculus that the Taylor polynomial,

$$q(x) = \sum_{j=0}^{n} \frac{f^{(j)}(x_0)}{j!}(x - x_0)^j,$$

also has the property that $q^{(j)}(x_0) = f^{(j)}(x_0)$, $0 \leq j \leq n$; and so, by the uniqueness of $p(x)$, $q(x) = p(x)$. Thus truncated Taylor's expansions are actually special cases of this generalized interpolation problem. Since the error of the Taylor approximation at x is

$$\frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1},$$

this approximation is quite good if x is "close" to x_0 and usually worsens as x moves away from x_0. Note the similarity of this error formula to that of (5.27).

There are obviously many ways of matching derivatives of a polynomial to respective derivatives of $f(x)$ at various points in order to obtain a linear system of $(n + 1)$ equations in $(n + 1)$ unknowns [the coefficients of $p(x)$]. This problem in its most general form is known as *Hermite-Birkoff interpolation*. For example, we might ask for a third-degree polynomial, $p(x)$, such that $p(x_0) = f(x_0)$, $p'''(x_0) = f'''(x_0)$, $p'(x_1) = f'(x_1)$ and $p''(x_1) = f''(x_1)$. In this general setting, the problem need not always have a solution, and the question of finding conditions under which solutions exist remains unanswered and the subject of extensive research.

We wish now to consider a special form of this problem known as *Hermite* or *osculatory* interpolation. This form of interpolation not only is important as an approximation technique, but is useful in the derivation of the cubic spline approximations of the next section. For simplicity we assume that n is even and define $N = n/2$. We consider $N + 1$ distinct points $\{x_j\}_{j=0}^{N}$, and impose the conditions that $p(x_j) = f(x_j)$ *and* $p'(x_j) = f'(x_j)$, $0 \le j \le N$. Thus $p(x)$ matches $f(x)$ and $p'(x)$ matches $f'(x)$ at the interpolating points. Since we have $2(N + 1) = n + 2$ equations; so we will be searching for $p(x)$ in $\mathcal{P}_{2N+1} = \mathcal{P}_{n+1}$ in order that we have $2N + 2$ unknowns, $\{a_j\}_{j=0}^{2N+1}$. As an illustration of the problem, consider the following example.

EXAMPLE 5.10. Let $x_0 = \alpha$ and $x_1 = \beta$; and assume that $f(x)$ is given such that $f(\alpha) = y_0, f'(\alpha) = y_0', f(\beta) = y_1$, and $f'(\beta) = y_1'$. Then $N = 1$ and so $2N + 1 = 3$. Find $p(x) \in \mathcal{P}_3$ such that $p(x_j) = f(x_j) = y_j$ and $p'(x_j) = f'(x_j) = y_j'$ for $j = 0$ and 1. If $p(x) = a_0 + a_1x + a_2x^2 + a_3x^3$, then the equations above become (in matrix form)

$$\begin{bmatrix} 1 & \alpha & \alpha^2 & \alpha^3 \\ 0 & 1 & 2\alpha & 3\alpha^2 \\ 1 & \beta & \beta^2 & \beta^3 \\ 0 & 1 & 2\beta & 3\beta^2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} y_0 \\ y_0' \\ y_1 \\ y_1' \end{bmatrix}.$$

The determinant of the coefficient matrix for this system equals $(\beta - \alpha)^4$. Then when $\alpha \ne \beta$, $p(x)$ exists and is unique. We could solve this system for the a_i's to obtain $p(x)$, but we shall show for $\alpha < \beta$ that $p(x)$ is given by

$$p(x) = y_0\left[\frac{(x-\beta)^2}{(\beta-\alpha)^2} + 2\frac{(x-\alpha)(x-\beta)^2}{(\beta-\alpha)^3}\right] + y_1\left[\frac{(x-\alpha)^2}{(\beta-\alpha)^2} - 2\frac{(x-\beta)(x-\alpha)^2}{(\beta-\alpha)^3}\right]$$
$$+ y_0'\left[\frac{(x-\alpha)(x-\beta)^2}{(\beta-\alpha)^2}\right] + y_1'\left[\frac{(x-\alpha)^2(x-\beta)}{(\beta-\alpha)^2}\right]. \qquad (5.37)$$

We could approach the Hermite interpolation problem in the manner of undetermined coefficients as in Example 5.10, but we choose a Lagrangian-type approach instead. For each j, $0 \le j \le N$, let $\ell_j(x)$ be the Nth-degree polynomial, as defined by (5.2), with N substituted for n. Now for each j, $0 \le j \le N$, define the $(2N + 1)$st-degree polynomials

$$A_j(x) \equiv [1 - 2(x - x_j)\ell_j'(x_j)]\ell_j^2(x)$$
$$B_j(x) \equiv (x - x_j)\ell_j^2(x).$$

We leave to the reader to verify that $A_j(x_i) = \delta_{ij}$, $B_j(x_i) = 0$, $A_j'(x_i) = 0$, and $B_j'(x_i) = \delta_{ij}$ where $0 \le i, j \le N$, and δ_{ij} is again the Kronecker delta. Let $\{y_0, y_1, \ldots, y_N, y_0', y_1', \ldots, y_N'\}$ be any set of $(2N + 2)$ values, and define $p(x) \in \mathcal{P}_{2N+1}$ by

$$p(x) = \sum_{j=0}^{N} (y_j A_j(x) + y_j' B_j(x)). \tag{5.38a}$$

Because of the properties of $A_j(x)$ and $B_j(x)$, it is easily seen that

$$p(x_i) = \sum_{j=0}^{N} (y_j \cdot \delta_{ij} + y_j' \cdot 0) = y_i, \qquad 0 \le i \le N,$$

and

$$p'(x_i) = \sum_{j=0}^{N} (y_j \cdot 0 + y_j' \cdot \delta_{ij}) = y_i', \qquad 0 \le i \le N.$$

We can also show that the Hermite interpolating polynomial, $p(x)$, in (5.38a) is unique. Assume that $q(x) \in \mathcal{P}_{2N+1}$, $q(x) \ne p(x)$, and $q(x)$ also has the property that $q(x_j) = y_j$ and $q'(x_j) = y_j'$, $0 \le j \le N$. Let $r(x) \equiv p(x) - q(x)$. Then $r(x) \in \mathcal{P}_{2N+1}$ and $r(x_j) = r'(x_j) = 0$, $0 \le j \le N$. Thus, counting multiplicities, $r(x)$ has $2N + 2$ zeros, a result which contradicts the Fundamental Theorem of Algebra. Therefore no such $q(x)$ exists and $p(x)$ is unique. Hence we have just proved Theorem 5.7.

Theorem 5.7.

Let $\{x_j\}_{j=0}^N$ be a set of $(N + 1)$ distinct points and $\{y_j\}_{j=0}^N$ and $\{y_j'\}_{j=0}^N$ be any two sets of $N + 1$ values. Then there exists a unique $p(x) \in \mathcal{P}_{2N+1}$ such that $p(x_j) = y_j$ and $p'(x_j) = y_j'$, $0 \le j \le N$, and $p(x)$ is given by (5.38a).

Usually the y_j and y_j' are given by some function $f(x)$ where $f(x_j) = y_j$ and $f'(x_j) = y_j'$, $0 \le j \le N$. In this case, (5.38a) becomes

$$p(x) = \sum_{j=0}^{N} (f(x_j)A_j(x) + f'(x_j)B_j(x)). \tag{5.38b}$$

We leave as a problem for the reader that (5.38b) reduces to (5.37) in the case in which $N = 1$.

In the case in which $f'(x)$ does not exist or is unknown, Hermite interpolation is still sometimes used as in (5.38a) where we set $y_j = f(x_j)$ and $y_j' = 0$, $0 \le$

$j \leq n$. This is known as *Hermite-Fejér interpolation*, and the following result was proved by Fejér.

For any N, let $\{x_j\}_{j=0}^N$ be the zeros of the Chebyshev polynomial $T_{N+1}(x)$. Let $f(x)$ be any function in $C[-1, 1]$ and let $p_{2N+1}(x)$ be the Hermite-Fejér interpolating polynomial in $\mathcal{P}_{2N+1}$ for $f(x)$. Then

$$\lim_{N\to\infty} \left\{ \max_{-1 \leq x \leq 1} |f(x) - p_{2N+1}(x)| \right\} = 0.$$

This theorem of Fejér is a curious result since we ask for the condition $p'_{2N+1}(x_j) = 0$ [which is unrelated to $f(x)$] and still obtain uniform convergence of the Hermite interpolating polynomials. The result is doubly curious in that no matter what set of interpolation points are chosen, there are functions $f \in C[-1, 1]$ such that $\lim_{N\to\infty} \|p_N - f\|_\infty = +\infty$ where $p_N(x)$ denotes the Lagrange interpolating polynomial as in (5.4) (c.f., Natanson 1965).

The error analysis for Hermite interpolation, (5.38b), can be carried out in much the same way as in Theorem 5.5. This time, however, we define the auxiliary function as $F(t) = f(t) - p(t) - CW(t)^2$ with $C = (f(x) - p(x))/W(x)^2$, $W(t) = \Pi_{j=0}^N (t - x_j)$, and $p(t)$ as in (5.38b). Using Rolle's theorem as before, we obtain an expression for the error:

$$f(x) - p(x) = \frac{f^{(2N+2)}(\xi)W(x)^2}{(2N+2)!}. \tag{5.39}$$

5.2.6. Interpolation by Cubic Splines

In some practical problems, interpolating polynomials are not suitable for use as an approximation. For example, in order to obtain a "good" approximation to a function $f(x)$ by an nth-degree interpolating polynomial, it may be necessary to use a fairly large value of n. Unfortunately, polynomials of high degree often have a very oscillatory behavior, which is not desirable in approximating functions that are reasonably smooth. This disadvantage of polynomial interpolation becomes particularly apparent when interpolating tabular data as is shown in the example below. Also, computational problems arise when the number of data points is large. For instance, given 100 data points (x_0, y_0), (x_1, y_1), . . . , (x_{99}, y_{99}), it would usually be foolhardy to find the 99th-degree polynomial $p(x)$ such that $p(x_i) = y_i$, $i = 0, 1, \ldots, 99$. One alternative to interpolation, as we shall see in Section 5.3, would be to find a polynomial of a low degree that "best fits" the data. Unfortunately, the polynomial that best fits the data will not generally interpolate the data as well. When it is desirable to have an approximation $q(x)$ such that $q(x_i) = y_i$, $i = 0, 1, \ldots, n$, then *piecewise* polynomial interpolation is an attractive alternative and the one that we focus on in this section. In piecewise polynomial interpolation, several lower-degree polynomials are joined together in a continuous fashion so that the resulting piecewise polynomial, $q(x)$, interpolates the data. The extreme case of piecewise polynomial interpolation is the "broken line" as in Example 5.11, in which the approximating function $q(x)$ is obtained by joining (x_i, y_i) and (x_{i+1}, y_{i+1}) by a straight line for $i = 0, 1, \ldots, n-1$. In this case, $q(x_i) = y_i$, and $q(x)$ is a linear polynomial in each interval $[x_i, x_{i+1}]$.

EXAMPLE 5.11. The table below represents data taken from a hypothetical experiment.

x_i	0	1	2	3	4	5	6	7	8	9	10
y_i	3	2	3	5	3	4	3	2	2	3	2

In Fig. 5.6, we show the piecewise linear polynomial $q(x)$ (the broken line) and the tenth-degree interpolating polynomial, $p(x)$. Clearly, unless the function from which the data were taken is strange indeed, the broken line is a more acceptable approximation than is the interpolating polynomial. (Note that as mentioned before, the interpolating polynomial appears to be more reasonable near the center of the interval.)

Although polynomial interpolation may not give an acceptable approximation, as indicated by Fig. 5.6, the broken line approximation also has its disad-

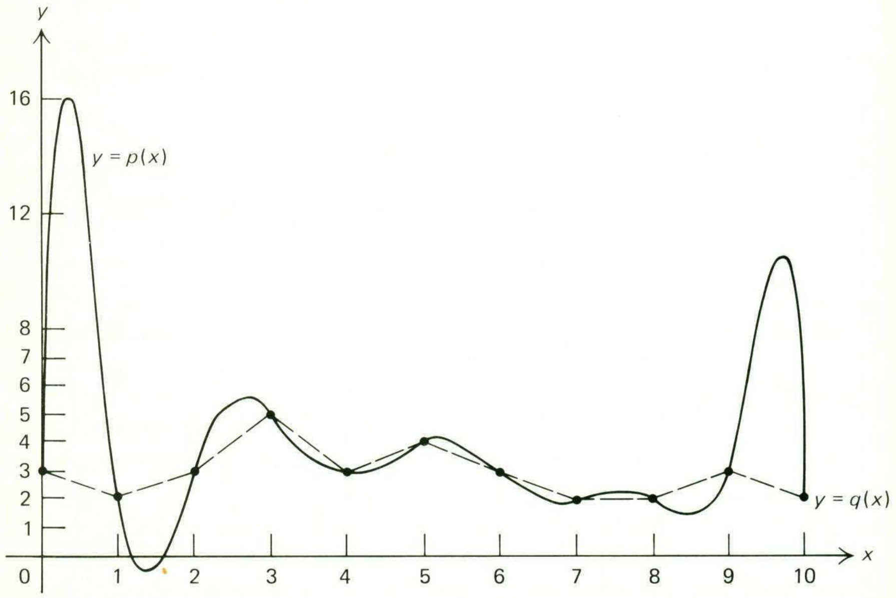

Figure 5.6 Piecewise linear and interpolating approximations of Example 5.11.

vantages. In particular, the broken line suffers from a lack of smoothness and has a discontinuous derivative. Thus, except to give a rough idea of the shape of the graph, the broken line is not well suited to approximate most of the functions that arise in physical problems since such functions are usually fairly smooth. To overcome the oscillatory behavior of polynomials and still provide a smooth approximation, an approximation technique using *splines* was presented in a paper by Schoenberg (1946), and has been the subject of extensive research ever since. This approximation technique resembles a physical process that has been used by drafters for many years. Given a set of points, $X_n = \{x_j\}_{j=0}^n$ where $a = x_0 < x_1 < \cdots < x_n = b$ and a set of functional values, $\{f(x_j)\}_{j=0}^n$, drafters will plot the data points $P_j = (x_j, f(x_j))$, $0 \le j \le n$. They will then take a thin elastic rod (called a *spline*) and a set of weights and will place the weights on the rod so that the rod must pass over each point P_j, $0 \le j \le n$. The resulting curve traced out by the spline then interpolates $f(x)$ at each x_j, and furthermore smooths out as much as possible between the points because of the elasticity of the rod. Thus the oscillatory behavior of the approximation is minimized as much as possible, but the approximation still retains the interpolation property.

We shall now consider a mathematical procedure that models the drafters' technique. First consider the set of all functions $Sp(X_n)$ such that if $S(x) \in Sp(X_n)$ then $S(x)$ satisfies the following three properties.

$S(x) \in C^2[a, b]$; that is, $S(x)$, $S'(x)$, and $S''(x)$ are continuous on $[a, b]$. (5.40a)

$S(x_j) = f(x_j) \equiv f_j$, $0 \le j \le n$; that is, $S(x)$ interpolates $f(x)$ on $[a, b]$. (5.40b)

$S(x)$ is a cubic polynomial on each subinterval $[x_j, x_{j+1}]$, $0 \le j \le n - 1$. (5.40c)

The function $S(x)$ defined by the three conditions of (5.40) is a piecewise cubic polynomial and is called a *cubic spline*. Higher-order splines are defined similarly. For example, a quartic spline would be a function $S(x)$ where $S(x) \in C^3[a, b]$ and $S(x)$ is a fourth-degree polynomial on each subinterval $[x_j, x_{j+1}]$. Note the dependence of the set of functions, $Sp(X_n)$, on the particular function, $f(x)$, being interpolated. Also note that $S(x)$ may be a *different* cubic on each subinterval; so we let $S_j(x)$ denote the cubic polynomial such that $S(x) = S_j(x)$ for x in $[x_j, x_{j+1}]$, $0 \le j \le n - 1$.

Before beginning the mathematical derivation needed to construct cubic splines, we pause to consider intuitively why we expect that there is any such function, $S(x)$, satisfying the three conditions of (5.40). For simplicity we let $n = 3$ so that $S(x)$ is represented by the cubic polynomials $S_0(x)$ on $[x_0, x_1]$, $S_1(x)$ on $[x_1, x_2]$, and $S_2(x)$ on $[x_2, x_3]$. Since each $S_j(x) \in \mathcal{P}_3$, it has four coefficients that we would like to choose to satisfy the conditions of (5.40). Therefore, in this case, we have twelve coefficients or unknowns at our disposal. Since each $S_j(x) \in \mathcal{P}_3$, all of its derivatives are continuous for any $x \in (x_j, x_{j+1})$—the *open interval*. Thus we need worry about the continuity conditions only at the interior interpolating points, x_1 and x_2, where the cubics must "patch together" with second-order continuity. With this point in mind, let us work from left to right on $[a, b] \equiv [x_0, x_3]$ and determine the number of equations that must be satisfied. Since $S(x)$ must interpolate $f(x)$ at each x_j, and since $S(x)$ must be continuous at x_1 and x_2, we have the following six equations to be satisfied:

$$S_0(x_0) = f_0; \qquad S_1(x_1) = f_1; \qquad S_0(x_1) = S_1(x_1); \qquad S_2(x_2) = f_2;$$
$$S_1(x_2) = S_2(x_2); \qquad \text{and} \qquad S_2(x_3) = f_3.$$

Finally since $S'(x)$ and $S''(x)$ must also be continuous at x_1 and x_2, we must have $S_0'(x_1) = S_1'(x_1)$, $S_0''(x_1) = S_1''(x_1)$, $S_1'(x_2) = S_2'(x_2)$, and $S_1''(x_2) = S_2''(x_2)$—four more equations. Therefore we have a total of ten equations in twelve unknowns. Thus we not only expect just one solution, but we hope that we can specify two of the unknowns to be designated arbitrarily and still have a solution for the remaining ten equations in ten unknowns. So, intuitively we expect that $Sp(X_n)$ is an infinite two-parameter family of functions. Two logical ways to obtain a good approximation for $f(x)$ would be to choose the two free parameters by specifying either $S'(x_0)$ and $S'(x_n)$ or $S''(x_0)$ and $S''(x_n)$. We choose the second alternative in the following mathematical derivation, but will also show how the first alternative may be implemented.

To construct a function $S(x)$ that satisfies the three conditions of (5.40), we first introduce some notation and then use the three conditions to obtain a linear system of equations that will enable us to determine the coefficients of each cubic polynomial, $S_j(x)$. We define $h_j \equiv \Delta x_j \equiv x_{j+1} - x_j$ and $\Delta f_j \equiv f(x_{j+1}) - f(x_j) \equiv f_{j+1} - f_j$. We next define $S''(x_j) \equiv y_j''$ and note that the quantities $y_0'', y_1'', \ldots, y_n''$ will appear as *unknowns* in a linear system of equations that will define the cubic spline $S(x)$. [The reader should note that usually $S''(x_j) \neq f''(x_j)$ and even $S'(x_j) \neq f'(x_j)$. All that (5.40) requires is that $S(x_j) = f(x_j)$, $0 \leq j \leq n$.]

The construction of the cubic spline proceeds roughly as follows. Suppose we choose any $n + 1$ values $y_0'', y_1'', \ldots, y_n''$ and then let $q(x)$ be the broken line such that $q(x_j) = y_j''$, $0 \leq j \leq n$. If we integrate $q(x)$ twice, we obtain a function $S(x)$ in $C^2[a, b]$, which is a piecewise cubic polynomial. The question is whether we can choose $y_0'', y_1'', \ldots, y_n''$ so that $S(x_j) = f_j$, $0 \leq j \leq n$. If we can, then $S(x)$ interpolates $f(x)$; so by (5.40), $S(x) \in Sp(X_n)$. Following this idea, we define the first-degree polynomials, $S_j''(x)$, on $[x_j, x_{j+1}]$ by

$$S_j''(x) = y_j'' \frac{x_{j+1} - x}{h_j} + y_{j+1}'' \frac{x - x_j}{h_j}, \qquad 0 \leq j \leq n - 1. \tag{5.41}$$

[In (5.41) we have not yet specified $y_0'', y_1'', \ldots, y_n''$; they can be any set of values.] Note that $S_j''(x_j) = y_j''$, $S_j''(x_{j+1}) = y_{j+1}''$, $0 \le j \le n-1$. Note also that $S_{j+1}''(x_{j+1}) = S_j''(x_{j+1})$ for $0 \le j \le n-2$; so the function $S''(x)$, which has the value $S_j''(x)$ for $x \in [x_j, x_{j+1}]$, is a continuous function defined on $[a, b]$ [$S''(x)$ is a broken line.] Integrating (5.41) twice, we obtain for $x \in [x_j, x_{j+1}]$

$$S_j(x) = \frac{y_j''}{6h_j}(x_{j+1} - x)^3 + \frac{y_{j+1}''}{6h_j}(x - x_j)^3 + c_j(x - x_j) + d_j(x_{j+1} - x) \quad (5.42)$$

where c_j and d_j are constants of integration. To make $S(x)$ continuous and to satisfy the interpolation constraints, we need to choose the constants of integration so that $S_j(x_j) = f_j$, $S_j(x_{j+1}) = f_{j+1}$ for $0 \le j \le n-1$. Substituting these two conditions into (5.42), we obtain

$$c_j = \frac{f_{j+1}}{h_j} - \frac{y_{j+1}''h_j}{6} \qquad \text{and} \qquad d_j = \frac{f_j}{h_j} - \frac{y_j''h_j}{6};$$

so

$$\begin{aligned} S_j(x) = {} & \frac{y_j''}{6h_j}(x_{j+1} - x)^3 + \frac{y_{j+1}''}{6h_j}(x - x_j)^3 + \left(\frac{f_{j+1}}{h_j} - \frac{y_{j+1}''h_j}{6}\right)(x - x_j) \\ & + \left(\frac{f_j}{h_j} - \frac{y_j''h_j}{6}\right)(x_{j+1} - x), \qquad 0 \le j \le n-1. \end{aligned} \quad (5.43)$$

Now, $S_j(x)$ as defined in (5.43) was chosen to match $f(x)$ at x_j and x_{j+1} by adjusting the constants of integration, c_j and d_j. Since we do this adjustment independently in each interval $[x_j, x_{j+1}]$, we have no guarantee that $S_j'(x_j) = S_{j-1}'(x_j)$ for $j = 1, 2, \ldots, n-1$ [that is, that $S'(x)$ is continuous]. Differentiating (5.43) yields

$$S_j'(x) = -\frac{y_j''}{2h_j}(x_{j+1} - x)^2 + \frac{y_{j+1}''}{2h_j}(x - x_j)^2 + \frac{\Delta f_j}{h_j} - \frac{h_j}{6}(y_{j+1}'' - y_j''). \quad (5.44)$$

So the last condition that we must meet to satisfy all of (5.40) is to make $S'(x)$ continuous at the interior interpolating points. We do this by using (5.44) and choosing the values y_j'', $1 \le j \le n-1$, so that $S_j'(x_j) = S_{j-1}'(x_j)$ for $1 \le j \le n-1$. This procedure yields the following linear system of $(n-1)$ equations in the unknowns $\{y_j''\}_{j=0}^{n}$:

$$\begin{aligned} & h_{j-1}y_{j-1}'' + 2(h_j + h_{j-1})y_j'' + h_j y_{j+1}'' = b_j \\ & b_j \equiv 6\left(\frac{\Delta f_j}{h_j} - \frac{\Delta f_{j-1}}{h_{j-1}}\right), \qquad 1 \le j \le n-1. \end{aligned} \quad (5.45)$$

Since we have $(n+1)$ unknowns and $(n-1)$ equations in (5.45), we can specify fixed values for y_0'' and y_n'' and transfer the two terms involving these values to the right-hand side. We then have the following $(n-1) \times (n-1)$ linear system (written in matrix form and denoted by $A\mathbf{y} = \mathbf{b}$):

$$\begin{bmatrix} \gamma_1 & h_1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ h_1 & \gamma_2 & h_2 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & & & & & & & \\ 0 & 0 & 0 & 0 & \cdots & h_{n-3} & \gamma_{n-2} & h_{n-2} \\ 0 & 0 & 0 & 0 & \cdots & 0 & h_{n-2} & \gamma_{n-1} \end{bmatrix} \begin{bmatrix} y_1'' \\ y_2'' \\ \vdots \\ y_{n-2}'' \\ y_{n-1}'' \end{bmatrix} = \begin{bmatrix} b_1 - h_0 y_0'' \\ b_2 \\ \vdots \\ b_{n-2} \\ b_{n-1} - h_{n-1} y_n'' \end{bmatrix} \tag{5.46}$$

where $\gamma_i = 2(h_i + h_{i-1})$.

It is clear that the coefficient matrix A is tridiagonal and diagonally dominant and hence nonsingular (see Chapter 2). Thus, no matter what values we select for y_0'' and y_n'', (5.46) has a unique solution for $\{y_j''\}_{j=1}^{n-1}$ (the solution does depend on our choice of y_0'' and y_n'', of course). Furthermore the $\{y_j''\}_{j=1}^{n-1}$ can be easily found by the finite recurrence given for the LU-decomposition of a tridiagonal matrix. These values are then substituted into (5.43) and thus yield $S(x)$. It is common, as we shall see shortly, to set $y_0'' = y_n'' = 0$; and the unique cubic spline that results from this choice is called the *natural cubic spline*. A final point that should be made is that once the values $y_1'', y_2'', \ldots, y_{n-1}''$ are determined from (5.46), we then use (5.43) to evaluate $S(x)$ at $x = \alpha$. That is, if $\alpha \in [x_j, x_{j+1}]$, then $S(\alpha) = S_j(\alpha)$.

EXAMPLE 5.12. In this example, we consider approximating $f(x) = |x|$ by a cubic spline and contrast the spline approximation with an interpolating polynomial. Let $x_0 = -2$, $x_1 = -1$, $x_2 = 0$, $x_3 = 1$, and $x_4 = 2$. In this case, $h_j = 1$ for $j = 0, 1, 2, 3$ and $b_j = 6(f_{j+1} - 2f_j + f_{j-1}) = 6\Delta^2 f_{j-1}$ for $j = 1, 2, 3$. Thus we are led to the linear system [as in Eq. (5.46)]

$$\begin{aligned} 4y_1'' + y_2'' \qquad &= 0 \\ y_1'' + 4y_2'' + y_3'' &= 12 \\ y_2'' + 4y_3'' &= 0. \end{aligned}$$

Solving this system, we find $y_1'' = -6/7$, $y_2'' = 24/7$, and $y_3'' = -6/7$. Using these values in (5.43), we can evaluate the cubic spline in any subinterval $[x_j, x_{j+1}]$.

The fourth-degree interpolating polynomial for $f(x)$ at the points x_j above is easily seen to be given by $p(x) = \frac{7}{6}x^2 - \frac{1}{6}x^4$. We will compare these two approximations in the interval $[1, 2]$, where to evaluate the cubic spline in $[1, 2]$ we use [from (5.43)]

$$S_3(x) = -(2-x)^3/7 + 2(x-1) + 8(2-x)/7.$$

Since both of these approximations and $f(x)$ as well have such simple forms, it is easy to verify that the maximum value of $p(x) - |x|$ occurs at approximately $x = 1.6$. The maximum value of $(S_3(x) - |x|)$ occurs approximately at $x = 1.423$ and $|S_3(x) - |x|| \le 0.055$ for all x in $[1, 2]$. In Table 5.3, we have listed values of $S_3'(x)$ and $p'(x)$ as well. [As we shall see, the derivative of the cubic spline provides a fairly good approximation to

TABLE 5.3

x	$f(x)$	$S_3(x)$	$S_3'(x)$	$p(x)$	$p'(x)$
1.0	1.0	1.000	1.286	1.000	1.667
1.2	1.2	1.241	1.131	1.334	1.648
1.4	1.4	1.455	1.011	1.646	1.437
1.6	1.6	1.648	0.926	1.894	1.003
1.8	1.8	1.827	0.874	2.030	0.312
2.0	2.0	2.000	0.857	2.000	−0.667

$f'(x)$. On the other hand, derivatives of the interpolating polynomial are not generally good approximations to the derivatives of $f(x)$.] Since $f'(x) \equiv 1$ for $1 \leq x \leq 2$, this example shows how well the cubic spline can duplicate the general shape of $f(x)$. We shall return to this point when we discuss numerical differentiation in the next chapter.

If the derivatives of $f(x)$ at the endpoints are known, that is, if $f'(x_0) \equiv y_0'$ and $f'(x_n) \equiv y_n'$ are given, then we may suspect that we will get a better cubic spline approximation to $f(x)$ if we choose the two free parameters, y_0'' and y_n'', in a manner such that the resulting cubic spline, $S(x)$, satisfies $S'(x_0) = f'(x_0) = y_0'$ and $S'(x_n) = f'(x_n) = y_n'$ as well as the interpolation property, $S(x_j) = f_j$, $0 \leq j \leq n$. From Eq. (5.44) we see for $j = n - 1$ that

$$f'(x_n) = S_{n-1}'(x_n) = \frac{y_n''}{2h_{n-1}}(x_n - x_{n-1})^2 + \frac{\Delta f_{n-1}}{h_{n-1}} - \frac{h_{n-1}}{6}(y_n'' - y_{n-1}''),$$

or

$$\frac{h_{n-1}}{3} y_n'' + \frac{h_{n-1}}{6} y_{n-1}'' = y_n' - \frac{\Delta f_{n-1}}{h_{n-1}}. \tag{5.47a}$$

Similarly, from (5.44) with $j = 0$, we obtain

$$\frac{h_0}{6} y_1'' + \frac{h_0}{3} y_0'' = \frac{\Delta f_0}{h_0} - y_0'. \tag{5.47b}$$

The continuity of $S'(x)$ must be maintained so that all of the equations of (5.45) must still hold. If Eqs. (5.45) are written together with (5.47), we get an $(n + 1) \times (n + 1)$ tridiagonal diagonally dominant linear system of equations. These may be easily solved as before and again substituted into (5.43) to obtain $S(x)$. (There are other ways of deriving this particular cubic spline, but they involve a different derivation. See Rivlin, 1969.)

For brevity of notation we shall denote the cubic spline above as $S^{(1)}(x)$ and the natural cubic spline as $S^{(2)}(x)$. These two particular cubic splines are the ones that are most often used in practice because of the so-called *extremal*