1 Polynomial approximation and interpolation

1.1 Jackson theorems

1.1.1 Polynomials $P_n$ and trigonometric polynomials $T_n$

In order to state the approximation problem we define the functions which we want to approximate and the functions which we want to use for approximation:

**Definition 1.1** We denote by $C^k[a,b]$ for $k = 0,1,2,\ldots$ the space of functions which have derivatives $f^{(1)},\ldots,f^{(k)}$ that are continuous on the closed interval $[a,b]$.

We denote by

$$P_n = \{ c_0 + c_1 x + \cdots + c_n x^n \mid c_k \in C \}$$

the space of polynomials of degree less or equal to $n$ ($n = 0,1,2,\ldots$).

Then the approximation problem is: Given $f \in C^k[a,b]$, what is the rate with which error of the best approximation

$$\inf_{p_n \in P_n} \|f - p_n\|_\infty$$

converges to zero as $n$ goes to infinity?

The so-called **Jackson theorems** shows that the decay rate of the error depends on the smoothness of the function $f$. E.g. for $f \in C^1[a,b]$ we will prove an approximation rate of $O(1/n)$, and for $f \in C^2[a,b]$ we will obtain an approximation rate of $O(1/n^2)$.

The problem of approximating functions on intervals by polynomials is closely related to the problem of approximating periodic functions by trigonometric polynomials:

**Definition 1.2** The space $C_{2\pi}$ of $2\pi$-periodic functions consists of all functions $f \in C(\mathbb{R})$ which satisfy

$$\forall x \in \mathbb{R} \quad f(x) = f(x + 2\pi).$$

The space of $k$ times continuously differentiable $2\pi$-periodic functions is defined as $C^k = C^k(\mathbb{R}) \cap C_{2\pi}$.

We denote by $T_n$ the space of trigonometric polynomials of degree less or equal to $n$ ($n = 0,1,2,\ldots$):

$$T_n = \{ a_0 + \sum_{k=1}^n (a_k \cos kx + b_k \sin kx) \mid a_k, b_k \in C \}$$

1.1.2 Interpolation with trigonometric polynomials

Note that the space $T_n$ has dimension $2n + 1$ (whereas $P_n$ has dimension $n + 1$). Therefore we can ask whether we can always find a trigonometric interpolation polynomial $p_n \in T_n$ which passes through $2n + 1$ given points $(x_j, y_j)$, $j = 0,\ldots,2n$. We first note that $T_n$ has a property similar to $P_n$:

**Lemma 1.3** A nonzero function $f \in T_n$ has at most $2n$ zeros in $[0,2\pi)$.
Proof: Assume that \( f \in \mathcal{T}_n \) has \( 2n + 1 \) zeros \( \theta_0, \ldots, \theta_{2n} \) in \([0, 2\pi)\). Writing \( \sin k\theta \) and \( \cos k\theta \) in terms of \( e^{ik\theta} \) and \( e^{-ik\theta} \) and with \( z := e^{i\theta} \) we have that

\[
f(\theta) = \sum_{k=-n}^{n} c_k e^{ik\theta} = \sum_{k=-n}^{n} c_k z^k
\]
equals zero for \( z_j := e^{i\theta_j} \). Multiplying by \( z^n \) we obtain that \( \sum_{k=-n}^{n} c_k z^{k+n} \) is a polynomial of degree \( \leq 2n \) in \( z \) which has at least \( 2n + 1 \) zeros \( z_j \in \mathbb{C} \). Hence it must be the zero polynomial and all \( c_k = 0 \).

**Corollary 1.4** Let \( x_0, \ldots, x_{2n} \) be distinct values in \([0, 2\pi)\). Then for any given values \( y_0, \ldots, y_{2n} \) there exists a unique interpolating trigonometric polynomial \( p_n \in \mathcal{T}_n \) which satisfies \( p_n(x_j) = y_j, \ j = 0, \ldots, 2n \).

Proof: The interpolation problem leads to a linear system of \( 2n + 1 \) equations for the \( 2n + 1 \) unknowns \( a_0, \ldots, a_n, b_1, \ldots, b_n \) with the right hand side vector \( (y_0, \ldots, y_{2n})^\top \). This system has a unique solution for every right hand side if the matrix is nonsingular. To show that the matrix is nonsingular consider the problem with the zero right hand side vector. Any solution of this linear system corresponds to a function \( p_n \in \mathcal{T}_n \) which is zero in \( x_0, \ldots, x_{2n} \). By Lemma 1.3 \( p_n \) must be zero. Hence the homogeneous linear system has only the zero solution and the matrix is nonsingular.

### 1.1.3 An auxiliary approximation problem

As a first step toward proving the Jackson theorems let us consider the \( 2\pi \) periodic function \( f \) with \( f(x) = x \) for \( x \in (-\pi, \pi] \). In order to approximate it by a function in \( \mathcal{T}_n \) we can use the interpolation \( p_n \) through the \( 2n + 1 \) nodes \( k\frac{\pi}{n+1}, \ k = -n, \ldots, n \) which exists and is unique due to Corollary 1.4. Since the function \( f \) is odd, i.e., \( f(-x) = -f(x) \) (for \( x \neq k\pi \)), the function \( -p_n(-x) \in \mathcal{T}_n \) is also an interpolation. By the uniqueness of the interpolation, \( p_n(x) \) and \( -p_n(-x) \) must have the same coefficients and so we have that

\[
p_n(x) = \sum_{k=1}^{n} b_k \sin kx. \tag{1.1}
\]

Now we consider the interpolation error \( e(x) := f(x) - p_n(x) \). Since \( p_n \) interpolates \( f \) in \( 2n + 1 \) points in \((-\pi, \pi]\), \( e(x) \) has at least \( 2n + 1 \) simple zeros in \((-\pi, \pi]\). There cannot be more zeros in \((-\pi, \pi]\): If \( e(x) \) has \( 2n + 2 \) zeros in \((-\pi, \pi]\), then \( e'(x) = 1 - p'_n(x) \in \mathcal{T}_n \) has at least \( 2n + 1 \) zeros by Rolle’s theorem. By Corollary 1.3 we then have \( e'(x) = 0 \) which is a contradiction. The same argument also shows that the \( 2n + 1 \) zeros of \( e \) in \((-\pi, \pi]\) are simple, i.e., \( e'(x) \neq 0 \) in those points. Hence the function \( e(x) \) changes its sign in \((-\pi, \pi]\) only at the interpolation points \( k\frac{\pi}{n+1}, \ k = -n, \ldots, n \). Let \( s(x) \) denote the function which is alternatingly 1 and -1 between the nodes, i.e.,

\[
s(x)|_{k\frac{\pi}{n+1} \to (k+1)\frac{\pi}{n+1}} = (-1)^k,
\]
then we have

\[
\int_{-\pi}^{\pi} |f(x) - p_n(x)| \, dx = |\int_{-\pi}^{\pi} (f(x) - p_n(x))s(x) \, dx| = |\int_{-\pi}^{\pi} f(x)s(x) \, dx - \int_{-\pi}^{\pi} p_n(x)s(x) \, dx|.
\]
Let us show that the second term on the right hand side must be 0: Consider
\[ A := \int_{-\pi}^{\pi} s(x) \sin kx \, dx = -\int_{-\pi}^{\pi} s(x + \frac{\pi}{n+1}) \sin kx \, dx = -\int_{-\pi}^{\pi} s(x) \sin k(x - \frac{\pi}{n+1}) \, dx \]
\[ = -\int_{-\pi}^{\pi} s(x) \sin kx \cos(-k \frac{\pi}{n+1}) \, dx - \int_{-\pi}^{\pi} s(x) \cos kx \sin(-k \frac{\pi}{n+1}) \, dx \]
\[ = -\cos(k \frac{\pi}{n+1}) A + 0 \]

Since \( \cos(k \frac{\pi}{n+1}) \neq -1 \) we obtain \( A = 0 \) and also \( \int_{-\pi}^{\pi} p_n(x) s(x) \, dx = 0 \) because of (1.1). It remains to evaluate \( \int_{-\pi}^{\pi} f(x) s(x) \, dx \): It can be easily seen geometrically that we have
\[ \left| \int_{-\pi}^{\pi} x s(x) \, dx \right| = 2 \left| \int_{-\pi}^{0} x s(x) \, dx \right| = 2 \left( \frac{\pi}{n+1} \right)^2 \left| \frac{1}{2} + \frac{3}{2} \cdots \pm (n + \frac{1}{2}) \right| = 2 \left( \frac{\pi}{n+1} \right)^2 \frac{n+1}{2} = \frac{\pi^2}{n+1}. \]

Hence we obtain the following result for the interpolation error:
\[ \int_{-\pi}^{\pi} \left| f(x) - p_n(x) \right| \, dx = \frac{\pi^2}{n+1}. \]

### 1.1.4 Jackson theorems for periodic functions

Now we can prove the Jackson theorem for \( f \in C^1_{2\pi} \): Using integration by parts we see that
\[ \int_{-\pi}^{\pi} \theta f'(\theta + x + \pi) \, d\theta = [\theta f(\theta + x + \pi)]_{-\pi}^{\pi} - \int_{-\pi}^{\pi} 1 f(\theta + x + \pi) \, d\theta = 2\pi f(x) - \int_{-\pi}^{\pi} f(\theta) \, d\theta \]
yielding
\[ f(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) \, d\theta + \frac{1}{2\pi} \int_{-\pi}^{\pi} \theta f'(\theta + x + \pi) \, d\theta \]

In order to approximate \( f(x) \) we replace \( \theta \) in the second integral by the interpolation \( p_n \in \mathcal{T}_n \) from above and obtain
\[ q(x) = a_0 + \frac{1}{2\pi} \int_{-\pi}^{\pi} p_n(\theta) f'(\theta + x + \pi) \, d\theta = a_0 + \frac{1}{2\pi} \int_{-\pi}^{\pi} p_n(\theta - \pi - x) f'(\theta) \, d\theta \quad (1.3) \]
where
\[ p_n(\theta - \pi - x) = \sum_{k=1}^{n} b_k \sin k(\theta - \pi - x) = \sum_{k=1}^{n} b_k \left( \sin k(\theta - \pi) \cos kx - \cos k(\theta - \pi) \sin kx \right). \quad (1.4) \]

If we insert this expression for \( p_n(\theta - \pi - x) \) into the integral in (1.3) we see that the underbraced terms can be pulled out of the integral and that \( q \in \mathcal{T}_n \).
It remains to estimate the approximation error:

\[
\|f - q\|_\infty = \sup_x \left\| \frac{1}{2\pi} \int_{-\pi}^{\pi} (\theta - p_n(\theta)) f'(\theta + x + \pi) \, d\theta \right\| \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |\theta - p_n(\theta)| |f'(\theta + x + \pi)| \, d\theta
\]

\[
\leq \frac{1}{2\pi} \|f'|_\infty \int_{-\pi}^{\pi} |\theta - p_n(\theta)| \, d\theta = \frac{1}{2\pi} \|f'|_\infty \frac{\pi^2}{n + 1}
\]

We have therefore proved the following theorem:

**Theorem 1.5 (Jackson theorem for \(C^1_{2\pi}\))** For \(f \in C^1_{2\pi}\) there holds

\[
\inf_{p \in \mathcal{T}_n} \|f - p\|_\infty \leq \frac{\pi}{2(n + 1)} \|f'|_\infty
\]  

(1.5)

If the function \(f(x)\) has mean 0, i.e., \(\int_{-\pi}^{\pi} f(x) \, dx = 0\), then the approximating function \(q(x)\) has the same property: The constant \(a_0\) in (1.2) is 0, and (1.3), (1.4) show that \(q(x)\) is of the form \(\sum_{k=1}^{n}(a_k \cos kx + b_k \sin kx)\). Let us define \(\mathcal{T}_N^0\) as the space of trigonometric polynomials of degree less or equal \(n\) with mean 0:

\[
\mathcal{T}_N^0 := \{ p \in \mathcal{T}_N \mid \int_{-\pi}^{\pi} p(x) \, dx = 0 \} = \{ \sum_{k=1}^{n}(a_k \cos kx + b_k \sin kx) \mid \text{\(a_k, b_k \in \mathbb{C}\)}\}
\]

We then have

**Lemma 1.6** Let \(f \in C^1_{2\pi}\) with \(\int_{-\pi}^{\pi} f(x) \, dx = 0\). Then

\[
\inf_{p \in \mathcal{T}_N^0} \|f - p\|_\infty \leq \frac{\pi}{2(n + 1)} \|f'|_\infty
\]  

(1.6)

Next, we would like to investigate the approximation rates for functions with more smoothness, e.g., \(f \in C^3_{2\pi}\). We can use the following “bootstrap” argument: Assume that \(f \in C^1_{2\pi}\). Let \(\tilde{p}\) be an arbitrary trigonometric polynomial in \(\mathcal{T}_n\). Then we have obviously

\[
\inf_{p \in \mathcal{T}_n} \|f - p\|_\infty = \inf_{p \in \mathcal{T}_n} \|(f - \tilde{p}) - p\|_\infty \leq \frac{\pi}{2(n + 1)} \|f' - \tilde{p}'\|_\infty
\]  

(1.7)

using (1.5) for \((f - \tilde{p})\). Note that (1.7) holds for any function \(\tilde{p} \in \mathcal{T}_n\), so we can choose it in such a way that the right hand side becomes as small as possible. Since

\[
\{ \tilde{p}' \mid \tilde{p} \in \mathcal{T}_n \} = \mathcal{T}_n^0,
\]

equation (1.7) implies

\[
\inf_{p \in \mathcal{T}_n} \|f - p\|_\infty \leq \frac{\pi}{2(n + 1)} \inf_{p \in \mathcal{T}_n^0} \|f' - p\|_\infty
\]  

(1.8)

For approximation errors in \(\mathcal{T}_n^0\) we can use a similar argument: Assume \(f \in C^1_{2\pi}\) with \(\int_{-\pi}^{\pi} f(x) \, dx = 0\). Using (1.6) and

\[
\{ \tilde{p}' \mid \tilde{p} \in \mathcal{T}_n^0 \} = \mathcal{T}_n^0,
\]

we obtain analogously

\[
\inf_{p \in \mathcal{T}_n^0} \|f - p\|_\infty \leq \frac{\pi}{2(n + 1)} \inf_{p \in \mathcal{T}_n^0} \|f' - p\|_\infty
\]  

(1.9)

This allows us to prove
Theorem 1.7 (Jackson theorem for $C^{k_{2\pi}}$) Let $k \in \{1, 2, \ldots\}$ and $f \in C^{k_{2\pi}}$. Then

$$\inf_{p \in T_n} \| f - p \|_\infty \leq \left( \frac{\pi}{2(n+1)} \right)^k \| f^{(k)} \|_\infty$$

(1.10)

Proof: Note that $f', \ldots, f^{(k-1)}$ have mean 0 since $f$ is periodic. Applying (1.8), (1.9) ($k - 2$ times) and finally (1.6) we get

$$\inf_{p \in T_n} \| f - p \|_\infty \leq \left( \frac{\pi}{2(n+1)} \right)^2 \inf_{p \in T_n} \| f' - p \|_\infty$$

(1.8)

$$\leq \cdots \leq \left( \frac{\pi}{2(n+1)} \right)^k \inf_{p \in T_n} \| f^{(k-1)} - p \|_\infty \leq \left( \frac{\pi}{2(n+1)} \right)^k \| f^{(k)} \|_\infty.$$

\[\square\]

1.1.5 Jackson theorems on an interval

For simplicity we first consider the interval $[-1, 1]$ and consider a function $f \in C^1[-1, 1]$. We can transform this function to a $2\pi$-periodic function $g$ by the change of variables

$$g(\theta) = f(\cos \theta) \quad (1.11)$$

Note that the function $g$ is even. This transformation is useful since polynomials $f(x)$ of degree less or equal $n$ are transformed into even trigonometric polynomials $g(\theta)$ and vice versa. If $g \in T_m$ and $h \in T_n$ then the product $g \cdot h$ is in $T_{m+n}$ as can be seen, e.g., by writing $\cos x$ and $\sin x$ in terms of $e^{\pm ix}$. If $f \in P_n$ then we see that $g$ given by (1.11) is in $T_{n+1}$ and even. It is also clear that $g$ can only be the zero function if $f$ is the zero function. Since $\dim T_n^{\text{even}} = \dim P_n = n+1$ it follows that the transformation (1.11) gives a one-to-one linear mapping between $P_n$ and $T_n^{\text{even}}$.

The functions $\cos kx$, $k = 0, \ldots, n$ form a basis of $T_n^{\text{even}}$. These functions are transformed by (1.11) to certain polynomials $T_k(x)$ with

$$\cos k\theta = T_k(\cos \theta).$$

(1.12)

These are the so-called Chebyshev polynomials. Obviously, $T_0(x) = 1$ and $T_1(x) = x$. If we add the formulae $\cos(k \pm 1)x = \cos kx \cos x \mp \sin kx \sin x$ we obtain $\cos(k + 1)x = 2 \cos kx \cos x - \cos(k - 1)x$ which gives the recursion formula

$$T_{k+1}(x) = 2T_k(x)x - T_{k-1}(x).$$

(1.13)

We can now prove an approximation result for polynomials on an interval:

Theorem 1.8 (Jackson theorem for $C^1[-1, 1]$) For $f \in C^1[-1, 1]$ there holds

$$\inf_{p \in P_n} \| f - p \|_\infty \leq \frac{\pi}{2(n+1)} \| f' \|_\infty$$

(1.14)
Proof: Define \( g \) by (1.11). Then \( g'(\theta) = -f'(\cos \theta) \sin \theta \). Therefore the limits \( \lim_{\theta \to \pm \pi} g'(\theta) \) exist and are equal (both are 0), hence \( g \in C^1_{2\pi} \). We also see that
\[
\|g'\|_\infty \leq \|f'\|_\infty. \tag{1.15}
\]
Consider an approximation \( p \in T_n \) for \( g \). Since \( p \) need not be even we consider the symmetrized function \( \tilde{p}(\theta) = \frac{p(\theta) + p(-\theta)}{2} \). Then we see
\[
\|g(\theta) - \tilde{p}(\theta)\|_\infty = \frac{1}{2}\|g(\theta) + g(-\theta) - p(\theta) - p(-\theta)\|_\infty
\leq \frac{1}{2}(\|g(\theta) - p(\theta)\|_\infty + \|g(-\theta) - p(-\theta)\|_\infty) = \|g(\theta) - p(\theta)\|_\infty
\]
and this implies
\[
\inf_{p \in T_{n,\text{even}}} \|g - p\|_\infty = \inf_{p \in T_n} \|g - p\|_\infty \leq \frac{\pi}{2(n + 1)} \|g'\|_\infty \tag{1.16}
\]
The change of variables \( q(\cos \theta) = p(\theta) \) defines for \( p \in T_{n,\text{even}} \) a function \( q \in \mathcal{P}_n \) such that
\[
\max_{x \in [-1,1]} |f(x) - q(x)| = \max_{\theta \in [-\pi,\pi]} |g(\theta) - p(\theta)|. \tag{1.17}
\]
Equations (1.17), (1.16), and (1.15) together yield (1.14). \( \Box \)

In the same way as we proved Theorem 1.7 using Theorem 1.5 we can use Theorem 1.8 to prove an approximation result for \( f \in C^k[-1,1] \):

**Theorem 1.9 (Jackson theorem for \( C^k[-1,1] \))** Let \( n, k \) be integers with \( n \geq k - 1 \geq 0 \) and \( f \in C^k[-1,1] \). Then
\[
\inf_{p \in \mathcal{P}_n} \|f - p\|_\infty \leq \left(\frac{\pi}{2}\right)^k \frac{1}{(n + 1)n \cdots (n - k + 2)} \|f^{(k)}\|_\infty \tag{1.18}
\]

### 1.2 Polynomial interpolation

The proofs for the Jackson theorems are constructive. But the constructions of the approximating functions are of no practical value since they are very complicated, involving various differentiations and integrations (especially for higher values of \( k \)) and are therefore not very efficient.

In this section we will see that we can achieve an approximation which is almost as good as the optimal approximation using interpolation, provided the interpolation nodes are correctly chosen.

#### 1.2.1 The error formula and Chebyshev nodes

**Theorem 1.10** Let \( f \in C^{n+1}[a,b] \) and let \( p_n \in \mathcal{P}_n \) be the interpolating polynomial through the distinct interpolation nodes \( x_0, \ldots, x_n \in [a,b] \). Then for \( t \in [a,b] \) there exists \( \xi \in [a,b] \) such that
\[
f(t) - p_n(t) = \frac{f^{(n+1)}(\xi)}{(n + 1)!} (t - x_0) \cdots (t - x_n) \tag{1.19}
\]
Proof: Let \( \omega(x) := (x-x_0) \cdots (x-x_n) \). Fix \( t \in [a,b] \) with \( t \neq x_j, j = 0, \ldots, n \) (otherwise (1.19) is trivial). Then \( \omega(t) \neq 0 \) and the definitions

\[
A_t := (f(t) - p_n(t))/\omega(t), \quad g(x) := f(x) - p_n(x) - A_t \omega(x)
\]

imply that \( g(t) = 0 \). Since also \( g(x_j) = 0 \) \( (p_n \) interpolates \( f \) and \( \omega(x_j) = 0 \)) we see that \( g(x) \) has at least \( n+2 \) zeros in \( [a,b] \). By Rolle’s theorem, \( g' \) has at least \( n+1 \) zeros, \( \ldots \), \( g^{(n+1)} \) has at least 1 zero which we will call \( \xi \). Therefore

\[
0 = g^{(n+1)}(\xi) = f^{(n+1)}(\xi) - p_n^{(n+1)}(\xi) - A_t \omega^{(n+1)}(\xi) = f^{(n+1)}(\xi) - 0 - A_t (n+1)!
\]

since \( p_n \in P_n \) and \( \omega(x) = x^{n+1} + q(x), q \in P_n \). Therefore \( A_t = f^{(n+1)}(\xi)/(n+1)! \) which together with \( 0 = g(t) = f(t) - p_n(t) - A_t \omega(t) \) yields (1.19).

Estimate (1.19) implies

\[
\|f - p_n\|_\infty \leq \frac{1}{(n+1)!} \|f^{(n+1)}\|_\infty \|\omega(x)\|_\infty \quad (1.20)
\]

where \( \omega(x) \) denotes the node polynomial

\[
\omega(x) := (x-x_0) \cdots (x-x_n).
\]

We see that only the last factor in (1.20) depends on the choice of nodes. Therefore it is a good idea to choose the \( n+1 \) nodes in \( [a,b] \) in such a way that \( \|\omega(x)\|_\infty \) becomes minimal.

Let us first consider the interval \([-1,1]\). The Chebyshev polynomial \( T_k(x) \) has by its definition (1.12) the property that it oscillates between 1 and \( \pm 1 \) and has all its \( k \) zeros

\[
x_j = \cos \left( \frac{j+1/2}{k} \pi \right), \quad j = 0, \ldots, k-1
\]

in the interval \([-1,1]\). Furthermore, by the recursion formula (1.13) the polynomial \( T_k \) has the leading coefficient \( 2^{k-1} \) for \( k \geq 1 \). Hence

\[
T_k(x) = 2^{k-1}(x-x_0) \cdots (x-x_k). \quad (1.21)
\]

If we choose the \( n+1 \) interpolation nodes as the zeros of the polynomial \( T_{n+1} \) we have therefore

\[
\omega(x) = 2^{-n}T_{n+1}(x) \quad (1.22)
\]

and because of \( \|T_k\|_\infty = 1 \) we obtain

\[
\|\omega\|_\infty = 2^{-n}.
\]

One cannot achieve a smaller value of \( \|\omega\|_\infty \) than \( 2^{-n} \) with any arrangement of nodes \( x_0, \ldots, x_n \): Assume \( \|\tilde{\omega}\|_\infty < 2^{-n} \). Since \( \omega \) alternatingly assumes \( n+2 \) extrema \( 2^{-n} \) and \( -2^{-n} \), the polynomial \( q := \omega - \tilde{\omega} \) must have a zero between two subsequent extrema of \( \omega \) by the intermediate value theorem. Thus \( q = \omega - \tilde{\omega} \) has at least \( n+1 \) zeros. But since both \( \omega \) and \( \tilde{\omega} \) are of the form \( x^{n+1} + \) lower order terms we see that \( q \in P_n \) and hence \( q = 0 \) which contradicts the assumption.
Therefore using the Chebyshev nodes \( x_k = \cos\left(\frac{k+1/2}{n+1} \pi\right), \ k = 0, \ldots, n \) leads to the best estimate in (1.20). Using a linear transformation, we see that the best choice of nodes for the interval \([a, b]\) is

\[
\frac{a + b}{2} + \cos\left(\frac{k + 1/2}{n + 1} \pi\right) \frac{a - b}{2}, \quad k = 0, \ldots, n
\]

yielding \( \|\omega\|_\infty = 2^{-n} \left(\frac{b-a}{2}\right)^{n+1} = 2 \left(\frac{b-a}{4}\right)^{n+1} \) and

\[
\|f - p_n\|_\infty \leq 2 \left(\frac{b-a}{4}\right)^{n+1} \frac{\|f^{(n+1)}\|_\infty}{(n+1)!}
\]

For an arbitrary choice of nodes we have using (1.20) and \(|x - x_j| \leq (b-a)\) the estimate

\[
\|f - p_n\|_\infty \leq (b-a)^{n+1} \frac{\|f^{(n+1)}\|_\infty}{(n+1)!}.
\]

These formulas show that the interpolation error converges to 0 if the derivatives do not grow too fast. For example, assume that the function \( f \) is analytic on \([a, b]\), i.e. the Taylor series converges for every \( x \in [a, b] \) for all \( y \) with \(|y - x| < \rho\) with some \( \rho > 0 \). Equivalently, the function \( f \) can be extended to a holomorphic function in all points in \( C \) which have a distance of less than \( \rho \) to \([a, b]\) (\( \rho \) can be chosen as the distance of the closest singularity of \( f \) in \( C \) to the interval \([a, b]\)). Then we know (e.g., expressing \( f^{(n)} \) with the Cauchy theorem) that for any \( \hat{\rho} < \rho \) there exists \( C \) such that

\[
\text{for all } n \geq 0 \quad \frac{1}{n!} \max_{x \in [a, b]} |f^{(n)}(x)| \leq C \frac{1}{\hat{\rho}^n}.
\]

If \((b-a) < \rho\) then we see that we have exponential convergence for any choice of nodes.

But the formulae (1.24) and (1.25) do not give any information about the convergence if \( f \) is less smooth. In fact, interpolation with equally spaced nodes can diverge even if \( f \) is analytic on \([a, b]\), as the Runge example \( f(x) = \frac{1}{1+x^2} \) on \([-5, 5]\) shows. (Here we have singularities at \( \pm i \) and hence \( \rho = 1 \).) On the other hand, interpolation at the Chebyshev nodes converges if \( f \) is only marginally smoother than a continuous function: The so-called Dini-Lipschitz condition

\[
\log \varepsilon \max_{|x-y| \leq \varepsilon, \ x, y \in [a, b]} |f(x) - f(y)| \to 0 \quad \text{for } \varepsilon \to 0
\]

is sufficient. This includes continuous, piecewise differentiable functions, as well as functions like \( \sqrt{|x|} \). However, there exist continuous functions for which the Chebyshev approximations do not uniformly converge. But one can show that this is true for any sequence of interpolation nodes.

### 1.2.2 The Lagrange form and the Lebesgue function

Let \( x_0, \ldots, x_n \) be distinct nodes in the interval \([a, b]\). For a function \( f \in C[a, b] \) let us denote the interpolating polynomial in \( \mathcal{P}_n \) by \( P_n f \). We can give an explicit formula for this polynomial: The **Lagrange polynomials**

\[
l_k(x) := \prod_{\substack{j=0, \ldots, n \ \text{if} \ j \neq k}} \frac{x - x_j}{x_k - x_j}
\]
obviously satisfy
\[ l_k \in \mathcal{P}_n, \quad l_k(x_j) = \begin{cases} 0 & \text{for } k \neq j \\ 1 & \text{for } k = j. \end{cases} \] (1.26)

Therefore we can write the interpolating polynomial in the so-called Lagrange form
\[ P_n f = \sum_{k=0}^{n} f(x_k) l_k. \]

Then we get the following estimate
\[ |(P_n f)(x)| = \left| \sum_{k=0}^{n} f(x_k) l_k(x) \right| \leq (\max_{k=0,\ldots,n} |f(x_k)|) \sum_{k=0}^{n} |l_k(x)| \leq \|f\|_{\infty} \lambda_n(x) \]
where \( \lambda_n(x) := \sum_{k=0}^{n} |l_k(x)| \) is the so-called Lebesgue function. We also have
\[ \|P_n f\|_{\infty} \leq \|\lambda_n\|_{\infty} \|f\|_{\infty} \] (1.27)
and this estimate is sharp, i.e., there exist \( f \in C[a,b] \) where we have equality in (1.27). Since \( \lambda_n(x_k) = 1 \) by (1.26) there holds \( \|\lambda_n\|_{\infty} \geq 1 \). In order to have a good approximation we would like to choose the interpolation points in such a way that \( \|\lambda_n\|_{\infty} \) does not become too large. Otherwise there exist functions with small values which have interpolating polynomials with very large values.

The size of the Lebesgue constant \( \|\lambda_n\|_{\infty} \) also characterizes the relation between the interpolation error and the best possible approximation error: Let \( q_n \in \mathcal{P}_n \) be an arbitrary polynomial, then \( P_n q = q \) and
\[ \|f - P_n f\|_{\infty} \leq \|f - q\|_{\infty} + \|q - P_n f\| = \|f - q\|_{\infty} + \|P_n(q - f)\|_{\infty} \leq \|f - q\|_{\infty} + \|\lambda_n\|_{\infty} \|f - q\|_{\infty} \]
Therefore
\[ \|f - P_n f\|_{\infty} \leq (1 + \|\lambda_n\|_{\infty}) \inf_{q \in \mathcal{P}_n} \|f - q\|_{\infty} \] (1.28)

1.2.3 Estimates for the Lebesgue constant

For uniformly spaced interpolation points one can prove
\[ \|\lambda_n\|_{\infty} \geq Ce^{n/2} \]
This is illustrated by the Runge example where \( \|P_n f\|_{\infty} \) grows exponentially. Unfortunately one can never achieve a bounded Lebesgue constant: For any sequence of interpolation nodes there holds
\[ \|\lambda_n\|_{\infty} \geq \frac{2}{\pi} \log n - c. \]
However, the Chebyshev nodes give a Lebesgue constant which has the same growth rate as this lower bound: For Chebyshev nodes we have

**Theorem 1.11** The Lebesgue function for Chebyshev nodes (1.23) satisfies
\[ \|\lambda_n\|_{\infty} \leq \frac{2}{\pi} \log(n + 1) + \frac{4}{\pi} \] (1.29)
Proof: Note that we can write the Lagrange polynomial \( l_k \) in terms of the node polynomial \( \omega(x) = (x - x_0) \cdots (x - x_n) \) as

\[
l_k(x) = \frac{\omega(x)}{(x - x_k)\omega'(x_k)}
\]

Using (1.22) and (1.12) and the change of variables \( x = \cos \theta, \; x_k = \cos \theta_k \) with \( \theta, \theta_k \in [0, \pi] \) we have

\[
l_k(\cos \theta) = \frac{\cos(n + 1) \theta}{(\cos \theta - \cos \theta_k)(n + 1) \sin(n + 1) \theta_k \frac{d\theta}{dx}|_{x=x_k}}
\]

Because of \( \sin(n + 1) \theta_k = \pm 1 \) and \( \frac{dx}{d\theta} = -\sin \theta \) this yields

\[
\mu(\theta) := \lambda_n(\cos \theta) = \frac{|\cos(n + 1) \theta|}{n + 1} \sum_{k=0}^{n} \left| \frac{\sin \theta_k}{\cos \theta - \cos \theta_k} \right| \tag{1.30}
\]

or, using the trigonometric identity \( \sin a/(\cos a - \cos b) = (\cot \frac{a-b}{2} - \cot \frac{a+b}{2})/2 \)

\[
\mu(\theta) = \frac{|\cos(n + 1) \theta|}{2(n + 1)} \sum_{k=0}^{n} \left| \cot \frac{\theta + \theta_k}{2} - \cot \frac{\theta - \theta_k}{2} \right| \tag{1.31}
\]

We want to show that \( \max_{x \in [-1,1]} \lambda_n(x) = \max_{\theta \in [0,\pi]} \mu(\theta) \) attains its maximal value in the interval \( [0, \theta_0] = [0, \frac{\pi}{2(n+1)}] \). Let \( \phi \in [0, \pi] \) arbitrary, then we can write \( \phi = \phi_0 + l\pi/(n+1) \) with \( \phi_0 \in [-\frac{\pi}{2(n+1)}, \frac{\pi}{2(n+1)}] \) and integer \( l \). Since \( |\cos t| \) is a \( \pi \)-periodic function we have \( |\cos(n + 1) \phi| = |\cos(n + 1) \phi_0| \). For the sum we claim

\[
\sum_{k=0}^{n} \left| \cot \frac{\phi + \theta_k}{2} - \cot \frac{\phi - \theta_k}{2} \right| \leq \sum_{k=0}^{n} \left| \cot \frac{\phi + \theta_k}{2} \right| + \left| \cot \frac{\phi - \theta_k}{2} \right| = \sum_{k=0}^{n} \left| \cot \frac{\phi_0 + \theta_k}{2} - \cot \frac{\phi_0 - \theta_k}{2} \right| \tag{1.32}
\]

The \( 2(n+1) \) summands in the second and third sum are the same, but in different combinations since \( \cos \) is \( \pi \)-periodic and

\[
\{ \phi + \theta_k, \phi - \theta_k \mid k = 0, \ldots, n \} = \{ \phi + \frac{\pi}{2(n+1)} + k \frac{\pi}{n+1} \mid k = -n-1, \ldots, n \}.
\]

The last equality in (1.32) follows from

\[
0 \leq \frac{\phi_0 + \theta_k}{2} \leq \frac{\pi}{2}, \quad -\frac{\pi}{2} \leq \frac{\phi_0 - \theta_k}{2} \leq 0 \quad \Rightarrow \quad \cot \frac{\phi_0 + \theta_k}{2} \geq 0, \quad \cot \frac{\phi_0 - \theta_k}{2} \leq 0
\]

Therefore we have (using that \( \mu \) is an even function)

\[
\mu(\phi) \leq \mu(\phi_0) = \mu(|\phi_0|) \quad \Rightarrow \quad \max_{\phi \in [0,\pi]} \mu(\phi) = \max_{\phi_0 \in [0,\pi]} \mu(\phi_0).
\]

Now we consider \( \theta \in [0, \frac{\pi}{2(n+1)}] \). Then \( \cos(n + 1) \theta \geq 0, \cos \theta - \cos \theta_k \geq 0 \) and hence we obtain

\[
\mu(\theta) = \lambda_n(\cos \theta) = \frac{1}{n + 1} \sum_{k=0}^{n} \sin \theta_k \frac{\cos(n + 1) \theta}{\cos \theta - \cos \theta_k} = \frac{1}{n + 1} \sum_{k=0}^{n} (\sin \theta_k) 2^n \prod_{j=0 \atop j \neq k}^{n} (\cos \theta - \cos \theta_j)
\]
using (1.21). Since $\cos \theta - \cos \theta_j \geq 0$, each term in the sum is a nonnegative decreasing function of $\theta$, therefore
\[
\|\lambda_n\|_\infty = \max_{\theta \in [0, \pi/2]} \mu(\theta) = \mu(0) \overset{(1.31)}{=} \frac{1}{2(n+1)} \sum_{k=0}^{n} 2 \cot \frac{\theta_k}{2}
\] (1.33)

We see that the last expression is exactly the composite midpoint rule applied to the integral $\frac{1}{\pi} \int_0^\pi \cot(t/2) \, dt$ (which is infinite). We split off the first term and note that the integrand is convex ($f'' \geq 0$), hence the integral is larger than the midpoint rule approximation:
\[
\|\lambda_n\|_\infty = \frac{1}{n+1} \cot \frac{\theta_0}{2} + \frac{1}{n+1} \sum_{k=1}^{n} \cot \frac{\theta_k}{2} \leq \frac{1}{n+1} \cot \frac{\pi}{4(n+1)} + \frac{1}{\pi} \int_{\pi/(n+1)}^{\pi} \cot \frac{t}{2} \, dt
\]
\[
= \frac{1}{n+1} \cot \frac{\pi}{4(n+1)} + \frac{1}{\pi} \left[ 2 \log \left( \tan \frac{t}{2} \right) \right]_{\pi/(n+1)}^{\pi} = \frac{1}{n+1} \cot \frac{\pi}{4(n+1)} - \frac{2}{\pi} \log \left( \sin \frac{\pi}{2(n+1)} \right)
\]

Finally the inequalities $\tan t \geq t$ ($0 \leq t < \pi/2$) and $\sin t \geq \frac{2}{\pi} t$ ($0 \leq t \leq \pi/2$) give
\[
\|\lambda_n\|_\infty \leq \frac{4}{\pi} + \frac{2}{\pi} \log(n+1)
\] \hfill \Box

Table 1 shows the values of $\|\lambda_n\|_\infty$ computed directly from (1.33) and the upper bounds

<table>
<thead>
<tr>
<th>$n$</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>25</th>
</tr>
</thead>
<tbody>
<tr>
<td>$|\lambda_n|_\infty$</td>
<td>2.104</td>
<td>2.489</td>
<td>2.728</td>
<td>2.901</td>
<td>3.037</td>
</tr>
<tr>
<td>$\frac{2}{\pi} \log(n+1) + \frac{4}{\pi}$</td>
<td>2.414</td>
<td>2.800</td>
<td>3.038</td>
<td>3.211</td>
<td>3.347</td>
</tr>
</tbody>
</table>

Table 1: Lebesgue constants for Chebyshev nodes

from (1.29).

Because of (1.28) the interpolation $P_n f$ at Chebyshev nodes satisfies for $n \leq 20$
\[
\|f - P_n f\|_\infty \leq 4 \inf_{q \in P_n} \|f - q\|_\infty.
\]

Therefore the Chebyshev interpolation gives almost the best possible approximation. Trying to find the best possible approximation (which can only be done approximately by an iterative procedure) can at most decrease the error by a factor of 4.

One can do even better by using the expanded Chebyshev nodes
\[
\tilde{x}_j := x_j / x_0
\]
which stretch the Chebyshev nodes so that the endpoints $\tilde{x}_0 = 1$ and $\tilde{x}_n = -1$ are nodes. In this case we have $\|\lambda_n\|_\infty \leq 2.01$ for $n \leq 9$ and $\|\lambda_n\|_\infty \leq 3$ for $n \leq 47$. Furthermore one can show that $\|\lambda_n\|_\infty$ is never more than 0.02 larger than the best possible value of the Lebesgue constant with $n+1$ nodes.
2 PIECEWISE POLYNOMIAL INTERPOLATION

2 Piecewise polynomial interpolation

2.1 Introduction

One problem with polynomial or trigonometric interpolation (even with a good choice of interpolation nodes) is the following: If the function is not smooth at an isolated point (e.g., \( f'(x) \) is discontinuous) then this introduces large oscillations and pollutes the error everywhere (Gibbs phenomenon).

This is related to the fact that algebraic and trigonometric polynomials are globally defined functions, and the Lagrange polynomials \( l_i(x) \) have oscillations which decay only slowly away from \( x_i \). Often one would prefer an approximation which shows a more “local” behavior. This is possible if one uses approximating functions which are no longer analytic functions everywhere, but which are piecewise defined and have only a limited degree of smoothness across the so-called breakpoints.

2.2 Piecewise linear functions

2.2.1 Interpolation with piecewise linears

A simple example for piecewise polynomial interpolation is given by piecewise linear functions. We define on the interval \([a, b]\) the breakpoints \( x_0 = a < x_1 < \cdots < x_{N-1}, x_N = b \) and define the space \( S^2_N \) of functions which are linear on each subinterval \( I_j := [x_{j-1}, x_j], j = 1, \ldots, N \) and continuous on \([a, b]\). Let us define the lengths of the intervals \( h_i := x_i - x_{i-1}, i = 1, \ldots, N \) and the largest length \( h_{\text{max}} := \max_{i=1,\ldots,N} h_i \). For a given function \( f \in C[a, b] \) we let \( I^N_2 f \in S^2_N \) be the function which interpolates \( f \) in \( x_0, \ldots, x_N \). If \( f \in C^2[a, b] \), then Theorem 1.10 implies the estimates

\[
\| f - I^N_2 f \|_{\infty, I_j} \leq \frac{h_j^2}{8} \| f'' \|_{\infty, I_j}
\]

\[
\| f - I^N_2 f \|_{\infty} \leq \frac{1}{8} \max_{j=1,\ldots,N} \left( h_j^2 \| f'' \|_{\infty, I_j} \right) \leq \frac{h_{\text{max}}^2}{8} \| f'' \|_{\infty}
\]

(2.1)

(2.2)

If we use the uniform mesh \( x_j = a + j(b-a)/N, j = 0, \ldots, N \), we therefore obtain convergence \( \| f - I^N_2 f \|_{\infty} \leq O(N^{-2}) \) for \( N \to \infty \).

Note that we can write (analogous to the Lagrange form of interpolating polynomials) the interpolating polynomials as a linear combination of the “hat functions” \( \phi_j \in S^2_N \) which satisfy

\[
\phi_j(x_k) = \begin{cases} 1 & \text{for } j = k \\ 0 & \text{for } j \neq k \end{cases}
\]

Then the interpolating polynomial \( p = I^N_2 f \) can be written as

\[
p(x) = \sum_{j=0}^{N} f(x_j)\phi_j(x).
\]
2.2.2 Least squares approximation with piecewise linears

Let us consider a different approximation method with piecewise linear polynomials: We define for functions \( f, g \in C[a, b] \) the inner product \( \langle f, g \rangle \) and the \( L_2 \)-norm \( \| f \|_2 \) by

\[
\langle f, g \rangle = \int_a^b f(x)\overline{g(x)}\, dx, \quad \| f \|_2 = \langle f, f \rangle^{1/2} = \left( \int_a^b |f(x)|^2\, dx \right)^{1/2}.
\]

The least squares approximation \( p \in S^N_2 \) of \( f \) is characterized by the so-called orthogonality conditions or normal equations

\[
\langle f - p, q \rangle = 0, \quad \text{for all } q \in S^N_2. \tag{2.3}
\]

If \( p \) satisfies this condition, then it minimizes the error \( \| f - p \|_2 \): Replacing \( p \) by \( p + q \) with \( q \in S^N_2 \) gives

\[
\| f - (p + q) \|_2^2 = \langle f - p - q, f - p - q \rangle \tag{2.3} = \langle f - p, f - p \rangle + \langle q, q \rangle \geq \| f - p \|_2^2
\]

The normal equations always have a unique solution: Using the basis \( \{ \phi_0, \ldots, \phi_N \} \), (2.3) is equivalent to

\[
\langle f - p, \phi_k \rangle = 0 \quad \text{for } k = 0, \ldots, N,
\]

and expressing \( p \) as \( p = \sum_{j=0}^N \alpha_j \phi_j \) we obtain a linear system of \( N + 1 \) equations for \( N + 1 \) unknowns \( \alpha_0, \ldots, \alpha_N \):

\[
\sum_{j=0}^N \langle \phi_j, \phi_k \rangle \alpha_j = \langle f, \phi_k \rangle \quad \text{for } k = 0, \ldots, N \tag{2.4}
\]

A solution of the homogeneous linear system corresponds to a function \( p \in S^N_2 \) with \( \langle p, q \rangle = 0 \) for all \( q \in S^N_2 \) and hence \( 0 = \langle p, p \rangle = \| p \|_2^2 \) and \( p = 0 \). The unique solution \( p \in S^N_2 \) of the normal equations is also known as the \( L_2 \)-projection of \( f \) onto \( S^N_2 \), we will denote it by \( \Pi^N_2 f \).

Let us compute the entries \( \langle \phi_j, \phi_k \rangle \) of the matrix in the linear system (2.4): Obviously, \( \langle \phi_j, \phi_k \rangle = 0 \) if \( |j - k| \geq 2 \) since \( \phi_j \) and \( \phi_k \) have disjoint support. Therefore the matrix is tridiagonal. As

\[
\phi_j(x) = \begin{cases} \frac{x-x_{j-1}}{h_j} & \text{for } x_{j-1} \leq x \leq x_j \\ \frac{x_j-x}{h_{j+1}} & \text{for } x_{j-1} \leq x \leq x_j \\ 0 & \text{otherwise} \end{cases}
\]

and

\[
\int_{x_{j-1}}^{x_j} \left( \frac{x-x_{j-1}}{h_j} \right)^2 dx = h_j \int_0^1 t^2\, dt = h_j \frac{1}{3}, \quad \int_{x_{j-1}}^{x_j} \frac{x_j-x-x_{j-1}}{h_j} \, dx = h_j \int_0^1 (1-t)t\, dt = h_j \frac{1}{6}
\]

we obtain with \( h_0 := 0, h_{N+1} := 0 \)

\[
\langle \phi_j, \phi_j \rangle = \frac{h_j + h_{j+1}}{3} \quad \text{for } j = 0, \ldots, N \quad \langle \phi_{j-1}, \phi_j \rangle = \frac{h_j}{6} \quad \text{for } j = 1, \ldots, N. \tag{2.5}
\]
and the linear system (2.4) becomes

\[
\frac{1}{3} \begin{pmatrix}
  h_1 & h_1/2 & 0 \\
  h_1/2 & h_1 + h_2 & h_2/2 \\
  & & \ddots \\
  & & & h_{N-1}/2 & h_{N-1} + h_N & h_N/2 \\
  0 & & & & h_N/2 & h_N \\
\end{pmatrix}
\begin{pmatrix}
  \alpha_0 \\
  \alpha_1 \\
  \vdots \\
  \alpha_{N-1} \\
  \alpha_N \\
\end{pmatrix} = \begin{pmatrix}
  \langle f, \phi_0 \rangle \\
  \langle f, \phi_1 \rangle \\
  \vdots \\
  \langle f, \phi_{N-1} \rangle \\
  \langle f, \phi_N \rangle \\
\end{pmatrix}
\]

(2.6)

We see that the matrix is strictly diagonally dominant. Therefore we can solve it using Gaussian elimination without pivoting.

The diagonal dominance of the matrix yields the following estimate:

**Lemma 2.1** Let \( f \in C[a, b] \). Then the least squares approximation \( \Pi_N^2 f \) of \( f \) in \( S_N^2 \) satisfies

\[
\| \Pi_N^2 f \|_\infty \leq 3 \| f \|_\infty
\]

(2.7)

Proof: Let \( p = \Pi_N^2 f = \sum_{j=0}^N \alpha_j \phi_j \). The normal equations (2.6) have the form

\[
\frac{1}{3} h_j \alpha_{j-1} + \frac{1}{3} (h_j + h_{j+1}) \alpha_j + \frac{1}{3} h_{j+1} \alpha_{j+1} = \langle f, \phi_j \rangle
\]

(2.8)

if we define \( h_0 = h_{N+1} = 0 \). Solving this equation for \( \alpha_j \) gives

\[
\frac{1}{3} (h_j + h_{j+1}) |\alpha_j| \leq |\langle f, \phi_j \rangle| + \frac{1}{3} h_j |\alpha_{j-1}| + \frac{1}{3} h_{j+1} |\alpha_{j+1}| \leq \| f \|_\infty \frac{1}{2} (h_j + h_{j+1}) + \frac{1}{6} (h_j + h_{j+1}) \| \alpha \|_\infty.
\]

since \( |\langle f, \phi_j \rangle| \leq \| f \|_\infty \langle 1, \phi_j \rangle = \| f \|_\infty \frac{1}{2} (h_j + h_{j+1}) \). Therefore

\[
\frac{1}{3} \| \alpha \|_\infty \leq \frac{1}{2} \| f \|_\infty + \frac{1}{6} \| \alpha \|_\infty
\]

which implies (2.7).

We can get more localized information by using weights:

**Lemma 2.2** Let \( f \in C[a, b] \) and let \( p := \Pi_N^2 f = \sum_{j=0}^N \alpha_j \phi_j \) denote the least squares approximation of \( f \) in \( S_N^2 \).

1. Let \( \tilde{w}_j > 0 \) for \( j = 0, \ldots, N \) and satisfy \( r^{-1} \leq \frac{\tilde{w}_j}{\tilde{w}_{j-1}} \leq r \) for \( j = 1, \ldots, N \) with \( r < 2 \). Then we have with \( w_j := \max\{ \tilde{w}_{j-1}, \tilde{w}_j \} \)

\[
\max_{j=0, \ldots, N} (\tilde{w}_j |\alpha_j|) \leq \frac{3}{2 - r} \max_{j=1, \ldots, N} (w_j \| f \|_{\infty, I_j})
\]

(2.9)

2. Let \( w_j > 0 \) for \( j = 1, \ldots, N \) and satisfy \( r^{-1} \leq \frac{w_j}{w_{j-1}} \leq r \) for \( j = 2, \ldots, N \) with \( r < 2 \). Then

\[
\max_{j=1, \ldots, N} (w_j \| p \|_{\infty, I_j}) \leq \frac{3r}{2 - r} \max_{j=1, \ldots, N} (w_j \| f \|_{\infty, I_j})
\]

(2.10)
Proof: Multiplying (2.8) by $\tilde{w}_j$ gives

$$1/6 h_j \frac{\tilde{w}_j}{\tilde{w}_{j-1}} \tilde{w}_{j-1} \alpha_{j-1} + 1/3 (h_j + h_{j+1}) \tilde{w}_j \alpha_j + 1/6 h_{j+1} \frac{\tilde{w}_j}{\tilde{w}_{j+1}} \tilde{w}_{j+1} \alpha_{j+1} = \tilde{w}_j \langle f, \phi_j \rangle \quad (2.11)$$

Using $\tilde{w}_{j-1} \leq r$, $\tilde{w}_j / \tilde{w}_{j+1} \leq r$ we obtain

$$\frac{1}{3} \tilde{w}_j |\alpha_j| \leq \frac{1}{2} \tilde{w}_j \|f\|_{\infty, I_j \cup I_{j+1}} + \frac{r}{6} \max_{k=0, \ldots, N} (\tilde{w}_k |\alpha_k|)$$

with $I_0 = I_{N+1} := \emptyset$ which implies (2.10).

For (2.10) we use (2.9) with $(\tilde{w}_0, \ldots, \tilde{w}_N) = (w_1, w_2, \ldots, w_N)$ and then with $(w_1, w_2, \ldots, w_N, w_N)$. We take the maximum of the two estimates and use $w_j / w_{j-1} \leq r$. \[\square\]

**Corollary 2.3** Assume $f \in C[a, b]$. Let $w_j > 0$ for $j = 1, \ldots, N$ so that $r^{-1} \leq w_j / w_{j-1} \leq r$ with $r < 2$ and define

$$\|g\|_w := \max_{j=1, \ldots, N} (w_j \|g\|_{\infty, I_j}).$$

Then

$$\|f - \Pi^N_2 f\|_{\infty} \leq 4 \inf_{q \in S_2^N} \|f - q\|_{\infty} \quad (2.12)$$

$$\|f - \Pi^N_2 f\|_w \leq \frac{2 + 2r}{2 - r} \inf_{q \in S_2^N} \|f - q\|_w \quad (2.13)$$

If $f \in C^2[a, b]$ then

$$\|f - \Pi^N_2 f\|_{\infty} \leq \frac{1}{2} \max_{j=1, \ldots, N} \left( h_j^2 \|f''\|_{\infty, I_j} \right) \leq \frac{h^2_{\max}}{2} \|f''\|_{\infty} \quad (2.14)$$

$$\|f - \Pi^N_2 f\|_w \leq \frac{1 + r}{8 - 4r} \max_{j=1, \ldots, N} \left( w_j h_{\max}^2 \|f''\|_{\infty, I_j} \right) \leq \frac{1 + r}{8 - 4r} h_{\max}^2 \|f''\|_w \quad (2.15)$$

Proof: Equation (2.12) follows from $\Pi^N_2 q = q$ for $q \in S_2^N$ and

$$\|f - \Pi^N_2 f\|_{\infty} \leq \|f - q\|_{\infty} + \|\Pi^N_2 f - q\|_{\infty} = \|f - q\|_{\infty} + \|\Pi^N_2 (f - q)\|_{\infty} \quad (2.7)$$

Using $q := I^N_2$ and (2.1) gives equation (2.14). The other estimates follow in the same way using $\|\cdot\|_w$. \[\square\]

Note that the $L^2$-projection on $\Pi^N_2$ is not a strictly local approximation method: Even if the function $f$ is nonzero only in a small part of the interval $[a, b]$, the approximation $\Pi^N_2 f$ will in general be nonzero everywhere in $[a, b]$. However, the function $\Pi^N_2 f$ will decrease very rapidly away from the support of $f$. 

2.3 Cubic Splines

The space $S_N^1$ of piecewise constant (discontinuous) functions and the space $S_N^2$ of piecewise linear, continuous functions are the two simplest examplest of so-called smoothest spline spaces. These are piecewise polynomials with a maximal degree of smoothness across the breakpoints: Let $x_0, \ldots, x_N$ be a set of breakpoints. Then we denote by $S_k^N$ the space of functions which are polynomials of degree $\leq k - 1$ in each interval and which are contained in $C^{k-2}[a, b]$ for $k \geq 2$. E.g., $S_3^N$ is the space of piecewise quadratics which have continuous derivatives across the breakpoints. With higher values of $k$ the splines become smoother and can approximate smooth functions with a higher rate. However, for high $k$ the splines become less “flexible” and show oscillations for functions with isolated singularities (similar to polynomials or trigonometric polynomials).

2.3.1 Definition and properties of cubic splines

In practice the cubic splines $S_4^N$ are used most often: These are piecewise cubic polynomials which are two times continuously differentiable across the breakpoints:

$$S_4^N := \{ p \in C^2[a, b] \mid p|_{I_j} \in \mathcal{P}_3 \text{ for } j = 1, \ldots, N \}.$$ 

Therefore we have $\dim S_4^N = 4N - 3(N - 1) = N + 3$.

For a given function $f \in C^1[a, b]$ we choose the breakpoints as interpolation points, this gives $N + 1$ conditions. This leaves 2 additional conditions for which we will specify “boundary conditions” at the points $x = a$ and $x = b$: We additionally require that $p'(a) = f'(a)$ and $p'(b) = f'(b)$ (“clamped boundary conditions”). A function $p \in S_4^N$ is called the complete cubic spline interpolation of the function $f$ if it satisfies the conditions

$$p(x_j) = f(x_j) \quad \text{for } j = 1, \ldots, N, \quad p'(a) = f'(a), \quad p'(b) = f'(b) \quad (2.16)$$

The following property is the key for understanding cubic splines:

**Theorem 2.4** Let $e \in C^1[a, b]$ with $e|_{I_j} \in C^2(I_j)$ satisfy

$$e(x_j) = 0, \quad j = 0, \ldots, N, \quad e'(a) = 0, \quad e'(b) = 0 \quad (2.17)$$

Then there holds

$$\langle e'', q \rangle = 0 \quad \text{for all } q \in S_2^N. \quad (2.18)$$

Proof: We use integration by parts, note that $\overline{q}$ is piecewise constant:

$$\sum_{j=1}^{N} \int_{x_{j-1}}^{x_j} e''(x) \overline{q}(x) \, dx = \sum_{j=1}^{N} \left( \left[ e'(x) \overline{q}(x) \right]_{x_{j-1}}^{x_j} - \int_{x_{j-1}}^{x_j} e'(x) \overline{q}'(x) \, dx \right)$$

$$= \sum_{j=1}^{N} \left( \left[ e'(x) \overline{q}(x) \right]_{x_{j-1}}^{x_j} - \left[ \underbrace{e(x) \overline{q}(x)}_{0 \text{ by }(2.17)} \right]_{x_{j-1}}^{x_j} \right)$$

$$= e'(b) \overline{q}(b) - e'(a) \overline{q}(a) \quad (2.17)$$

$$= 0$$
since \( e'(x) \overline{q}(x) \) is continuous across the breakpoints, hence all terms in the sum except the ones at \( x_0 \) and \( x_N \) cancel out.

For a function \( f \in C^1[a,b] \) with \( f|_{I_j} \in C^2(I_j) \) let \( p \in S_4^N \) be a complete cubic spline interpolation satisfying (2.16). Then \( \varepsilon := f - p \) satisfies 2.17 and the theorem gives \( \langle f'' - p'', q \rangle = 0 \) for all \( q \in S_2^N \). Note that \( p \in S_4^N \) implies \( p'' \in S_2^N \), hence we have that \( p'' = \Pi_2^N f'' \), i.e., \( p'' \) is the least squares approximation in \( S_2^N \) of \( f'' \).

**Corollary 2.5** For \( f \in C^1[a,b] \) there exists a unique spline interpolation \( p = I_4^N f \) satisfying (2.17).

*Proof:* The spline function \( p \) is a cubic polynomial on each of the \( N \) intervals and can therefore be written in terms of \( 4N \) parameters. The continuity conditions at the breakpoints give \( 3(N - 1) \) equations, and the interpolation conditions (2.17) give another \( N + 3 \) equations, yielding a linear system of \( 4N \) equations for \( 4N \) unknowns. The linear system with a zero right hand side corresponds to a spline interpolation of the zero function, hence by (2.18) (with \( q := p \)) \( \langle p'', p'' \rangle = 0 \) and \( p'' = 0 \). Thus \( p(x) \) is a linear function with \( p(a) = p(b) = 0 \) and therefore \( p = 0 \).

**Corollary 2.6** Let \( f \in C^1[a,b] \). The spline interpolation \( I_4^N f \) is the unique function minimizing \( \int_a^b |g''(x)|^2 \, dx \) among all functions \( g \in C^1[a,b] \) with \( g|_{I_j} \in C^2(I_j) \) satisfying the conditions

\[
g(x_j) = f(x_j), \quad j = 0, \ldots, N, \quad g'(a) = f'(a), \quad g'(b) = f'(b).
\]

(2.19)

*Proof:* Let \( p = I_4^N f \). Then any function \( g \) satisfying the assumptions is of the form \( g = p + e \) with \( e \) satisfying the assumptions of Theorem 2.4 and

\[
\langle g'', g'' \rangle = \langle p'' + e'', p'' + e'' \rangle \stackrel{(2.18)}{=} \langle p'', p'' \rangle + \langle e'', e'' \rangle \geq \langle p'', p'' \rangle
\]

and we have equality only if \( e''|_{I_j} = 0 \) for \( j = 1, \ldots, N \). This implies that \( e \) is linear on \( I_j \), and \( e(x_{j-1}) = e(x_j) = 0 \) gives \( e = 0 \).

By using the density of continuous function in \( L^2 \) one can obtain that \( I_4^N f \) minimizes \( \int_a^b |g''(x)|^2 \, dx \) among all functions \( g \) with \( g'' \in L^2(a,b) \) satisfying (2.19).

This minimization property is one motivation for the use of spline interpolation and also for the name “spline”: Given the points \((x_j, y_j), j = 0, \ldots, N\) and the derivatives \( y'_0, y'_N \) we want to find the “smoothest” curve \((x(s), y(s))\) (where \( s \) denotes the arc length) which passes through the given points and satisfies the derivative conditions. One definition of “smoothest” is given by the behavior of an elastic rod which is required to pass through the given points, but otherwise free: The rod will assume a configuration of minimal potential energy

\[
E := \int \kappa(s)^2 \, ds
\]

where \( \kappa(s) \) denotes the curvature at \( s \). If we assume that \( y' = \frac{dy}{dx} \) is small everywhere, then \( s \approx x \) and \( \kappa(s) = y''/(1 + y'^2)^{3/2} \approx y''(x) \) ("geometric linearization") and we obtain

\[
E \approx E_0 := \int_a^b y''(x)^2 \, dx.
\]
Theorem 2.7 Let \( f \in C^2[a, b] \). Then the complete cubic spline interpolation \( p = I_4^N f \) satisfies
\[
\| f - p \|_\infty \leq \frac{h_{\text{max}}^2}{2} \| f'' \|_\infty
\]  
(2.20)

If \( f \in C^4[a, b] \) then
\[
\| f - p \|_\infty \leq \frac{h_{\text{max}}^2}{16} \max_{j=1, \ldots, N} \left( h_j^2 \| f^{(4)} \|_{\infty, I_j} \right) \leq \frac{h_{\text{max}}^4}{16} \| f^{(4)} \|_\infty
\]  
(2.21)

Assume that \( r^{-1} \leq \frac{h_j}{N} \leq r \) with \( r < \sqrt{2} \). Then we have for \( f \in C^2[a, b] \)
\[
\| f - p \|_\infty \leq \frac{1 + r^2}{8 - 4r^2} \max_{j=1, \ldots, N} \left( h_j^2 \| f'' \|_{\infty, I_j} \right).
\]  
(2.22)

If \( f \in C^4[a, b] \) then
\[
\| f - p \|_\infty \leq \frac{1 + r^2}{8(8 - 4r^2)} \max_{j=1, \ldots, N} \left( h_j^4 \| f^{(4)} \|_{\infty, I_j} \right)
\]  
(2.23)

Proof: Let \( e := f - p \) denote the interpolation error. Then \( I_2^N e = 0 \) and (2.1) yields
\[
\| e \|_\infty \leq \max_{j=1, \ldots, N} \left( h_j^2 \| e'' \|_{\infty, I_j} \right) \leq \frac{h_{\text{max}}^2}{8} \| e'' \|_\infty
\]

By (2.18) we have \( p'' = \Pi_2^N (f'') \), hence \( e'' = f'' - \Pi_2^N (f'') \). Therefore (2.20) follows from (2.12) and \( q = 0 \), and (2.21) follows from (2.14).

If we use \( w_j := h_j^2 \) we obtain (2.22) from (2.13), and (2.23) follows from (2.15).

\[ \square \]

2.3.2 Computation of cubic splines

In order to compute the cubic spline \( p = I_4^N f \) from the given values \( y_0 := f(x_0), \ldots, y_N := f(x_N) \), \( y'_0 := f'(a), y'_N := f'(b) \) we use the orthogonality relation (2.18): This relation states that the piecewise linear function \( p'' \in S_2^N \) is the least squares approximation of \( f'' \), i.e., \( p'' = \Pi_2^N f'' \).

We can write \( p'' \) as linear combination of the basis functions \( \phi_j \) of \( S_2^N \):
\[
p'' = \sum_{j=0}^{N} \alpha_j \phi_j
\]

Then the coefficients \( \alpha_0, \ldots, \alpha_N \) satisfy the normal equations (2.6) (with \( f'' \) instead of \( f \)):
\[
\frac{1}{3} \begin{pmatrix}
  h_1 & h_1/2 & 0 \\
  h_1/2 & h_1 + h_2 & h_2/2 \\
  \vdots & \vdots & \vdots \\
  0 & h_{N-1}/2 & h_{N-1} + h_N \\
  h_N/2 & h_N & h_N/2
\end{pmatrix}
\begin{pmatrix}
  \alpha_0 \\
  \alpha_1 \\
  \vdots \\
  \alpha_{N-1} \\
  \alpha_N
\end{pmatrix}
=
\begin{pmatrix}
  \langle f'', \phi_0 \rangle \\
  \langle f'', \phi_1 \rangle \\
  \vdots \\
  \langle f'', \phi_{N-1} \rangle \\
  \langle f'', \phi_N \rangle
\end{pmatrix}
\]  
(2.24)
The right hand side vector $\langle f'', \phi_j \rangle$ can be expressed in terms of the given values $y_0, \ldots, y_N, y'_0, y'_N$ using integration by parts as in Theorem 2.4: For $j = 1, \ldots, N - 1$

$$\langle f'', \phi_j \rangle = \left[ f' \phi_j \right]_{x_{j-1}}^{x_{j+1}} - \int_{x_{j-1}}^{x_{j+1}} f'(x) \phi'_j(x) \, dx = -\frac{1}{h_j} [f]_{x_{j-1}}^{x_j} + \frac{1}{h_{j+1}} [f]_{x_j}^{x_{j+1}} - \frac{y_j - y_{j-1}}{h_j} + \frac{y_{j+1} - y_j}{h_{j+1}}$$

$$\langle f'', \phi_0 \rangle = \left[ f' \phi_0 \right]_{x_0}^{x_1} - \int_{x_0}^{x_1} f'(x) \phi_0(x) \, dx = -f'(x_0) + \frac{y_1 - y_0}{h_1}$$

$$\langle f'', \phi_N \rangle = \left[ f' \phi_N \right]_{x_{N-1}}^{x_N} - \int_{x_{N-1}}^{x_N} f'(x) \phi_N(x) \, dx = f'(x_N) - \frac{y_N - y_{N-1}}{h_N}$$

We can solve the linear system (2.24) using Gaussian elimination without pivoting. We then know $\alpha_0, \ldots, \alpha_N$ and $p'' = \sum_{j=0}^{N} \alpha_j \phi_j$ and it remains to find $p$. For $x \in [x_{j-1}, x_j]$ we have

$$p''(x) = \alpha_{j-1} \frac{x_j - x}{h_j} + \alpha_{j} \frac{x - x_{j-1}}{h_j}$$

By taking antiderivatives we obtain for $x \in [x_{j-1}, x_j]$

$$p'(x) = \alpha_{j-1} \frac{-1}{2} \frac{(x_j - x)^2}{h_j} + \alpha_{j} \frac{1}{2} \frac{(x - x_{j-1})^2}{h_j} + C_j$$

$$p(x) = \alpha_{j-1} \frac{1}{6} \frac{(x_j - x)^3}{h_j} + \alpha_{j} \frac{1}{6} \frac{(x - x_{j-1})^3}{h_j} + \ell_j(x)$$

where $C_j$ is a constant and $\ell_j(x)$ is a linear function. By inserting the conditions $p(x_{j-1}) = y_{j-1}$ and $p(x_j) = y_j$ in the last equation we obtain $\ell_j(x_{j-1}) = y_{j-1} - \alpha_{j-1} h_j^2 / 6$, $\ell_j(x_j) = y_j - \alpha_{j} h_j^2 / 6$ yielding

$$\ell_j(x) = \left( y_{j-1} - \alpha_{j-1} \frac{h_j^2}{6} \right) \frac{x_j - x}{h_j} + \left( y_j - \alpha_{j} \frac{h_j^2}{6} \right) \frac{x - x_{j-1}}{h_j}$$

$$p(x) = \alpha_{j-1} \frac{(x_j - x)^3}{6h_j} + \alpha_{j} \frac{(x - x_{j-1})^3}{6h_j}$$

$$+ \left( y_{j-1} - \alpha_{j-1} \frac{h_j^2}{6} \right) \frac{x_j - x}{h_j} + \left( y_j - \alpha_{j} \frac{h_j^2}{6} \right) \frac{x - x_{j-1}}{h_j} \quad \text{for} \ x \in [x_{j-1}, x_j] \quad (2.25)$$

We see that cubic spline interpolation $I_4^N$ is not a strictly local approximation method, similar to $\Pi_2^N$: Even if the function $f$ is nonzero only in a small part of the interval $[a, b]$, the approximation $I_4^N f$ will in general be nonzero everywhere in $[a, b]$. However, the function $I_4^N f$ will decrease very rapidly away from the support of $f$.

### 2.4 Other types of piecewise polynomial interpolation

Another type of piecewise cubic interpolation is the so-called piecewise cubic Hermite interpolation. Here one constructs on each subinterval the cubic function with the same function
values and derivatives at the left and right endpoint. The resulting function has therefore a
continuous derivative, but the second derivative is in general discontinuous at the breakpoints.
The number of degrees of freedom is \(2N + 2\).

In order to find an error estimate we consider first cubic Hermite interpolation on an interval
\([\alpha, \beta]\). The error formula gives for the interpolation error

\[
\|f - p\|_\infty \leq \frac{\|f^{(4)}\|_\infty}{4!} (x - \alpha)^2 (x - \beta)^2 \leq \frac{\|f^{(4)}\|_\infty}{24} 2^{-4}(\beta - \alpha)^4
\]

since the maximum of the node polynomial occurs at the center of the interval. Now we use
this estimate for the piecewise cubic interpolation on each of the \(N\) intervals of length \((b - a)/N\)
and obtain

\[
\|f - p\|_\infty \leq \frac{1}{384} \frac{(b - a)^4}{N^4} \|f^{(4)}\|_\infty,
\]

i.e., the error decreases with a rate of \(O(N^{-4})\) if \(f \in C^4\). Note that the Hermite interpolation
is local: The interpolation on each subinterval is only determined by the data on the endpoints
of the subinterval.

This property makes cubic Hermite interpolation useful for computer design applications: To
specify a curve \((x(t), y(t), t \in [0, 1])\) in \(\mathbb{R}^2\), one chooses \(t_j = j/N, j = 0, \ldots, N\) and specifies the
points \((x(t_j), y(t_j)) = (x_j, y_j)\) and the derivatives \((x'(t_j), y'(t_j))\). Then computes the piecewise
Hermite polynomial \(p\) for the data \(x(t_j), x'(t_j)\) and the piecewise Hermite polynomial \(q\) for
the data \(y(t_j), y'(t_j)\), yielding an interpolating curve \((p(t), q(t)), t \in [0, 1]\) in \(\mathbb{R}^2\). Usually,
the polynomials on each interval are expressed as a Bernstein polynomials ("Beziers"), and one
specifies the location of the Bernstein points for each part of the curve instead of the derivatives
("handles on curve").

In general, a piecewise polynomial interpolation method is determined by the choice of a
certain space of piecewise polynomials and the choice of certain interpolation conditions.

For the space of piecewise polynomials we have the following choices:
(i) the degree of polynomials on each subinterval: A larger degree of polynomials gives better
convergence where the function is very smooth, but it does not help very much near points
where the function shows singular behavior. It also increases the number of degrees of freedom.
(ii) the degree of smoothness of the function across breakpoints: using a high degree of smooth-
ness gives a less localized behavior, but decreases the number of degrees of freedom. For
applications in computer aided design, one often desires approximations which are at least, say
2 times continuously differentiable.

One also has the possibility of choosing a nonuniform distribution of breakpoints in order
to adapt the approximating functions to certain features of the functions to be approximated,
e.g., singular behavior like \(|x|^\alpha\). As illustrated by a homework problem, a properly chosen
nonuniform mesh can compensate singular behavior of the function and restore the optimal
convergence rate.

After we have chosen a certain space of piecewise polynomial, we must then choose a set
of interpolation conditions such that the number of conditions matches the dimension of the
space of piecewise polynomials. One also has to make sure that the homogeneous interpolation
problem has only the zero solution.