APPROXIMATION THEORY

Preliminaries

- Let $X \equiv [a,b]$ be a closed interval on R, with $a \neq b$, and let $C[X]$ be the **set of all continuous real functions $f: X \to R$**.
- $C[X]$ is a vector space under the usual pointwise operations on functions.

Definition *Being continuous, all $f \in C[X]$ attain a supremum in X. We can therefore define:*

$$\|f\| \equiv \max_{x \in X} |f(x)|$$

Remark *$\|f\|$ is a norm on $C[X]$, and is known as the uniform, sup, L_∞, or Chebyshev norm. Therefore, $C[X]$ is a normed vector space.*

Definition *A sequence of functions* $\{f_n\} \in C[X]$ *converges uniformly to* $f \in C[X]$ *in the sup norm if:*

$$\lim_{n\to\infty} \|f_n - f\| = 0$$

Definition *A sequence of functions* $\{f_n\} \in C[X]$ *is a Cauchy sequence if:*

$$\lim_{m,n\to\infty} \|f_n - f_m\| = 0$$

Theorem (Weierstrass) *Any Cauchy sequence* $\{f_n\} \in C[X]$ *converges to some element of* $C[X]$*. Therefore,* $C[X]$ *is a complete normed vector space, i.e. a Banach space.*

Definition *Let* $\Phi = \{\phi_j\}$ *be a distinguished subset of* $C[X]$, *i.e.* $\Phi \subset C[X]$. *A linear combination* $p(x) \equiv \sum_{j=0}^{n} c_j \phi_j(x)$ *where* $c \equiv [c_0, c_1, \ldots, c_n] \in R^{n+1}$ *and* $\phi_j(x) \in \Phi$ *for* $j = 0, 1, \ldots n$, *is called a polynomial in the elements of* Φ, *or a polynomial in the* ϕ_j.

Definition *The linear combinations of the first* $n + 1$ *elements in* $\{x^j\}$, $p_n(x) \equiv \sum_{j=0}^{n} c_j x^j$, *are called algebraic polynomials of degree n.*

Definition *Let* $f \in C[X]$ *and* $p \in P_\Phi$ *where* $\Phi \subset C[X]$. *If the infimum in* $E_n(f) \equiv \inf_{c \in R^n} \| f - p \|$ *is attained for some c, the polynomial* $p(x)$ *is called polynomial of best approximation (in the sup norm).*

Theorem *Let* $f \in C[X]$. *For a given n, there exists a unique* $p_n^* \in P_n$ *of best approximation.*

Theorem (Jackson) *There exists a constant* $\xi > 0$ *such that, for every* $f \in C^k[X]$ *where* $k \geq 1$ *there exist a* $p_n \in P_n$ *with* $n \geq k$ *such that:*

$$E_n(f) \leq \| f - p_n \| \leq \frac{\xi}{n^k} \| f^{(k)} \|$$

Definition *Let again* $\Phi \subset C[X]$. *A function* $f \in C[X]$ *is approximable by polynomials in the* ϕ_j *if for each* $\varepsilon > 0$ *there is a* $p(x)$ *such that* $\|f - p\| < \varepsilon$.

Theorem (Weierstrass) *Each* $f \in C[X]$ *is approximable by algebraic polynomials.*

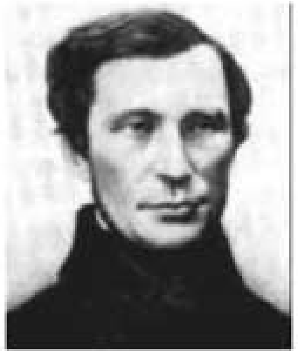
- In other words, *Weierstrass' Theorem* states that any f in $C[X]$ can be expressed as an infinite sum of powers.
- However, note that, so far, no result guaranteed that $lim_{j\to\infty}|c_j|\to 0$;
- Hence, low-order polynomials are not necessarily good approximations: the Weierstrass Theorem holds only in the limit.

Chebyshev systems

Definition 1 *Let $\Phi \subset C[X]$ contain at least $n+1$ elements. The set Φ is a Chebyshev system if the functions ϕ_j are linearly independent, and the exactly identified linear system $\sum_{j=0}^{n} c_j \phi_j(x_k) = 0$, for $k = 0, 1, ..., n$, admits only the trivial solution $\mathbf{c} = \mathbf{0}$, where $\mathbf{c} \equiv [c_0, c_1, ..., c_n]'$.*

Claim *The sequence of powers $\{x^j\}_{j=0}^{n}$ is a Chebyshev system on any nontrivial closed interval of R.*

Theorem *Let Φ be a Chebyshev system. Given n, there exists a unique polynomial in the ϕ_j of best approximation for each $f \in C[X]$.*



Theorem (Chebyshev) *Let Φ be a Chebyshev system, $p \in P_\Phi$, and $f \in C[X]$ with $p \neq f$. The polynomial $p(x) \equiv \sum_{j=0}^{n} c_j \phi_j(x)$ is the polynomial of best approximation for f if and only if there are $n+1$ points in X such that:*

$$f(x_j) - p(x_j) = m(-1)^j \|f - p\|, \quad j = 0, 1, \ldots, n$$

where $m = 1$ or $m = -1$.

Chebyshev polynomials

Consider the sequence $\Phi \equiv \{x^j\}_{j=0}^{n-1}$ for $n \geq 1$: we already know that Φ is a Chebyshev system on any nontrivial subset X of R. The power function x^n is an element of $C[X]$ but not of Φ.

Consider now the problem of finding the algebraic polynomial of best approximation for x^n on $X = [-1,+1]$:

$$\min_{\mathbf{c}\in R^{n-1}} \| x^n - p_{n-1} \| = \min_{\mathbf{c}\in R^{n-1}} \max_{x\in X} |x^n - p_{n-1}(x)|$$

The previous problem is equivalent to finding the *polynomial of least deviation from zero* among all *monic* $p_n \in P_n$, i.e. all algebraic polynomials of degree n with leading coefficient 1:

$$\min_{\mathbf{c}\in R^{n-1}} \| \hat{p}_n(x) \|$$

where $\hat{p}_n(x) \equiv x^n + \sum_{j=0}^{n-1} c_j x^j$.

Chebyshev's Theorem implies that $\hat{p}_n(x)$ is a polynomial of best approximation if and only if there are $n+1$ points in X where:

$$\hat{p}_n(x_j) = m(-1)^j \|\hat{p}_n\|, \quad j = 0, 1, \ldots, n$$

for $m = -1$ or $m = +1$.

In other words, $\hat{p}_n(x)$ has to reach $n+1$ times its absolute maximum distance from zero, but with alternating sign.

- Note that $cos:[0,\pi]\rightarrow[-1,+1]$ is a continuous and invertible function: for each $x\in[-1,+1]$ there is a unique t in $[0,\pi]$ such that $x=cos(t)$.

- Therefore, we can define a function $T_n:[-1,+1]\rightarrow[-1,+1]$ such that:

$$T_n(x) \equiv \cos[n \arccos(x)], \quad n = 0, 1, \ldots$$

- Note furthermore that $T_0(x)=1$, $T_1(x)=x$, and that (for $n\geq 2$):

$$T_n(x) = 2xT_{n-1} - T_{n-2}(x)$$

- Each $T_n(x)$ is an **algebraic polynomial** of degree n with leading coefficient equal to 2^{n-1}.

- These polynomials are called ***Chebyshev polynomials***.

Theorem *The monic algebraic polynomial of degree n and of least deviation from zero on* $[-1,+1]$ *is* $2^{1-n}T_n(x)$.

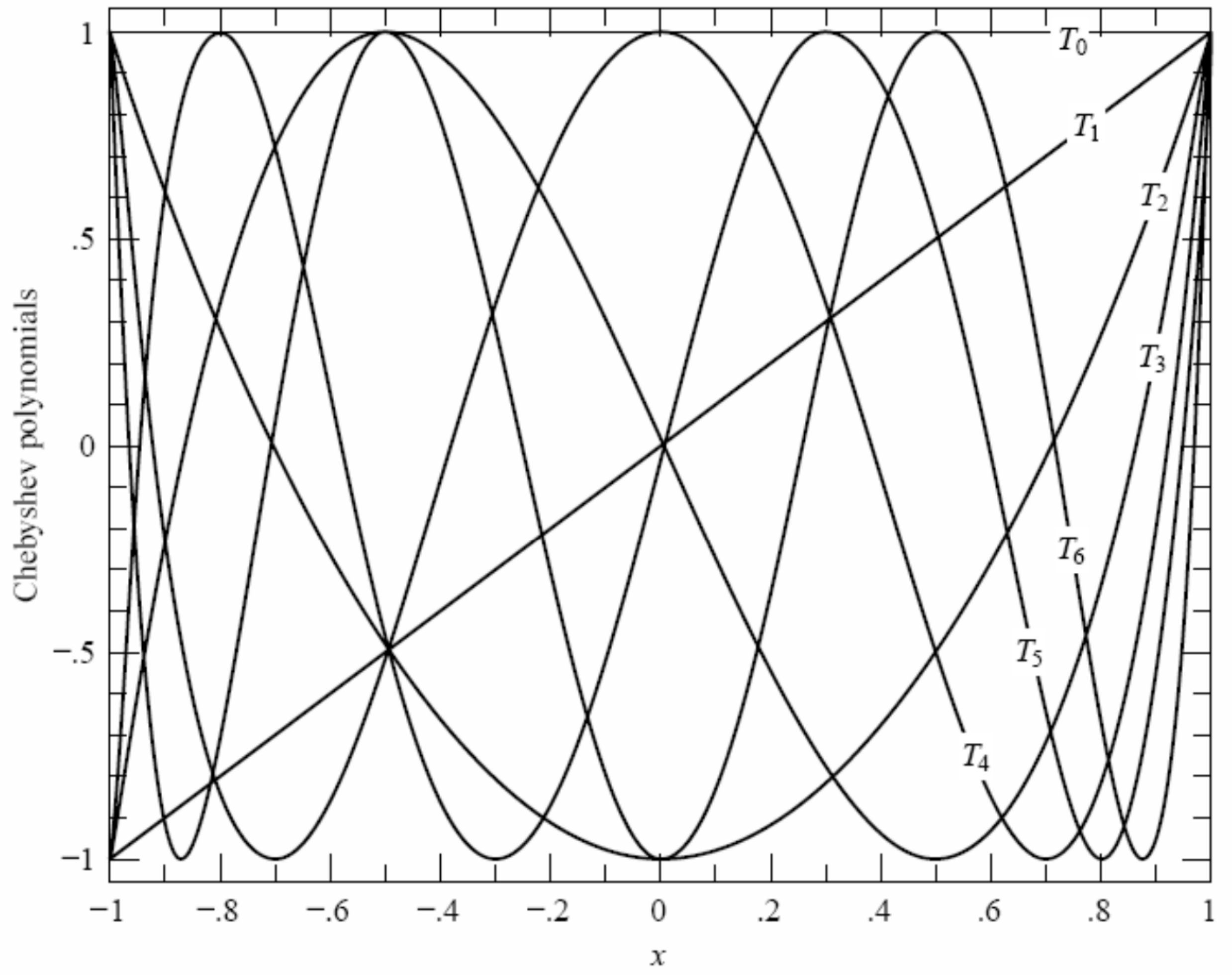


Figure 5.8.1. Chebyshev polynomials $T_0(x)$ through $T_6(x)$. Note that T_j has j roots in the interval $(-1,1)$ and that all the polynomials are bounded between ± 1.

- We can easily show that:

$$T'_n(x) \equiv \frac{n\sin[n\arccos(x)]}{\sin[\arccos(x)]}, \quad n = 0, 1, \ldots$$

- Given that $cos(nt)=0$ for $t=[(2j-1)\pi]/(2n)$, where $j=1,\ldots,n$, the points:

$$x_j = \cos\left[\left(j - \frac{1}{2}\right)\frac{\pi}{n}\right]$$

are the n zeros of $T_n(x)$ in decreasing order.

- They are all real, simple, and lie in $(-1,+1)$. Furthermore, $T_n(x)$ and $T_{n-1}(x)$ have no common zeros.

- Any algebraic polynomial of degree n can be uniquely expressed as a **finite sum** of Chebyshev polynomials:

$$p(x) = \sum_{j=0}^{n} c_j x^j = \sum_{j=0}^{n} \alpha_j T_j(x)$$

- We can extend the domain of Chebyshev polynomials to general intervals $X \equiv [a,b]$ in R by using the change of variable $y=2(x-a)/(b-a)-1$ where y *is* in $[-1,+1]$.
- The corresponding ***extended Chebyshev polynomials*** are denoted by:

$$\tilde{T}_n(x) \equiv T_n\left(2\frac{x-a}{b-a} - 1\right)$$

Interpolation

Definition *Let* $\{x_j\}_{j=0}^{n}$ *be a set of* $n+1$ *distinct points in* X*, called nodes, and* $\{y_j\}_{j=0}^{n}$ *a set of* $n+1$ *real numbers. A polynomial* $p(x) \equiv \sum_{j=0}^{n} c_j \phi_j(x) \in P_\Phi$ *interpolates the values* y_j *at the nodes* x_j *if* $p(x_j) = y_j$ *for* $j = 0, 1, \ldots, n$.

Theorem *Let* Φ *be a Chebyshev system. If* $\{x_j\}_{j=0}^{n}$ *are distinct nodes in* X*, and* $\{y_j\}_{j=0}^{n}$ *arbitrary real numbers, there is a unique interpolating polynomial in the* ϕ_j.

Lagrange interpolation

- Given $n+1$ distinct nodes in X and as many real numbers y_j, the unique interpolating algebraic polynomial of degree n can be written as:

$$L_n(x) = \sum_{j=0}^{n} y_j l_{j,n}(x)$$

where:

$$l_{j,n}(x) = \prod_{\substack{k=0 \\ k \neq j}}^{n} \frac{x - x_k}{x_j - x_k}, \quad j = 0, 2, \ldots, n$$

- The polynomials $l_{k,n}$ are algebraic polynomials of degree n too, and are known as ***Lagrange fundamental polynomials***.

- When $n=3$ the previous formula simplifies to:

$$l_{0,3}(x) = \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)}$$

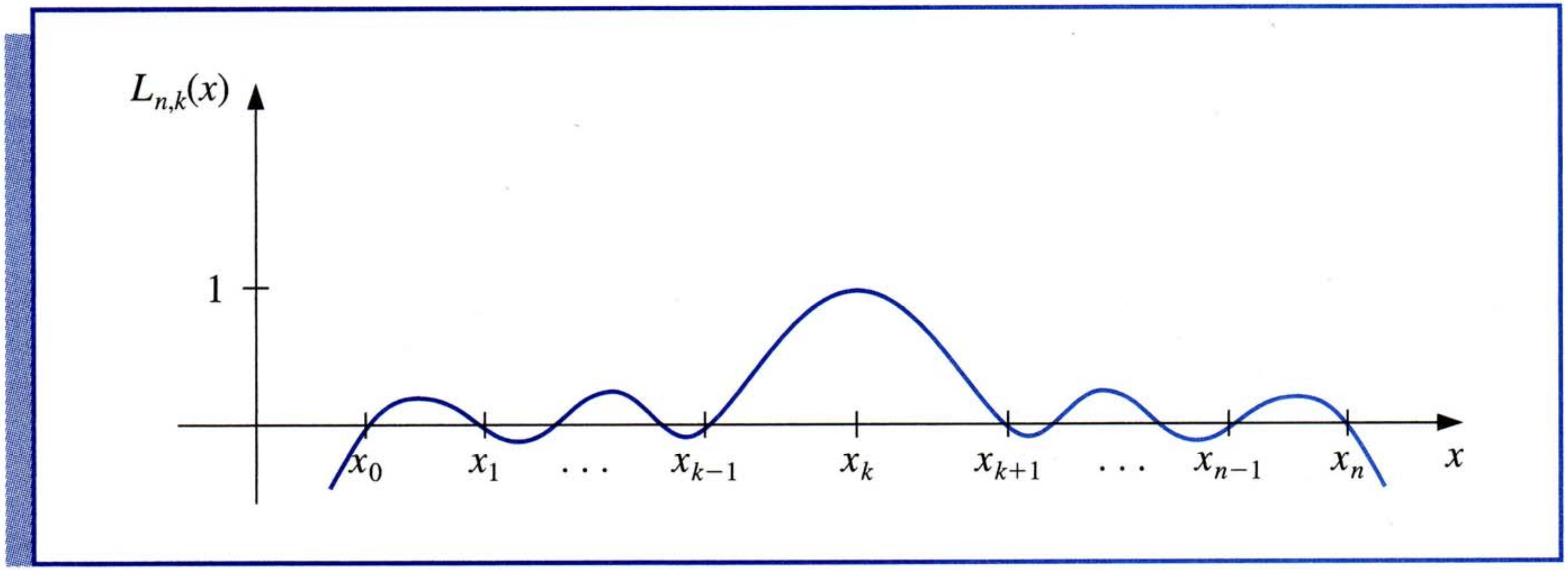
$$l_{1,3}(x) = \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)}$$

$$l_{2,3}(x) = \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)}$$

- Note that:

$$\begin{cases} l_{j,n}(x) = 1 \text{ if } x = x_j \\ l_{j,n}(x) = 0 \text{ if } x = x_k, \text{ where } k \neq j \end{cases}$$

A sketch of the graph of a typical $L_{n,k}$ is shown in Figure 3.5.



1. The *Lagrange interpolating polynomial* can be expressed as $L_n(x) = \sum_{j=0}^{n} c_j x^j$. The vector $c \in R^n$ solves the linear system $Vc = y$ where $y \equiv [y_0, y_1, \ldots, y_n]'$ and:

$$V \equiv \begin{bmatrix} 1 & x_0 & \cdots & x_0^n \\ 1 & x_1 & \cdots & x_1^n \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \cdots & x_n^n \end{bmatrix}$$

is the corresponding VanderMonde's matrix.

2. If the function to interpolate is an algebraic polynomial of degree $m \leq n$, then the Lagrange interpolating polynomial fits exactly and $L_n(p_m)(x) = \sum_{j=0}^{n} p_m(x_j) l_{j,n}(x) = p_m(x) = \sum_{j=0}^{n} c_j x^j$, where $c_j = 0$ for $m < j \leq n$.

Theorem (Lebesgue) *Let $L_n(f)$ be the Lagrange interpolating polynomial that interpolates a function $f: C[X] \to R$ at some $n+1$ nodes in X. Then:*

$$\|f - L_n(f)\| \leq E_n(f)(1 + \Lambda_n)$$

where $\Lambda_n \equiv \left\| \sum_{j=0}^{n} |l_{j,n}(x)| \right\|$ is the Lebesgue constant (it depends only on the nodes) of order n.

Theorem (Erdos) *For any set of $n+1$ nodes in X, there is a $c > 0$ such that:*

$$\Lambda_n > \frac{2}{\pi} \ln(n+1) - c$$

- Note that for any set of nodes $lim_{n\to\infty}\Lambda_n=\infty$, and that the rate of growth of Λ_n is at least logarithmic.
- Hence, even if $lim_{n\to\infty}E_n(f)=0$ for Weierstrass Theorem, we cannot conclude in general that $lim_{n\to\infty}||f-L_n(f)||=0$.

Theorem (Faber) *For any set of nodes in X, there is a $f \in C[X]$ such that:*

$$\lim_{n\to\infty} \| f - L_n(f) \| = \infty$$

- Faber's theorem is disruptive: Lagrange interpolation does not necessarily produce a uniformly convergent approximation under general assumptions. Example: Runge's phenomenon.

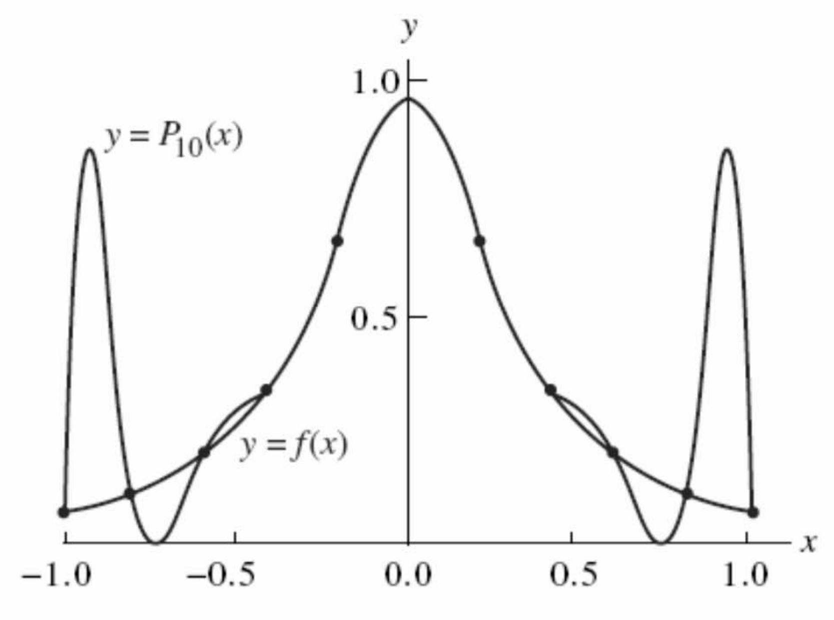


Figure 4.17 (a) The polynomial approximation to $y = 1/(1+12x^2)$ based on 11 equally spaced nodes over $[-1, 1]$.

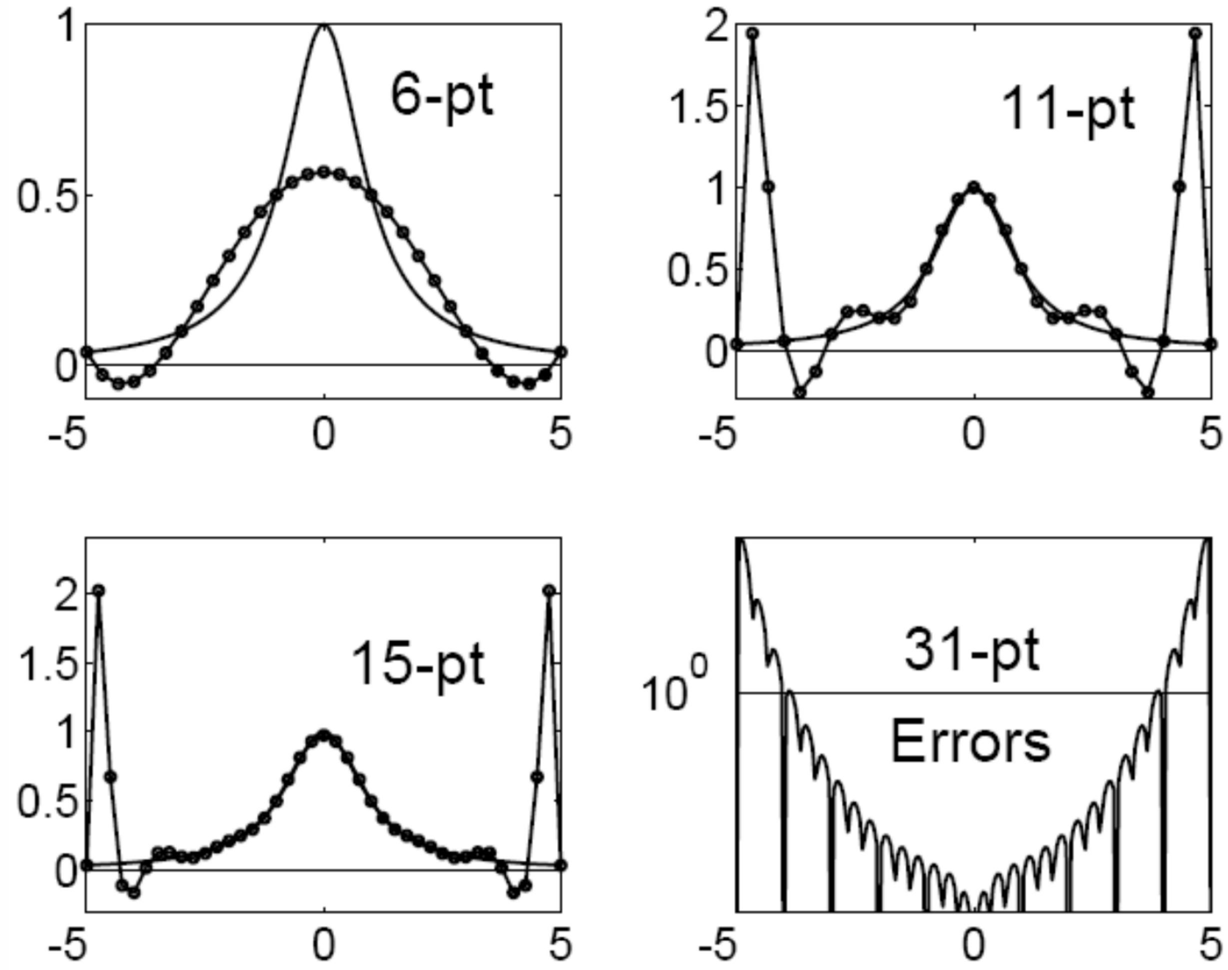


Figure 4.3: An example of the Runge phenomenon.
a. Solid curve without symbols: $f(x) \equiv 1/(1+x^2)$, known as the "Lorentzian" or "witch of Agnesi". Disks-and-solid curve: fifth-degree polynomial interpolant on $x \in [-5, 5]$. The six evenly spaced interpolation points are the locations where the dashed and the solid curves intersect.
b. Interpolating polynomial [disks] of tenth degree.
c. The interpolating polynomial of degree *fifteen* is the dashed curve. d. Same as previous parts except that only the error, $\log_{10}(|f(x) - P_{30}|)$, is shown where $P_{30}(x)$ is the polynomial of degree *thirty* which interpolates $f(x)$ at 31 evenly spaced points on $x \in [-5, 5]$. Because the error varies so wildly, it is plotted on a logarithmic scale with limits of 10^{-3} and 10^2.

- However, the following theorem shows that uniform convergence can be achieved if the set of nodes is chosen appropriately:

Theorem *For any $f \in C[X]$ there is a set of nodes in X such that:*

$$\lim_{n\to\infty} \| f - L_n(f) \| = 0$$

- **Question: how can we choose the set of nodes appropriately??**

Chebyshev interpolation

Theorem *Let* $\{x_j\}_{j=0}^{n}$ *be a set of* $n+1$ *distinct points in* X*, and let* $L_n(f)(x)$ *be the algebraic polynomial that interpolates a function* $f \in C^{n+1}[X]$ *at these points. Then:*

$$\|f - L_n(f)\| \leq \frac{\|f^{(n+1)}\|}{(n+1)!}\|W\|$$

where $W(x) \equiv \prod_{j=0}^{n}(x - x_j)$ *is an algebraic polynomial of order* $n+1$.

- The only element of the error bound that is directly under our control is $\|W\|$.
- Actually, by choosing the interpolation nodes we indirectly chose the upper bound for the interpolation error.

Theorem (Chebyshev) *If the nodes are the zeros of $\tilde{T}_n(x)$, the norm of $W(x)$ is minimized, and $\|W\| = 2^{1-n}$.*

- Note that $W(x)$ is itself a polynomial of degree $n+1$, and that the zeros of $W(x)$ correspond to the nodes x_j.
- The leading term in $W(x)$ is x^{n+1}, and we know that:

$$\frac{(b-a)^{n+1}}{2^{2(n+1)-1}} T_{n+1}\left(2\frac{x-a}{b-a}-1\right)$$

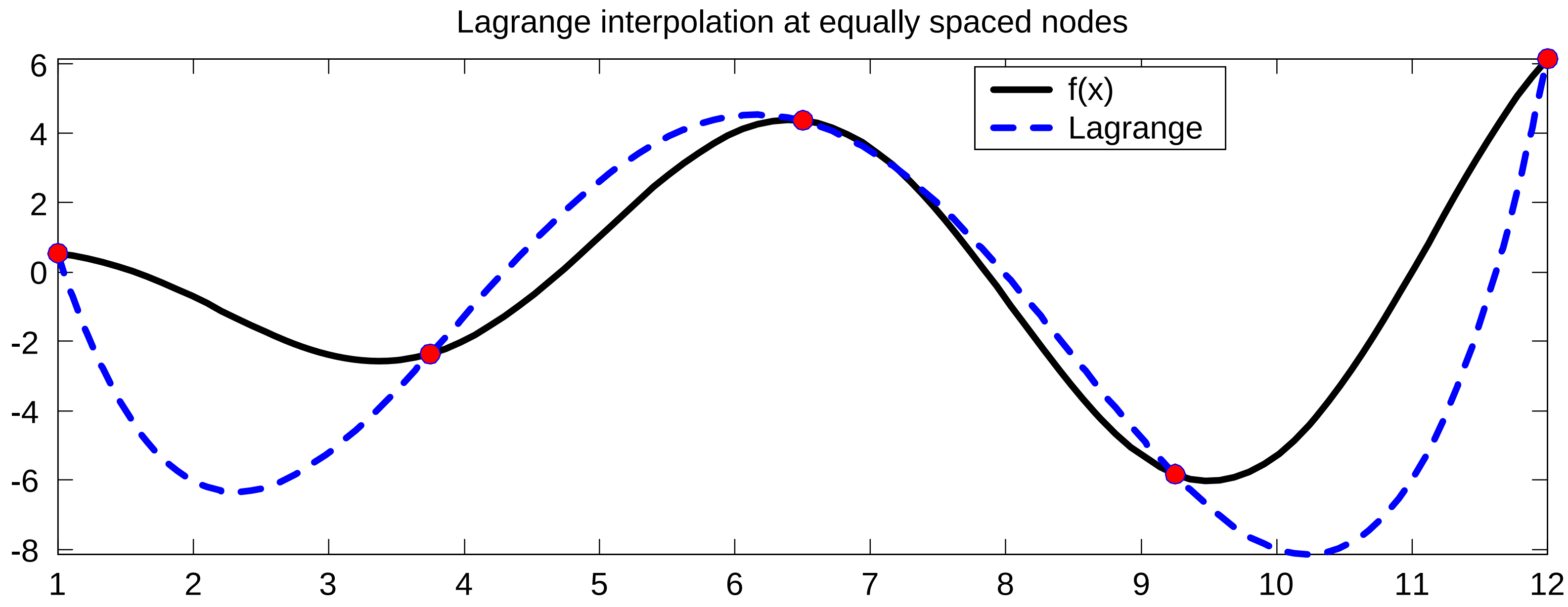
is the algebraic polynomial of degree $n+1$ of least deviation from zero with leading coefficient 1 on $[a,b]$.

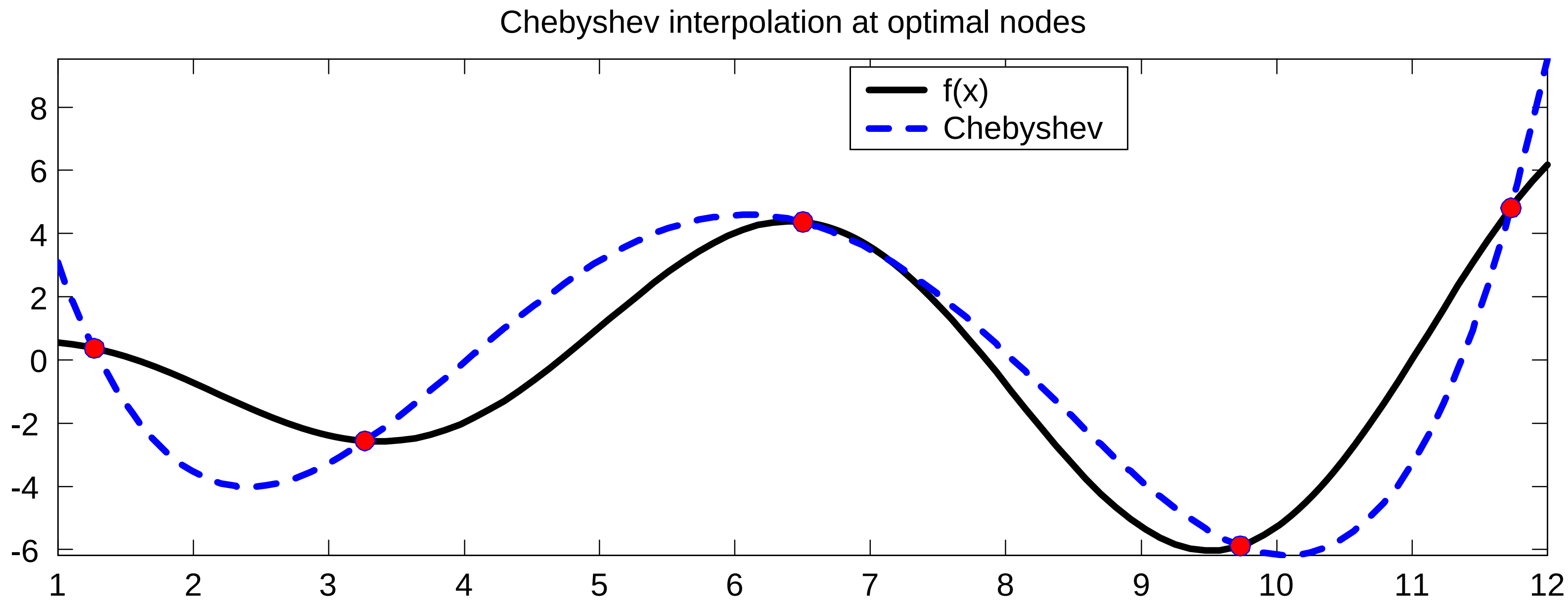
- By setting the points x_j equal to the zeros of T_{n+1}, we are transforming $W(x)$ in the polynomial of least deviation from zero, and therefore we are minimizing its sup norm.

Theorem *Let $f \in C^k[X]$ where $k \geq 1$. If the nodes are the zeros of $\tilde{T}_n(x)$ and $n \geq 1$, then:*

$$\lim_{n\to\infty} \| f - L_n(f) \| = 0$$

- We obtained two very important results:
  - Chebyshev interpolation, i.e. Lagrange interpolation at optimally chosen Chebyshev nodes, minimizes the approximation error for a given n.
  - the Chebyshev interpolating polynomial certainly convergences to the objective function for $n \to \infty$.


Lagrange interpolation at equally spaced nodes
f(x)
Lagrange


Chebyshev interpolation at optimal nodes
f(x)
Chebyshev

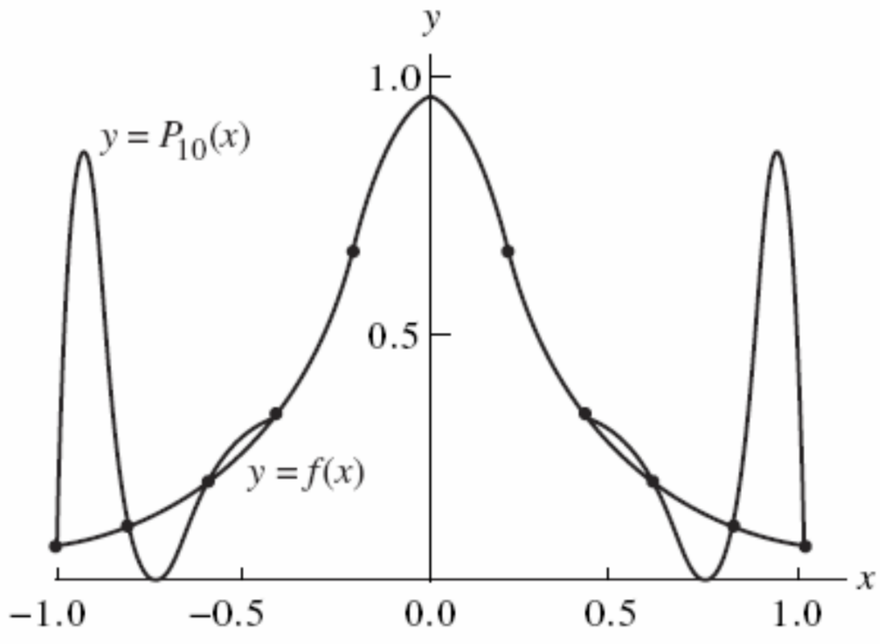


Figure 4.17 (a) The polynomial approximation to $y = 1/(1 + 12x^2)$ based on 11 equally spaced nodes over [−1, 1].

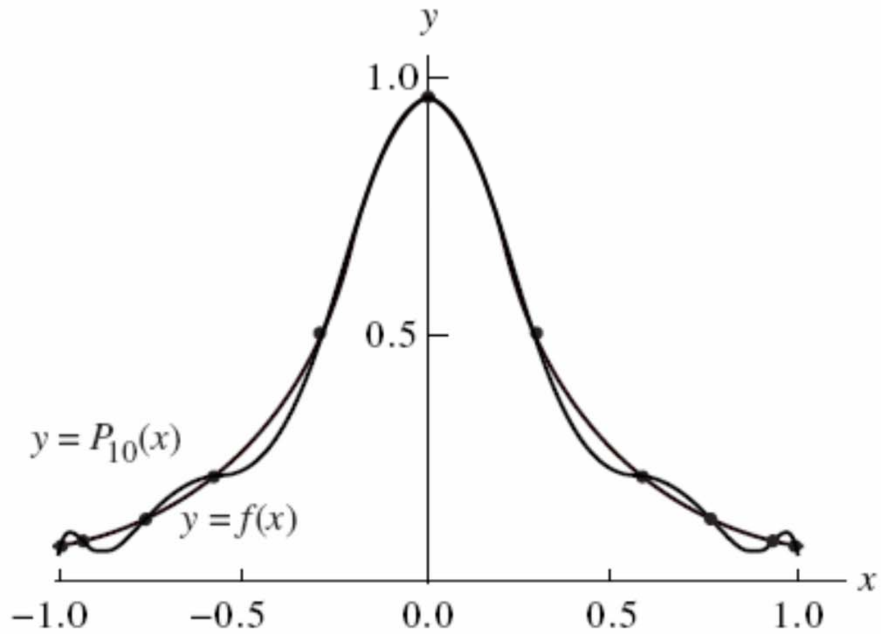


Figure 4.17 (b) The polynomial approximation to $y = 1/(1 + 12x^2)$ based on 11 Chebyshev nodes over [−1, 1].