For an even grid function, $f_m = f_{-m}$, formulae (3.51)–(3.54) transform into:

$$\begin{aligned}
\tilde{a}_0 &= \frac{1}{N}(f_0 + f_n) + \frac{2}{N}\sum_{m=1}^{n-1} f_m, \\
\tilde{a}_k &= \frac{2}{N}(f_0 + (-1)^k f_n) + \frac{4}{N}\sum_{m=1}^{n-1} f_m \cos\frac{2\pi km}{N}, \qquad k = 1,2,\dots,n-1, \\
\tilde{a}_n &= \frac{1}{N}(f_0 + (-1)^n f_n), \\
\tilde{b}_k &= 0, \qquad k = 1,2,\dots,n-1,
\end{aligned}$$

and the polynomial (3.50) reduces to:

$$\tilde{Q}_n\left(\cos\frac{2\pi}{L}x, \sin\frac{2\pi}{L}x, f\right) = \sum_{k=0}^{n} \tilde{a}_k \cos\frac{2\pi k}{L}x.$$

Note that the arguments which are very similar to those used when proving the key properties of the trigonometric interpolating polynomial $Q_n\left(\cos\frac{2\pi}{L}x, \sin\frac{2\pi}{L}x, f\right)$ in Theorems 3.4 and 3.5, also apply to the polynomial $\tilde{Q}_n\left(\cos\frac{2\pi}{L}x, \sin\frac{2\pi}{L}x, f\right)$ defined by formulae (3.50)–(3.54). Namely, this polynomial has slowly growing Lebesgue constants and as such, is basically stable with respect to the perturbations of the grid function f_m. Moreover, it converges to the interpolated function $f(x)$ as $n \longrightarrow \infty$ with the rate determined by the smoothness of $f(x)$, i.e., there is no saturation.

REMARK 3.1 If the interpolated function $f(x)$ has derivatives of all orders, then the rate of convergence of the trigonometric interpolating polynomials to $f(x)$ will be faster than any inverse power of n. In the literature, this type of convergence is often referred to as *spectral*. ☐

3.2 Interpolation of Functions on an Interval. Relation between Algebraic and Trigonometric Interpolation

Let $f = f(x)$ be defined on the interval $-1 \le x \le 1$, and let it have there a bounded derivative of order $r+1$. We have chosen this specific interval $-1 \le x \le 1$ as the domain of $f(x)$, rather than an arbitrary interval $a \le x \le b$, for the only reason of simplicity and convenience. Indeed, the transformation $x = \frac{a+b}{2} + t\frac{b-a}{2}$ renders a transition from the function $f(x)$ defined on an arbitrary interval $a \le x \le b$ to the function $F(t) \equiv f\left(\frac{a+b}{2} + t\frac{b-a}{2}\right)$ defined on the interval $-1 \le t \le 1$.

3.2.1 Periodization

According to Theorem 3.5 of Section 3.1, trigonometric interpolation is only suitable for the reconstruction of smooth periodic functions from their tables of values.

Therefore, to be able to apply it to the function $f(x)$ given on $-1 \leq x \leq 1$, one should first equivalently replace $f(x)$ by some smooth periodic function. However, a straightforward extension of the function $f(x)$ from its domain $-1 \leq x \leq 1$ to the entire real axis may, generally speaking, yield a discontinuous periodic function with the period $L = 2$, see Figure 3.1.

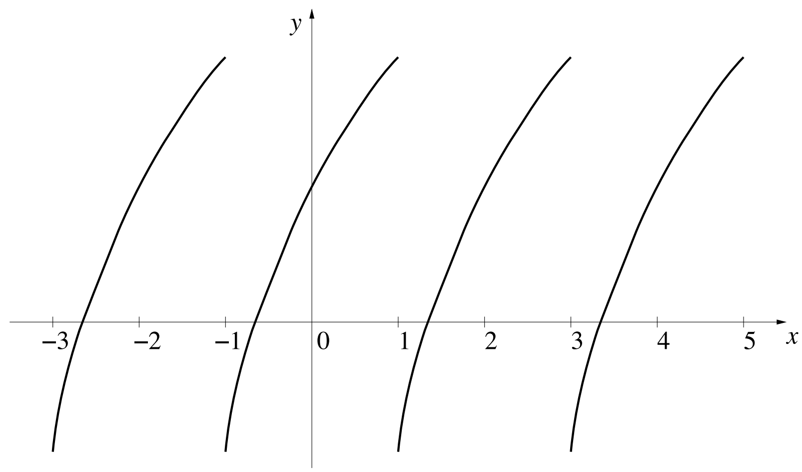

FIGURE 3.1: Straightforward periodization.

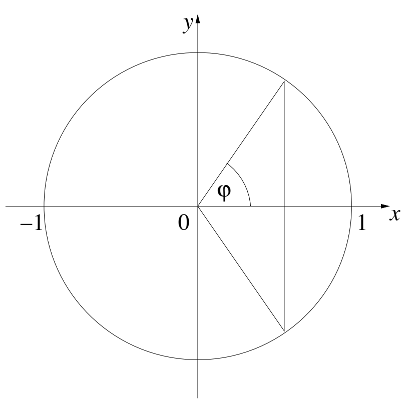

FIGURE 3.2: Periodization according to formula (3.55).

Therefore, instead of the function $f(x)$, $-1 \leq x \leq 1$, let us consider a new function

$$F(\varphi) = f(\cos\varphi), \quad x = \cos\varphi. \quad (3.55)$$

It will be convenient to think that the function $F(\varphi)$ of (3.55) is defined on the unit circle as a function of the polar angle φ. The value of $F(\varphi)$ is obtained by merely translating the value of $f(x)$ from the point $x \in [-1, 1]$ to the corresponding point $\varphi \in [0, \pi]$ on the unit circle, see Figure 3.2. In so doing, one can interpret the resulting function $F(\varphi)$ as even, $F(-\varphi) = F(\varphi)$, 2π-periodic function of its argument φ. Moreover, it is easy to see from definition (3.55) that the derivative $\frac{d^{r+1}F(\varphi)}{d\varphi^{r+1}}$ exists and is bounded.

3.2.2 Trigonometric Interpolation

Let us choose the following interpolation nodes:

$$\varphi_m = \frac{2\pi}{N}m + \frac{\pi}{N}, \qquad m = 0, \pm 1, \dots, \pm n, -(n+1), \qquad N = 2(n+1). \tag{3.56}$$

According to (3.55), the values $F_m = F(\varphi_m)$ of the function $F(\varphi)$ at the nodes φ_m of (3.56) coincide with the values $f_m = f(x_m)$ of the original function $f(x)$ at the points $x_m = \cos\varphi_m$. To interpolate a 2π-periodic even function $F(\varphi)$ using its tabulated values at the nodes (3.56), one can employ formula (3.25) of Section 3.1:

$$Q_n(\cos\varphi, \sin\varphi, F) = \sum_{k=0}^{n} a_k \cos k\varphi. \tag{3.57}$$

As $F_m = f_m$ for all m, the coefficients a_k of the trigonometric interpolating polynomial (3.57) are given by formulae (3.26), (3.27) of Section 3.1:

$$\begin{aligned} a_0 &= \frac{1}{n+1}\sum_{m=0}^{n} f_m, \\ a_k &= \frac{2}{n+1}\sum_{m=0}^{n} f_m \cos k\varphi_m, \qquad k = 1, 2, \dots, n. \end{aligned} \tag{3.58}$$

3.2.3 Chebyshev Polynomials. Relation between Algebraic and Trigonometric Interpolation

Let us use the equality $\cos\varphi = x$ and introduce the functions:

$$T_k(x) = \cos k\varphi = \cos(k \arccos x), \qquad k = 0, 1, 2, \dots. \tag{3.59}$$

THEOREM 3.7

The functions $T_k(x)$ defined by formula (3.59) are polynomials of degree $k = 0, 1, 2, \dots$. Specifically, $T_0(x) = 1$, $T_1(x) = x$, and all other polynomials: $T_2(x)$, $T_3(x)$, etc., can be obtained consecutively using the recursion formula

$$T_{k+1}(x) = 2xT_k(x) - T_{k-1}(x). \tag{3.60}$$

PROOF It is clear that $T_0(x) = \cos 0 = 1$ and $T_1(x) = \cos\arccos x = x$. Then, we employ a well-known trigonometric identity

$$\cos(k+1)\varphi = 2\cos\varphi\cos k\varphi - \cos(k-1)\varphi, \qquad k = 1, 2, \dots,$$

which immediately yields formula (3.60) when $\varphi = \arccos x$. It only remains to prove that $T_k(x)$ is a polynomial of degree k; we will use induction with respect to k to do that. For $k = 0$ and $k = 1$ it has been proven directly. Let us fix some $k > 1$ and assume that for all $j = 0, 1, \dots, k$ we have already shown that $T_j(x)$ are polynomials of degree j. Then, the expression on the right-hand side of (3.60), and as such, $T_{k+1}(x)$, is a polynomial of degree $k+1$. ∎

The polynomials $T_k(x)$ were first introduced and studied by Chebyshev. We provide here the formulae for a first few Chebyshev polynomials, along with their graphs, see Figure 3.3:

$$T_0(x) = 1, \qquad T_1(x) = x, \qquad T_2(x) = 2x^2 - 1, \qquad T_3(x) = 4x^3 - 3x.$$

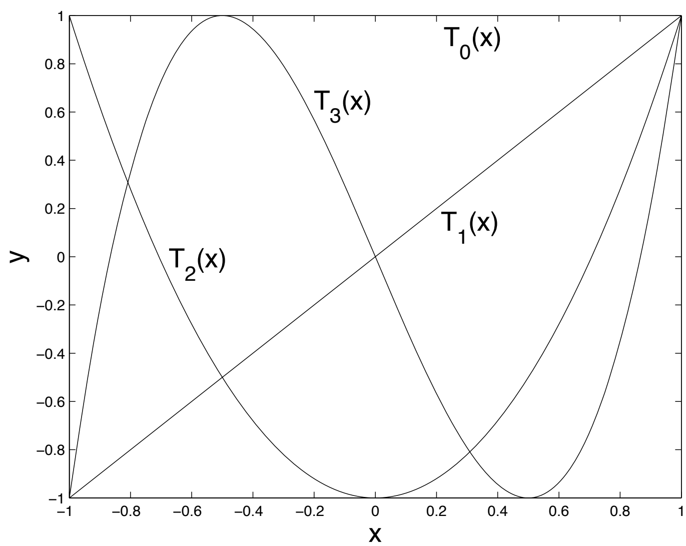

FIGURE 3.3: Chebyshev polynomials.

Next, by substituting $\varphi = \arccos x$ into the right-hand side of formula (3.57), we can recast it as a function of x, thus obtaining:

$$Q_n(\cos\varphi, \sin\varphi, F) \equiv P_n(x, f), \tag{3.61}$$

where

$$P_n(x, f) = \sum_{k=0}^{n} a_k T_k(x), \tag{3.62}$$

and

$$\begin{aligned} a_0 &= \frac{1}{n+1}\sum_{m=0}^{n} f_m = \frac{1}{n+1}\sum_{m=0}^{n} f_m T_0(x_m), \\ a_k &= \frac{2}{n+1}\sum_{m=0}^{n} f_m \cos k\varphi_m = \frac{2}{n+1}\sum_{m=0}^{n} f_m T_k(x_m). \end{aligned} \tag{3.63}$$

Therefore, we can conclude using formulae (3.61), (3.62) and Theorem 3.7, that $P_n(x, f)$ is an algebraic polynomial of degree no greater than n that coincides with the given function values $f(x_m) = f_m$ at the interpolation nodes $x_m = \cos\varphi_m$. In accordance with (3.56), these interpolation nodes can be defined by the formula:

$$x_m = \cos\varphi_m = \cos\frac{\pi(2m+1)}{2(n+1)}, \quad m = 0, 1, \dots, n, \tag{3.64}$$

that basically coincides with formula (2.20) of Chapter 2. We are schematically showing the nodes (3.64) in Figure 3.4.

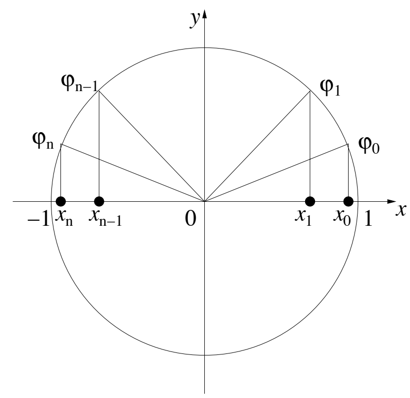

FIGURE 3.4: Chebyshev interpolation nodes.

Note that the points $\varphi_m = \frac{\pi}{n+1}m + \frac{\pi}{2(n+1)}$, $m = 0, 1, \dots, n$, defined by formula (3.56) are actually zeros of the function $\cos(n+1)\varphi$. Accordingly, the points $x_m = \cos\varphi_m$ defined by formula (3.64) are roots of the Chebyshev polynomial $T_{n+1}(x) = \cos(n+1)\varphi$. In the literature, this particular choice of nodes for the Chebyshev grid, see Figure 3.4, is often referred to as the Chebyshev-Gauss or simply Gauss nodes (an alternative choice of nodes is discussed in Section 3.2.6).

In other words, the polynomial $P_n(x, f)$ specified by formulae (3.62), (3.63) renders algebraic interpolation of the function $f(x)$ based on its values f_m sampled at the roots x_m, see (3.64), of the Chebyshev polynomial $T_{n+1}(x)$.

3.2.4 Properties of Algebraic Interpolation with Roots of the Chebyshev Polynomial $T_{n+1}(x)$ as Nodes

Equality $Q_n(\cos\varphi, \sin\varphi, F) = P_n(x, f)$ implies that the properties of the trigonometric interpolating polynomial $Q_n(\cos\varphi, \sin\varphi, F)$ established by Theorems 3.4 and 3.5 of Section 3.1 do carry over to the algebraic interpolating polynomial $P_n(x, f)$ defined by formula (3.62). In particular, the Lebesgue constants L_n that characterize the sensitivity of the polynomial $P_n(x, f)$ to the perturbations of f_m, satisfy estimate (3.33) from Section 3.1:

$$L_n \leq 4(n+1),$$

and the interpolation error

$$R_n(x) = f(x) - P_n(x, f)$$

uniformly converges to zero as $n \longrightarrow \infty$ with the rate automatically determined by the number of derivatives $r+1$ that the function $f(x)$ has:

$$\max_{-1 \leq x \leq 1} |R_n(x)| \leq \frac{\zeta_n}{n^{r-1/2}}, \quad \text{where} \quad \zeta_n = o(1), \;\; n \longrightarrow \infty. \tag{3.65}$$

In other words, similarly to the trigonometric interpolation (see Section 3.1.3), algebraic interpolation on the Chebyshev nodes *does not get saturated by smoothness*. In other words, the interpolation error self-adjusts to the regularity of the interpolated function without having to change anything in the construction of the method.

REMARK 3.2 Estimate (3.33) for the Lebesgue constants that we proved in Theorem 3.4 of Section 3.1 can be substantially improved. In fact, the following equality holds [see the bibliography quoted in Section 3.2.7, and cf. formula (2.21) of Section 2.1, Chapter 2]:

$$L_n = \frac{2}{\pi}\ln(n+1) + 1 - \theta_n, \qquad 0 \le \theta_n \le \frac{1}{4}. \tag{3.66}$$

Accordingly, the estimate for the interpolation error can also be improved, and instead of (3.65) we will obtain:

$$\max_{-1\le x\le 1} |R_n(x)| = o\left(\frac{\ln n}{n^{r+1/2}}\right) \quad \text{as} \quad n \longrightarrow \infty, \tag{3.67}$$

where $r+1$ is the maximum number of derivatives that the function $f(x)$ has. Further improvements of estimate (3.67) can be obtained with the help of the Jackson inequality, see Section 3.2.7. □

In contradistinction to the Chebyshev nodes (3.64), when a uniform grid is used for interpolation, the Lebesgue constants rapidly grow as n increases, see inequalities (2.18) of Chapter 2, and convergence of the interpolating polynomial to the function $f(x)$ may break down even for infinitely smooth functions, see Section 2.1.5. These are precisely the considerations that make the algebraic interpolation of high degree inappropriate, and prompt the use of piecewise polynomial or spline interpolation on uniform or arbitrary non-uniform grids (see Chapter 2).

3.2.5 An Algorithm for Evaluating the Interpolating Polynomial

To obtain the coefficients a_k of the polynomial $P_n(x,f)$ of (3.62) with the help of formulae (3.63), as well as to actually evaluate this polynomial itself at a given $x \in [-1, 1]$, one needs to be able to compute the values of the polynomials $T_k(x)$, $k = 0, 1, 2, \ldots$, for $-1 \le x \le 1$. We will show that it is appropriate to use formula (3.60) for this purpose. This formula is obviously easy to use and its computational efficacy is also apparent. We only need to demonstrate that the computations according to this formula are stable with respect to the round-off errors.

Consider a difference equation of the type:

$$y_{k+1} = 2xy_k - y_{k-1}, \tag{3.68}$$

where y_k is the unknown sequence parameterized by the integer quantity k. We will be looking for a solution of equation (3.68) in the form $y_k = q^k$, where q is a fixed number. Substituting the latter expression into the difference equation (3.68), we obtain the following algebraic equation for q:

$$q^2 - 2xq + 1 = 0.$$

It is often called the characteristic equation, and it has two roots:

$$q_{1,2} = x \pm \sqrt{x^2 - 1}.$$

Due to the linearity of equation (3.68), its general solution can be written in the form

$$y_k = c_1 q_1^k + c_2 q_2^k, \tag{3.69}$$

where c_1 and c_2 are arbitrary constants. Let us choose these constants c_1 and c_2 so that to satisfy the following conditions:

$$y_0 = T_0(x) = 1, \qquad y_1 = T_1(x) = x,$$

or equivalently,

$$c_1 + c_2 = 1, \qquad c_1 q_1 + c_2 q_2 = x.$$

This implies $c_1 = c_2 = 1/2$, and then formula (3.69) yields the following solution of equation (3.68):

$$T_k(x) = \frac{1}{2}\left(x + \sqrt{x^2-1}\right)^k + \frac{1}{2}\left(x - \sqrt{x^2-1}\right)^k. \tag{3.70}$$

According to formula (3.60), $T_k(x)$ defined by (3.70) shall be interpreted for a given fixed x as a (discrete) function of k that solves equation (3.68).

Next, note that when $|x| < 1$ the roots $q_{1,2} = x \pm \sqrt{x^2-1}$ of the characteristic equation are complex conjugate and have unit moduli. Consequently, the quantities q_1^k and q_2^k will remain equal to one by their absolute value as k increases. An error committed for some $k = k_0$ would cause a perturbation in the values of c_1 and c_2 that enter into formula (3.69) for $k > k_0$. However, due to the equalities $|q_1^k| = |q_2^k| = 1$, $k = 1, 2, \dots$, this error will not get amplified as k increases. This implies numerical stability of the computations according to formula (3.60) for $|x| < 1$.

3.2.6 Algebraic Interpolation with Extrema of the Chebyshev Polynomial $T_n(x)$ as Nodes

To interpolate the function $F(\varphi) = f(\cos\varphi)$, let us now use the nodes:

$$\tilde{\varphi}_m = \frac{\pi}{n} m, \qquad m = 0, 1, \dots, n.$$

In accordance with Theorem 3.6 of Section 3.1, and the discussion on page 72 that follows this theorem, we obtain the trigonometric interpolating polynomial:

$$\tilde{Q}_n(\cos\varphi, \sin\varphi, F) = \sum_{k=0}^{n} \tilde{a}_k \cos k\varphi,$$

$$\tilde{a}_0 = \frac{1}{2n}(f_0 + f_n) + \frac{1}{n}\sum_{m=1}^{n-1} f_m, \qquad \tilde{a}_n = \frac{1}{2n}(f_0 + (-1)^n f_n),$$

$$\tilde{a}_k = \frac{1}{n}(f_0 + (-1)^k f_n) + \frac{2}{n}\sum_{m=1}^{n-1} f_m \cos k\varphi_m, \qquad k = 1, 2, \dots, n-1.$$

Changing the variable to $x = \cos\varphi$ and denoting $\tilde{Q}_n(\cos\varphi, \sin\varphi, F) = \tilde{P}_n(x, f)$, we have:

$$\tilde{P}_n(x, f) = \sum_{k=0}^{n} \tilde{a}_k T_k(x),$$

$$\tilde{a}_0 = \frac{1}{2n}(f_0 + f_n) + \frac{1}{n}\sum_{m=1}^{n-1} f_m, \qquad \tilde{a}_n = \frac{1}{2n}(f_0 + (-1)^n f_n),$$

$$\tilde{a}_k = \frac{1}{n}(f_0 + (-1)^k f_n) + \frac{2}{n}\sum_{m=1}^{n-1} f_m T_k(\tilde{x}_m), \qquad k = 1, 2, \dots, n-1.$$

Similarly to the polynomial $P_n(x, f)$ of (3.62), the algebraic interpolating polynomial $\tilde{P}_n(x, f)$ built on the grid:

$$\tilde{x}_m = \cos\tilde{\varphi}_m = \cos\frac{\pi}{n}m, \qquad m = 0, 1, \dots, n, \tag{3.71}$$

also inherits the two foremost advantageous properties from the trigonometric interpolating polynomial $\tilde{Q}_n(\cos\varphi, \sin\varphi, F)$. They are the slow growth of the Lebesgue constants as n increases (that translates into the numerical stability with respect to the perturbations of f_m), as well convergence with the rate that automatically takes into account the smoothness of $f(x)$, i.e., no susceptibility to saturation.

Finally, we notice that the Chebyshev polynomial $T_n(x)$ reaches its extreme values on the interval $-1 \le x \le 1$ precisely at the interpolation nodes $\tilde{x}_m$ of (3.71): $T_n(\tilde{x}_m) = \cos\pi m = (-1)^m$, $m = 0, 1, \dots, n$. In the literature, the grid nodes $\tilde{x}_m$ of (3.71) are known as the Chebyshev-Gauss-Lobatto nodes or simply the Gauss-Lobatto nodes.

3.2.7 More on the Lebesgue Constants and Convergence of Interpolants

In this section, we discuss the problem of interpolation from the general perspective of approximation of functions by polynomials. Our considerations, in a substantially abridged form, follow those of [LG95], see also [Bab86]. We quote many of the fundamental results without a proof (the theorems of Jackson, Weierstrass, Faber-Bernstein, and Bernstein). The justification of these results, along with a broader and more comprehensive account of the subject, can be found in the literature on the classical theory of approximation, see, e.g., [Jac94, Ber52, Ber54, Ach92, Nat64, Nat65a, Nat65b, Lor86, Che66, Riv74]. In the numerical analysis literature, some of these issues are addressed in [Wen66]. In these books, the reader will also find references to research articles. The material of this section is more advanced, and can be skipped during the first reading.

The meaning of Lebesgue's constants L_n introduced in Chapter 2 as minimum numbers that for each n guarantee the estimate [see formula (2.17)]:

$$\max_{a \le x \le b} |P_n(x, \delta f)| \le L_n \max_j |\delta f(x_j)|$$

is basically that of an operator norm. Indeed, interpolation by means of the polynomial $P_n(x, f)$ can be interpreted as a linear operator that maps the finite-dimensional space of vectors $[f_0, f_1, \dots, f_n]$ into the space $C[a, b]$ of all continuous functions $f(x)$ defined on $[a, b]$. The space $C[a, b]$ is equipped with the maximum norm $\|f\| = \max_{a \le x \le b} |f(x)|$. Likewise, the space of vectors $\vec{f} = [f_0, f_1, \dots, f_n]$ can also be equipped with a maximum norm, but discrete rather than continuous: $\|\vec{f}\| = \max_{0 \le j \le n} |f_j|$. Then, L_n of (2.17) appears to be the induced norm of the foregoing linear operator:

$$L_n = \sup_{\|\vec{f}\|=1} \|P_n(x, f)\|. \tag{3.72}$$

However, for subsequent analysis it will be more convenient to use a slightly different definition of L_n — as norm of an operator that would rather map $C[a, b]$ onto itself.

DEFINITION 3.1 *The operator $\mathscr{P}_n = \mathscr{P}_n(x_0, x_1, \dots, x_n) : \ C[a, b] \longmapsto \{P_n(x)\} \subset C[a, b]$ takes a function $f \in C[a, b]$, samples its values at a given set of nodes $\{x_0, x_1, \dots, x_n\} \in [a, b]$ thus creating the table $\{f_0, f_1, \dots, f_n\}$, and subsequently builds the polynomial $P_n(x, f, x_0, x_1, \dots, x_n) \in C[a, b]$.*

LEMMA 3.1

The operator $\mathscr{P}_n$ introduced by Definition 3.1 is linear and continuous.

PROOF The linearity of $\mathscr{P}_n$ is obvious. To show the continuity, we use the Lagrange formula (2.1) of Section 2.1, Chapter 2, and obtain:

$$|\mathscr{P}_n[f](x)| = |P_n(x, f, x_0, x_1, \dots, x_n)| \le \sum_{k=0}^{n} |f_k||l_k(x)| \le \|f\| \sum_{k=0}^{n} |l_k(x)|.$$

Next, we introduce a new quantity

$$\lambda_n \stackrel{\text{def}}{=} \sup_{[a,b]} \sum_{k=0}^{n} |l_k(x)| = \max_{[a,b]} \sum_{k=0}^{n} |l_k(x)|, \tag{3.73}$$

where the second equality in (3.73) holds because $[a,b]$ is a compact set, and $\sum_{k=0}^{n} |l_k(x)|$ is a continuous function. Then clearly,

$$|\mathscr{P}_n[f](x)| \le \lambda_n \|f\|.$$

Consequently, $\mathscr{P}_n$ is a bounded operator, $\mathscr{P}_n : C[a,b] \mapsto C[a,b]$, and therefore, it is continuous. Moreover, $\|\mathscr{P}_n\| \le \lambda_n$. ∎

DEFINITION 3.2 *The norm of the operator $\mathscr{P}_n$ introduced by Definition 3.1 is called the Lebesgue constant of the polynomial interpolation based on the nodes $x_0, x_1, \dots, x_n$:*

$$L_n = \|\mathscr{P}_n\|. \tag{3.74}$$

Recall that the operator norm on the right-hand side of formula (3.74) is given by:

$$\|\mathscr{P}_n\| = \sup_{\|f\|=1} \|\mathscr{P}_n[f](x)\| = \sup_{\|f\|=1} \|P_n(x,f)\|. \tag{3.75}$$

We have therefore formulated two alternative definitions of the Lebesgue constants — by means of formula (3.72) and by means of formulae (3.74), (3.75). We will now prove that these definitions are, in fact, equivalent. In other words, we will show that the right-hand side of formula (3.75) coincides with the right-hand side of formula (3.72). The difference between these right-hand sides is that in (3.75) the smallest upper bound is taken across the unit sphere in the space $C[a,b]$ that has infinite dimension, whereas in (3.72) it is taken across the unit sphere in the $n+1$-dimensional space of vectors $\vec{f} = [f_0, f_1, \dots, f_n]$.

LEMMA 3.2
The Lebesgue constant defined by formulae (3.74), (3.75) is the same as the Lebesgue constant defined by formula (3.72).

PROOF For every vector $\vec{f} = [f_0, f_1, \dots, f_n]$, consider a piecewise linear function defined as:

$$f(x) = f_{j+1}\frac{x - x_j}{x_{j+1} - x_j} + f_j \frac{x_{j+1} - x}{x_{j+1} - x_j} \quad \text{for} \quad x \in [x_j, x_{j+1}], \quad j = 0, 1, \dots, n-1.$$

Clearly, $f(x) \in C[a,b]$, and also if $\|\vec{f}\| = 1$ then $\|f\| = 1$. In other words, every unit vector $\vec{f} = [f_0, f_1, \dots, f_n]$ gives rise to a continuous (piecewise linear) function that belongs to the unit sphere in $C[a,b]$. Therefore, one can say that the smallest upper bound on the right-hand side of (3.75) is taken across a wider set than that on the right-hand side of (3.72). Consequently,

$$\sup_{\|f\|=1} \|P_n(x,f)\| \ge \sup_{\|\vec{f}\|=1} \|P_n(x,f)\|. \tag{3.76}$$

On the other hand, let $f(x) \in C[a,b]$ be a particular function that realizes the smallest upper bound on the right-hand side of (3.75). By construction, $\|f\| = 1$. Let us sample the values of $f(x)$ at the nodes $x_0, x_1, \ldots, x_n$. This yields the table $\{f_0, f_1, \ldots, f_n\}$, or equivalently, the vector $\vec{f} = [f_0, f_1, \ldots, f_n]$. Assume that $\|\vec{f}\| < 1$. Then, denote $\alpha = \|\vec{f}\|^{-1} > 1$ and stretch the vector $\vec{f}$: $\vec{f} \longmapsto \alpha\vec{f} = [\alpha f_0, \alpha f_1, \ldots, \alpha f_n]$ so that $\|\alpha\vec{f}\| = 1$. As the interpolation by means of the polynomials P_n is a linear operator, we obviously have $P_n(x, \alpha f) = \alpha P_n(x, f)$, and consequently, $\|P_n(x, \alpha f)\| > \|P_n(x, f)\|$. We have therefore found a unit vector $\alpha\vec{f}$, for which the norm of the corresponding interpolating polynomial will be greater than the left-hand side of (3.76). The contradiction proves that the two definitions of the Lebesgue constants are indeed equivalent. $\square$

LEMMA 3.3

The Lebesgue constant of (3.74), (3.75) is equal to

$$L_n = \lambda_n, \tag{3.77}$$

where the quantity λ_n is defined by formula (3.73).

PROOF When proving Lemma 3.1, we have seen that $\|\mathscr{P}_n\| \le \lambda_n$. We therefore need to show that $\lambda_n \le \|\mathscr{P}_n\|$.

As has been mentioned, the function $\psi(x) \stackrel{\text{def}}{=} \sum_{k=0}^{n} |l_k(x)|$ is continuous on $[a,b]$. Consequently, $\exists x^* \in [a,b]: \psi(x^*) = \lambda_n$. Let us now consider a function $f_0 \in C[a,b]$ such that $f_0(x_k) = \operatorname{sign} l_k(x^*)$, $k = 0, 1, 2, \ldots, n$, and also $\|f_0\| = 1$. For this function we have:

$$|\mathscr{P}_n[f_0](x^*)| = \left| \sum_{k=0}^{n} f_0(x_k) l_k(x^*) \right| = \left| \sum_{k=0}^{n} (\operatorname{sign} l_k(x^*))\, l_k(x^*) \right| = \sum_{k=0}^{n} |l_k(x^*)| = \psi(x^*) = \lambda_n.$$

On the other hand,

$$|\mathscr{P}_n[f_0](x^*)| \le \|\mathscr{P}_n[f_0]\| \le \|\mathscr{P}_n\| \cdot \|f_0\| = \|\mathscr{P}_n\|,$$

which implies $\lambda_n \le \|\mathscr{P}_n\|$. It only remains to construct a specific example of $f_0 \in C[a,b]$. This can be done easily by taking $f_0(x)$ as a piecewise linear function with the values $\operatorname{sign} l_k(x^*)$ at the points x_k, $k = 0, 1, 2, \ldots, n$. $\square$

We can therefore conclude that

$$L_n = \max_{a \le x \le b} \sum_{k=0}^{n} |l_k(x)|.$$

We have used a somewhat weaker form of this result in Section 2.1.4 of Chapter 2.

The Lebesgue constants of Definition 3.2 play a fundamental role when studying the convergence of interpolating polynomials. To actually see that, we will first need to introduce another key new concept and formulate some important results.

DEFINITION 3.3 *The quantity*

$$\varepsilon(f,P_n) = \min_{P_n(x)} \max_{a \le x \le b} |P_n(x) - f(x)| \tag{3.78}$$

is called the best approximation of a given function $f(x)$ by polynomials of degree no greater than n on the interval $[a,b]$.

Note that the minimum in formula (3.78) is taken with respect to all algebraic polynomials of degree no greater than n on the interval $[a,b]$, not only the interpolating polynomials. In other words, the polynomials in (3.78) do not, generally speaking, have to coincide with $f(x)$ at any given point of $[a,b]$. It is possible to show existence of a particular polynomial that realizes the best approximation (3.78). In most cases, however, this polynomial is difficult to obtain constructively. In general, polynomials of the best approximation can only be built using sophisticated iterative algorithms of non-smooth optimization. On the other hand, their theoretical properties are well studied. Perhaps the most fundamental property is given by the Jackson inequality.

THEOREM 3.8 (Jackson inequality)
Let $f = f(x)$ be defined on the interval $[a,b]$, let it be $r-1$ times continuously differentiable, and let the derivative $f^{(r-1)}(x)$ be Lipshitz-continuous:

$$\forall x_1, x_2 \in [a,b]: \quad |f^{(r-1)}(x_1) - f^{(r-1)}(x_2)| \le M|x_1 - x_2|, \qquad M > 0.$$

Then, for any $n \ge r$ the following inequality holds:

$$\varepsilon(f,P_n) < C_r \left(\frac{b-a}{2}\right)^r \frac{M}{n^r}, \tag{3.79}$$

where $C_r = \left(\frac{\pi r}{2}\right)^r \frac{1}{r!}$ are universal constants that depend neither on f, nor on n, nor on M.

The Jackson inequality [Jac94] reinforces, for sufficiently smooth functions, the result of the following classical theorem established in real analysis (see, e.g., [Rud87]):

THEOREM 3.9 (Weierstrass)
Let $f \in C[a,b]$. Then, for any $\varepsilon > 0$ there is an algebraic polynomial $P_\varepsilon(x)$ such that $\forall x \in [a,b]: \quad |f(x) - P_\varepsilon(x)| \le \varepsilon$.

A classical proof of Theorem 3.9 is based on periodization of f that preserves its continuity (the period should obviously be larger than $[a,b]$) and then on the approximation by partial sums of the Taylor series that converges uniformly. The Weierstrass theorem implies that for $f \in C[a,b]$ the best approximation defined by (3.78) converges to zero: $\varepsilon(f,P_n) \longrightarrow 0$ when $n \longrightarrow \infty$. This is basically as much as one can tell regarding the behavior of $\varepsilon(f,P_n)$ if nothing else is known about $f(x)$ except that it is continuous. On the other hand, the Jackson inequality specifies the rate of decay for the best approximation as a particular inverse power of n, see formula (3.79), provided that $f(x)$ is smooth.

Let us also note that the value of $\frac{\pi}{2}$ that enters the expression for $C_r = \left(\frac{\pi r}{2}\right)^r \frac{1}{r!}$ in the Jackson inequality (3.79) may, in fact, be replaced by smaller values:

$$K_0 = 1, \quad K_1 = \frac{\pi}{2}, \quad K_2 = \frac{\pi^2}{8}, \quad K_3 = \frac{\pi^3}{24}, \quad K_4 = \frac{5\pi^4}{384}, \dots$$

known as the Favard constants. The Favard constants can be obtained explicitly for all $r = 0, 1, 2, \dots$, and it is possible to show that all $K_r < \frac{\pi}{2}$. The key consideration regarding the Favard constants is that substituting them into (3.79) makes this inequality sharp.

The main result that connects the properties of the best approximation (Definition 3.3) and the quality of interpolation by means of algebraic polynomials is given by the following

THEOREM 3.10 (Lebesgue inequality)
Let $f \in C[a,b]$ and let $\{x_0, x_1, \dots, x_n\}$ be an arbitrary set of distinct interpolation nodes on $[a,b]$. Then,

$$\varepsilon(f,P_n) \le \|f - \mathscr{P}_n[f]\| \le (L_n + 1)\varepsilon(f,P_n). \tag{3.80}$$

Note that according to Definition 3.1, the operator $\mathscr{P}_n$ in formula (3.80) generally speaking depends on the choice of the interpolation nodes.

PROOF It is obvious that we only need to prove the second inequality in (3.80), i.e., the upper bound. Consider an arbitrary polynomial $Q(x) \in \{P_n(x)\}$ of degree no greater than n. As the algebraic interpolating polynomial is unique (Theorem 2.1 of Chapter 2), we obviously have $\mathscr{P}_n[Q] = Q$. Next,

$$\begin{aligned}\|f - \mathscr{P}_n[f]\| &= \|f - Q + \mathscr{P}_n[Q] - \mathscr{P}_n[f]\| \\ &\le \|f - Q\| + \|\mathscr{P}_n[Q] - \mathscr{P}_n[f]\| = \|f - Q\| + \|\mathscr{P}_n[Q - f]\| \\ &\le \|f - Q\| + \|\mathscr{P}_n\|\|f - Q\| = (1 + L_n)\|f - Q\|.\end{aligned}$$

Let us now introduce $\delta > 0$ and denote by $Q_\delta(x) \in \{P_n(x)\}$ a polynomial for which $\|f - Q_\delta\| < \varepsilon(f,P_n) + \delta$. Then,

$$\|f - \mathscr{P}_n[f]\| \le (1 + L_n)\|f - Q_\delta\| < (1 + L_n)(\varepsilon(f,P_n) + \delta).$$

Finally, by taking the limit $\delta \longrightarrow 0$, we obtain the desired inequality (3.80). ∎

The Lebesgue inequality (3.80) essentially provides an upper bound for the interpolation error $\|f - \mathscr{P}_n[f]\|$ in terms of a product of the best approximation (3.78) times the Lebesgue constant (3.74). Often, this estimate allows one to judge the convergence of algebraic interpolating polynomials as n increases. It is therefore clear that the behavior of the Lebesgue constants is of central importance for the convergence study.

THEOREM 3.11 (Faber-Bernstein)
For any choice of interpolation nodes $x_0, x_1, \dots, x_n$ on the interval $[a,b]$, the following inequality holds:

$$L_n > \frac{1}{8\sqrt{\pi}} \ln(n+1). \tag{3.81}$$

Theorem 3.11 shows that the Lebesgue constants always grow as the grid dimension n increases. As such, the best one can generally hope for is to be able to place the interpolation nodes in such a way that this growth will be optimal, i.e., logarithmic.

As far as the problem of interpolation is concerned, if, for example, nothing is known about the function $f \in C[a,b]$ except that it is continuous, then nothing can be said about the behavior of the error beyond the estimate given by the Lebesgue inequality (3.80). The Weierstrass theorem (Theorem 3.9) indicates that $\varepsilon(f, P_n) \longrightarrow 0$ as $n \longrightarrow \infty$, and the Faber-Bernstein theorem (Theorem 3.11) says that $L_n \longrightarrow \infty$ as $n \longrightarrow \infty$. We therefore have the uncertainty $0 \cdot \infty$ on the right-hand side of the Lebesgue inequality; and the behavior of this right-hand side is determined by which of the two processes dominates — the decay of the best approximations or the growth of the Lebesgue constants. In particular, if $\lim_{n\to\infty} L_n \varepsilon(f, P_n) = 0$, then the interpolating polynomials uniformly converge to $f(x)$.

If the function $f(x)$ is sufficiently smooth (as formulated in Theorem 3.8), then combining the Lebesgue inequality (3.80) and the Jackson inequality (3.79) we obtain the following error estimate:[3]

$$\|f - \mathscr{P}_n[f]\| < (L_n + 1) C_r \left(\frac{b-a}{2}\right)^r \frac{M}{n^r}, \tag{3.82}$$

which implies that the convergence rate (if there is convergence) will depend on the behavior of L_n when n increases. If the interpolation grid is uniform (equidistant nodes), then the Lebesgue constants grow exponentially as n increases, see inequalities (2.18) of Section 2.1, Chapter 2. In this case, the limit (as $n \longrightarrow \infty$) on the right-hand side of (3.82) is infinite for any finite value of r. This does not necessarily mean that the sequence of interpolating polynomials $\mathscr{P}_n[f](x)$ diverges, because inequality (3.82) only provides an upper bound for the error. It does mean though that in this case convergence of the interpolating polynomials simply cannot be judged using the arguments based on the inequalities of Lebesgue and Jackson.

[3] Note that in estimate (3.82) the function $f(x)$ is assumed to have a maximum of $r-1$ derivatives, and the derivative $f^{(r-1)}(x)$ is required to be Lipshitz-continuous (Theorem 3.8), which basically makes $f(x)$ "almost" r times differentiable. Previously, we have used a slightly different notation and in the error estimate (3.67) the function was assumed to have a maximum of $r+1$ derivatives.

On the other hand, for the Chebyshev interpolation grid (3.64) the following theorem asserts that the asymptotic behavior of the Lebesgue constants is optimal:

THEOREM 3.12 (Bernstein)

Let the interpolation nodes $x_0, x_1, \dots, x_n$ *on the interval* $[-1,1]$ *be given by roots of the Chebyshev polynomial* $T_{n+1}(x)$. *Then,*

$$L_n < 8 + \frac{4}{\pi}\ln(n+1). \tag{3.83}$$

Therefore, according to (3.82) and (3.83), if the derivative $f^{(r-1)}(x)$ of the function $f(x)$ is Lipshitz-continuous, then the sequence of algebraic interpolating polynomials built on the Chebyshev nodes converges uniformly to $f(x)$ with the rate

$$\mathcal{O}(n^{-r}\ln(n+1)) \quad \text{as} \quad n \longrightarrow \infty.$$

Let us now recall that a Lipshitz-continuous function $f(x)$ on the interval $[-1,1]$:

$$|f(x') - f(x'')| \le \text{const}|x' - x''|, \quad x', x'' \in [-1,1],$$

is absolutely continuous on $[-1,1]$, and as such, according to the Lebesgue theorem [KF75], its derivative is integrable, i.e., exists in $L_1[-1,1]$. In this sense, we can say that Lipshitz-continuity is "not very far" from differentiability, although this is, of course, not sufficient to claim that the derivative $f^{(r)}(x)$ is bounded. On the other hand, if $f(x)$ is r times differentiable on $[-1,1]$ and $f^{(r)}(x)$ is bounded, then $f^{(r-1)}(x)$ is Lipshitz-continuous. Therefore, for a function with its r-th derivative bounded, the rate of convergence of Chebyshev interpolating polynomials is at least as fast as the inverse of the grid dimension raised to the power r (smoothness of the interpolated function), times an additional slowly increasing factor $\sim \ln n$. At the same time, recall that the unavoidable error of reconstructing a function with r derivatives on a uniform grid with n nodes is $\mathcal{O}(n^{-r})$. This is also true for the Chebyshev grid, because Chebyshev grid on the diameter is equivalent to a uniform grid on the circle, see Figure 3.4. Consequently, accuracy of the Chebyshev interpolating polynomials appears to be only a logarithmic factor away from the level of the unavoidable error. As such, we have shown that Chebyshev interpolation is not saturated by smoothness and practically reaches the intrinsic accuracy limit.

Altogether, we see that the type of interpolation grid may indeed have a drastic effect on convergence, which corroborates our previous observations. For the Bernstein example $f(x) = |x|$, $-1 \le x \le 1$ (Section 2.1.5 of Chapter 2), the sequence of interpolating polynomials constructed on a uniform grid diverges. On the Chebyshev grid we have seen experimentally that it converges. Now, using estimates (3.82) and (3.83) we can say that the rate of this convergence is at least $\mathcal{O}(n^{-1}\ln(n+1))$.

To conclude, let us also note that strictly speaking the behavior of L_n on the Chebyshev grid is only asymptotically optimal rather than optimal, because the constants in the lower bound (3.81) and in the upper bound (3.83) are different. Better values of these constants than those guaranteed by the Bernstein theorem (Theorem 3.12) have been obtained more recently, see inequality (3.66). However, there is still a gap between (3.81) and (3.66).

REMARK 3.3 Formula (3.78) that introduces the best approximation according to Definition 3.3 can obviously be recast as

$$\varepsilon(f,P_n) = \min_{P_n(x)} \|P_n(x) - f(x)\|_C,$$

where the norm on the right-hand side is taken in the sense of the space $C[a,b]$. In general, the notion of the best approximation admits a much broader interpretation, when both the norm (currently, $\|\cdot\|_C$) and the class of approximating functions (currently, polynomials $P_n(x)$) may be different. In fact, one can consider the problem of approximating a given element of the linear space by linear combinations of a pre-defined set of elements from the same space in the sense of a selected norm. This is the same concept as exploited when introducing the Kolmogorov diameters, see Remark 2.3 on page 41.

For example, consider the space $L_2[a,b]$ of all square integrable functions $f(x)$, $x \in [a,b]$, equipped with the norm:

$$\|f\|_2 \stackrel{\text{def}}{=} \left[\int_a^b f^2(x)dx\right]^{\frac{1}{2}}.$$

This space is known to be a Hilbert space. Let us take an arbitrary $f \in L_2[a,b]$ and consider a set of all trigonometric polynomials $Q_n(x)$ of type (3.6), where $L = b - a$ is the length of the interval. Similarly to the algebraic polynomials $P_n(x)$ employed in Definition 3.3, the trigonometric polynomials $Q_n(x)$ do not have to be interpolating polynomials. Then, it is known that the best approximation in the sense of L_2:

$$\varepsilon_2(f,Q_n) = \min_{Q_n(x)} \|f(x) - Q_n(x)\|_2$$

is, in fact, realized by the partial sum $S_n(x)$, see formula (3.38), of the Fourier series for $f(x)$ with the coefficients defined by (3.40). An upper bound for the actual magnitude of the L_2 best approximation is then given by estimate (3.41) for the remainder $\delta S_n(x)$ of the series, see formula (3.39):

$$\varepsilon_2(f,Q_n) \le \frac{\zeta_n}{n^{r+\frac{1}{2}}}, \quad \text{where} \quad \zeta_n = o(1), \quad n \longrightarrow \infty,$$

where $r+1$ is the maximum smoothness of $f(x)$. Having identified what the best approximation in the sense of L_2 is, we can easily see now that both the Lebesgue inequality (Theorem 3.10) and the error estimate for trigonometric interpolation (Theorem 3.5) are, in fact, justified using the same argument. It employs uniqueness of the corresponding interpolating polynomial, the estimate for the best approximation, and the estimate of sensitivity to perturbations given by the Lebesgue constants. □

Exercises

1. Let the function $f = f(x)$ be defined on an arbitrary interval $[a, b]$, rather than on $[-1, 1]$. Construct the Chebyshev interpolation nodes for $[a, b]$, and write down the interpolating polynomials $P_n(x, f)$ and $\tilde{P}_n(x, f)$ similar to those obtained in Sections 3.2.3 and 3.2.6, respectively.
2. For the function $f(x) = \frac{1}{x^2+1/4}$, $-1 \leq x \leq 1$, construct the algebraic interpolating polynomial $P_n(x, f)$ using roots of the Chebyshev polynomial $T_{n+1}(x)$, see (3.64), as interpolation nodes. Plot the graphs of $f(x)$ and $P_n(x, f)$ for $n = 5, 10, 20, 30$, and 40. Do the same for the interpolating polynomial $P_n(x, f)$ built on the equally spaced nodes $x_k = -1 + 2k/n$, $k = 0, 1, 2, \ldots, n$ (Runge example of Section 2.1.5, Chapter 2). Explain the observable qualitative difference between the two interpolation techniques.
3. Introduce the normalized Chebyshev polynomial $\hat{T}_n(x)$ of degree n by setting $\hat{T}_n(x) = 2^{1-n}T_n(x)$.

   a) Show that the coefficient in front of x^n in the polynomial $\hat{T}_n(x)$ is equal to one.

   b) Show that the deviation $\max\limits_{-1\leq x\leq 1} |\hat{T}_n(x)|$ of the polynomial $\hat{T}_n(x)$ from zero on the interval $-1 \leq x \leq 1$ is equal to 2^{1-n}.

   c)* Show that among all the polynomials of degree n with the leading coefficient (i.e., the coefficient in front of x^n) equal to one, the normalized Chebyshev polynomial $\hat{T}_n(x)$ has the smallest deviation from zero on the interval $-1 \leq x \leq 1$.

   d) How can one choose the interpolation nodes $t_0, t_1, \ldots, t_n$ on the interval $[-1, 1]$, so that the polynomial $(t-t_0)(t-t_1)\ldots(t-t_n)$, which is a part of the formula for the interpolation error (2.23), Chapter 2, would have the smallest possible deviation from zero on the interval $[-1, 1]$?

4. Find a set of interpolation nodes for an even 2π-periodic function $F(\varphi)$, see formula (3.55), for which the Lebesgue constants would coincide with the Lebesgue constants of algebraic interpolation on equidistant nodes.