9 Polynomial Interpolation

In this chapter, we begin the study of *constructive approximation theory*, which, as its name suggests, is concerned with methods to approximate a function (which may only be known approximately) by a simpler one. One of the recurring features in our discussion will be the interplay between the smoothness of a function, measured in an appropriate sense, and the quality of approximation that we are able to produce. Applications of approximation theory are plentiful, and we will see several of these throughout this book. For instance, in Chapter 14, we will see how this is used to approximate the value of integrals. It will also play a central role in Part V, where the performance of a numerical scheme for the approximation of the solution to a boundary value problem depends in a fundamental way not only on the method of choice but also on the *smoothness of the solution*.

We begin approximation theory with the topic of polynomial approximation. Given a, usually continuous, function we wish to construct a polynomial satisfying certain properties that, in a very definite sense, approximates the given function. In fact, most of this part of the book is about generating approximations with polynomials, of both the ordinary and the trigonometric kind.

Let us immediately remark that, since our discussion is now concerned with functions and their properties, the reader should be familiar with some basic facts in real analysis: continuity, compactness, etc. We refer to Appendix B for a review and guide to notation. Some facts about spaces of smooth functions will also be necessary, and Appendix D provides an overview. With this in mind, our discussion is started by first presenting a cornerstone approximation result, which the reader may have seen before. For the moment, this is stated without proof.

Theorem (Weierstrass Approximation Theorem[1]). *Let $[a, b] \subset \mathbb{R}$ be a compact interval and $f \in C([a, b])$. For every $\varepsilon > 0$, there exists an $n \in \mathbb{N}$ and a polynomial $p_n \in \mathbb{P}_n$ such that*

$$\|f - p_n\|_{L^\infty(a,b)} \leq \varepsilon.$$

In other words, there is a sequence of polynomials, $\{p_n\}_{n=0}^{\infty}$ with $p_n \in \mathbb{P}_n$, that converges uniformly to f as $n \to \infty$.

This theorem tells us that a continuous function on a compact interval can be well approximated by polynomials, but it does not tell us how to construct such a polynomial and it does not tell us the convergence rate.

[1] Named in honor of the German mathematician Karl Theodor Wilhelm Weierstrass (1815–1897).

https://doi.org/10.1017/9781108942607.013 Published online by Cambridge University Press

On the other hand, Taylor's Theorem B.31 gives a precise way to construct a polynomial approximation, but requires the function f to be n-times differentiable to build an approximating polynomial of degree n. Moreover, we have to know the values of all the derivatives at a certain point, which is not always practical.

Suppose, for example, that one only knows function values at a finite collection of points in the domain. What can we do? The celebrated theorems mentioned above do not give any indication. For constructive approximation, another place to start is interpolation. Interpolation works by demanding that a simple function — usually, but not always, a polynomial — exactly matches the values of a function of interest at a given number of points. Typically, we only require that the function of interest is continuous, or piecewise continuous, in its domain of definition. Interpolation is simple and often works well. But sometimes it can go badly wrong. Paradoxically, this can happen when we try to interpolate a large number of points. One must be careful, as we will see.

9.1 The Vandermonde Matrix and the Vandermonde Construction

To understand interpolation, we need some basic definitions.

Definition 9.1 (nodal set)**.** Let $[a, b] \subset \mathbb{R}$ be a compact interval. X is called a **nodal set** of size $n+1 \in \mathbb{N}$ in $[a, b]$ if and only if $X = \{x_i\}_{i=0}^{n} \subset [a, b]$ is a set of distinct elements. The elements of X, x_i are called **nodes**.

Observe that we usually enumerate the nodes in a very particular way, starting with 0 and ending with n. Also note that the points of any nodal set are always, by design, distinct. Not all finite sets that we shall introduce will have this property. Typically the nodes are numbered in increasing order, for convenience; but this is not always the case.

Definition 9.2 (interpolating polynomial)**.** Suppose that $X = \{x_i\}_{i=0}^{n}$ is a nodal set of size $n+1 \in \mathbb{N}$ contained in the compact interval $[a, b] \subset \mathbb{R}$ and $f\colon [a, b] \to \mathbb{R}$ is a function. The function $I\colon [a, b] \to \mathbb{R}$ is called an **interpolant of** f subordinate to X if and only if $I(x_i) = f(x_i)$, $i = 0, \dots, n$. In this case, we write $I(X) = f(X)$, for short. Suppose that $Y = \{y_i\}_{i=0}^{n} \subset \mathbb{R}$ is a set of not necessarily distinct points. Define the set of ordered pairs

$$O = \{(x_i, y_i) \mid x_i \in X,\ \ y_i \in Y,\ \ i = 0, \dots, n\}.$$

We say that I is an **interpolant of** O if and only if $I(x_i) = y_i$, $i = 0, \dots, n$. Often, we write $I(X) = Y$ as a shorthand. If the interpolant I is a polynomial, it is called an **interpolating polynomial**.

In short, an interpolant is a function that agrees with another, given, function at a finite number of points in its domain.

Definition 9.3 (Vandermonde matrix[2])**.** Suppose that $X = \{x_i\}_{i=0}^{n}$ is a nodal set

[2] Named in honor of the French mathematician, musician and chemist Alexandre-Théophile Vandermonde (1735–1796).

https://doi.org/10.1017/9781108942607.013 Published online by Cambridge University Press

of size $n+1 \in \mathbb{N}$ in the compact interval $[a, b] \subset \mathbb{R}$. The **Vandermonde matrix** subordinate to the nodal set X, denoted $\mathsf{V} = [v_{i,j}] \in \mathbb{R}^{(n+1)\times(n+1)}$, is the matrix with entries

$$v_{i,j} = x_{i-1}^{j-1}, \quad i, j = 1, \dots, n+1,$$

where the superscript represents exponentiation. In other words,

$$\mathsf{V} = \begin{bmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{bmatrix}.$$

Theorem 9.4 (Vandermonde). *Suppose that $X = \{x_i\}_{i=0}^n$ is a nodal set in the compact interval $[a, b] \subset \mathbb{R}$ and $\mathsf{V} = [v_{i,j}] \in \mathbb{R}^{(n+1)\times(n+1)}$ is the Vandermonde matrix subordinate to X. Then V is invertible and, in particular,*

$$\det(\mathsf{V}) = \prod_{0 \le i < j \le n} (x_j - x_i) \neq 0.$$

Proof. See Problem 9.1. Of course, if one can establish the determinant formula, the invertibility of V follows because nodes are distinct. □

With this at hand we can establish the existence and uniqueness of an interpolating polynomial.

Proposition 9.5 (existence and uniqueness). *Suppose that $X = \{x_i\}_{i=0}^n$ is a nodal set in the compact interval $[a, b] \subset \mathbb{R}$ and $Y = \{y_i\}_{i=0}^n \subset \mathbb{R}$. There is a unique polynomial $p \in \mathbb{P}_n$ with the property that $p(X) = Y$.*

Proof. We can set this problem up in the following way. Express p as $p(x) = \sum_{j=0}^n c_j x^j$, where the coefficients $c_i \in \mathbb{R}$ must be determined. Observe that p satisfies $p(X) = Y$ if and only if, for all $i = 0, \dots, n$,

$$p(x_i) = \sum_{j=0}^n c_j x_i^j = y_i.$$

In matrix form, this can be rewritten as

$$\begin{bmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{bmatrix}, \tag{9.1}$$

where the vector $\boldsymbol{c} = [c_0, \dots, c_n]^\intercal \in \mathbb{R}^{n+1}$ is unknown. Since the coefficient matrix is the Vandermonde matrix, which is invertible, the result follows. □

The interpolating polynomial construction represented by the linear system (9.1) is called the *Vandermonde construction*.

https://doi.org/10.1017/9781108942607.013 Published online by Cambridge University Press

Example 9.1 While, in theory, all that is needed to construct an interpolating polynomial is to solve system (9.1), it turns out that the Vandermonde matrix tends to be severely ill-conditioned. For example, define $X_n = \{i/n\}_{i=0}^n$, a set of uniform nodes in $[0, 1]$. The following table shows how the spectral condition number of V grows with n:

n	$\kappa_2(\mathsf{V})$
4	6.86×10^{02}
8	2.01×10^{06}
16	2.42×10^{13}

Clearly, one should avoid using the Vandermonde matrix in practical applications when n is large. We will look for other more practical ways of constructing interpolants.

Interpolation, from the theoretical perspective, can be thought of as a linear projection operator.

Definition 9.6 (interpolation operator)**.** Suppose that $X = \{x_i\}_{i=0}^n$ is a nodal set in the compact interval $[a, b] \subset \mathbb{R}$. The **interpolation operator** subordinate to X, denoted

$$\mathcal{I}_X \colon C([a, b]) \to \mathbb{P}_n,$$

is defined as follows: for $f \in C([a, b])$, $\mathcal{I}_X[f] \in \mathbb{P}_n$ is the unique interpolating polynomial satisfying $\mathcal{I}_X[f](X) = f(X)$.

The reader can easily prove the following result.

Proposition 9.7 (projection)**.** *Suppose that $X = \{x_i\}_{i=0}^n$ is a nodal set in the compact interval $[a, b] \subset \mathbb{R}$ and $\mathcal{I}_X \colon C([a, b]) \to \mathbb{P}_n$ is the interpolation operator subordinate to X. $\mathcal{I}_X$ is a linear projection operator, meaning that*

$$\mathcal{I}_X[\alpha f + \beta g] = \alpha \mathcal{I}_X[f] + \beta \mathcal{I}_X[g], \quad \forall f, g \in C([a, b]), \quad \forall \alpha, \beta \in \mathbb{R}$$

and

$$\mathcal{I}_X[p] = p, \quad \forall p \in \mathbb{P}_n.$$

Proof. See Problem 9.2. □

The norm of the interpolation operator has a special name.

Definition 9.8 (Lebesgue constant[3])**.** Suppose that $X = \{x_i\}_{i=0}^n$ is a nodal set in the compact interval $[a, b] \subset \mathbb{R}$ and $\mathcal{I}_X \colon C([a, b]) \to \mathbb{P}_n$ is the interpolation

[3] Named in honor of the French mathematician Henri Léon Lebesgue (1875–1941).

https://doi.org/10.1017/9781108942607.013 Published online by Cambridge University Press

operator subordinate to X. The **Lebesgue constant subordinate to** X, denoted $\Lambda(X)$, is the operator norm of $\mathcal{I}_X$, i.e., $\Lambda(X) = \|\mathcal{I}_X\|_\infty$, where

$$\|\mathcal{I}_X\|_\infty = \sup_{0\neq f\in C([a,b])} \frac{\|\mathcal{I}_X[f]\|_{L^\infty(a,b)}}{\|f\|_{L^\infty(a,b)}} = \sup_{\substack{f\in C([a,b])\\ \|f\|_{L^\infty(a,b)}=1}} \|\mathcal{I}_X[f]\|_{L^\infty(a,b)}. \tag{9.2}$$

In the next section, we will show how one might compute the Lebesgue constant. At this point, just note that the constant depends upon our choice of the nodal set.

9.2 Lagrange Interpolation and the Lagrange Nodal Basis

It is sometimes more convenient, from the viewpoint of representation and construction, to use the so-called Lagrange basis to build our interpolating polynomials.

Definition 9.9 (Lagrange nodal basis[4])**.** Suppose that $X = \{x_i\}_{i=0}^n$ is a nodal set of size $n+1 \in \mathbb{N}$ in the compact interval $[a,b] \subset \mathbb{R}$. The **Lagrange nodal basis** subordinate to X is the set of polynomials $\mathcal{L}_X = \{L_\ell\}_{\ell=0}^n \subset \mathbb{P}_n$ defined via

$$L_\ell(x) = \prod_{\substack{i=0\\ i\neq\ell}}^n \frac{x - x_i}{x_\ell - x_i}. \tag{9.3}$$

Proposition 9.10 (properties of $\mathcal{L}_X$)**.** *Suppose that $X = \{x_i\}_{i=0}^n$ is a nodal set of size $n+1 \in \mathbb{N}$ in the compact interval $[a,b] \subset \mathbb{R}$ and $\mathcal{L}_X = \{L_i\}_{i=0}^n \subset \mathbb{P}_n$ is the Lagrange basis subordinate to X. Then $\mathcal{L}_X$ is a basis of $\mathbb{P}_n$ with the properties*

$$L_i(x_j) = \delta_{i,j}, \quad i,j = 0,\ldots,n \tag{9.4}$$

and

$$\sum_{i=0}^n L_i(x) = 1, \quad \forall x \in \mathbb{R}. \tag{9.5}$$

Proof. See Problem 9.3 □

Example 9.2 Figure 9.1 shows the Lagrange nodal basis subordinate to the nodal set $X = \{0, 0.30, 0.42, 0.71, 1.00\} \subset [0,1]$. The figure also shows the sum of the basis functions, $\sum_{i=0}^4 L_i(x)$, which confirms (9.5) for this case.

Theorem 9.11 (interpolation polynomial)**.** *Suppose that $X = \{x_i\}_{i=0}^n$ is a nodal set in the compact interval $[a,b] \subset \mathbb{R}$, and $Y = \{y_i\}_{i=0}^n \subset \mathbb{R}$. Let $\mathcal{L}_X = \{L_i\}_{i=0}^n \subset \mathbb{P}_n$*

[4] Named in honor of the Italian, later naturalized French, mathematician and astronomer Joseph-Louis Lagrange (1736–1813).

https://doi.org/10.1017/9781108942607.013 Published online by Cambridge University Press

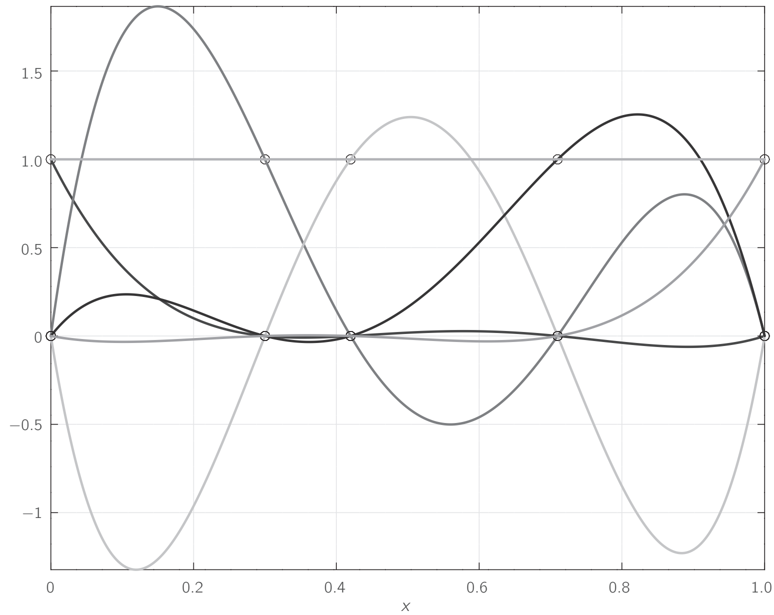

Figure 9.1 The Lagrange nodal basis of degree $n = 4$, subordinate to the nodal set $X = \{0, 0.30, 0.42, 0.71, 1.00\} \subset [0, 1]$. We also plot the sum of the basis functions, $\sum_{i=0}^{4} L_i(x)$, which confirms (9.5) for this case.

be the Lagrange nodal basis subordinate to X. The unique polynomial $p \in \mathbb{P}_n$, with the property that $p(X) = Y$, has the form

$$p(x) = \sum_{i=0}^{n} y_i L_i(x) \in \mathbb{P}_n. \tag{9.6}$$

Proof. From definition (9.6), we have, using property (9.4), that $p(x_i) = y_i$. $\square$

Definition 9.12 (Lagrange interpolating polynomial)**.** Suppose that $X = \{x_i\}_{i=0}^{n}$ is a nodal set in the compact interval $[a, b] \subset \mathbb{R}$, $\mathcal{L}_X = \{L_i\}_{i=0}^{n} \subset \mathbb{P}_n$ is the Lagrange nodal basis subordinate to X, and $f : [a, b] \to \mathbb{R}$. The **Lagrange interpolating polynomial** of the function f, subordinate to the nodal set X, is the polynomial

$$p(x) = \sum_{i=0}^{n} f(x_i) L_i(x) \in \mathbb{P}_n. \tag{9.7}$$

Observe that, by uniqueness, the Lagrange interpolating polynomial coincides with the interpolant we obtained in Proposition 9.5 via the Vandermonde construction. For historical reasons, the interpolating polynomial constructed by matching point values of the given function is called the Lagrange interpolating polynomial. We will give one more construction in Section 9.6 based on Newton's basis.

Definition 9.13 (nodal polynomial)**.** Suppose that $X = \{x_i\}_{i=0}^n$ is a nodal set in the compact interval $[a, b] \subset \mathbb{R}$. The polynomial $\omega_{n+1} \in \mathbb{P}_{n+1}$, defined by

$$\omega_{n+1}(x) = \prod_{i=0}^{n} (x - x_i), \tag{9.8}$$

is called the **nodal polynomial subordinate to** X.

The following alternate formula for the elements of the Lagrange nodal basis is often useful.

Proposition 9.14 (Lagrange nodal basis)**.** *Suppose that $X = \{x_i\}_{i=0}^n$ is a nodal set in the compact interval $[a, b] \subset \mathbb{R}$, $\mathcal{L}_X = \{L_i\}_{i=0}^n \subset \mathbb{P}_n$ is the Lagrange nodal basis subordinate to X, ω_{n+1} is the nodal polynomial subordinate to X, and $f\colon [a, b] \to \mathbb{R}$. Then, for all $i = 0, \ldots, n$, we have*

$$L_i(x) = \frac{\omega_{n+1}(x)}{(x - x_i)\omega'_{n+1}(x_i)}. \tag{9.9}$$

Consequently, the Lagrange interpolating polynomial, $p \in \mathbb{P}_n$, of the function f subordinate to the nodal set X is

$$p(x) = \sum_{i=0}^{n} f(x_i) \frac{\omega_{n+1}(x)}{(x - x_i)\omega'_{n+1}(x_i)}. \tag{9.10}$$

Proof. One can show that

$$\omega'_{n+1}(x_i) = \prod_{\substack{j=0 \\ j \neq i}}^{n} (x_i - x_j).$$

This and further details are left to the reader as an exercise; see Problem 9.5. $\square$

Before we move on to the error analysis for Lagrange interpolation, let us compute the Lebesgue constant.

Theorem 9.15 ($\Lambda(X)$)**.** *Suppose that $X = \{x_i\}_{i=0}^n$ is a nodal set in the compact interval $[a, b] \subset \mathbb{R}$. The Lebesgue constant satisfies*

$$\Lambda(X) = \max_{a \le x \le b} \lambda_X(x), \quad \lambda_X(x) = \sum_{i=0}^{n} |L_i(x)| \ge 1,$$

where $\mathcal{L}_X = \{L_i\}_{i=0}^n \subset \mathbb{P}_n$ is the Lagrange nodal basis subordinate to X. The function λ_X is called the Lebesgue function subordinate to X*.*

Proof. Suppose that $f \in C([a, b])$ is arbitrary. Then

$$|\mathcal{I}_X[f](x)| = \left| \sum_{i=0}^{n} f(x_i) L_i(x) \right| \le \max_{0 \le i \le n} |f(x_i)| \sum_{i=0}^{n} |L_i(x)| \le \|f\|_{L^\infty(a,b)} \sum_{i=0}^{n} |L_i(x)|.$$

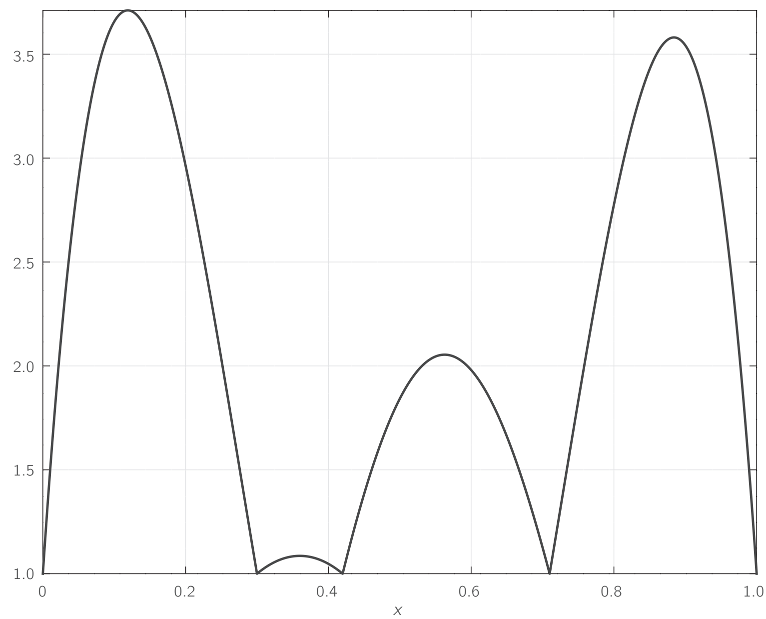

Figure 9.2 The Lebesgue function, $\lambda_X(x) = \sum_{i=0}^{4} |L_i(x)|$, subordinate to the nodal set $X = \{0, 0.30, 0.42, 0.71, 1.00\} \subset [0, 1]$. The basis functions, L_i, $i = 0, \ldots, 4$, are plotted in Figure 9.1, for comparison.

Thus,

$$\|\mathcal{I}_X\|_\infty \le \max_{a \le x \le b} \sum_{i=0}^{n} |L_i(x)| .$$

Now, on the other hand, there is a function $f \in C([a, b])$ with $\|f\|_{L^\infty(a,b)} = 1$ satisfying

$$\max_{a \le x \le b} |\mathcal{I}_X[f](x)| = \max_{a \le x \le b} \left| \sum_{i=0}^{n} f(x_i) L_i(x) \right| = \max_{a \le x \le b} \sum_{i=0}^{n} |L_i(x)| ,$$

which proves that

$$\|\mathcal{I}_X\|_\infty = \max_{a \le x \le b} \sum_{i=0}^{n} |L_i(x)| .$$

We leave it to the reader as an exercise to find the appropriate function f; see Problem 9.6. □

Example 9.3 The Lebesgue function, $\lambda_X(x) = \sum_{i=0}^{4} |L_i(x)|$, subordinate to the nodal set $X = \{0, 0.30, 0.42, 0.71, 1.00\} \subset [0, 1]$ is shown in Figure 9.2. The basis functions, L_i, $i = 0, \ldots, 4$, are plotted in Figure 9.1, for comparison.

https://doi.org/10.1017/9781108942607.013 Published online by Cambridge University Press

In Chapter 10, we will give an error estimate for Lagrange interpolation that involves the Lebesgue constant and an object known as the minimax polynomial. For now, we give a more standard error formula.

Theorem 9.16 (Lagrange interpolation error)**.** *Suppose that $X = \{x_i\}_{i=0}^n$ is a nodal set in the compact interval $[a, b] \subset \mathbb{R}$, $f \in C^{n+1}([a, b])$, and $p \in \mathbb{P}_n$ is the Lagrange interpolating polynomial of f subordinate to X. Then, for every $x \in [a, b]\backslash X$, there is a point $\xi = \xi(x) \in (a, b)$ with*

$$\min\{x_0, \dots, x_n, x\} < \xi < \max\{x_0, \dots, x_n, x\}$$

such that

$$f(x) - p(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}\omega_{n+1}(x), \tag{9.11}$$

where ω_{n+1} is the nodal polynomial introduced in Definition 9.13.

Proof. Fix $x \in [a, b]\backslash X$. Let us consider cases.

($n = 0$) In this case, $p(x) = f(x_0)$ for all $x \in [a, b]$. By the Mean Value Theorem B.30, there is a point ξ between x and x_0 such that

$$f(x) - p(x) = f(x) - f(x_0) = f'(\xi)(x - x_0),$$

which is the same as (9.11), since, in this case, $(n + 1)! = 1$ and $\omega_1(x) = x - x_0$.

($n \geq 1$) Define, for any $s \in [a, b]$,

$$e(s) = f(s) - p(s) - \frac{f(x) - p(x)}{\omega_{n+1}(x)}\omega_{n+1}(s).$$

Observe that, for all $i = 0, \dots, n$,

$$e(x_i) = 0 - \frac{f(x) - p(x)}{\omega_{n+1}(x)} \cdot 0 = 0.$$

Furthermore, $e(x) = 0$. Thus, $e(s)$ is zero at at least $n+2$ distinct points in $[a, b]$. By Rolle's Theorem B.29, there are $n+1$ distinct points in (a, b) where e' vanishes. Applying Rolle's Theorem B.29 again, there are n distinct points in (a, b) where e'' vanishes. Continuing in this fashion, there is one point (at least) ξ in (a, b) for which $e^{(n+1)}(\xi) = 0$. Now observe that

$$0 = e^{(n+1)}(\xi) = f^{(n+1)}(\xi) - \frac{f(x) - p(x)}{\omega_{n+1}(x)}(n + 1)!.$$

Rearranging terms we get (9.11). □

The interpolating polynomial, p, can become a surrogate for the original function f. If what is wanted is a derivative of f, we can approximate that with the derivative of p. If what is wanted is the integral of f, we can integrate p instead. This will be much further explored in Chapter 14, where we discuss numerical integration.

For the present, with a technique similar to that used in the last proof, we can also establish an error formula for derivatives.

Theorem 9.17 (error formula for derivatives)**.** *Suppose that* $X = \{x_i\}_{i=0}^n$ *is a nodal set in the compact interval* $[a, b] \subset \mathbb{R}$, $f \in C^{n+1}([a, b])$, *and* $p \in \mathbb{P}_n$ *is the Lagrange interpolating polynomial of* f *subordinate to* X. *Then there are distinct points* $\zeta_i \in (a, b)$, $i = 1, \ldots, n$, *whose values depend on* f *and* X, *such that, for every* $x \in [a, b] \backslash X$, *there is a point* $\xi = \xi(x) \in (a, b)$ *with*

$$\min\{x_0, \ldots, x_n, x\} < \xi < \max\{x_0, \ldots, x_n, x\}$$

for which

$$f'(x) - p'(x) = \frac{f^{(n+1)}(\xi)}{n!}\psi_n(x), \tag{9.12}$$

where

$$\psi_n(x) = \prod_{i=1}^{n}(x - \zeta_i). \tag{9.13}$$

Proof. By Rolle's Theorem B.29, there are points $\zeta_i \in (x_{i-1}, x_i)$, $i = 1, \ldots, n$ such that $f'(\zeta_i) - p'(\zeta_i) = 0$. Clearly, all these ζ_i are distinct and only depend on f and the distribution of the nodes, x_i.

Suppose that $x \in [a, b] \backslash \{\zeta_i\}_{i=1}^n$. Define, for all $s \in [a, b]$,

$$e(s) = f'(s) - p'(s) - \frac{f'(x) - p'(x)}{\psi_n(x)}\psi_n(s),$$

where $\psi_n(x) = \prod_{i=1}^n (x - \zeta_i)$. Clearly, e vanishes at $n + 1$ distinct points: ζ_i, $i = 1, \ldots, n$, and x. Repeated application of Rolle's Theorem B.29 gives the result, as before. □

9.3 The Runge Phenomenon

It is natural to think that, by increasing the size of the nodal set and, therefore, the degree of the interpolating polynomial, we should be able to get better and better approximations of the function of interest. But this is not always the case.

Proposition 9.18 (conditional convergence)**.** *Suppose that* $n \in \mathbb{N}$, $[a, b]$ *is a compact interval,* $h = \frac{b-a}{n}$, *and* $X = \{x_i\}_{i=0}^n$ *is the uniformly spaced nodal set*

$$x_i = a + ih, \quad i = 0, 1, \ldots, n.$$

Let $f \in C^\infty([a, b])$ *and* $p_n \in \mathbb{P}_n$ *be the Lagrange interpolating polynomial of* f *subordinate to* X. *Then*

$$\|f - p_n\|_{L^\infty(a,b)} = \max_{a \le x \le b} |f(x) - p_n(x)| \to 0, \qquad n \to \infty,$$

provided that

$$\lim_{n\to\infty} \frac{F_{n+1}\Omega_{n+1}}{(n+1)!} = 0,$$

https://doi.org/10.1017/9781108942607.013 Published online by Cambridge University Press

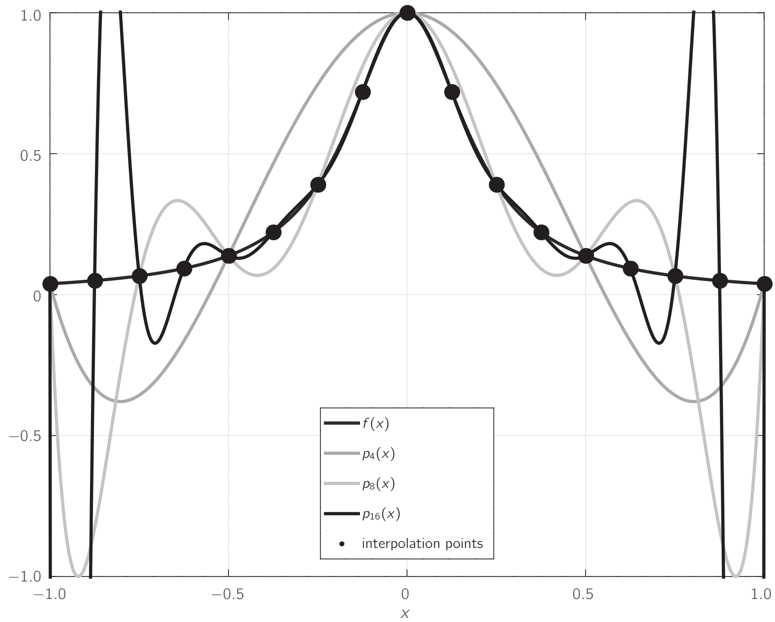

Figure 9.3 The Runge phenomenon: with uniformly spaced interpolation nodes, interpolation polynomials can oscillate wildly with increasing n, yielding inaccurate approximations.

where

$$F_{n+1} = \left\| f^{(n+1)} \right\|_{L^\infty(a,b)}, \qquad \Omega_{n+1} = \frac{n!}{4} h^{n+1}.$$

In other words, the sequence $\{p_n\}_{n\in\mathbb{N}}$ *converges uniformly to* f*, provided that*

$$\lim_{n\to\infty} \frac{F_{n+1} h^{n+1}}{4(n+1)} = 0.$$

Proof. It suffices to use (9.11) and prove that

$$\max_{a\le x\le b} |\omega_{n+1}(x)| \le \Omega_{n+1}.$$

We leave the remaining steps to the reader as an exercise; see Problem 9.8. □

The convergence condition of the previous result is not a trivial one, as the following example shows.

https://doi.org/10.1017/9781108942607.013 Published online by Cambridge University Press

Example 9.4 The following is commonly referred to as the *Runge phenomenon*.[5] Suppose that $n \in \mathbb{N}$, $h = \frac{2}{n}$, and $X_n = \{x_i\}_{i=0}^n$ is the uniformly spaced nodal set

$$x_i = -1 + ih, \quad i = 0, 1, \dots, n.$$

Suppose that $p_n \in \mathbb{P}_n$ is the Lagrange interpolating polynomial subordinate to X_n of the function

$$f(x) = \frac{1}{1 + 25x^2}, \qquad x \in [-1, 1].$$

Figure 9.3 shows the interpolating polynomials p_4, p_8, and p_{16}. It appears that the error is growing in the $L^\infty(-1,1)$ norm as n is getting larger. In particular, the error in the tails is increasing. In fact, it is possible to prove that, for this problem,

$$\|f - p_n\|_{L^\infty(-1,1)} \to \infty, \qquad n \to \infty.$$

This is due to the fact that, in this case, the growth in F_{n+1} cannot be controlled by $\frac{h^{n+1}}{4(n+1)}$. This has to do with the structure of f and the fact that our points are uniformly spaced. However, if we modify the spacing of the points in a smart way, we can do better, as we will show when we discuss Chebyshev interpolation.

9.4 Hermite Interpolation

One way to generalize Lagrange interpolation is to match not only the values of a given function at a nodal set but also the values of a certain number of derivatives. Matching point values and first derivatives of a function leads to Hermite interpolation.

Definition 9.19 (Hermite interpolating polynomial[6]). Suppose that $X = \{x_i\}_{i=0}^n$ is a nodal set of size $n+1 \in \mathbb{N}$ in the compact interval $[a, b] \subset \mathbb{R}$ and $f \in C^1([a, b])$. The polynomial $p \in \mathbb{P}_{2n+1}$ is called a **Hermite interpolating polynomial** of f if and only if

$$p(x_i) = f(x_i), \qquad p'(x_i) = f'(x_i), \qquad i = 0, \dots, n.$$

Theorem 9.20 (existence and uniqueness). *Suppose that $X = \{x_i\}_{i=0}^n$ is a nodal set of size $n+1 \in \mathbb{N}$ in the compact interval $[a, b] \subset \mathbb{R}$ and $f \in C^1([a, b])$. Then the Hermite interpolating polynomial is well defined, i.e., it exists and is unique.*

Proof. Since the case $n = 0$ is trivial, let us suppose that $n \geq 1$. Define, for $0 \leq \ell \leq n$,

$$F_{0,\ell}(x) = (L_\ell(x))^2(1 - 2L'_\ell(x_\ell)(x - x_\ell)), \qquad F_{1,\ell} = (L_\ell(x))^2(x - x_\ell), \quad (9.14)$$

[5] Named in honor of the German mathematician, physicist, and spectroscopist Carl David Tolmé Runge (1856–1927).
[6] Named in honor of the French mathematician Charles Hermite (1822–1901).

https://doi.org/10.1017/9781108942607.013 Published online by Cambridge University Press

where L_ℓ is as in (9.3), and

$$p(x) = \sum_{\ell=0}^{n} \left[F_{0,\ell}(x) f(x_\ell) + F_{1,\ell}(x) f'(x_\ell) \right]. \tag{9.15}$$

We leave it to the reader as an exercise to prove that p has the desired properties and is unique; see Problem 9.9. $\square$

Theorem 9.21 (Hermite interpolation error). *Suppose that $X = \{x_i\}_{i=0}^n$ is a nodal set of size $n+1 \in \mathbb{N}$ in the compact interval $[a,b] \subset \mathbb{R}$, $f \in C^{(2n+2)}([a,b])$, and $p \in \mathbb{P}_{2n+1}$ is the unique Hermite interpolating polynomial of f. Then, for every $x \in [a,b] \backslash X$, there is a point $\xi = \xi(x) \in (a,b)$, with*

$$\min\{x_0, \dots, x_n, x\} < \xi < \max\{x_0, \dots, x_n, x\}$$

such that

$$f(x) - p(x) = \frac{f^{(2n+2)}(\xi)}{(2n+2)!} (\omega_{n+1}(x))^2, \tag{9.16}$$

where ω_{n+1} is given in 9.8.

Proof. Fix $x \in [a,b] \backslash X$. Define, for all $s \in [a,b]$,

$$e(s) = f(s) - p(s) - \frac{f(x) - p(x)}{(\omega_{n+1}(x))^2} (\omega_{n+1}(s))^2.$$

As before, e vanishes at $n+2$ distinct points in $[a,b]$: the $n+1$ nodes x_i and the point x. By Rolle's Theorem B.29, e' vanishes at $n+1$ distinct points, which are distinct from the $n+2$ points in $X \cup \{x\}$. Furthermore, by construction, e' vanishes at the $n+1$ points of X. Thus, e' vanishes at a total of (at least) $2n+2$ distinct points in $[a,b]$. This is the key fact we need.

By repeated application of Rolle's Theorem B.29, the function $e^{(2n+2)}$, which exists and is continuous by assumption, vanishes at (at least) one point $\xi \in (a,b)$. Of course, since $p \in \mathbb{P}_{2n+1}$, $p^{(2n+2)} \equiv 0$ and, therefore,

$$0 = e^{(2n+2)}(\xi) = f^{(2n+2)}(\xi) - \frac{f(x) - p(x)}{(\omega_{n+1}(x))^2} (2n+2)!,$$

which proves the result. $\square$

9.5 Complex Polynomial Interpolation

We now discuss the interpolation of functions that are holomorphic in a region of the complex plane. In doing so, we gain some further insight into the curious example provided by Runge; see Example 9.4.

https://doi.org/10.1017/9781108942607.013 Published online by Cambridge University Press

9.5.1 Some Facts of Complex Analysis

In what follows, we will make use of some tools from complex analysis to make the discussion rigorous, but we supply these as needed. For further results, see, for example, [3, 24, 35, 60, 77].

Definition 9.22 (holomorphic function)**.** Let $D \subset \mathbb{C}$ be a simply connected, open region and $g\colon D \to \mathbb{C}$. We say that g is **complex differentiable** at $a \in D$ if the limit

$$\lim_{z \to a} \frac{g(z) - g(a)}{z - a}$$

exists. In this case, we denote by $g'(a)$ the value of this limit and call it the **derivative** of g at a. If g is complex differentiable at every point of D, then we say that g is **holomorphic** in D.

As the reader may be aware, being complex differentiable is a very strong condition. It implies, in particular, that if g is complex differentiable in a neighborhood of a point $a \in \mathbb{C}$, then it is infinitely differentiable (again in the complex sense) at this point. Moreover, this also implies that this function is *analytic*, i.e., its power series representation

$$\sum_{k=0}^{\infty} g^{(k)}(a) \frac{(z-a)^k}{k!}$$

converges uniformly in a neighborhood of a. This can be proved with the help of Cauchy's integral theorem, which we discuss next. We will first need to define path integrals, and for that we need several definitions.

Definition 9.23 (path)**.** Let $[a, b] \subset \mathbb{R}$ be a compact interval. A function $\gamma \in C([a, b]; \mathbb{C})$ is called a **path** or **curve**. We say that a path is **simple** or a **Jordan path**[7] if

$$\gamma(t_1) \neq \gamma(t_2), \quad \forall t_1, t_2 \in (a, b).$$

We say that a path is **closed**, or a **contour**, if $\gamma(a) = \gamma(b)$.

A powerful result known as the *Jordan Curve Theorem* (see [39, Section 2.B]), states that a simple contour divides the complex plane $\mathbb{C}$ into two parts: an "interior" region bounded by the path and an "exterior," so that every other path that goes from a point of the interior to the exterior must intersect the given contour. With the help of this, we can talk about points *inside* or *outside* a simple closed path.

Definition 9.24 (smooth path)**.** Let $\gamma \in C([a, b]; \mathbb{C})$ be a path. We say that this path is **piecewise smooth** if there is $S \subset [a, b]$ of finite cardinality such that $\gamma \in C^1([a, b] \backslash S; \mathbb{C})$.

[7] Named in honor of the French mathematician Marie Ennemond Camille Jordan (1838–1922).

https://doi.org/10.1017/9781108942607.013 Published online by Cambridge University Press

For piecewise smooth paths, the derivative exists at all, but possibly a finite number of, points. Thus, we can define the orientation of a path. This will be mostly used for contours, so we only define them this way.

Definition 9.25 (orientation)**.** Let $\gamma \in C([a,b];\mathbb{C})$ be a simple, piecewise smooth contour and let $\Gamma = \gamma([a,b])$. Denote by $D \subset \mathbb{C}$ the collection of points inside γ, i.e., $\partial D = \Gamma$. Let $\boldsymbol{n}_D(t) = [n_1(t), n_2(t)]^\intercal$, $t \in [a,b]$ be the unit exterior normal to D at the point $\gamma(t)$. We say that γ is **traversed counterclockwise** if

$$\det \begin{bmatrix} n_1(t) & n_2(t) \\ \Re\gamma'(t) & \Im\gamma'(t) \end{bmatrix} > 0$$

for all $t \in [a,b]$.

Finally, we can define integrals over a path.

Definition 9.26 (path integral)**.** Let $\gamma \in C([a,b];\mathbb{C})$ be a simple, piecewise smooth path; and let $\Gamma = \gamma([a,b])$. For $g \colon \Gamma \to \mathbb{C}$, the **path integral** of g over γ is defined as

$$\int_\gamma g(z)\mathrm{d}z = \int_a^b g(\gamma(t))\frac{\mathrm{d}\gamma(t)}{\mathrm{d}t}\,\mathrm{d}t.$$

It turns out that values of a holomorphic function can be obtained using path integrals.

Theorem 9.27 (Cauchy Integral Theorem[8])**.** *Suppose that $D \subset \mathbb{C}$ is a simply connected, open region and $g \colon D \to \mathbb{C}$ is holomorphic. Suppose that $\gamma \colon [0, 2\pi] \to D$ is a simple closed contour traversed counterclockwise and the point $z_0 \in D$ is inside γ. Then*

$$g(z_0) = \frac{1}{2\pi\mathrm{i}}\int_\gamma \frac{g(z)}{z - z_0}\,\mathrm{d}z.$$

Let us now briefly discuss what can be said about functions that are holomorphic except at an isolated point.

Definition 9.28 (Laurent expansion[9])**.** Suppose that $D \subset \mathbb{C}$ is a simply connected, open region and $z_0 \in D$. Assume that $g \colon D\backslash\{z_0\} \to \mathbb{C}$ is holomorphic, i.e., g is holomorphic on D, except for the isolated singularity at z_0. In this setting, g admits the **Laurent expansion**

$$\begin{aligned} g(z) &= \sum_{k=1}^{\infty} \frac{b_k}{(z - z_0)^k} + \sum_{k=0}^{\infty} a_k (z - z_0)^k \\ &= \cdots + \frac{b_2}{(z - z_0)^2} + \frac{b_1}{(z - z_0)} + a_0 + a_1(z - z_0) + a_2(z - z_0)^2 + \cdots, \end{aligned}$$

where $b_j \in \mathbb{C}$, $j = 1, 2, \ldots$, (not all of which are zero) and $a_j \in \mathbb{C}$, $j = 0, 1, \ldots$.

[8] Named in honor of the French mathematician Augustin-Louis Cauchy (1789–1857).
[9] Named in honor of the French mathematician and engineer Pierre Alphonse Laurent (1813–1854).

https://doi.org/10.1017/9781108942607.013 Published online by Cambridge University Press

Points like z_0 in the previous definition are called isolated singularities of a function.

Definition 9.29 (isolated singularity)**.** Suppose that $D \subset \mathbb{C}$ is a simply connected, open region and $z_0 \in D$. Assume that $g\colon D\backslash\{z_0\} \to \mathbb{C}$ is holomorphic. Then we call z_0 an **isolated singularity** of g. Moreover, we say that z_0 is a **removable singularity** if the limit

$$\lim_{z \to z_0} g(z) \tag{9.17}$$

exists and it is finite. If the limit (9.17) exists but equals ∞, then we say that the point z_0 is a **pole**. Finally, if (9.17) does not exist, we call z_0 an **essential singularity**.

It turns out that poles and Laurent expansions are closely related.

Definition 9.30 (poles)**.** Suppose that $D \subset \mathbb{C}$ is a simply connected, open region and $z_0 \in D$. Assume that $g\colon D\backslash\{z_0\} \to \mathbb{C}$ is holomorphic and that z_0 is a pole. Let g admit a Laurent expansion. If $b_1 \neq 0$, and $b_j = 0$ for $j > 1$, then we say that g has a **simple pole** at the singularity $z = z_0$. Similarly, if there is $k \in \mathbb{N}$ such that $b_k \neq 0$, but $b_j = 0$ for $j > k$, then we say that g has a **pole of degree** k at the singularity $z = z_0$. Finally, the **residue of** g **at** z_0, denoted $\operatorname{Res}(g, z_0)$, is the coefficient b_1, i.e.,

$$\operatorname{Res}(g, z_0) = b_1.$$

We comment that a function that is holomorphic in an open region except at a finite number of poles is called a *meromorphic* in this region.

Theorem 9.31 (Residue Theorem)**.** *Let $D \subset \mathbb{C}$ be a simply connected, open, bounded region and $Z = \{z_i\}_{i=0}^n \subset D$ be a set of $n+1$ distinct points. Assume that $g\colon D\backslash Z \to \mathbb{C}$ is holomorphic, i.e., g is holomorphic on D, except for isolated singularities at $z_0, \dots, z_n$. Suppose that $\gamma\colon [0, 2\pi] \to D$ is a simple closed contour traversed counterclockwise and the points of Z are inside γ. Then*

$$\int_\gamma g(z)\,\mathrm{d}z = 2\pi\mathrm{i} \sum_{i=0}^n \operatorname{Res}(g, z_i),$$

where $\operatorname{Res}(g, z_i)$ *is the residue of g at z_i.*

The following result is useful for calculating residues.

Proposition 9.32 (computation of residues)**.** *Suppose that $D \subset \mathbb{C}$ is a simply connected, open region and $z_0 \in D$. Assume that $g, h\colon D \to \mathbb{C}$ are holomorphic, $g(z_0) \neq 0$, $h(z_0) = 0$, and $h'(z_0) \neq 0$. Then the function $f(z) = \frac{g(z)}{h(z)}$ is holomorphic in $D\backslash\{z_0\}$ and has a simple pole at z_0. Furthermore,*

$$\operatorname{Res}(f, z_0) = \frac{g(z_0)}{h'(z_0)}.$$

We conclude our slight detour into complex analysis by stating a result that will be used in Chapter 21 and is commonly referred to as the *Maximum Modulus Principle*.

https://doi.org/10.1017/9781108942607.013 Published online by Cambridge University Press

Theorem 9.33 (Maximum Modulus Principle)**.** *Let $D \subset \mathbb{C}$ be an open and connected region and $g\colon D \to \mathbb{C}$ be holomorphic. If there is a point $z_0 \in D$ such that $z \mapsto |g(z)|$ attains a (local) maximum at z_0, then g is constant.*

In other words, if $g \in C(\bar{D}; \mathbb{C})$ is holomorphic in D, then it attains its maximum at some point of the boundary ∂D.

9.5.2 Lagrange Interpolation

We are now ready to study the problem of Lagrange interpolation in the complex case. The setting is similar to the real case: we let $D \subset \mathbb{C}$ be a simply connected, open region; $n \in \mathbb{N}$; and the nodal set is $Z = \{z_i\}_{i=0}^n \subset D$, where all the nodes are distinct. Given $f\colon D \to \mathbb{C}$, we need to find a (complex-valued) polynomial $p \in \mathbb{P}_n(\mathbb{C})$ such that

$$p(z_i) = f(z_i), \quad i = 0, \dots, n. \tag{9.18}$$

The following result is obtained by small variations of the real case.

Theorem 9.34 (Lagrange interpolation error)**.** *Suppose that $D \subset \mathbb{C}$ is a simply connected, open region, $f\colon D \to \mathbb{C}$, and $Z = \{z_i\}_{i=0}^n \subset D$ is a set of $n+1 \in \mathbb{N}$ distinct points. There exists a unique polynomial $p \in \mathbb{P}_n(\mathbb{C})$ that satisfies* (9.18)*. Suppose, in addition, that f is holomorphic, $\gamma\colon [0, 2\pi] \to D$ is a simple closed contour traversed counterclockwise, and the points of Z are inside γ. Then the error formula may be expressed as*

$$f(\zeta) - p(\zeta) = \frac{1}{2\pi \mathrm{i}} \int_\gamma \frac{f(z)}{z - \zeta} \prod_{i=0}^n \frac{\zeta - z_i}{z - z_i} \, \mathrm{d}z \tag{9.19}$$

for all ζ inside γ, but $\zeta \notin Z$.

Proof. The existence and uniqueness of the interpolating polynomial $p \in \mathbb{P}_n(\mathbb{C})$ follows from similar arguments to those used for the real case. One can either appeal to a Vandermonde-type construction, in which case the Vandermonde matrix is invertible; or one can appeal to a Lagrange-type construction. In fact, the complex analogue of the Lagrange interpolation formula (9.10) is valid, i.e., the interpolating polynomial is

$$p(z) = \sum_{i=0}^n f(z_i) \frac{\omega_{n+1}(z)}{(z - z_i)\omega_{n+1}'(z_i)}, \tag{9.20}$$

where

$$\omega_{n+1}(z) = \prod_{i=0}^n (z - z_i).$$

To get the error formula, consider now the function

$$g(z) = \frac{f(z)}{z - \zeta} \prod_{i=0}^n \frac{\zeta - z_i}{z - z_i} = \frac{f(z)}{z - \zeta} \cdot \frac{\omega_{n+1}(\zeta)}{\omega_{n+1}(z)},$$

https://doi.org/10.1017/9781108942607.013 Published online by Cambridge University Press

which is the integrand in (9.19). Observe that the function g is meromorphic on D, i.e., it is holomorpic on D except at the isolated singularities $Z \cup \{\zeta\}$. Thus, by the residue Theorem 9.31, the integral

$$\frac{1}{2\pi\mathrm{i}} \int_\gamma g(z)\mathrm{d}z = \frac{1}{2\pi\mathrm{i}} \int_\gamma \frac{f(z)}{z-\zeta} \prod_{i=0}^{n} \frac{\zeta - z_i}{z - z_i} \mathrm{d}z$$

is the sum of the residues of g contained inside of γ. Using Proposition 9.32, we find

$$\mathrm{Res}(g, z_i) = \frac{f(z_i)}{z_i - \zeta} \cdot \frac{\omega_{n+1}(\zeta)}{\omega'_{n+1}(z_i)}, \quad i = 0, \ldots, n,$$

and

$$\mathrm{Res}(g, \zeta) = f(\zeta) \cdot \frac{\omega_{n+1}(\zeta)}{\omega_{n+1}(\zeta)} = f(\zeta).$$

Thus,

$$\frac{1}{2\pi\mathrm{i}} \int_\gamma \frac{f(z)}{z-\zeta} \prod_{i=0}^{n} \frac{\zeta - z_i}{z - z_i} \mathrm{d}z = f(\zeta) - \sum_{i=0}^{n} \frac{f(z_i)}{\zeta - z_i} \cdot \frac{\omega_{n+1}(\zeta)}{\omega'_{n+1}(z_i)} = f(\zeta) - p(\zeta),$$

where in the last step we used (9.20). $\square$

This last result provides a useful means to estimate errors.

Theorem 9.35 (error estimate)**.** *Let $[a, b] \subset \mathbb{R}$ be a compact interval and $\varepsilon > 0$ be fixed. Define*

$$r = b - a + \varepsilon, \quad C_\varepsilon = \{z \in \mathbb{C} \mid \mathrm{dist}(z, [a, b]) = r\}.$$

Suppose that $\gamma\colon [0, 2\pi] \to C_\varepsilon$ is the simple closed contour traversing the set C_ε in a counterclockwise fashion. Define D_ε as the open, simply connected set of all points inside γ. Suppose that $f\colon \overline{D_\varepsilon} \to \mathbb{C}$ is holomorphic on D_ε and there is $M > 0$ such that

$$|f(z)| \le M, \quad \forall x \in C_\varepsilon = \partial D_\varepsilon.$$

Let $n \in \mathbb{N}_0$, $X_n = \{x_i\}_{i=0}^n$ be a nodal set in the interval $[a, b]$, and $p_n \in \mathbb{P}_n(\mathbb{C})$ be the interpolating polynomial of f subordinate to the nodal set X_n. Then, for all $x \in [a, b] \backslash X_n$,

$$|f(x) - p_n(x)| \le \frac{(b - a + \pi r)M}{\pi} \left(\frac{b-a}{r}\right)^{n+1}.$$

Consequently, $p_n \to f$ uniformly on $[a, b]$, as $n \to \infty$, no matter how the nodes in X_n are chosen.

Proof. Observe that the length of γ is precisely $2(b - a) + 2\pi r$ and apply the last theorem. The details are left to the reader as an exercise; see Problem 9.10. $\square$

https://doi.org/10.1017/9781108942607.013 Published online by Cambridge University Press

Example 9.5 Let us revisit the Runge phenomenon. The function

$$f(x) = \frac{1}{1+25x^2}, \quad x \in [-1, 1],$$

which we have previously considered, and the uniformly spaced nodal set

$$X_n = \left\{-1 + \frac{2i}{n}\right\}_{i=0}^{n} \subset [-1, 1], \quad n \in \mathbb{N}.$$

As we mentioned in Example 9.4, if $p_n \in \mathbb{P}_n$ is the Lagrange interpolating polynomial of f subordinate to X_n, then one can show that $\|f - p_n\|_{L^\infty(-1,1)} \to \infty$, as $n \to \infty$. In light of Theorem 9.35, why do we fail to obtain uniform convergence? Clearly, one or more of the hypotheses of Theorem 9.35 must fail. In fact, as a complex function, f is not holomorphic, for any $\varepsilon > 0$, in the region D_ε defined in the hypotheses of Theorem 9.35. In fact, writing

$$f(z) = \frac{1}{1+25z^2} = \frac{1}{(1+5\mathrm{i}z)(1-5\mathrm{i}z)},$$

we observe that f has isolated singularities and simple poles at $z = \pm\frac{\mathrm{i}}{5}$. This explains is why Theorem 9.35 cannot be applied to guarantee uniform convergence in $[-1, 1]$ with the uniformly spaced nodes X_n.

9.6 Divided Differences and the Newton Construction

In this section, we give another method for computing interpolating polynomials, using the so-called *Newton construction*. This method is based on an alternate basis for $\mathbb{P}_n$, namely the Newton basis.

Definition 9.36 (Newton basis[10]). Let $n \in \mathbb{N}_0$ and $X = \{x_i\}_{i=0}^n$ be a nodal set in the compact interval $[a, b] \subset \mathbb{R}$. The polynomial set $B_n = \{\omega_j\}_{j=0}^n$, defined as $\omega_0 \equiv 1$ and, if $n \geq 1$,

$$\omega_j(x) = \prod_{k=0}^{j-1}(x - x_k), \quad j = 1, \ldots, n$$

is called the **Newton basis**. The polynomial $\omega_j \in \mathbb{P}_j$ is called the **nodal polynomial of order** j with respect to X.

The following result shows that B_n is indeed a basis.

Proposition 9.37 (basis). *For any compact interval $[a, b] \subset \mathbb{R}$, any $n \in \mathbb{N}_0$, and all nodal sets X in $[a, b]$, the Newton basis B_n is a basis of $\mathbb{P}_n$.*

Proof. See Problem 9.16. □

[10] Named in honor of the British mathematician, physicist, astronomer, theologian, and natural philosopher Sir Isaac Newton (1642–1726/27).

https://doi.org/10.1017/9781108942607.013 Published online by Cambridge University Press

Since, as the previous result shows, B_n is a basis we can expand any interpolating polynomial with respect to it. It turns out that the coefficients in this basis expansion carry very useful information.

Definition 9.38 (divided differences)**.** Suppose that $n \in \mathbb{N}_0$ and $X = \{x_i\}_{i=0}^n$ is a nodal set in the compact interval $[a, b]$. Let $f \in C([a, b])$ and $p_n \in \mathbb{P}_n$ be the (unique) interpolating polynomial of f subordinate to X. Set

$$p_n(x) = \sum_{k=0}^{n} a_k \omega_k(x). \tag{9.21}$$

The coefficient,

$$a_k = f[x_0, x_1, \ldots, x_k], \quad k = 0, \ldots, n,$$

in the expansion with respect to the Newton basis $B_n = \{\omega_j\}_{j=0}^n$ is called the ***k*th order divided difference**.

The choice of the name, divided difference, will be clear only later when we discuss the properties of these objects.

Example 9.6 The Newton form of the interpolating polynomial lends itself to a very efficient algorithm for computation. Suppose, for example, that $n = 4$ and we wish to evaluate p_4 in (9.21) at x. Then, once the coefficients $\{a_k\}_{k=0}^4$ are known, we can express $p_4(x)$ as

$$p_4(x) = a_0 + (x - x_0)\{a_1 + (x - x_1)[a_2 + (x - x_2)\{a_3 + (x - x_3)[a_4]\}]\}.$$

This method of evaluation is known as *Horner's method*.[11]

While this representation is, in theory, completely equivalent to (9.21), its evaluation is much cheaper. In fact, it can be shown that evaluating a polynomial of degree n using Horner's method only requires $\mathcal{O}(n)$ arithmetic operations, which is optimal.

Proposition 9.39 (recursion)**.** *Suppose that $n \in \mathbb{N}_0$ and $X_n = \{x_i\}_{i=0}^n$ is a nodal set in the compact interval $[a, b] \subset \mathbb{R}$. Let $f \in C([a, b])$ and, for $k \in \{0, \ldots, n\}$, $p_k \in \mathbb{P}_k$ be the (unique) interpolating polynomial of f subordinate to $X_k = \{x_i\}_{i=0}^k$. Then*

$$p_k(x) = p_{k-1}(x) + b_k \omega_k, \quad k = 1, \ldots, n, \tag{9.22}$$

holds if and only if

$$b_k = \frac{f(x_k) - p_{k-1}(x_k)}{\omega_k(x_k)}, \quad k = 1, \ldots, n. \tag{9.23}$$

[11] Named in honor of the Britsh mathematician William George Horner (1786–1837).

https://doi.org/10.1017/9781108942607.013 Published online by Cambridge University Press

As a consequence, it follows that, using the Newton basis representation (9.21), *we have* $f[x_0] = f(x_0)$ *and*

$$f[x_0, x_1, \dots, x_k] = a_k = b_k = \frac{f(x_k) - p_{k-1}(x_k)}{\omega_k(x_k)}, \quad k = 1, \dots, n.$$

Proof. We will sketch the proof and leave the details to the reader as an exercise; see Problem 9.17. Suppose that $1 \le k \le n$. Then

$$p_{k-1}(x_j) = f(x_j), \quad j = 0, \dots, k-1,$$

and

$$p_k(x_j) = f(x_j), \quad j = 0, \dots, k.$$

In general, however, $p_k(x_k) \neq p_{k-1}(x_k)$.

($\Longrightarrow$) Suppose that (9.22) holds. Then

$$f(x_j) = p_k(x_j) = p_{k-1}(x_j) + b_k\omega_k(x_j), \quad j = 0, \dots, k.$$

This last equation holds identically for $j = 0, \dots, k-1$, since $\omega_k(x_j) = 0$, for $j = 0, \dots, k-1$. We get new information only from $j = k$. Thus,

$$f(x_k) = p_k(x_k) = p_{k-1}(x_k) + b_k\omega_k(x_k),$$

which implies, since $\omega_k(x_k) \neq 0$, that

$$b_k = \frac{f(x_k) - p_{k-1}(x_k)}{\omega_k(x_k)}.$$

($\Longleftarrow$) Suppose that (9.23) holds. Then, it is easy to see that (9.22) holds. $\square$

This last result shows that the construction of the interpolating polynomial $p_n \in \mathbb{P}_n$ can be done in a particular order. First, set $p_0 \equiv f(x_0)$. Then p_1, the interpolant using the points x_0 and x_1, is constructed from p_0 by computing b_1. p_2, the interpolant using the points x_0, x_1, and x_2, is constructed from p_1 by computing b_2, and so on. The construction proceeds according to our labeling of the nodes: $x_0, x_1, \dots, x_n$. But, as we have noted, there is no special numbering assigned to X_n. Thus, we could choose another order for our construction; for example, reverse order construction: $x_n, x_{n-1}, \dots, x_0$.

Proposition 9.40 (divided differences)**.** *Suppose that* $n \in \mathbb{N}_0$, $X_n = \{x_i\}_{i=0}^n$ *is a nodal set in the compact interval* $[a, b] \subset \mathbb{R}$, *and* $f\colon [a, b] \to \mathbb{R}$. *Then*

$$a_n = f[x_0, x_1, \dots, x_n] = \sum_{j=0}^{n} \frac{f(x_j)}{\prod_{\substack{k=0 \\ k\neq j}}^{n} (x_j - x_k)}. \tag{9.24}$$

Proof. This follows by equating the Lagrange, Newton, and canonical forms of the interpolating polynomial. Suppose that $p_n \in \mathbb{P}_n$ is the unique interpolating polynomial of f subordinate to $X_n = \{x_i\}_{i=0}^n$. Then

$$p_n(x) = \sum_{j=0}^{n} L_j(x) f(x_j) = \sum_{k=0}^{n} a_k\omega_k(x) = \sum_{j=0}^{n} c_j x^j,$$

https://doi.org/10.1017/9781108942607.013 Published online by Cambridge University Press

where $L_j \in \mathbb{P}_n$ is the jth Lagrange basis element; see (9.3). Let us figure out what are the coefficients of x^n when the Lagrange and Newton forms are expanded to canonical form. For the Newton form, clearly the coefficient of x^n is

$$c_n = a_n = f[x_0, x_1, \dots, x_n].$$

On the other hand, for the Lagrange form,

$$c_n = \sum_{j=0}^{n} \frac{f(x_j)}{\prod_{\substack{k=0\\k\neq j}}^{n}(x_j - x_k)}.$$ □

Corollary 9.41 (invariance)**.** *Suppose that $n \in \mathbb{N}_0$, $X_n = \{x_i\}_{i=0}^n$ is a nodal set in the compact interval $[a, b] \subset \mathbb{R}$, and $f\colon [a, b] \to \mathbb{R}$. Assume that $(i_0, i_1, \dots, i_n)$ is a permutation of $(0, 1, \dots, n)$. Define $\tilde{\omega}_0 \equiv 1$ and, if $n \geq 1$,*

$$\tilde{\omega}_k(x) = \prod_{j=0}^{k-1}(x - x_{i_j}), \quad k = 1, \dots, n.$$

Suppose that $p_n \in \mathbb{P}_n$ is the interpolating polynomial of f subordinate to X_n, so that

$$p_n = \sum_{j=0}^{n} a_j\omega_j(x) = \sum_{j=0}^{n} \tilde{a}_j\tilde{\omega}_j(x). \tag{9.25}$$

Then

$$a_n = \tilde{a}_n$$

and

$$f[x_0, x_1, \dots, x_n] = f\left[x_{i_0}, x_{i_1}, \dots, x_{i_n}\right].$$

Proof. This follows directly from the right-hand side of (9.24), since the sum and product can rearranged arbitrarily. Alternately, one can simply expand the representations in (9.25) and consider the coefficient of x^n. □

Remark 9.42 (invariance)**.** Let us consider what the expansions in (9.25) represent. $\sum_{j=0}^n a_j\omega_j$ represents the construction of p_n in the order $x_0, x_1, \dots, x_n$, whereas $\sum_{j=0}^n \tilde{a}_j\tilde{\omega}_j$ represents the construction of p_n in the permuted order $x_{i_0}, x_{i_1}, \dots, x_{i_n}$. Though the order of the constructions may differ, the resulting interpolating polynomial must be the same if all of the same interpolation points are ultimately used.

Theorem 9.43 (divided difference formula)**.** *Suppose that $n \in \mathbb{N}$, $X_n = \{x_i\}_{i=0}^n$ is a nodal set in the compact interval $[a, b] \subset \mathbb{R}$, and $f\colon [a, b] \to \mathbb{R}$. Assume that $p_n \in \mathbb{P}_n$ is the interpolating polynomial of f subordinate to X_n. When constructed in the order $x_0, x_1, \dots, x_n$, let us write, as in* (9.21),

$$p_n(x) = \sum_{j=0}^{n} a_j\omega_j(x).$$

https://doi.org/10.1017/9781108942607.013 Published online by Cambridge University Press

When p_n is constructed in the reverse order $x_n, x_{n-1}, \ldots, x_0$, let us write

$$p_n(x) = \sum_{j=0}^{n} \hat{a}_j \hat{\omega}_j(x),$$

where $\hat{\omega}_0 \equiv 1$ and, if $n \geq 1$,

$$\hat{\omega}_j(x) = \prod_{k=0}^{j-1} (x - x_{n-k}), \quad j = 1, \ldots, n,$$

and where, by definition,

$$\hat{a}_j = f[x_n, x_{n-1}, \ldots, x_{n-j}], \quad j = 0, \ldots, n.$$

Then it follows that

$$a_n = \hat{a}_n = \frac{a_{n-1} - \hat{a}_{n-1}}{x_0 - x_n},$$

which, in turn, implies that

$$f[x_0, x_1, \ldots, x_n] = \frac{f[x_0, x_1, \ldots, x_{n-1}] - f[x_1, x_2, \ldots, x_n]}{x_0 - x_n}, \quad \forall n \in \mathbb{N}. \tag{9.26}$$

Proof. Clearly, for all $x \in \mathbb{R}$,

$$\Delta(x) = \sum_{j=0}^{n} a_j \omega_j(x) - \sum_{j=0}^{n} \hat{a}_j \hat{\omega}_j(x) = 0$$

and, by Corollary 9.41, $a_n = \hat{a}_n$. Using $a_n = \hat{a}_n$, the function Δ may be re-expressed as

$$0 = \Delta(x) = a_n[(x - x_0) - (x - x_n)](x - x_1)\cdots(x - x_{n-1}) + (a_{n-1} - \hat{a}_{n-1})x^{n-1} + p_{n-2}(x),$$

where $p_{n-2} \in \mathbb{P}_{n-2}$. The coefficient of x^{n-1} in the polynomial Δ is precisely

$$(x_n - x_0)a_n + (a_{n-1} - \hat{a}_{n-1}),$$

which must be zero since $\Delta \equiv 0$:

$$(x_n - x_0)a_n + (a_{n-1} - \hat{a}_{n-1}) = 0.$$

Equation 9.26 then follows immediately from the last equation and another application of Corollary 9.41, which guarantees that

$$\hat{a}_{n-1} = f[x_n, x_{n-1}, \ldots, x_1] = f[x_1, x_2, \ldots, x_n]. \qquad \square$$

We have the following, as a simple consequence of the last result.

Corollary 9.44 (divided difference formula)**.** *Let $n \in \mathbb{N}$. Suppose $X = \{x_j, \ldots, x_{j+n}\}$ is a nodal set in the compact interval $[a, b] \subset \mathbb{R}$. If $f\colon [a, b] \to \mathbb{R}$, then*

$$f[x_j, \ldots, x_{j+n}] = \frac{f[x_{j+1}, \ldots, x_{j+n}] - f[x_j, \ldots, x_{j+n-1}]}{x_{j+n} - x_j}. \tag{9.27}$$

https://doi.org/10.1017/9781108942607.013 Published online by Cambridge University Press

Proof. See Problem 9.18. □

x_i	$f[x_i]$	$f[x_i, x_{i+1}]$	$f[x_i, x_{i+1}, x_{i+2}]$	$f[x_i, \ldots, x_{i+3}]$	$f[x_i, \ldots, x_{i+4}]$
x_0	$f[x_0]$				
x_1	$f[x_1]$	$f[x_0, x_1]$			
x_2	$f[x_2]$	$f[x_1, x_2]$	$f[x_0, x_1, x_2]$		
x_3	$f[x_3]$	$f[x_2, x_3]$	$f[x_1, x_2, x_3]$	$f[x_0, x_1, x_2, x_3]$	
x_4	$f[x_4]$	$f[x_3, x_4]$	$f[x_2, x_3, x_4]$	$f[x_1, x_2, x_3, x_4]$	$f[x_0, x_1, x_2, x_3, x_4]$

Table 9.1 A table of divided differences for computing up to quartic (fourth degree) interpolating polynomials.

Example 9.7 Suppose that $[a, b] \subset \mathbb{R}$ is a compact interval, $f : [a, b] \to \mathbb{R}$, and we wish to compute a quartic interpolating polynomial for f with the nodes $X_4 = \{x_0, \ldots, x_4\} \subset [a, b]$. Table 9.1 depicts the construction of the divided differences, where the order is assumed to proceed according to the numbering of the nodes: $0, 1, \ldots, 4$. The construction can be accomplished by recursively computing all of the divided differences in the table. For example, using 9.27, we compute

$$f[x_1, x_2, x_3] = \frac{f[x_2, x_3] - f[x_1, x_2]}{x_3 - x_1}.$$

Note that all of the values in the table are needed to compute the last entry, namely

$$f[x_0, x_1, x_2, x_3, x_4].$$

When all of the values are computed, the interpolating polynomial is

$$\begin{aligned} p_4(x) = f[x_0] & \\ &+ (x - x_0)f[x_0, x_1] \\ &+ (x - x_0)(x - x_1)f[x_0, x_1, x_2] \\ &+ (x - x_0)(x - x_1)(x - x_2)f[x_0, x_1, x_2, x_3] \\ &+ (x - x_0)(x - x_1)(x - x_2)(x - x_3)f[x_0, x_1, x_2, x_3, x_4]. \end{aligned}$$

If, on the other hand, one wished to stop with the cubic interpolating polynomial, for example, the last row of Table 9.1 need not be calculated. One will obtain

$$\begin{aligned} p_3(x) = f[x_0] & \\ &+ (x - x_0)f[x_0, x_1] \\ &+ (x - x_0)(x - x_1)f[x_0, x_1, x_2] \\ &+ (x - x_0)(x - x_1)(x - x_2)f[x_0, x_1, x_2, x_3]. \end{aligned}$$

Example 9.8 As a concrete example, consider the data in Table 9.2. There are five interpolation nodes, which allows us to construct an interpolating polynomial of

https://doi.org/10.1017/9781108942607.013 Published online by Cambridge University Press

degree up to four (4) for the function of interest, f. The interpolating polynomials of degrees three (3) and four (4) are plotted in Figure 9.4. Specifically,

$$\begin{aligned} p_3(x) = \mathcal{I}_{X_3}[f](x) = 5 & \\ &+ (x-0.1)(-30) \\ &+ (x-0.1)(x-0.3)\left(\frac{400}{3}\right) \\ &+ (x-0.1)(x-0.3)(x-0.4)\left(-\frac{2125}{9}\right) \end{aligned}$$

and

$$p_4(x) = \mathcal{I}_{X_4}[f](x) = p_3(x) + (x-0.1)(x-0.3)(x-0.4)(x-0.7)\left(\frac{5125}{18}\right).$$

---

x_i	$f[x_i]$	$f[x_i, x_{i+1}]$	$f[x_i, x_{i+1}, x_{i+2}]$	$f[x_i, \dots, x_{i+3}]$	$f[x_i, \dots, x_{i+4}]$
0.1	5				
0.3	-1	-30			
0.4	0	10	$\frac{400}{3}$		
0.7	2	$\frac{20}{3}$	$-\frac{25}{3}$	$-\frac{2125}{9}$	
0.9	2	0	$-\frac{40}{3}$	$-\frac{25}{3}$	$\frac{5125}{18}$

Table 9.2 A table of divided differences for computing up to quartic (fourth degree) interpolating polynomials with real data. See Figure 9.4 for the interpolating polynomials of degrees three (3) and four (4) constructed from these data.

It turns out that divided differences are not only useful in practical computation, They can also be used to obtain an alternate error formula for interpolation.

Theorem 9.45 (Newton–Lagrange interpolation error)**.** *Suppose that $n \in \mathbb{N}$, $X = \{x_i\}_{i=0}^n$ is a nodal set in the compact interval $[a, b] \subset \mathbb{R}$, and $f\colon [a, b] \to \mathbb{R}$. Let $p_n \in \mathbb{P}_n$ be the interpolating polynomial of f subordinate to X. Then, for any $x \in [a, b]\backslash X$,*

$$f(x) - p_n(x) = \omega_{n+1}(x) f[x_0, x_1, \dots, x_n, x]. \tag{9.28}$$

Proof. Since $x \notin X$, then $X' = X \cup \{x\}$ is a nodal set in $[a, b]$. The unique interpolating polynomial of f subordinate to X', which we label $p_{n+1} \in \mathbb{P}_{n+1}$, is

$$p_{n+1}(t) = p_n(t) + \omega_{n+1}(t) f[x_0, x_1, \dots, x_n, x], \quad \forall t \in [a, b],$$

according to Proposition 9.39. Since $x \in [a, b]$ is an interpolation node,

$$f(x) = p_{n+1}(x)$$

and the result follows. □

Comparing the Lagrange and Newton error formulas, we get the following result.

https://doi.org/10.1017/9781108942607.013 Published online by Cambridge University Press

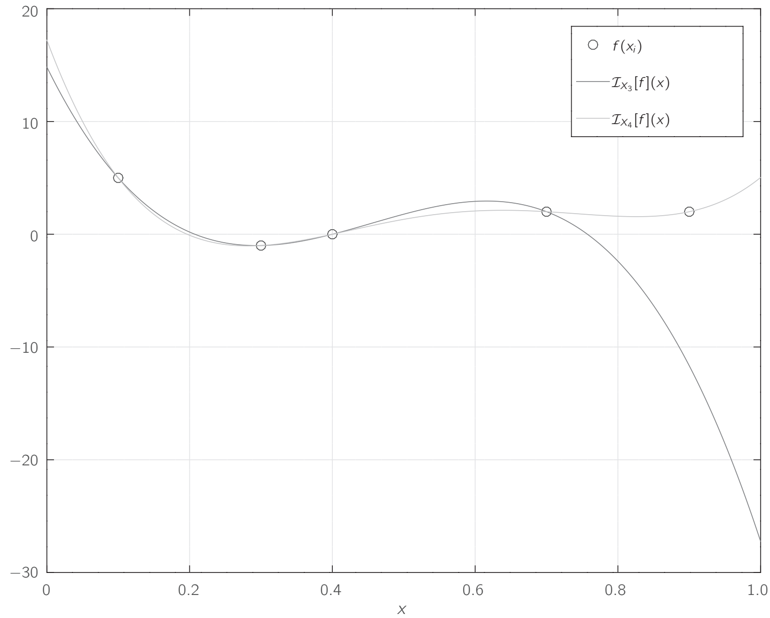

Figure 9.4 The third- and fourth-order interpolating polynomials constructed using Newton's divided differences and the data from Table 9.2.

Corollary 9.46 (divided differences). *Suppose that $n \in \mathbb{N}$, $X = \{x_i\}_{i=0}^n$ is a nodal set in the compact interval $[a, b] \subset \mathbb{R}$, and $f \in C^{n+1}([a, b])$. Let $x \in [a, b]\backslash X$ be arbitrary. Then there is a point $\xi = \xi(x_0, \dots, x_n, x) \in (a, b)$ with*

$$\min\{x_0, x_1, \dots, x_n, x\} < \xi < \max\{x_0, x_1, \dots, x_n, x\}$$

such that

$$f[x_0, x_1, \dots, x_n, x] = \frac{f^{(n+1)}(\xi)}{(n+1)!}.$$

Proof. Compare the error representations in 9.28 from Theorem 9.45 and 9.11 from Theorem 9.16. □

The following is just a cosmetic, but quite useful, reformulation of the last result.

Corollary 9.47 (divided differences). *Suppose that $n \in \mathbb{N}$ and $X = \{x_i\}_{i=0}^n$ is a nodal set in the compact interval $[a, b] \subset \mathbb{R}$. Let $f \in C^n([a, b])$. Then there is a point $\xi = \xi(x_0, \dots, x_n) \in (a, b)$ with*

$$\min\{x_0, x_1, \dots, x_n\} < \xi < \max\{x_0, x_1, \dots, x_n\}$$

https://doi.org/10.1017/9781108942607.013 Published online by Cambridge University Press

such that

$$f[x_0, x_1, \dots, x_n] = \frac{f^{(n)}(\xi)}{n!}.$$

Proof. See Problem 9.19. □

9.7 Extended Divided Differences

In this section, we give an alternate characterization of the divided difference function that will help us to extend its definition to the case where the points are not distinct, as well as to yield some other useful properties. First, we need a definition.

Definition 9.48 (n-simplex)**.** Suppose that $n \in \mathbb{N}$. The **canonical n-simplex** is the set

$$T_n = \{\boldsymbol{\tau} \in \mathbb{R}^n \mid \boldsymbol{\tau} \cdot \mathbf{1} \le 1,\ \boldsymbol{\tau} \cdot \boldsymbol{e}_i \ge 0,\ i = 1, \dots, n\}, \tag{9.29}$$

where we recall that $\{\boldsymbol{e}_j\}_{j=1}^n$ is the canonical basis of $\mathbb{R}^n$ and $\mathbf{1} = [1, \dots, 1]^\intercal \in \mathbb{R}^n$.

Lemma 9.49 (volume)**.** *Suppose that $n \in \mathbb{N}$. The n-dimensional volume of the canonical n-simplex is*

$$\operatorname{vol}(T_n) = \int_{T_n} \mathrm{d}^n\boldsymbol{\tau} = \frac{1}{n!}, \tag{9.30}$$

where $\mathrm{d}^n\boldsymbol{\tau} = \mathrm{d}\tau_1 \mathrm{d}\tau_2 \cdots \mathrm{d}\tau_n$.

Proof. The proof proceeds by induction. The crucial step is to show that, for $2 \le k \le n$,

$$\int_{T_k} \mathrm{d}^k\boldsymbol{\tau} = \int_{T_{k-1}} \int_{\tau_k=0}^{\tau_k = 1 - \sum_{j=1}^{k-1} \tau_j} \mathrm{d}\tau_k \mathrm{d}^{k-1}\boldsymbol{\tau}.$$

The details are left for the reader as an exercise; see Problem 9.22. □

Theorem 9.50 (Hermite–Genocchi Theorem[12])**.** *Suppose that $n \in \mathbb{N}$, $X = \{x_i\}_{i=0}^n$ is a nodal set in the compact interval $[a, b] \subset \mathbb{R}$, and $f \in C^n([a, b])$. We have*

$$f[x_0, x_1, \dots, x_n] = \int_{T_n} f^{(n)}(\tau_0 x_0 + \tau_1 x_1 + \cdots + \tau_n x_n) \mathrm{d}^n\boldsymbol{\tau}, \tag{9.31}$$

where $\boldsymbol{\tau} = [\tau_1, \dots, \tau_n]^\intercal \in T_n$, $\tau_0 = 1 - \boldsymbol{\tau} \cdot \mathbf{1}$, and $\mathrm{d}^n\boldsymbol{\tau} = \mathrm{d}\tau_1 \mathrm{d}\tau_2 \cdots \mathrm{d}\tau_n$.

Proof. The proof is by induction on n.

[12] Named in honor of the French mathematician Charles Hermite (1822–1901) and the Italian mathematician Angelo Genocchi (1817–1889).

https://doi.org/10.1017/9781108942607.013 Published online by Cambridge University Press

($n = 1$) Suppose that $f \in C^1([a,b])$. Clearly, $T_1 = [0,1]$ and

$$\begin{aligned}\int_{T_1} f'(\tau_0 x_0 + \tau_1 x_1)\mathrm{d}\tau_1 &= \int_0^1 f'(x_0 + \tau_1(x_1 - x_0))\mathrm{d}\tau_1 \\ &= \frac{1}{x_1 - x_0} f(x_0 + \tau_1(x_1 - x_0))\Big|_{\tau_1=0}^{\tau_1=1} \\ &= \frac{f(x_1) - f(x_0)}{x_1 - x_0} \\ &= f[x_0, x_1].\end{aligned}$$

($n = 2$) Suppose that $f \in C^2([a,b])$. The integration region T_2 is the triangle in $(\tau_1, \tau_2) \in \mathbb{R}^2$ with the vertices $(0,0)$, $(1,0)$, and $(0,1)$. Then

$$\begin{aligned}&\int_{T_2} f''(\tau_0 x_0 + \tau_1 x_1 + \tau_2 x_2)\mathrm{d}^2\boldsymbol{\tau} \\ &\quad= \int_0^1 \int_0^{1-\tau_1} f''(x_0 + \tau_1(x_1 - x_0) + \tau_2(x_2 - x_0))\mathrm{d}\tau_2 \mathrm{d}\tau_1 \\ &\quad= \frac{1}{x_2 - x_0} \int_0^1 [f'(x_0 + \tau_1(x_1 - x_0) + \tau_2(x_2 - x_0))]\Big|_{\tau_2=0}^{\tau_2=1-\tau_1} \mathrm{d}\tau_1 \\ &\quad= \frac{1}{x_2 - x_0}\left[\int_0^1 f'(x_2 + \tau_1(x_1 - x_2))\mathrm{d}\tau_1 - \int_0^1 f'(x_0 + \tau_1(x_1 - x_0))\mathrm{d}\tau_1\right] \\ &\quad= \frac{1}{x_2 - x_0}[f[x_2, x_1] - f[x_0, x_1]] \\ &\quad= \frac{1}{x_2 - x_0}[f[x_1, x_2] - f[x_0, x_1]] \\ &\quad= f[x_0, x_1, x_2].\end{aligned}$$

($n = k - 1$) Let $k \in \mathbb{N}$, $k \geq 2$ be arbitrary. Suppose that $f \in C^{k-1}([a,b])$. Assume that $\{z_0, z_1, \ldots, z_{k-1}\}$ is an arbitrary set of k distinct points in the compact interval $[a,b]$. For the induction hypothesis, let us assume that the formula holds for $n = k - 1$:

$$f[z_0, z_1, \ldots, z_{k-1}] = \int_{T_{k-1}} f^{(k-1)}(\tau_0 z_0 + \tau_1 z_1 + \cdots + \tau_{k-1} z_{k-1})\mathrm{d}^{k-1}\boldsymbol{\tau}.$$

($n = k$) Suppose that $f \in C^k([a,b])$. Using the induction hypothesis, Corollary 9.41, and 9.26 from Theorem 9.43, we have

https://doi.org/10.1017/9781108942607.013 Published online by Cambridge University Press

$$
\begin{aligned}
&\int_{T_k} f^{(k)}(\tau_0 x_0 + \tau_1 x_1 + \cdots + \tau_k x_k)\mathrm{d}^k\boldsymbol{\tau} \\
&\quad = \int_{T_{k-1}} \int_{\tau_k=0}^{\tau_k = 1-\sum_{j=1}^{k-1}\tau_j} f^{(k)}\left(x_0 + \sum_{j=1}^{k} \tau_j(x_j - x_0)\right) \mathrm{d}\tau_k \mathrm{d}^{k-1}\boldsymbol{\tau} \\
&\quad = \frac{1}{x_k - x_0} \int_{T_{k-1}} \left[f^{(k-1)}\left(x_0 + \sum_{j=1}^{k} \tau_j(x_j - x_0)\right)\right]\Bigg|_{\tau_n=0}^{\tau_n = 1-\sum_{j=1}^{k-1}\tau_j} \mathrm{d}^{k-1}\boldsymbol{\tau} \\
&\quad = \frac{1}{x_k - x_0} \left[\int_{T_{k-1}} f^{(k-1)}\left(x_k + \sum_{j=1}^{k-1} \tau_j(x_j - x_k)\right) \mathrm{d}^{k-1}\boldsymbol{\tau} \right. \\
&\qquad \left. - \int_{T_{k-1}} f^{(k-1)}\left(x_0 + \sum_{j=1}^{k-1} \tau_j(x_j - x_0)\right) \mathrm{d}^{k-1}\boldsymbol{\tau} \right] \\
&\quad = \frac{1}{x_k - x_0} \left[f[x_k, x_1, x_2, \ldots, x_{k-1}] - f[x_0, x_1, \ldots, x_{k-1}] \right] \\
&\quad = \frac{1}{x_k - x_0} \left[f[x_1, x_2, \ldots, x_k] - f[x_0, x_1, \ldots, x_{k-1}] \right] \\
&\quad = f[x_0, x_1, \ldots, x_k].
\end{aligned}
$$

The proof is complete. □

Let us now extend the definition of the divided difference function to the case where the nodes are not necessarily distinct.

Definition 9.51 (extended divided differences)**.** Suppose that $n \in \mathbb{N}$ and $[a,b] \subset \mathbb{R}$ is a compact interval. Set

$$
[a,b]^{n+1} = \left\{ [z_0, z_1, \ldots, z_n]^\intercal \in \mathbb{R}^{n+1} \;\middle|\; z_j \in [a,b],\ j = 0, \ldots, n \right\}.
$$

Let $f \in C^n([a,b])$. For every point $\boldsymbol{z} \in [a,b]^{n+1}$, we define

$$
f[\![\boldsymbol{z}]\!] = f[\![z_0, z_1, \ldots, z_n]\!] = \int_{T_n} f^{(n)}(\tau_0 z_0 + \tau_1 z_1 + \cdots + \tau_n z_n)\mathrm{d}^n\boldsymbol{\tau}, \tag{9.32}
$$

where T_n is the canonical n-simplex defined in (9.29), $\boldsymbol{\tau} = [\tau_1, \ldots, \tau_n]^\intercal \in T_n$, $\tau_0 = 1 - \boldsymbol{\tau} \cdot \mathbf{1}$, and $\mathrm{d}^n\boldsymbol{\tau} = \mathrm{d}\tau_1 \mathrm{d}\tau_2 \cdots \mathrm{d}\tau_n$. The function $f[\![\cdot]\!]\colon [a,b]^{n+1} \to \mathbb{R}$ is called the **extended divided difference of order** n.

Theorem 9.52 (continuity)**.** *Suppose that $n \in \mathbb{N}$, $[a,b]$ is a compact interval, and $f \in C^n([a,b])$. The extended finite difference function defined by* (9.32) *satisfies $f[\![\cdot]\!] \in C([a,b]^{n+1})$.*

Proof. Let $\boldsymbol{\tau} \in T_n$ and $\tau_0 = 1 - \boldsymbol{\tau} \cdot \mathbf{1}$. Define $\vec{\boldsymbol{\tau}} = [\tau_0, \boldsymbol{\tau}]^\intercal \in T_{n+1}$.

https://doi.org/10.1017/9781108942607.013 Published online by Cambridge University Press

Notice now that $f^{(n)} \in C([a,b])$ and, consequently, it is uniformly continuous on $[a,b]$. In other words, given $\varepsilon > 0$, there is a $\delta > 0$, such that, if x and y are any points in $[a,b]$ satisfying $|x-y| \le \delta$, then it follows that

$$\left| f^{(n)}(x) - f^{(n)}(y) \right| \le \varepsilon n!.$$

Now suppose that $\boldsymbol{z}_1, \boldsymbol{z}_2 \in [a,b]^{n+1}$ satisfy $\|\boldsymbol{z}_1 - \boldsymbol{z}_2\|_\infty \le \delta$. Then

$$|\vec{\boldsymbol{\tau}} \cdot \boldsymbol{z}_1 - \vec{\boldsymbol{\tau}} \cdot \boldsymbol{z}_2| = |\vec{\boldsymbol{\tau}} \cdot (\boldsymbol{z}_1 - \boldsymbol{z}_2)| \le \|\vec{\boldsymbol{\tau}}\|_1 \|\boldsymbol{z}_1 - \boldsymbol{z}_2\|_\infty \le \delta.$$

Therefore, uniformly with respect to $\vec{\boldsymbol{\tau}}$, we have

$$\left| f^{(n)}(\vec{\boldsymbol{\tau}} \cdot \boldsymbol{z}_1) - f^{(n)}(\vec{\boldsymbol{\tau}} \cdot \boldsymbol{z}_2) \right| \le \varepsilon n!,$$

as long as $\boldsymbol{z}_1, \boldsymbol{z}_2 \in [a,b]^{n+1}$ satisfy $\|\boldsymbol{z}_1 - \boldsymbol{z}_2\|_\infty \le \delta$. Then,

$$\begin{aligned}
|f[\![\boldsymbol{z}_1]\!] - f[\![\boldsymbol{z}_2]\!]| &= \left| \int_{T_n} f^{(n)}(\vec{\boldsymbol{\tau}} \cdot \boldsymbol{z}_1) \mathrm{d}^n\boldsymbol{\tau} - \int_{T_n} f^{(n)}(\vec{\boldsymbol{\tau}} \cdot \boldsymbol{z}_2) \mathrm{d}^n\boldsymbol{\tau} \right| \\
&\le \int_{T_n} \left| f^{(n)}(\vec{\boldsymbol{\tau}} \cdot \boldsymbol{z}_1) - f^{(n)}(\vec{\boldsymbol{\tau}} \cdot \boldsymbol{z}_2) \right| \mathrm{d}^n\boldsymbol{\tau} \\
&\le \int_{T_n} \varepsilon n! \mathrm{d}^n\boldsymbol{\tau} \\
&= \varepsilon.
\end{aligned}$$

This proves the result. □

We now prove a result for this extended definition, which is analogous to Corollary 9.47, but does not require the nodes to be distinct.

Proposition 9.53 (extended finite differences). *Let $n \in \mathbb{N}$, $[a,b] \subset \mathbb{R}$ be a compact interval, and $f \in C^n([a,b])$. Suppose that*

$$\{x_j\}_{j=0}^n \subset [a,b].$$

Then there is a point $\xi = \xi(x_0, \ldots, x_n) \in (a,b)$ with

$$\min\{x_0, x_1, \ldots, x_n\} < \xi < \max\{x_0, x_1, \ldots, x_n\}$$

such that

$$f[\![x_0, x_1, \ldots, x_n]\!] = \frac{f^{(n)}(\xi)}{n!}.$$

Proof. Set

$$\hat{a} = \min\{x_0, x_1, \ldots, x_n\}, \qquad \hat{b} = \max\{x_0, x_1, \ldots, x_n\},$$

and

$$m = \min_{\hat{a} \le x \le \hat{b}} f^{(n)}(x), \qquad M = \max_{\hat{a} \le x \le \hat{b}} f^{(n)}(x).$$

Then, using (9.30), we deduce

$$\frac{m}{n!} \le f[\![x_0, x_1, \ldots, x_n]\!] \le \frac{M}{n!},$$

https://doi.org/10.1017/9781108942607.013 Published online by Cambridge University Press

where we have again used $\mathrm{vol}(T_n) = \int_{T_n} \mathrm{d}^n \boldsymbol{\tau} = \frac{1}{n!}$. In other words,

$$m \leq f[\![x_0, x_1, \ldots, x_n]\!] n! \leq M.$$

By the Intermediate Value Theorem B.27, there is a point $\xi \in [\hat{a}, \hat{b}]$ such that

$$f^{(n)}(\xi) = f[\![x_0, x_1, \ldots, x_n]\!] n!. \qquad \square$$

The following is a simple consequence.

Corollary 9.54 (extended divided differences). *Suppose that $[a, b]$ is a compact interval and $f \in C^n([a, b]; \mathbb{R})$. Suppose that $x \in [a, b]$. Then*

$$f[\![\underbrace{x, x, \ldots, x}_{n+1}]\!] = \frac{f^{(n)}(x)}{n!}.$$

Proof. See Problem 9.23. $\square$

We showed that divided differences are invariant to permutations. Let us show that this is the case for extended finite differences as well. First, we need an approximation result.

Lemma 9.55 (density). *Let $n \in \mathbb{N}$. Suppose that $\boldsymbol{z} \in [a, b]^{n+1}$. There exists a sequence $\{\boldsymbol{z}_\ell\}_{\ell=1}^{\infty} \subset [a, b]^{n+1}$ with the following properties.*

1. *For each $\ell \in \mathbb{N}$ the coordinates of $\boldsymbol{z}_\ell$ are all distinct.*
2. *The sequence converges to $\boldsymbol{z}$ as $\ell \to \infty$.*

Proof. See Problem 9.24. $\square$

Proposition 9.56 (invariance). *Suppose that $n \in \mathbb{N}$, $[a, b] \subset \mathbb{R}$ is a compact interval, and $f \in C^n([a, b])$. Let*

$$\boldsymbol{z} = [z_0, z_1, \ldots, z_n]^\intercal \in [a, b]^{n+1}$$

be arbitrary, and $(i_0, i_1, \ldots, i_n)$ be a permutation of $(0, 1, \ldots, n)$. Then

$$f[\![z_{i_0}, z_{i_1}, \ldots, z_{i_n}]\!] = f[\![z_0, z_1, \ldots, z_n]\!].$$

Proof. If the coordinates of $\boldsymbol{z}$, i.e., $z_0, z_1, \ldots, z_n$, are all distinct, there is nothing to prove.

Assume, then, that the coordinates of $\boldsymbol{z}$ are not distinct. Let $\{\boldsymbol{z}_j\}_{j=1}^{\infty}$ be the sequence of Lemma 9.55. Since the divided difference function and the extended divided difference function agree for distinct points, by Corollary 9.41,

$$f[\![z_{\ell,i_0}, z_{\ell,i_1}, \ldots, z_{\ell,i_n}]\!] = f[\![z_{\ell,0}, z_{\ell,1}, \ldots, z_{\ell,n}]\!], \quad \forall \ell \in \mathbb{N},$$

where $\boldsymbol{z}_\ell = [z_{\ell,0}, z_{\ell,1}, \ldots, z_{\ell,n}]^\intercal$. Now Theorem 9.52 showed that $f[\![\cdot]\!]$ is continuous, i.e.,

$$\begin{aligned} f[\![z_{i_0}, z_{i_1}, \ldots, z_{i_n}]\!] &= \lim_{\ell \to \infty} f[\![z_{i_0,\ell}, z_{i_1,\ell}, \ldots, z_{i_n,\ell}]\!] \\ &= \lim_{\ell \to \infty} f[\![z_{0,\ell}, z_{1,\ell}, \ldots, z_{n,\ell}]\!] \\ &= f[\![z_0, z_1, \ldots, z_n]\!], \end{aligned}$$

and the claimed invariance follows. $\square$

https://doi.org/10.1017/9781108942607.013 Published online by Cambridge University Press

The following result will be needed for our analysis of numerical integration methods in Chapter 14.

Theorem 9.57 (differentiation)**.** *Suppose that $n \in \mathbb{N}$, $[a, b] \subset \mathbb{R}$ is a compact interval, and $f \in C^{n+1}([a, b])$. Let $X = \{x_i\}_{i=0}^{n} \subset [a, b]$ be a nodal set in $[a, b]$. Then $f[\![x_0, x_1, \ldots, x_n]\!]$ is continuously differentiable with respect to x_n and*

$$\frac{\partial}{\partial x_n} f[\![x_0, x_1, \ldots, x_n]\!] = f[\![x_0, x_1, \ldots, x_n, x_n]\!]. \tag{9.33}$$

Proof. Using (9.27),

$$\begin{aligned}
\frac{\partial}{\partial x_n} f[\![x_0, x_1, \ldots, x_n]\!] &= \lim_{h\to 0} \frac{f[\![x_0, x_1, \ldots, x_n + h]\!] - f[\![x_0, x_1, \ldots, x_n]\!]}{h} \\
&= \lim_{h\to 0} \frac{f[\![x_0, x_1, \ldots, x_n + h]\!] - f[\![x_n, x_0, x_1, \ldots, x_{n-1}]\!]}{h} \\
&= \lim_{h\to 0} f[\![x_n, x_0, x_1, \ldots, x_n + h]\!] \\
&= f[\![x_n, x_0, x_1, \ldots, x_n]\!] \\
&= f[\![x_0, x_1, \ldots, x_n, x_n]\!],
\end{aligned}$$

where the last equality came from Proposition 9.56. Now observe that

$$f[\![x_0, x_1, \ldots, x_n, x_n]\!] = \int_{T_{n+1}} f^{(n+1)}(\tau_0 z_0 + \tau_1 z_1 + \cdots + \tau_n z_n + \tau_{n+1} z_n) \mathrm{d}^{n+1}\boldsymbol{\tau}.$$

As long as $f \in C^{n+1}([a, b])$, $f[\![x_0, x_1, \ldots, x_n, x_n]\!]$ is a continuous function of its arguments. $\square$

We conclude this chapter with alternate formulas for the Hermite interpolating polynomial and its error in terms of divided differences.

Theorem 9.58 (Newton–Hermite interpolation error)**.** *Suppose that $n \in \mathbb{N}$, $X = \{x_j\}_{j=0}^{n}$ is a nodal set in the compact interval $[a, b] \subset \mathbb{R}$, and $f \in C^{2n+2}([a, b])$. Then $p \in \mathbb{P}_{2n+1}$, defined by*

$$\begin{aligned}
p(x) = f(x_0) &+ (x - x_0) f[\![x_0, x_0]\!] + (x - x_0)^2 f[\![x_0, x_0, x_1]\!] \\
&+ (x - x_0)^2 (x - x_1) f[\![x_0, x_0, x_1, x_1]\!] \\
&+ (x - x_0)^2 (x - x_1)^2 f[\![x_0, x_0, x_1, x_1, x_2]\!] \\
&+ \cdots + (x - x_0)^2 \cdots (x - x_{n-1})^2 (x - x_n) f[\![x_0, x_0, x_1, x_1, \ldots, x_n, x_n]\!],
\end{aligned}$$

is the unique Hermite interpolating polynomial satisfying

$$p(x_j) = f(x_j), \qquad p'(x_j) = f'(x_j), \qquad j = 0, \ldots, n.$$

For any $x \in [a, b]$, the interpolation error has the representation

$$f(x) - p(x) = \omega_{n+1}^2(x) f[\![x_0, x_0, x_1, x_1, \ldots, x_n, x_n, x]\!]. \tag{9.34}$$

https://doi.org/10.1017/9781108942607.013 Published online by Cambridge University Press

Proof. Suppose that $Z = \{z_0, \ldots, z_{2n+1}\}$ is a nodal set in $[a, b]$. Then $p \in \mathbb{P}_{2n+1}$, defined by

$$\begin{aligned} p(z) = f(z_0) &+ (z - z_0)f[\![z_0, z_1]\!] + (z - z_0)(z - z_1)f[\![z_0, z_1, z_2]\!] \\ &+ (z - z_0)(z - z_1)(z - z_2)f[\![z_0, z_1, z_2, z_3]\!] \\ &+ \cdots + (z - z_0)\cdots(z - z_{2n})f[\![z_0, z_1, \ldots, z_{2n+1}]\!], \end{aligned}$$

is the unique Lagrange interpolating polynomial satisfying

$$p(z_j) = f(z_j), \quad j = 0, \ldots, 2n + 1.$$

Owing to Theorem 9.45, if $f \in C^{2n+2}([a, b])$, the error may be expressed as follows: for any $x \in [a, b]$,

$$f(x) - p(x) = (x - z_0)\cdots(x - z_{2n+1})f[\![z_0, z_1, \ldots, z_{2n+1}, x]\!].$$

Use now the continuity of the extended divided differences to take the limits

$$z_0, z_1 \to x_0, \qquad z_2, z_3 \to x_1, \qquad \ldots, \qquad z_{2n}, z_{2n+1} \to x_n$$

and show that the desired properties hold. The details are left to the reader as an exercise; see Problem 9.25. $\square$

Problems

9.1 Prove Theorem 9.4. In fact, prove the following more general result. Suppose that $Z = \{z_i\}_{i=0}^n \subset \mathbb{C}$ is a set of $n + 1 \in \mathbb{N}$ distinct points. Then the matrix

$$\mathsf{V} = \begin{bmatrix} 1 & z_0 & z_0^2 & \cdots & z_0^n \\ 1 & z_1 & z_1^2 & \cdots & z_1^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & z_n & z_n^2 & \cdots & z_n^n \end{bmatrix} \in \mathbb{C}^{(n+1)\times(n+1)}$$

is invertible and, in particular,

$$\det(\mathsf{V}) = \prod_{0 \le i < j \le n} (z_j - z_i) \neq 0.$$

9.2 Prove Proposition 9.7.

9.3 Prove Proposition 9.10.

9.4 Complete the proof of Theorem 9.11.

9.5 Prove Proposition 9.14.

9.6 Complete the proof of Theorem 9.15.

9.7 Suppose that $X = \{x_i\}_{i=0}^n$ is a nodal set of size $n + 1 \in \mathbb{N}$ in the compact interval $[a, b] \subset \mathbb{R}$ and $\mathcal{L}_X = \{L_i\}_{i=0}^n \subset \mathbb{P}_n$ is the Lagrange basis subordinate to X. Prove that if $0 \le m \le n$, then

$$\sum_{i=0}^n x_i^m L_i(x) = x^m, \quad \forall x \in \mathbb{R}.$$

https://doi.org/10.1017/9781108942607.013 Published online by Cambridge University Press

9.8 Prove Proposition 9.18 by showing the following. Let $n \in \mathbb{N}$. Define

$$h = \frac{b-a}{n},$$

and, for $i = 0, \dots, n$, set $x_i = a + ih$. Show that the nodal polynomial ω_{n+1}, introduced in Definition 9.13, satisfies

$$\max_{a \le x \le b} |\omega_{n+1}(x)| \le \Omega_{n+1},$$

where

$$\Omega_{n+1} = \frac{n!}{4} h^{n+1}.$$

9.9 Prove Theorem 9.20 by the following steps. Let $n \in \mathbb{N}$. Suppose that $X = \{x_i\}_{i=0}^n$ is a nodal set in the compact interval $[a, b] \subset \mathbb{R}$ and $f \in C^1([a, b])$. Define, for $0 \le \ell \le n$,

$$F_{0,\ell}(x) = (L_\ell(x))^2(1 - 2L'_\ell(x_\ell)(x - x_\ell)), \quad F_{1,\ell} = (L_\ell(x))^2(x - x_\ell),$$

where L_ℓ is the Lagrange basis polynomial of order ℓ, defined in (9.3), and

$$p(x) = \sum_{\ell=0}^{n} \left[F_{0,\ell}(x)f(x_\ell) + F_{1,\ell}(x)f'(x_\ell)\right].$$

a) Prove that

$$F_{0,\ell}(x_k) = \delta_{k,\ell}, \quad F'_{0,\ell}(x_k) = 0, \quad F_{1,\ell}(x_k) = 0, \quad F'_{1,\ell}(x_k) = \delta_{k,\ell},$$

and, therefore, that, for all $i = 0, \dots, n$, we have

$$f(x_i) = p(x_i), \qquad f'(x_i) = p'(x_i).$$

b) Prove that p is the unique polynomial in $\mathbb{P}_{2n+1}$ with the property above.

9.10 Prove Theorem 9.35.

9.11 Let $f \in C([-1, 1])$. Construct the Lagrange interpolation polynomial $p_1 \in \mathbb{P}_1$ for f using the interpolation nodes $x_0 = -1$ and $x_1 = 1$. Show that if $f \in C^2([-1, 1])$, then

$$|f(x) - p_1(x)| \le \frac{F_2}{2}\left(1 - x^2\right) \le \frac{F_2}{2}, \quad \forall x \in [-1, 1],$$

where $F_2 = \|f''\|_{L^\infty(-1,1)}$. Find a function f and a point x for which equality is achieved in the estimates above.

9.12 This problem is concerned with the estimate provided in Theorem 9.16.

a) Compute the Lagrange interpolating polynomial of degree one for the function $f(x) = x^3$ using the interpolation nodes $x_0 = 0$ and $x_1 = a$. Verify Theorem 9.16 by a direct calculation, showing that, in this case, ξ has the unique value $\xi = (x + a)/3$.

b) Repeat the calculation for the function $f(x) = (2x - a)^4$. Show that, in this case, there are two possible values for ξ and give their values.

https://doi.org/10.1017/9781108942607.013 Published online by Cambridge University Press

9.13 Let $n \in \mathbb{N}_0$. Given the nodal set $X = \{x_i\}_{i=0}^{n+1}$ and $Y = \{y_i\}_{i=0}^{n+1} \subset \mathbb{R}$, let $q_n, r_n \in \mathbb{P}_n$ be the Lagrange interpolating polynomials for the coordinate sets

$$Q = \{(x_i, y_i) \mid i = 0, 1, \ldots, n\}, \qquad R = \{(x_i, y_i) \mid i = 1, 2, \ldots, n+1\},$$

respectively. Define

$$p_{n+1}(x) = \frac{(x - x_0)r_n(x) - (x - x_{n+1})q_n(x)}{x_{n+1} - x_0}.$$

Show that p_{n+1} is the Lagrange interpolating polynomial of degree $n+1$ for the coordinate set $P = \{(x_i, y_i) \mid i = 0, 1, \ldots, n+1\}$.

9.14 Construct the Hermite interpolating polynomial of degree three (3) for the function $f(x) = x^5$, using the points $x_0 = 0$ and $x_1 = a$, and show that it has the form $p_3(x) = 3a^2x^3 - 2a^3x^2$. Verify Theorem 9.21 by direct calculation, showing that, in this case, ξ has the unique value $\xi = (x + 2a)/5$.

9.15 Let $n \in \mathbb{N}_0$ and $X = \{x_i\}_{i=0}^{n}$ be a nodal set in the compact interval $[a, b] \subset \mathbb{R}$. Let $f \in C^1([a, b])$ and $p \in \mathbb{P}_{2n+1}$ be the Hermite interpolating polynomial of f subject to X. State and prove an error estimate for $f'(x) - p'(x)$ with $x \in [a, b] \setminus X$.

9.16 Prove Proposition 9.37.

9.17 Complete the proof of Proposition 9.39.

9.18 Prove Corollary 9.44.

9.19 Prove Corollary 9.47.

9.20 Let $[a, b] \subset \mathbb{R}$ be a compact interval, $n \in \mathbb{N}_0$, and $f \in C([a, b])$. Show that $f \in \mathbb{P}_n$ if and only if for every nodal set $X = \{x_j\}_{j=0}^{n} \subset [a, b]$ we have

$$f[x_0, \ldots, x_n] = 0.$$

9.21 Let $[a, b] \subset \mathbb{R}$ be a compact interval, $n \in \mathbb{N}_0$, and $g, h \in C([a, b])$. Define $f(x) = g(x)h(x)$. Show that, for every nodal set $X = \{x_j\}_{j=0}^{n} \subset [a, b]$, we have

$$f[x_0, \ldots, x_n] = \sum_{j=0}^{n} g[x_0, \ldots, x_j]h[x_{j+1}, \ldots, x_n].$$

9.22 Prove Lemma 9.49.

9.23 Prove Corollary 9.54.

9.24 Prove Lemma 9.55.

9.25 Prove Theorem 9.58.