where $\det(A_\lambda^{(1)})$ are the corresponding subdeterminants. In view of the sign change properties of $f - p^{(0)}$, it follows that

$$\delta^{(1)} = \left[\sum_{\lambda=1}^{n+1} \det(A_\lambda^{(1)})\right]^{-1} \sum_{\lambda=1}^{n+1} \det(A_\lambda^{(1)})|(f-p^{(0)})(x_\lambda^{(1)})|.$$

This is a weighted average, and since we assumed $\delta^{(0)} < \|f-p^{(0)}\|_\infty$, we get $\delta^{(1)} > \delta^{(0)}$.

Iteration: The above steps are repeated until the best approximation $\tilde{p}$ is approximated to a sufficient accuracy. A complete discussion of the convergence of the exchange method can be found in the book of G. Meinardus [1967]. The problem of convergence does not arise in the case where the best approximation is to be determined on the discrete set x_ν, $1 \le \nu \le m$, with $m \ge n+1$. In this case there are only $\binom{m}{n+1}$ ways to choose $(n+1)$-tuples of points $\{x_1^{(j)}, x_2^{(j)}, \ldots, x_{n+1}^{(j)}\}$, while by the monotonicity $\delta^{(j)} < \delta^{(j+1)}$, and so the same $(n+1)$-tuple cannot appear twice.

Example. We illustrate the Remez algorithm with a simple example. Consider the problem of computing the best approximation of $f(x) := x^2$ for $x \in [0,1]$ from the space of linear polynomials P_1. As a starting alternant we choose $\{x_1^{(0)}, x_2^{(0)}, x_3^{(0)}\} = \{0, \frac{1}{3}, 1\}$.

Step 1: We determine $p^{(0)}$ from the equations

$$\begin{aligned} \rho^{(0)} + \alpha_0^{(0)} \qquad\qquad &= 0 \\ -\rho^{(0)} + \alpha_0^{(0)} + \alpha_1^{(0)}\tfrac{1}{3} &= \tfrac{1}{9} \\ \rho^{(0)} + \alpha_0^{(0)} + \alpha_1^{(0)} \quad &= 1. \end{aligned}$$

The solution gives $\alpha_0^{(0)} = -\frac{1}{9}$, $\alpha_1^{(0)} = 1$ and $\rho^{(0)} = \frac{1}{9}$ so that $p^{(0)}(x) = -\frac{1}{9} + x$. This is the best approximation on the set $\{0, \frac{1}{3}, 1\}$, and satisfies

$$\|f - p^{(0)}\|_\infty = \max_{x\in[0,1]} |x^2 - x + \frac{1}{9}| = \frac{5}{36} > \frac{1}{9}.$$

This value is assumed for $\xi^{(1)} = \frac{1}{2}$. Hence we replace the alternant point $x_2^{(0)}$ by $\xi^{(1)}$ so that the new alternant for $p^{(1)}$ is $\{x_1^{(1)}, x_2^{(1)}, x_3^{(1)}\} = \{0, \frac{1}{2}, 1\}$.

Step 2: $p^{(1)}$ and $\rho^{(1)}$ are computed from

$$\begin{aligned} \rho^{(1)} + \alpha_0^{(1)} \qquad\qquad &= 0 \\ -\rho^{(1)} + \alpha_0^{(1)} + \alpha_1^{(1)}\tfrac{1}{2} &= \tfrac{1}{4} \\ \rho^{(1)} + \alpha_0^{(1)} + \alpha_1^{(1)} \quad &= 1. \end{aligned}$$

This gives $\alpha_0^{(1)} = -\frac{1}{8}$, $\alpha_1^{(1)} = 1$ and $\rho^{(1)} = \frac{1}{8}$. The corresponding polynomial is $p^{(1)}(x) = -\frac{1}{8} + x$, and $\|f - p^{(1)}\|_\infty = \max_{x\in[0,1]} |x^2 - x + \frac{1}{8}| = \frac{1}{8}$. Since this

value is assumed at all three points $x_1^{(1)} = 0$, $x_2^{(1)} = \frac{1}{2}$ and $x_3^{(1)} = 1$, it follows that $p^{(1)}$ is the best approximation, and the algorithm stops.

In general, we cannot expect that the algorithm will lead to the exact solution in a finite number of steps as was the case for this simple example. In practice it is common to terminate the iteration when the bounds $\delta^{(k)}$ and $\|f - p^{(k)}\|_\infty$ at the k-th step are sufficiently close to each other.

4.7 Chebyshev Polynomials of the First Kind. We now consider the problem of finding a best possible uniform approximation to the monomial $f(x) := x^n$ on $[-1,+1]$ by a polynomial from P_{n-1}, $(n = 1, 2, \ldots)$. We next show that the solution of this problem can be found using the Alternation Theorem.

The problem is to find the unique polynomial $\tilde{p} \in P_{n-1}$ satisfying

$$\begin{aligned} &\max_{x\in[-1,+1]} |x^n - (\tilde{\alpha}_{n-1}x^{n-1} + \cdots + \tilde{\alpha}_0)| = \\ &= \min_{\alpha\in\mathbb{R}^n} \max_{x\in[-1,+1]} |x^n - (\alpha_{n-1}x^{n-1} + \cdots + \alpha_0)|. \end{aligned}$$

Solution: For $n = 1$,

$$\min_{\alpha_0\in\mathbb{R}} \max_{x\in[-1,+1]} |x - \alpha_0| = \min_{\alpha_0\in\mathbb{R}} \max(|1 - \alpha_0|, |-1-\alpha_0|) = 1,$$

and thus $\tilde{\alpha}_0 = 0$. It follows that $\tilde{p} = 0$ is the best approximation from P_0.

For $n = 2$, the solution can be found from the construction 4.4. In this case the best approximation $\tilde{p} \in P_1$ of $f(x) := x^2$ on $[-1,+1]$ is $\tilde{p}(x) = \frac{1}{2}$ since the difference $d(x) = x^2 - \frac{1}{2}$ satisfies $d(-1) = -d(0) = d(1) = \frac{1}{2}$, and thus the points $\{-1, 0, 1\}$ form an alternant with maximal deviation.

We claim that, in general, the solution is given by the polynomial $\tilde{p}(x) = x^n - \hat{T}_n(x)$ with $\hat{T}_n(x) := \frac{1}{2^{n-1}}T_n(x)$, $T_n(x) := \cos(n \arccos(x))$. The proof proceeds as follows:

1) $\tilde{p} \in P_{n-1}$. For $n = 1$ we directly compute $T_1(x) = \cos(\arccos(x)) = x$ and $\hat{T}_1(x) = x$ with $\tilde{p}(x) = 0$. For $n > 1$, we employ the substitution $\theta := \arccos(x)$ (or equivalently $x = \cos(\theta)$) which gives a mapping $\theta : [-1,+1] \to [-\pi, 0]$. Then $T_n(x(\theta)) = \cos(n\theta)$.

Now the trigonometric identity $\cos((n+1)\theta) + \cos((n-1)\theta) = 2\cos(\theta)\cos(n\theta)$ implies the recursion relation $T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$ for $n \in \mathbb{Z}_+$. Starting with $T_0(x) = 1$, we obtain

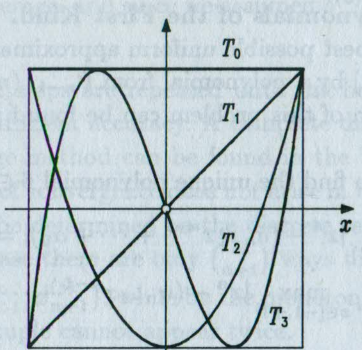
$$T_2(x) = 2x^2 - 1, \quad T_3(x) = 4x^3 - 3x, \quad \ldots, \quad T_n(x) = 2^{n-1}x^n - \cdots \quad \text{etc.}$$

The polynomials $\hat{T}_n$ are normalized so that the leading coefficient is 1, and thus $\tilde{p}(x) = x^n - \hat{T}_n(x)$ is a polynomial in P_{n-1}.

2) $\tilde{p} \in P_{n-1}$ is a best approximation. At the $n\theta_\nu := -(n - \nu + 1)\pi$, $1 \le \nu \le n+1$, we have $T_n(x(\theta_\nu)) = \cos(n\theta_\nu) = (-1)^{n-\nu+1}$. It follows that the points $x_\nu := \cos(-\frac{n-\nu+1}{n}\pi) = \cos((1 - \frac{\nu-1}{n})\pi)$ form an

alternant for $d(x) := \hat{T}_n(x) = x^n - \tilde{p}(x)$. Moreover, since $|\hat{T}_n(x_\nu)| = \frac{1}{2^{n-1}} = \|d\|_\infty$, the maximal deviation is assumed at these points; that is, $d(x_\nu) = \varepsilon(-1)^\nu \|d\|_\infty$ with $\varepsilon = \pm 1$ for $\nu = 1, \ldots, n+1$.

Clearly, the polynomial T_n has n simple zeros in the interval $(-1, +1)$ located at the points $x_\nu = \cos \frac{2\nu - 1}{2n}\pi$, $1 \le \nu \le n$.



The polynomials $T_n(x) = \cos(n \arccos(x))$ are called *Chebyshev polynomials of the first kind.* They are defined for all $n \ge 0$.

The approximation problem discussed in this section can be reinterpreted as follows: Find a polynomial of degree n, with leading coefficient one, whose maximum in $[-1, +1]$ is minimal. This is equivalent to finding a polynomial in the subset

$$\hat{\mathrm{P}}_n := \{p \in \mathrm{P}_n \mid p(x) = x^n + \alpha_{n-1}x^{n-1} + \cdots + \alpha_0\}$$

which best approximates the function $f = 0$ on $[-1, +1]$. The solution of this problem is given by $\tilde{p}(x) = x^n - \hat{T}_n(x)$. This says that the polynomial $\hat{T}_n(x) = x^n - \tilde{p}(x)$ is the unique polynomial in $\hat{\mathrm{P}}_n$ of minimal norm; i.e., $\|\hat{T}_n\|_\infty \le \|p\|_\infty$ for all $p \in \hat{\mathrm{P}}_n$.

This reformulation of the approximation problem of this section is a relatively simple example of a nonlinear approximation problem. Indeed, the subset $\hat{\mathrm{P}}_n$ is not a linear space, although it is an affine subspace of a linear space. We have been able to derive a remarkable minimal property of the Chebyshev polynomial of the first kind by considering an appropriate linear approximation problem.

4.8 Expansions in Chebyshev Polynomials. Using the fact that the Chebyshev polynomials of the first kind were defined in terms of trigonometric functions, it is easy to show that they form an orthogonal system of functions with respect to the weight function $w(x) := \frac{1}{\sqrt{1-x^2}}$. In fact,

$$\int_{-1}^{+1} T_k(x) T_\ell(x) \frac{dx}{\sqrt{1-x^2}} = \int_0^\pi \cos(k\theta)\cos(\ell\theta)\frac{\sin\theta}{\sin\theta}d\theta = 0 \quad \text{for} \quad k \ne \ell$$

and

$$\int_{-1}^{+1} T_k^2(x)\frac{dx}{\sqrt{1-x^2}} = \begin{cases} \pi & \text{for } k = 0 \\ \frac{\pi}{2} & \text{for } k \ne 0. \end{cases}$$

It is known from a theorem in analysis that every function $f \in \mathrm{C}[a, b]$ can be expanded in terms of a complete orthogonal system of functions. The partial sums of such a Fourier series expansion provide approximations to f which converge with respect to the norm $\|f\| := [\int_a^b f^2(x)w(x)dx]^{\frac{1}{2}}$ associated with the weight function w. In 5.5-5.8 we will discuss this fact further, especially for the case where the norm is $\|\cdot\|_2$. Here we proceed directly, and for given f, define approximations of f in terms of Chebyshev polynomials $T_0, T_1, \ldots$ by

$$\tilde{f}_n(x) = \frac{c_0}{2} + \sum_{k=1}^{n} c_k T_k(x),$$

where

$$c_k = \frac{2}{\pi}\int_{-1}^{+1} f(x)T_k(x)\frac{dx}{\sqrt{1-x^2}}, \quad k \in \mathbb{N},$$

or equivalently,

$$c_k = \frac{2}{\pi}\int_0^\pi f(\cos\theta)\cos(k\theta)d\theta = \frac{1}{\pi}\int_{-\pi}^{\pi} f(\cos\theta)\cos(k\theta)d\theta.$$

Under appropriate hypotheses, this sequence of approximations can even be shown to converge to f with respect to the uniform norm $\|\cdot\|_\infty$. Uniform convergence is especially interesting for the following reason. Suppose that a function $f \in \mathrm{C}[a, b]$ can be expanded in a uniformly convergent series in terms of a system of polynomials $\{\psi_0, \psi_1, \ldots\}$ which are normed so that $|\psi_k(x)| \le 1$ in $[a, b]$. Then

$$|f(x) - \tilde{f}_n(x)| = |\sum_{k=n+1}^{\infty} c_k\psi_k(x)| \le \sum_{k=n+1}^{\infty} |c_k|,$$

and thus if the coefficients c_k for $k \ge n+1$ are negligibly small, then $\tilde{f}_n$ provides a good approximation to the best approximating polynomial $\tilde{p} \in P_n$ of f with respect to the Chebyshev norm. The following theorem shows that this situation persists for the Chebyshev expansion, provided we restrict f to lie in $\mathrm{C}_2[-1, +1]$.

Expansion Theorem. *Let $f \in \mathrm{C}_2[-1, +1]$. Then the expansion of f in terms of Chebyshev polynomials T_k of the first kind for $x \in [-1, +1]$ converge uniformly on this interval to f. Moreover, the corresponding coefficients satisfy*

$$|c_k| \le \frac{A}{k^2},$$

where the constant A depends only on f.

Proof. From the formula for the coefficients given above, it follows, after integrating by parts twice and writing $\varphi(\theta) := f(\cos\theta)$, that

$$c_k = -\frac{2}{\pi k}\int_0^\pi \frac{d\varphi}{d\theta}\sin(k\theta)d\theta = \frac{2}{\pi k^2}\frac{d\varphi}{d\theta}\cos(k\theta)\Big|_0^\pi - \frac{2}{\pi k^2}\int_0^\pi \frac{d^2\varphi}{d\theta^2}\cos(k\theta)d\theta.$$

This immediately implies the estimate $|c_k| \leq \frac{A}{k^2}$. In addition, it follows that there exists a function $g \in C[-1,+1]$ with $\lim_{n\to\infty}\|\tilde{f}_n - g\|_\infty = 0$. Since $\lim_{n\to\infty}\|\tilde{f}_n - f\| = 0$ while

$$\|\tilde{f}_n - g\| = [\int_{-1}^{+1}(\tilde{f}_n(x)-g(x))^2\frac{dx}{\sqrt{1-x^2}}]^{\frac{1}{2}} \leq \|\tilde{f}_n - g\|_\infty\left(\int_{-1}^{+1}\frac{dx}{\sqrt{1-x^2}}\right)^{\frac{1}{2}},$$

the inequality

$$\|f-g\| \leq \|f-\tilde{f}_n\| + \|\tilde{f}_n - g\|$$

implies $f = g$ and the assertion is established. □

Practical Consequence. Given a function $f \in C_2[-1,+1]$, we can obtain a good approximation to the best approximation $\tilde{p} \in P_n$ by taking the partial sum $\tilde{f}_n = \sum_0^n c_k T_k$. This method is particularly applicable when the coefficients c_k are simple to compute.

Example. Consider the approximation of the function $f(x) := \sqrt{1-x^2}$ on $[-1,+1]$ by partial sums of the Chebyshev polynomial expansion of f. In this case

$$c_k = \frac{2}{\pi}\int_0^\pi \cos(kt)\sin t\,dt = \begin{cases}\frac{4}{\pi}\frac{1}{1-k^2} & \text{for } k = 2\kappa\\ 0 & \text{for } k = 2\kappa+1\end{cases},\ \kappa \in \mathbb{N}.$$

This leads to the approximations

$$\tilde{f}_0(x) = \frac{2}{\pi},\quad \tilde{f}_2(x) = \frac{2}{3\pi}(5-4x^2),$$

$$\tilde{f}_4(x) = \frac{2}{15\pi}(23-4x^2-16x^4),\quad \text{etc.}$$

We note that in this example the bound for $|c_k|$ given in the Expansion Theorem above is valid, even though here f is only two-times continuously differentiable on $(-1,+1)$.

In practice it is not usually possible to determine the coefficients c_k of the Chebyshev expansion by explicit integration, as was the case in this simple example. Generally, it will be necessary to use numerical quadrature formula to calculate these coefficients. An example is given in Problem 7 in 7.4.4.

4.9 Convergence of Best Approximations. Given a function $f \in C[a,b]$ and the sequence $(\tilde{p}_n)$ of best approximating polynomials, where $\tilde{p}_n \in P_n$ is best in the Chebyshev sense, it is natural to ask whether this sequence converges to f. This question can be answered with the help of the Weierstrass Approximation Theorem 2.2. Indeed, let $(p_n)_{n\in\mathbb{N}}$ be some sequence of polynomials $p_n \in P_n$ such that $\lim_{n\to\infty}\|f-p_n\|_\infty = 0$. Then since $\|f-\tilde{p}_n\|_\infty \leq \|f-p_n\|_\infty$ for all $n \in \mathbb{N}$, it follows immediately from $\lim_{n\to\infty}\|f-p_n\|_\infty = 0$ that $\lim_{n\to\infty}\|f-\tilde{p}_n\|_\infty = 0$. We have established the

Convergence Theorem. *Let $f \in \mathbb{C}[a,b]$. Then the sequence $(\tilde{p}_n)_{n\in\mathbb{N}}$ of best approximations $\tilde{p}_n \in P_n$ with respect to the norm $\|\cdot\|_\infty$ converges uniformly to f.*

4.10 Nonlinear Approximation. In this section we discuss an approximation problem involving a nonlinear subset of $(C[a,b],\|\cdot\|_\infty)$ which is especially important; namely, approximation by rational functions. We shall restrict our discussion primarily to the problem of the existence of a best approximation.

Let $R_{n,m}[a,b]$ be the set of continuous rational functions on the interval $[a,b]$ of the form $r(x) := \frac{p(x)}{q(x)}$, where $p \in P_n$, $q \in P_m$, $\|q\|_\infty = 1$ and $q(x) > 0$ for $x \in [a,b]$. In addition, suppose that p and q have no common factors, so that they have no common zeros anywhere on $\mathbb{C}$. The following theorem settles the problem of existence of a best approximation $\tilde{r} \in R_{n,m}[a,b]$.

Theorem. *Let $f \in C[a,b]$. Then in the set $R_{n,m}[a,b]$ of continuous rational functions, there always exists a best approximation $\tilde{r}$ of f.*

Proof. Let $(r_\nu)_{\nu\in\mathbb{N}}$ be a minimizing sequence for f in $R_{n,m}$, and suppose $r_\nu = \frac{p_\nu}{q_\nu}$, where $p_\nu \in P_n$ and $q_\nu \in P_m$ have no common factors. $\|q_\nu\|_\infty = 1$ implies that (q_ν) is a bounded sequence in P_m, and thus contains a convergent subsequence $(q_{\nu(\kappa)})$. Since P_m is finite dimensional, as $\kappa \to \infty$ this subsequence converges to some $q^* \in P_m$, $\|q^*\|_\infty = 1$.

By Lemma 3.4, the minimizing sequence (r_μ), $\mu := \nu(\kappa)$, is itself bounded. Now $\frac{|p_\mu(x)|}{|q_\mu(x)|} \leq C$ for $x \in [a,b]$ implies that $\|p_\mu\|_\infty \leq C$, which again gives the existence of a convergent subsequence $(p_{\mu(\kappa)})$ converging to some $p^* \in P_n$. Clearly, $|p^*(x)| \leq C|q^*(x)|$, and thus if $x_1,\cdots,x_k$ are zeros of q^*, $k \leq m$, then they are also zeros of p^*. Thus the common factors of $\frac{p^*}{q^*}$ can be cancelled to produce a rational function $\frac{\hat{p}}{\hat{q}} \in R_{n,m}$ with $\hat{q}(x) > 0$ for $x \in [a,b]$ and

$$|f(x) - \frac{\hat{p}(x)}{\hat{q}(x)}| = |f(x) - \frac{p^*(x)}{q^*(x)}| \leq |f(x) - \frac{p_{\mu(\kappa)}(x)}{q_{\mu(\kappa)}(x)}| + |\frac{p_{\mu(\kappa)}(x)}{q_{\mu(\kappa)}(x)} - \frac{p^*(x)}{q^*(x)}|$$

$$\Rightarrow \|f - \frac{p^*}{q^*}\|_\infty \leq \|f - \frac{p_{\mu(\kappa)}}{q_{\mu(\kappa)}}\|_\infty + \|\frac{p_{\mu(\kappa)}}{q_{\mu(\kappa)}} - \frac{p^*}{q^*}\|_\infty.$$

Since $\lim_{\kappa\to\infty}\|f - \frac{p_{\mu(\kappa)}}{q_{\mu(\kappa)}}\|_\infty = E_{R_{n,m}}(f)$ and $\lim_{\kappa\to\infty}\|\frac{p_{\mu(\kappa)}}{q_{\mu(\kappa)}} - \frac{p^*}{q^*}\|_\infty = 0$, we conclude that $\|f - \frac{p^*}{q^*}\|_\infty \leq E_{R_{n,m}}(f)$. Now since $\frac{p^*}{q^*} \in R_{n,m}[a,b]$, we know that $E_{R_{n,m}}(f) \leq \|f - \frac{p^*}{q^*}\|_\infty$, and so $\|f - \frac{p^*}{q^*}\|_\infty = E_{R_{n,m}}(f)$; i.e., $\frac{p^*}{q^*}$ is a best approximation of f in $R_{n,m}[a,b]$. □

It is beyond the scope of this book to develop fully further properties of approximation by rational functions, but we complete this section by mentioning a result on uniqueness and an outline of how best approximations can be computed.