JOURNAL OF APPROXIMATION THEORY 41, 257–278 (1984)

Rational Carathéodory–Fejér Approximation on a Disk, a Circle, and an Interval

MARTIN H. GUTKNECHT

Seminar für angewandte Mathematik,
ETH-Zentrum, CH-8092 Zürich, Switzerland

Communicated by G. Meinardus

Received November 14, 1982

A systematic description of the Carathéodory–Fejér method (CF method) is given for near-best uniform rational approximation of a (high degree) polynomial on a disk. On the basis of Takagi's extension of the Carathéodory–Fejér theorem degeneracies are characterized and it is proven that they appear in the CF table, which is introduced, in square blocks. Related method for complex and real trigonometric rational approximation, and for ordinary real rational approximation on an interval, are then derived. For each problem several types of CF approximation are defined depending on truncation. Certain weight functions are also allowed.

INTRODUCTION

In systems theory [10, 21, 28, 40], digital signal processing [24, 25], and numerical analysis [46, 47] there has been much interest recently in a new numerical method for computing near-best solutions of the rational Chebyshev approximation problem on the unit disk D. If the given function f is a polynomial (possibly of very high degree), this method can be based upon Takagi's extension [41, 42] of the classical theorem of Carathéodory and Fejér (abbreviated CF in the following) [11, 39]. It mainly requires then the singular value decomposition of a finite Hankel matrix. This version of the method was proposed and investigated by Trefethen [46, 47], who called it the CF method. For both the resulting CF approximation and the best approximation Trefethen presented strong theoretical and numerical results on the near-cicurlarity of the error curves. For polynomial approximants the method had been mentioned before by Hollenhorst [27] and by G. H. Elliott [19], both of whom derived it from a corresponding method for real polynomial and rational approximation on an interval. This latter method we will call Chebyshev–CF approximation (in analogy to Gragg's terminology in Padé approximation [22]). Its polynomial version was proposed by Darlington [15] and in [26], where an asymptotic error analysis is also

0021-9045/84 $3.00

given. A very special rational case of the method and similar approaches to the polynomial approximation problem had appeared implicitly or explicitly in a variety of papers, starting with the monograph of Bernstein [8]; see, e.g., Achieser [1, 2, 4, 5], Tabot [43, 44], Clenshaw [14], Meinardus [36], Lam and D. Elliott [18, 29–33], Hollenhorst [27], and G. H. Elliott [19]. However, none of those authors seem to have been aware of Takagi's theorem, and only a few mentioned the CF theorem. Some results on the general real rational case were given by Lam [29, 31], but the method she proposes is somewhat different from ours (even in the polynomial case), and her tratment can be simplified a great deal by transforming the problem from the interval to the unit disk, as we do here.

A comprehensive theory generalizing the CF approach to the approximation of an arbitrary $f \in L_\infty(\partial D)$ by a rational function without poles in D was derived by Adamjan *et al.* [6]. However, its numerical application seems limited to the case where f itself is rational [10, 28, 40]. Here we restrict ourselves to the simpler but still very useful case in which f is (essentially) a polynomial. By slightly extending Takagi's results [41, 42] we first present a complete discussion of the related minimal extension problem and of the related CF table (Sections 1 and 2). By applying a modification of the splitting technique previously used by Gragg [22, 23] and others in Padé approximation, we then introduce CF approximation by complex (Laurent–CF) and real (Fourier–CF) trigonometric rational functions as well as by real rational functions on a real interval (Chebyshev–CF), which now emerges as a special case (Section 3).

CF approximation and its extensions to rational functions f have a close connection with algebraically solvable examples of best polynomial or rational approximation. In conjunction with Talbot's theory [43, 44] many published examples are easily explained in a unified way. We hope to cover this in a forthcoming paper.

Recently, the principle of the CF method has also been applied to Faber series in order to compute near-best polynomial and rational approximations on "general" simply connected domains in the plane [16, 17].

1. Takagi's Extension of the CF Theorem

When investigating the singular value problem related to the CF problem, Takagi [41, 42] found that *all* singular values and singular vectors correspond to certain minimal meromorphic extensions of the given polynomial. His results were later partly rediscovered by Achieser [3–5], Lam [29, 31], and Talbot [43, 44]. Finally, they were generalized by Clark [12, 13], and Adamjan *et al.* [6], who by heavy use of results from functional analysis came up with an appealing and complete, but difficult, theory. For the derivation and the discussion of a numerical method based upon this theory it is nevertheless sufficient to study Takagi's case. In this section we summarize his results enriched by a few simple consequences.

Let $\mathscr{R}_{mn}$ denote the set of rational functions with numerator degree at most m, denominator degree at most n, *and no poles on the unit circle* ∂D. Let $\mathscr{P}_m := \mathscr{R}_{m0}$; for $p \in \mathscr{P}_m$, $p(z) = a_0 + a_1 z + \cdots + a_m z^m$ with $a_0 \neq 0$, $a_m \neq 0$ define the recipocal polynomial p^* by $p^*(z) = \bar{a}_m + \bar{a}_{m-1} z + \cdots + \bar{a}_0 z^m$ (the bar denotes complex conjugation). The numbers of zeros or poles we state always include multiplicity. Finally, $\|f\|$ denotes the Chebyshev norm (ess sup norm) of f on ∂D, which in the case where the function f is not defined on ∂D but has a radial limit almost everywhere on ∂D (e.g., if $f \in H_\infty$) is defined by a limit on growing circles (cf. Rudin [38]).

Our setting of the stage is summarized in

ASSUMPTION 1.0. *Let* $h_0 \neq 0$, $h_1, h_2, \ldots$ *be given complex numbers. Consider the Hankel matrices*

$$\mathbf{H}_k := \begin{pmatrix} h_k & h_{k-1} & \cdots & h_0 \\ h_{k-1} & & \iddots & 0 \\ \vdots & & & \vdots \\ h_0 & 0 & \cdots & 0 \end{pmatrix}, \qquad k = 0, 1, \ldots.$$

Fo fixed $k = K$ *let* $\mathbf{H}_K$ *have the singular values* $\sigma_0 \geqslant \sigma_1 \geqslant \cdots \geqslant \sigma_K$ (>0). *Set* $\sigma_{-1} := \infty$, $\sigma_{K+1} := 0$ *and assume that*

$$\sigma_{\nu-1} > \sigma_\nu = \cdots = \sigma_{\nu+\mu} > \sigma_{\nu+\mu+1}; \tag{1.1}$$

hence, if $\rho_k(\sigma_\nu) := \dim \ker(\mathbf{H}_k^H \mathbf{H}_k - \sigma_\nu^2 \mathbf{I}_k)$ *denotes the multiplicity of* σ_ν *as singular value of* $\mathbf{H}_k$, *we assume* $\rho_K(\sigma_\nu) = \mu + 1$.

Takagi's results in [41], which include extensions of the classical CF theorem [11, 39] and of a theorem due to Landau [34], are only valid in what he later [42] called the *regular case*, i.e., under the additional assumption $\rho_{K-1}(\sigma_\nu) = \mu$, which he missed in [41]. We call σ_ν a *regular singular value* of $\mathbf{H}_K$ if this assumption holds. Since the general results that include the irregular case are quite complicated, we state those for the regular case first.

THEOREM 1.1 [41, Theorems I, III, IV]. *If Assumption* 1.0 *holds and if* σ_ν *is a regular singular value of* $\mathbf{H}_K$, *then*:

(i) *There is a unique unimodular rational function Π of degree at most K such that for $z \to 0$*

$$\sigma_\nu \Pi(z) = h_0 + h_1 z + \cdots + h_K z^K + O(z^{K+1}). \tag{1.2}$$

(ii) $\Pi(z) = \Pi_w(z) := w(z)/w^*(z)$, *where $w \in \mathscr{P}_{K-\mu}$ has exactly ν zeros outside and $K-\mu-\nu$ zeros inside ∂D and where w and w^* are mutually prime and of exact degree $K-\mu$.*

(iii) *If $\boldsymbol{u} = (u_0, ..., u_K)^T$ is any left singular vector of $\mathbf{H}_K$ satisfying*

$$\mathbf{H}_K \bar{\mathbf{u}} = \sigma_\nu \mathbf{u}, \tag{1.3}$$

then

$$\frac{u_K + u_{K-1} z + \cdots + u_0 z^K}{\bar{u}_0 + \bar{u}_1 z + \cdots + \bar{u}_K z^K} = \frac{w(z)}{w^*(z)}$$

are representations of the same rational function.

(iv) *Every function h meromorphic in D, regular at 0, uniformly bounded in some annulus $R_h \leqslant |z| < 1$, with at most $\nu + \mu$ poles in D, for which*

$$h(z) = h_0 + h_1 z + \cdots + h_K z^K + O(z^{K+1}) \tag{1.4}$$

as $z \to 0$, satisfies $\|h\| \geqslant \sigma_\nu$. Equality holds only for $h = \sigma_\nu \Pi_w$.

(v) *Every function h meromorphic in D, regular at 0, with at most $K-\nu$ zeros in D, for which (1.4) holds as $z \to 0$, satisfies*

$$\lim_{R \uparrow 1} \min_{z \in \partial D} |h(Rz)| \leqslant \sigma_\nu. \tag{1.5}$$

Equality holds only for $h = \sigma_\nu \Pi_w$.

Remarks. (a) The set of vectors $\mathbf{u}$ satisfying (1.3) is not the full left-hand singular space $\mathscr{U}_K(\sigma_\nu)$ (spanned by the $\mu + 1$ columns corresponding to σ_ν of the unitary matrix $\mathbf{U}$ in any singular value decomposition $\mathbf{H}_K = \mathbf{U\Sigma V}^H$ of $\mathbf{H}_K$), but it spans $\mathscr{U}_K(\sigma_\nu)$. (If imbedded into $\mathbb{R}^{2K+2}$ the set mentioned is a *real* subspace of dimension $\mu + 1$.) The full space $\mathscr{U}_K(\sigma_\nu)$ is obtained if in Theorem 1.2 below the polynomial s is not required to be self-reciprocal.

(b) Takagi requires in (iv) and (v) that h be meromorphic in $\bar{D}$. But if h is meromorphic in D and $\inf |h(z)|$ is positive in some annulus $R_h \leqslant |z| < 1$ (otherwise (1.5) is trivial), then $1/h$ is the sum of a rational function with poles in D and a bounded analytic function (in H_∞); hence the limit in (1.5) is well defined, and it is easy to check that Takagi's proof (including the uniqueness statement based upon the classical CF theorem) remains valid. A similar argument applies to (iv).

Takagi's results for the irregular case are mainly based on an observation concerning $\rho_k(\sigma_\nu)$ for various $k \leqslant K$. This observation, which will also be basic for our description of the CF table (Section 2), allowed him to reduce the irregular to the regular case. It is itself based on two facts.

First, due to the special structure of $\mathbf{H}_k$, $\mathbf{H}_{k-1}^H \mathbf{H}_{k-1}$ may be obtained from $\mathbf{H}_k^H \mathbf{H}_k$ by deleting the first row and the first column of this latter matrix. Hence, the singular values $\sigma_0^{(k)} \geqslant \cdots \geqslant \sigma_k^{(k)}$ of $\mathbf{H}_k$ (whose squares are the eigenvalues of $\mathbf{H}_k^H \mathbf{H}_k$) interlace with those of $\mathbf{H}_{k-1}$ [20]:

$$\sigma_j^{(k)} \geqslant \sigma_j^{(k-1)} \geqslant \sigma_{j+1}^{(k)} \qquad (j = 0,..., k-1). \tag{1.6}$$

(This is a special case of a general result [45].) In particular,

$$\rho_k(\sigma_\nu) - \rho_{k-1}(\sigma_\nu) \in \{-1, 0, 1\}. \tag{1.7}$$

(We have $\rho_k(\sigma_\nu) := 0$ if σ_ν is not a singular value of H_k.)

Now, assume that $u_0 = u_1 = \cdots = u_{\gamma-1} = 0$ for every solution of (1.3), but $u_\gamma \neq 0$ for some solution. Then (1.3) implies $u_K = u_{K-1} = \cdots = u_{K-\gamma+1} = 0$, $u_{K-\gamma} \neq 0$, and

$$\sigma_\nu \frac{u_{K-\gamma} + \cdots + u_\gamma z^{K-2\gamma}}{\bar{u}_\gamma + \cdots + \bar{u}_{K-\gamma} z^{K-2\gamma}} = h_0 + h_1 z + \cdots + h_{K-\gamma} z^{K-\gamma} + O(z^{K-\gamma+1}). \tag{1.8}$$

Hence, σ_ν is a regular singular value of $H_{K-2\gamma}$. Moreover, since each of the $\mu + 1$ linearly independent singular vectors yields the same rational function in (1.8), one must be able to cancel a self-reciprocal factor $s = s^* \in \mathscr{P}_\mu$ in (1.8), cf. [42]; so the left-hand side of (1.8) reduces to $\sigma_\nu w(z)/w^*(z)$, where $w, w^* \in \mathscr{P}_{K-2\gamma-\mu}$ are mutually prime. On the other hand, if we multiply both w and w^* by *any* polynomial of the form $z^{j'} s(z)$ with $s = s^* \in \mathscr{P}_{j''}$ and $j' + j'' = j \leqslant \gamma + \mu$, the resulting polynomial $v(z) = v_{K-2\gamma-\mu+j} + \cdots + v_0 z^{K-2\gamma-\mu+j}$ still defines a solution of $\mathbf{H}_{K-2\gamma-\mu+j} \bar{\mathbf{v}} = \sigma_\nu \mathbf{v}$, and since we may choose $j' = 0$, σ_ν is in fact a regular singular value of $H_{K-2\gamma-\mu+j}$. We conclude that $\rho_{K-2\gamma-\mu+j}(\sigma_\nu) = \rho_{K-2\gamma-\mu}(\sigma_\nu) + j \geqslant 1 + j$; thus $\rho_{K-\gamma}(\sigma_\nu) \geqslant \gamma + \mu + 1$. On the other hand, $\rho_{K-\gamma}(\sigma_\nu) \leqslant \rho_K(\sigma_\nu) + \gamma = \gamma + \mu + 1$ due to (1.7), so equality must hold:

$$\rho_{K+l}(\sigma_\nu) = \begin{cases} 2\gamma + \mu + 1 + l, & l = -2\gamma - \mu - 1,..., -\gamma, \\ \mu + 1 - l, & l = -\gamma,..., 0. \end{cases} \tag{1.9}$$

This is Takagi's basic result on the irregular case [42, p. 16]. It is also clear now that the left-hand side of (1.8) cannot reproduce $h_0 + h_1 z + \cdots$ up to $O(z^{K-\gamma+2})$ if $\gamma > 0$. For this would imply $\rho_{K-\gamma+1}(\sigma_\nu) \geqslant \rho_{K-\gamma}(\sigma_\nu) + 1$ (by our previous argument), in contrast to (1.9).

Relation (1.9), which must also hold if $\gamma = 0$, means that for $k \in [K - 2\gamma - \mu - 1, K]$ the multiplicity $\rho_k(\sigma_\nu)$ increases from 0 linearly to

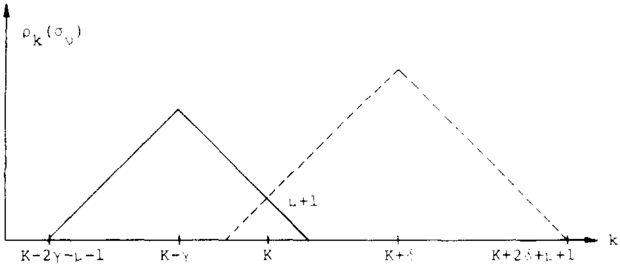

FIG. 1. Multiplicity $\rho_k(\sigma_\nu)$ of the singular value σ_ν: the two possibilities $\gamma > 0$ (——) and $\delta > 0$ (- - -).

$\mu + 1 + \gamma$ at $k = K - \gamma$ and then decreases linearly, cf. Fig. 1. Since this picture must also hold if we replace K by a larger value, we can immediately conclude that in extension of (1.9) $\rho_{K+l} = \mu + 1 - l$ if $-\gamma \leqslant l \leqslant \mu + 1$ and $\gamma > 0$. However, if $\gamma = 0$, $\rho_k(\sigma_\nu)$ may first increase beyond $k = K$. But then there is $\delta > 0$ such that

$$\rho_{K+l}(\sigma_\nu) = \begin{cases} \mu + 1 + l, & l = -\mu - 1, \ldots, \delta, \\ 2\delta + \mu + 1 - l, & l = \delta, \ldots, \mu + 2\delta + 1. \end{cases}$$

Let $\delta := 0$ if $\rho_{K+1}(\sigma_\nu) < \rho_K(\sigma_\nu) = \mu + 1$. (Thus $\gamma > 0$ implies $\delta = 0$, and $\delta > 0$ implies $\gamma = 0$.) Then we may formally define γ and δ by

$$\begin{aligned} \gamma &:= \max\{0, \min\{j: \rho_{K-j-1}(\sigma_\nu) < \rho_{K-j}(\sigma_\nu)\}\}, \\ \delta &:= \max\{0, \min\{j: \rho_{K+j+1}(\sigma_\nu) < \rho_{K+j}(\sigma_\nu)\}\}, \end{aligned} \tag{1.10}$$

and we can summarize the behaviour of $\rho_k(\sigma_\nu)$ by

$$\rho_{K+l}(\sigma_\nu) = \begin{cases} 2\gamma + \mu + 1 + l, & l = -2\gamma - \mu - 1, \ldots, \delta - \gamma, \\ 2\delta + \mu + 1 - l, & l = \delta - \gamma, \ldots, 2\delta + \mu + 1. \end{cases} \tag{1.11}$$

In particular, it is clear from Fig. 1 that σ_ν *is a regular value of* $\mathbf{H}_k$ *if* $K - 2\gamma - \mu \leqslant k \leqslant K - \gamma + \delta$, *while* σ_ν *is irregular if* $K - \gamma + \delta < k \leqslant K + 2\delta + \mu$. (Note that both cases $\gamma \geqslant 0$ and $\delta \geqslant 0$ are taken care of by this notation.) By applying Theorem 1.1 for these various regular cases and by summarizing some of the above results we finally get the following generalization of Theorem 1.1.

THEOREM 1.2. *Let Assumption* 1.0 *be satisfied and define* γ *and* δ *by* (1.10). *Then* $\gamma\delta = 0$, $\gamma \leqslant (K - \mu)/2$ *and in addition to* (1.11) *the following statements hold*:

(i) *For* $k = K - 2\gamma - \mu, \ldots, K - \gamma + \delta$ *there is a unique unimodular rational function* Π *of degree at most* k *such that for* $z \to 0$

$$\sigma_\nu \Pi(z) = h_0 + h_1 z + \cdots + h_k z^k + O(z^{k+1}). \tag{1.12}$$

(ii) $\Pi(z) = \Pi_w(z) := w(z)/w^*(z)$ *(independent of* k*), where* $w \in \mathscr{P}_{K-2\gamma-\mu}$ *has exactly* $\nu - \gamma$ *zeros outside and* $K - \gamma - \mu - \nu$ *zeros inside* ∂D *and where* w *and* w^* *are mutually prime.*

(iii) $\mathbf{u} = (u_0, \ldots, u_K)^T$ *is a left-hand singular vector satisfying* (1.3) *if and only if* $u(z) := u_K + u_{K-1} z + \cdots + u_0 z^K$ *is of the form*

$$u(z) = z^{\gamma + j} s(z)\, w(z), \tag{1.13}$$

where j *is an arbitrary integer satisfying* $0 \leqslant j \leqslant \mu$ *and* $s = s^*$ *is an arbitrary self-reciprocal polynomial of exact degree* $\mu - j$.

(iv) *For* $k = K - 2\gamma - \mu, \ldots, K - \gamma + \delta$ *every function* h *meromorphic in* D, *regular at* 0, *uniformly bounded in some annulus* $R_h \leqslant |z| < 1$, *with at most* $k - K + \gamma + \mu + \nu$ *poles in* D, *for which*

$$h(z) = h_0 + h_1 z + \cdots + h_k z^k + O(z^{k+1}) \tag{1.14}$$

as $z \to 0$, *satisfies* $\|h\| \geqslant \sigma_\nu$. *Equality holds only for* $h = \sigma_\nu \Pi_w$.

(v) *For* $k = K - 2\gamma - \mu, \ldots, K - \gamma + \delta$ *every function* h *meromorphic in* D, *regular at* 0, *with at most* $k + \gamma - \nu$ *zeros in* D, *for which* (1.14) *holds as* $z \to 0$, *satisfies* (1.5). *Equality holds only for* $h = \sigma_\nu \Pi_w$.

(vi) *Relation* (1.12) *does not hold for* $\Pi = \Pi_w$ *if* $k > K - \gamma + \delta$.

(vii) *For* $k = K - 2\gamma - \mu, \ldots, K - \gamma + \delta$ *the matrix* H_k *has exactly* $\nu - \gamma$ *singular values greater than* σ_ν *and* $K - \gamma - \mu - \nu$ *singular values smaller than* σ_ν *(independent of* k*). For* $k = K + \delta - \gamma, \ldots, K + 2\delta + \mu$ *the corresponding numbers are* $k - K - \delta + \nu$ *and* $k - \delta - \mu - \nu$.

(viii) $\Pi_w(\partial D)$ *has winding number* $\iota(\Pi_w) = K - \mu - 2\nu$ *with respect to* 0.

Proof. (i), (iv), and (v) follow from Theorem 1.1 with K replaced by k since σ_ν is a regular singular value of $\mathbf{H}_k$. (vii) is a consequence of (1.6) and (1.11). (ii), (iii), and (vi) emerged from the previous discussion; in particular, the number of zeros of w inside and outside ∂D, respectively, is derived by applying Theorem 1.1 with K replaced by k, $K - 2\gamma - \mu \leqslant k \leqslant K - \gamma + \delta$, and by using (vii). Finally, (viii) is a simple implication of (ii). ∎

Finally, let us state a related result of Takagi [42, p. 17] showing that the two cases $\nu = 0$ and $\nu + \mu = K$ are less complicated.

THEOREM 1.3 [42]. *If σ_ν is the greatest or the smallest singular value (i.e., if $\sigma_\nu = \sigma_0$ or $\sigma_\nu = \sigma_K$), then $\gamma = 0$.*

The proof is based on the fact that $\mathbf{H}_K^H \mathbf{H}_K - \sigma_\nu^2 \mathbf{I}_K$ is semi-definite and hence its rank $K + 1 - \rho_K(\sigma_\nu)$ cannot decrease by two (as it should in view of (1.11) if $\gamma > 0$) if we delete the first row and the first column of it in order to obtain $\mathbf{H}_{K-1}^H \mathbf{H}_{K-1} - \sigma_\nu^2 \mathbf{I}_{K-1}$.

Note that we might obtain a simpler formulation of Theorem 1.2 by assuming that K is *chosen* such that $\gamma = \delta = 0$. However, our formulation is the appropriate one for the application in the next section.

2. CF APPROXIMATION ON THE UNIT DISK

Let $\mathscr{H}$ denote the space of functions that are analytic and bounded outside the unit circle and vanish at ∞ (i.e., $f \in \mathscr{H}$ iff $z \mapsto f(1/z)/z \in H_\infty$). Let $\mathscr{R}_{mn}^0$ be the subset of functions in $\mathscr{R}_{mn}$ having all *poles* outside ∂D. In contrast let $\mathscr{P}_m^0 \subset \mathscr{P}_m$ consist of those polynomials with all *zeros* outside ∂D. Following Trefethen [47] we define, for $m, n \in \mathbb{Z}$, $n \geqslant 0$,

$$\tilde{\mathscr{R}}_{nn} := \mathscr{R}_{nn}^0 + \mathscr{H}, \qquad \tilde{\mathscr{R}}_{mn} := z^{m-n} \tilde{\mathscr{R}}_{nn}, \qquad \tilde{\mathscr{P}}_m := \tilde{\mathscr{R}}_{m0}.$$

$\tilde{r} \in \tilde{\mathscr{R}}_{mn}$ iff $\tilde{r}$ is meromorphic in $1 < |z| \leqslant \infty$, has $\nu \leqslant n$ poles in $1 < |z| < \infty$, is bounded in some annulus $1 < |z| \leqslant R_{\tilde{r}}$, and is of order (at most) $O(z^{m-\nu})$ as $z \to \infty$. Also, $\tilde{r} \in \tilde{\mathscr{R}}_{mn}$ iff $\tilde{r}$ can be written in the form

$$\tilde{r}(z) = \frac{1}{q(z)} \sum_{j=-\infty}^{m} \tilde{a}_j z^j, \tag{2.1}$$

where $q \in \mathscr{P}_n^0$ and the series converges for $|z| > 1$ and is bounded in $1 < |z| < R$ for some (and hence for every) $R > 1$.

The problem that can be solved on the basis of Theorem 1.2 is

Problem A. *Given $m < M$, $n \geqslant 0$, $\tilde{f} \in \tilde{\mathscr{P}}_M \setminus \tilde{\mathscr{P}}_{M-1}$ and $g \in \tilde{\mathscr{P}}_L \setminus \tilde{\mathscr{P}}_{L-1}$, g nonvanishing in $1 \leqslant |z| < \infty$, find $\tilde{r} \in \tilde{\mathscr{R}}_{mn}$ that minimizes $\| g\tilde{f} - g\tilde{r} \|$.*

First we note that—given g as above—we have $g\tilde{f} \in \tilde{\mathscr{P}}_{L+M}$, $g\tilde{r} \in \tilde{\mathscr{R}}_{L+m,n}$ iff $\tilde{f} \in \tilde{\mathscr{P}}_M$, $\tilde{r} \in \tilde{\mathscr{R}}_{mn}$; hence, the unweighted best approximation $\tilde{r}_g$ to $\tilde{f}_g := g\tilde{f}$ out of $\tilde{\mathscr{R}}_{L+m,n}$ yields the solution $\tilde{r} := \tilde{r}_g/g \in \tilde{\mathscr{R}}_{mn}$ of the weighted problem. Therefore we may assume $g(z) \equiv 1$ in our discussion. Furthermore, we assume $\tilde{f}$ given in the form

$$\tilde{f}(z) = \sum_{j=0}^{\infty} h_j z^{M-j} \qquad (|z| > 1), \tag{2.2}$$

and define $f \in z^{m-n+1}\mathscr{P}_K$, where $K := M - m + n - 1$, and $f^0 \in \tilde{\mathscr{P}}_{m-n}$ by

$$f(z) := \sum_{j=0}^{K} h_j z^{M-j}, \qquad f^0(z) := \tilde{f}(z) - f(z). \tag{2.3}$$

Now, $r^0 \in \tilde{\mathscr{R}}_{mn}$ implies $\tilde{r} := r^0 + f^0 \in \tilde{\mathscr{R}}_{mn}$ and $\|\tilde{f} - \tilde{r}\| = \|f - r^0\|$. Consequently, we are left with the problem of approximating $f \in z^{m-n+1}\mathscr{P}_K$ by $r^0 \in \tilde{\mathscr{R}}_{mn}$. This can be accomplished with Theorem 1.2: We choose ν such that $\nu \leqslant n \leqslant \nu + \mu$ and approximate

$$z^M f(1/z) = h_0 + h_1 z + \cdots + h_K z^K$$

by $z^M f(1/z) - \sigma_n \Pi_w(z)$. This means we approximate $f(z)$ by $f(z) - \sigma_n z^M \Pi_{w^*}(z)$ (where $\Pi_{w^*}(z) = w^*(z)/w(z)$) and $\tilde{f}(z) = f(z) + f^0(z)$ by

$$\tilde{r}^*(z) := \tilde{f}(z) - \sigma_n z^M \Pi_{w^*}(z). \tag{2.4}$$

We know according to (1.12) with $k = K - \gamma$ and $z := 1/z$ that

$$f(z) - \sigma_n z^M \Pi_{w^*}(z) = O(z^{m-n+\gamma}) \qquad \text{as} \quad z \to \infty. \tag{2.5}$$

But if $\delta > 0$ and if we take further coefficients $h_{K+1}, \ldots, h_{K+\delta}$ of $\tilde{f}$ into account, (1.12) holds even for $k = K - \gamma + \delta$. So in view of assertion (vi) of Theorem 1.2 we have

$$\tilde{r}^*(z) = O(z^{m-n+\gamma-\delta}) \qquad \text{but not} \qquad O(z^{m-n+\gamma-\delta-1}) \qquad \text{as} \quad z \to \infty. \tag{2.6}$$

(In other words, if we replaced m, K by $m - \delta$, $K + \delta$, respectively, we would end up with the same $\tilde{r}^*$.) Now Π_{w^*} and therefore $\tilde{r}^*$ have exactly $\nu - \gamma$ poles outside ∂D, cf. assertion (ii), so that (2.6) leads to

$$\begin{aligned} \tilde{r}^* &\in \tilde{\mathscr{R}}_{m',n'}, \qquad m' := m - n + \nu - \delta,\ n' := \nu - \gamma, \\ \tilde{r}^* &\notin \tilde{\mathscr{R}}_{m'',n''}, \qquad \text{if } m'' < m' \text{ or } n'' < n'. \end{aligned} \tag{2.7}$$

According to (2.4) the error function $\tilde{f} - \tilde{r}^*$ has constant modulus σ_n and winding number $M - \iota(\Pi_w)$, i.e.,

$$\iota(\tilde{f} - \tilde{r}^*) = m - n + \mu + 2\nu + 1 \geqslant m + \nu + 1. \tag{2.8}$$

Assume now that $\tilde{r} \in \mathscr{R}_{m'+l,n'+l}$ with $l := \delta + \gamma + \mu$ is a better approximation of f. Then by a Rouché-type argument $\iota(\tilde{r} - \tilde{r}^*) = \iota(\tilde{f} - \tilde{r}^*) = m' + n' + l + 1$, but on the other hand $\iota(\tilde{r} - \tilde{r}^*) \leqslant m' + n' + l$ since $\tilde{r} - \tilde{r}^* \in \tilde{\mathscr{R}}_{m'+n'+l,n'+l}$, cf. [47, Lemma 2.3]. Hence, $\tilde{r}^*$ is best out of $\tilde{\mathscr{R}}_{m'+l,n'+l}$.

However, Theorem 1.2 even implies uniqueness: Let l' denote the greatest integer with $\tilde{r} \in \mathscr{R}_{m'+l,n'+l-l'}$. Then, $\tilde{r}(z) = O(z^{m'-n'+l'}) = O(z^{m-n+\gamma-\delta+l'})$

as $z \to \infty$, and $z^M \tilde{r}(1/z) = O(z^{K+1-\gamma+\delta-l'})$ as $z \to 0$. Since $z^M \tilde{f}(1/z) = h_0 + h_1 z + \cdots$, the function $\tilde{h}$ defined by

$$\tilde{h}(z) := z^M [\tilde{f}(1/z) - \tilde{r}(1/z)], \tag{2.9}$$

which has exactly $n' + l - l' = \delta + \mu + \nu - l'$ poles, satisfies (1.14) with $k = K - \gamma + \delta - l'$, where $0 \leqslant l' \leqslant n' + l = \delta + \mu + \nu$. So, according to assertion (iv) of Theorem 1.2 we have $\|\tilde{h}\| > \sigma_n$ unless $\tilde{h} = \sigma_\nu \Pi_{w'}$; or, in view of (2.4) and (2.9), $\|\tilde{f} - \tilde{r}\| > \sigma_\nu$ unless $\tilde{r} = \tilde{r}^*$. Summarizing we get

THEOREM 2.1. *Problem* A *has a unique solution* $\tilde{r}^*$. *If we assume* $g(z)\tilde{f}(z) = h_0 z^{M+L} + h_1 z^{M+L-1} + \cdots$, *set* $K := M - m + n - 1$, *and adopt the notation of Assumption* 1.0 *and Theorem* 1.2, *choosing* ν *there such that* $\nu \leqslant n \leqslant \nu + \mu$, *then* $\tilde{r}^*$ *is given by*

$$\tilde{r}^*(z) := \tilde{f}(z) - \sigma_n z^{L+M} \Pi_{w'}(z)/g(z). \tag{2.10}$$

$\tilde{r}^*$ *satisfies* (2.7), *and is best out of* $\tilde{\mathscr{R}}_{m'+l, n'+l}$, *where* $l := \gamma + \delta + \mu$. *The weighted error curve* $(g\tilde{f} - g\tilde{r}^*)(\partial D)$ *is a circle of radius* $\sigma_n > 0$ *around the origin and has the winding number* $L + \iota(\tilde{f} - \tilde{r}^*)$, *where* (2.8) *holds for the second term.*

In particular, if $n = \nu$ (*i.e.*, $\sigma_{n-1} > \sigma_n$), $\tilde{r}^* \in \tilde{\mathscr{R}}_{m-\delta, n-\gamma}$ *is the best approximation out of* $\tilde{\mathscr{R}}_{m+\gamma+\mu, n+\delta+\mu}$ *and* $\iota(\tilde{f} - \tilde{r}^*) = m + n + \mu + 1$.

Remark. Adamjan *et al.* [6], who treat the corresponding problem with $\tilde{f} \in L_\infty(\partial D)$, $g(z) \equiv 1$, and $m = n$, do not give our details on the actual degrees of $\tilde{r}^*$.

Theorem 2.1 is best illustrated in terms of the *CF table*: For given $\tilde{f}$ and g the map $(m, n) \mapsto \tilde{r}^*$ induces a partition of the quadrant $m < M$, $n \geqslant 0$ of the (m, n)-plane into disjoint square blocks in each of which $\tilde{r}^*$ and σ_n are fixed, cf. Fig. 2.

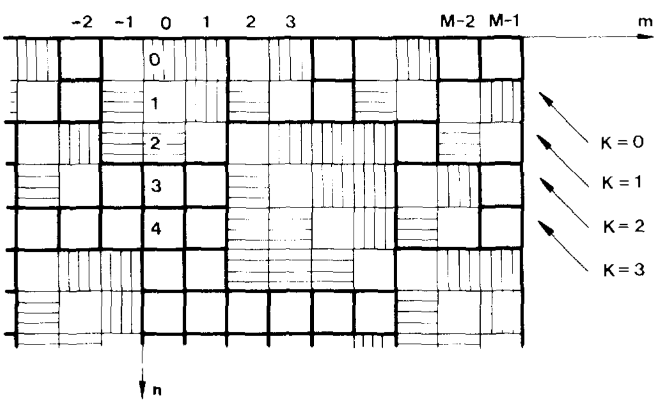

FIG. 2. The CF table.

Any diagonal of the table corresponds to a fixed K, and the $K+1$ not necessarily distinct elements $\tilde{r}^*$ on a diagonal have errors $\sigma_0^{(K)} \geqslant \sigma_1^{(K)} \geqslant \cdots \geqslant \sigma_K^{(K)}$. In any induced square block of length >1 all elements on the lth lower subdiagonal have $\gamma = l$ and all elements on the lth upper subdiagonal have $\delta = l$; on the diagonal itself $\gamma = \delta = 0$. The sum $\gamma + \delta + \mu$ is fixed in each block and equals its length. It is clear that $\gamma > 0$ is impossible in the top row $(n = 0)$ and the rightmost column $(m = M - 1)$. More generally, the assertion of Theorem 1.3 becomes evident. On the other hand, it seems clear that $\gamma > 0$ or $\delta > 0$ may occur even if all singular values of $\mathbf{H}_K$ are simple. Finally, the quadrant $m \geqslant M$, $n \geqslant 0$ of the table can be thought of as an infinite block with zero error $(\tilde{r}^* = \tilde{f})$.

Our original aim was to approximate $\tilde{f}$ by $r \in \mathscr{R}^0_{mn}$, where $m, n \geqslant 0$ are given. Once $\tilde{r}^*$ is known, there are at least four reasonable ways to accomplish this. The simplest one is evident from (2.1): If $\tilde{r}^*$ is written in this form, truncating the negative terms of the Laurent series (with respect to ∂D) of the "numerator" $\tilde{r}(z)\, q(z)$ yields

$$r_1^{\mathrm{cf}}(z) := \frac{1}{q(z)} \sum_{j=0}^{m} \tilde{a}_j z^j. \tag{2.11}$$

For simplicity reasons this was Trefethen's choice [47, p. 310]. We call r_1^{cf} the *truncated* or *type* 1 CF *approximation*. (In analogy to Padé approximation it could also be called the *Frobenius-type* CF *approximation*.) It is clear from (2.7) that $r_1^{\mathrm{cf}} \in \mathscr{R}^0_{m'n'}$ if $m' \geqslant 0$ and $r_1^{\mathrm{cf}}(z) \equiv 0$ if $m' < 0$. However, a further reduction of both degrees may be possible: we cannot exclude the possibility that $r_1^{\mathrm{cf}} \in \mathscr{R}^0_{m''n''}$ with $m' - m'' = n' - n'' > 0$. Moreover, we cannot exclude the possibility that another type 1 CF approximation with the same or smaller actual degrees is obtained when starting from m'', n'' instead of m, n.

The definition of the *type* 2 CF *approximation* r_2^{cf} is in general less straightforward. Let $\tilde{c}_k$ denote the Laurent coefficients (with respect to ∂D) of $\tilde{r}^*$, and let

$$\tilde{r}^0(z) := \sum_{k=0}^{\infty} \tilde{c}_k z^k$$

denote the analytic part of the Laurent series. We choose $p \in \mathscr{P}_m$ as the mth partial sum of the Maclaurin series of $\tilde{r}^0(z)\, q(z)$ and define $r_2^{\mathrm{cf}} := p/q$. Note that this is equivalent to requiring

$$\frac{p(z)}{q(z)} - \tilde{r}^0(z) = O(z^{m+1}) \qquad \text{as} \quad z \to 0. \tag{2.12}$$

Hence this definition can be thought of as a fixed denominator Padé-type approximation of the analytic part $\tilde{r}^0$ of $\tilde{r}^*$ [9]. However, in the case

$m \geqslant n-1$ it turns out that $\tilde{r}^0 \in \mathscr{R}^0_{mn}$ itself, so that trivially $r_2^{\rm cf} = \tilde{r}^0$. In fact, since

$$\tilde{\mathscr{R}}_{mn} = z^{m-n}(\mathscr{R}^0_{nn} + \mathscr{H}) = z^{m-n}(z\mathscr{R}^0_{n-1,n} + \mathbb{C} + \mathscr{H}),$$

it is easy to verify that $\tilde{r}^*$ can be written in the form

$$\tilde{r}^*(z) = \frac{1}{q(z)} \sum_{j=m-n+1}^{m} \tilde{\tilde{a}}_j z^j + \sum_{j=-\infty}^{m-n} \tilde{c}_j z^j. \tag{2.13}$$

Thus, if $m \geqslant n-1$, deleting the co-analytic part of the series leads to a function in $\mathscr{R}^0_{mn}$, cf. [47, p. 310]. Note that (2.13) still holds if m and n are replaced by m' and n', respectively. Hence, $\tilde{r}^0 \in \mathscr{R}^0_{m'n'}$ if $m' \geqslant n'-1$.

Our third choice is a natural extension of the second one: We define the *type* 3 CF *approximation* $r_3^{\rm cf}$ as the (m, n) Padé approximation of $\tilde{r}^0$. Note that an important property of the CF approximation is sacrificed: $r_3^{\rm cf}$ may have poles on or inside the unit circle.

Our fourth proposal is a pragmatic combination of type 1 and type 2 and is called *type* 2′ CF *approximation* here. Since $\tilde{r}^0 = r_2^{\rm cf} \in \mathscr{R}^0_{mn}$ for $m \geqslant n-1$, we let $r_{2'}^{\rm cf} := \tilde{r}^0$ in this case. If $m < n-1$, we define

$$r_{2'}^{\rm cf}(z) := \frac{1}{q(z)} \sum_{j=0}^{m} \tilde{\tilde{a}}_j z^j. \tag{2.14}$$

This is just a type 1 approximation of $s(z) := \sum_{j=m-n+1}^{m} \tilde{\tilde{a}}_j z^j / q(z)$ appearing in (2.13). Note that the second sum there, which is deleted first, is not the whole co-analytic part $\tilde{r}^* - \tilde{r}^0$ of $\tilde{r}^*$. However, s has the correct asymptotic behavior for $z \to \infty$, while $\tilde{r}^0$ has not, in general. In fact, from (2.13) one can see that $\tilde{r}^0 \in \mathscr{R}^0_{n-1,n}$ if $m \leqslant n-1$. Hence, $r_2^{\rm cf}$ is obtained from $\tilde{r}^0$ by deleting the terms of degree greater than m in the numerator of $\tilde{r}^0$, while $r_{2'}^{\rm cf}$ is obtained from s by deleting the negative powers in the "numerator" of s.

Yet another proposal that is worth pursuing is the *type* 4 CF *approximation* $r_4^{\rm cf}$ defined as the best L_2 approximation on ∂D of $\tilde{r}$ by elements of the form p/q with arbitrary $p \in \mathscr{P}_m$ but fixed q (as in (2.12)). Since $\tilde{r}^0$ is the best L_2 approximation of $\tilde{r}$ by functions analytic in D, we have again $r_4^{\rm cf} = \tilde{r}^0$ if $m \geqslant n-1$.

In order to maintain the structure of the CF table one may replace in all these definitions m and n by m' and n', respectively.

Numerical experiments performed by Trefethen (private commun.) indicate that for $f(z) = e^z$ the type 2 and the type 3 approximations yield error curves that are roughly twice as circular (and close to best) as those obtained by method 1. Hence, in this sense they are substantially better, though the error itself is diminished only very slightly, of course. However, for not so smooth functions the type 3 approximation may be completely unsatisfactory when the other two yield still reasonable results.

The derivation of bounds for the truncation errors $\|\tilde{r}^* - r_j^{\mathrm{cf}}\|$ $(j = 1,..., 4)$ is a difficult problem. It was addressed by Hollenhorst [27] in the unweighted polynomial case (where $j = 1,..., 4$ are all equivalent) and, in an asymptotic sense on small disks, by Trefethen [46, 47] in the unweighted rational case (where again j turns out to be insignificant, asymptotically).

Algorithmic details on how to compute $\tilde{r}^*$ and r_1^{cf} and numerical examples are given in [24, 25, 47]. The reduction of Π_{u^*} to Π_{w^*}, which is additionally needed in degenerate cases, can be done with Euclid's algorithm for the greatest common divisor.

3. Laurent–, Fourier–, and Chebyshev–CF Approximation

As in Padé approximation [22, 23], one can use an additive splitting of the Fourier series of $t \mapsto \tilde{f}(e^{it})$ to adapt the CF method to the case where $\tilde{f}$ is not analytic in D. (Though we did not assume in Section 2 that $\tilde{f}$ is analytic in D, it is clear that the co-analytic part of $\tilde{f}$ fully reappears in the error function $\tilde{f} - r^{\mathrm{cf}}$ if $m \geqslant n-1$.) Of course, polynomials and rational functions analytic in D are not suitable for approximating such a function $\tilde{f}$. Instead we approximate now by Laurent polynomials $p \in z^{-m}\mathscr{P}_{2m}$ of given degree m (i.e., p is a linear combination of $z^{-m}, z^{-m+1},..., 1,..., z^{m-1}, z^m$) and by quotients of Laurent polynomials. We denote these spaces by

$$\mathscr{E}_m := z^{-m}\mathscr{P}_{2m},$$
$$\mathscr{E}_{mn} := \{p/q;\ p \in \mathscr{E}_m, q \in \mathscr{E}_n, q(z) \neq 0 \text{ on } \partial D\},$$

respectively. ($\mathscr{E}_{mn} := \mathscr{E}_m := \{0\}$ if $m < 0$.)

The basic idea is the following one: Given a function $\tilde{f} \in L_\infty(\partial D)$ whose conjugate function is also bounded, there exist $\tilde{f}^+, \tilde{f}^- \in H_\infty$ such that

$$\tilde{f}(z) = \tilde{f}^+(z) + \tilde{f}^-(1/z) \qquad \text{a.e. on } \partial D, \qquad \tilde{f}^+(0) = \tilde{f}^-(0), \tag{3.1}$$

cf. [38, p. 264]. According to the theory of Adamjan *et al.* [6], there exist best approximations $\tilde{r}^\pm \in \tilde{\mathscr{R}}_{mn}$ to $\tilde{f}^\pm$. Truncating $\tilde{r}(z) := \tilde{r}^+(z) + \tilde{r}^-(1/z)$ to an element of $\mathscr{E}_{mn}$ yields an approximation of $\tilde{f}$ that one may hope is close to best. In practice one modifies this method by first truncating or, more generally, approximating $\tilde{f}^\pm$ by $f^\pm \in \mathscr{P}_M$, $M \gg m$; then, Theorem 2.1 is applicable to $f^\pm$. However, as in Padé approximation, this approach does not work if $m < n$: Typically, $\tilde{r}^\pm$ is close to an element $r^\pm \in \mathscr{R}_{mn}^0$, hence $\tilde{r}$ is close to $z \mapsto r^+(z) + r^-(1/z)$, but the latter is in $\mathscr{E}_{\max\{m,n\},n}$ and, in general, not in $\mathscr{E}_{mn}$ if $m < n$. We have been able to overcome this problem, but the resulting extended method is less obvious than the basic idea just described. (In particular, we have to redefine $f^{\pm}$, $\tilde{r}^{\pm}$, and $\tilde{r}$.)

Special attention is deserved by the case where the given function $\tilde{f}$ is real-valued on D. Then $\tilde{f}^{-}=\overline{\tilde{f}^{+}}$, i.e., the Maclaurin coefficients of $\tilde{f}^{-}$ are complex conjugate to those of $\tilde{f}^{+}$. If, in addition, these coefficients are real, the Fourier series of $t\mapsto\tilde{f}(e^{it})$ can be written as a cosine series and is equivalent to the Chebyshev series of $x\mapsto\tilde{f}(e^{i\arccos x})$, which is well defined on $I:=[-1,1]$. In these two cases the approximating family can be restricted by the same conditions. We define for $m, n\geqslant 0$

$$\mathscr{R}'_{mn}:=\{r\in\mathscr{R}_{mn};\, r(\partial D)\subset\mathbb{R}\},$$
$$\mathscr{R}''_{mn}:=\{r\in\mathscr{R}'_{mn};\, r(\mathbb{R})\subset\mathbb{R}\cup\{\infty\}\},$$
$$\mathscr{R}'_{m}:=\mathscr{R}'_{m0},\qquad \mathscr{R}''_{m}:=\mathscr{R}''_{m0}.$$

On ∂D the spaces $\mathscr{R}_m$ ($\mathscr{R}'_m$) and $\mathscr{R}_{mn}$ ($\mathscr{R}'_{mn}$) are isomorphic to the spaces of complex (real) trigonometric polynomials of degree at most m and complex (real) trigonometric rational functions with degrees at most m and n, respectively. $\mathscr{R}''_m$ corresponds to the subset of even trigonometric polynomials, which upon the substitution $x=\cos t$ is seen to be equivalent to the set of ordinary (algebraic) real polynomials (in x) of degree at most m on I. Similarly, $\mathscr{R}''_{mn}$ is equivalent to the set of real rational functions in x with numerator degree at most m, denominator degree at most n, and without poles on I. Though we will have these equivalent sets in mind, for simplicity we will always stick to the variable z.

We can again allow a weight function, which is now assumed to be the square root of a positive trigonometric polynomial. (Of course, the case of a positive trigonometric polynomial itself is included.)

Here is the general definition of $\tilde{r}$: *Given* $f\colon z\mapsto a_{-M^-}z^{-M^-}+\cdots+a_0+\cdots+a_{M^+}z^{M^+}$ (*with* $a_{-M^-}\neq 0$, $a_{M^+}\neq 0$, $M^-\geqslant 0$, $M^+\geqslant 0$), $g\in\mathscr{R}'_L\backslash\mathscr{R}'_{L-1}$ $(L\geqslant 0)$ *with* $g(z)>0$ *on* ∂D, *and* $m,n\geqslant 0$, *set* $f^+:=f$, $f^-(z):=f(1/z)$, *and determine* g^+ *and* $g^-:=\overline{g^+}$ *by spectral factorization such that all zeros of* g^+ *are inside* ∂D *and*

$$g(z)=g^+(z)\,\overline{g^-}(1/z). \tag{3.2}$$

Solve Problem A *twice for* $f^{\pm}$, $g^{\pm}$, m, *and* n, *and denote the solution by* $\tilde{r}^{\pm}$ (*if* $m\geqslant M^{\pm}$, *set* $\tilde{r}^{\pm}:=f^{\pm}$); *then let* $e^{\pm}:=f^{\pm}-\tilde{r}^{\pm}$, *and define the Laurent*–CF *extension* $\tilde{r}$ *of* f *by*

$$\tilde{r}(z):=f(z)-e^+(z)-e^-(1/z). \tag{3.3}$$

Remarks. (a) For the spectral factorization required in (3.2) there exist algorithms which directly produce the coefficients of g^+, see [7, 37, 49, 50]. There is no need to determine the zeros of g.

(b) If $f \in \mathscr{E}'_M(M := M^+ = M^-)$, then $f^- = \overline{f^+}$, and hence $\tilde{r}^- = \overline{\tilde{r}^+}$; so, only one Problem A has to be solved. If $f \in \mathscr{E}''_m$ and $g \in \mathscr{E}''_L$, then $f^- = f^+ = \overline{f^+}$ and $\tilde{r}^- = \tilde{r}^+ = \overline{\tilde{r}^+}$.

In order to get from $\tilde{r}$ to a function $r^{\text{cf}} \in \mathscr{E}_{mn}$ that is easily computed from $\tilde{r}$ but nevertheless in some sense an optimal approximation of $\tilde{r}$, we need some further insight into the structure of $\tilde{r}$. We proceed essentially as in [48], but we need superscripts $^+$ and $^-$ for all quantities appearing in the solutions of Problem A for $f^\pm$, $g^\pm$, m, and n. In particular, $q^\pm$ denotes the "denominator" of $\tilde{r}^\pm$, i.e., the monic polynomial of degree $\nu^\pm - \gamma^\pm$ whose zeros are the zeros of $w^\pm$ lying outside ∂D, and we set

$$q(z) := q^+(z)\, q^-(1/z).$$

In view of (2.6),

$$f^\pm(z) - e^\pm(z) = \tilde{r}^\pm(z) = O(z^{m-n+\gamma^\pm-\delta^\pm}) \qquad \text{as} \quad z \to \infty,$$

$$[f(z) - e^+(z)]\, q(z) = O(z^{m-n+\nu^+-\delta^+}) \qquad \text{as} \quad z \to \infty, \tag{3.4a}$$

$$[f(z) - e^-(1/z)]\, q(z) = O(z^{-(m-n+\nu^--\delta^-)}) \qquad \text{as} \quad z \to 0. \tag{3.4b}$$

$[f(z) - e^+(z)]\, q(z)$ is analytic outside ∂D except for a pole of order $m - n + \nu^+ - \delta^+ \leqslant m$ at ∞. By (3.4a) all terms of order greater than $m - n + \nu^+ - \delta^+$ in the Laurent series of

$$\tilde{r}(z)\, q(z) = [f(z) - e^+(z) - e^-(1/z)]\, q(z) \tag{3.5}$$

are due to $e^-(1/z)\, q(z)$. Likewise, by (3.4b) all terms of order less than $-(m - n + \nu^- - \delta^-)$ are due to $e^+(z)\, q(z)$. Let us denote the Laurent coefficients of $e^+(z)\, q(z)$ and $e^-(z)\, q(1/z)$ by e_k^+ and e_k^-, respectively, and let

$$e^T := \sum_{k=-m}^{M^-+\nu^--\delta^-} e_k^- z^{-k} + \sum_{k=-m}^{M^++\nu^+-\delta^+} e_k^+ z^k, \tag{3.6a}$$

$$e^R(z) := \sum_{k=-\infty}^{-m-1} (e_k^- z^{-k} + e_k^+ z^k), \tag{3.6b}$$

so that

$$[e^+(z) + e^-(1/z)]\, q(z) = e^T(z) + e^R(z), \tag{3.7}$$

and, by (3.5),

$$\tilde{r}(z)\, q(z) = f(z)\, q(z) - e^T(z) - e^R(z). \tag{3.8}$$

Then, if we define

$$p_1(z) := f(z)\, q(z) - e^T(z), \qquad r_1^{\text{cf}}(z) := \frac{p_1(z)}{q(z)}, \tag{3.9}$$

it follows in view of

$$m > \bar{m} := m - n + \max\{\nu^+ - \delta^+, \nu^- - \delta^-\}$$

that

$$p_1 \in \mathscr{P}_m, \qquad r_1^{\mathrm{cf}} \in \mathscr{R}_{m\bar{n}} \subset \mathscr{R}_{mn}, \tag{3.10}$$

where

$$\bar{n} := \max\{\nu^+ - \gamma^+, \nu^- - \gamma^-\}.$$

We call r_1^{cf} the *type* 1 or *Maehly type Laurent*–CF *approximation* (in analogy to Gragg's terminology in Padé approximation [22, 23]). Again there is another promising way to project $\tilde{r}$ into $\mathscr{R}_{mn}$: We choose $p_2 \in \mathscr{P}_m$ such that the *type* 2 or *Gragg-type Laurent*–CF *approximation* $r_2^{\mathrm{cf}} := p_2/q$ coincides with $\tilde{r}$ in as many low-order Laurent coefficients as possible (i.e., in the coefficients with indices $-\tilde{m}, -\tilde{m}+1, \ldots, 0, \ldots, \tilde{m}-1, \tilde{m}$, where $\tilde{m}$ is as large as possible). If $\tilde{r}$ is real-valued on ∂D, it can be seen that at least $2m+1$ coefficients can be matched (i.e., $\tilde{m} \geqslant m$) since the linear system to solve is Hermitian and positive definite, cf. [48]. The type 2 Laurent–CF approximation can be thought of as Laurent–Padé-type approximation of $\tilde{r}$ with fixed denominator q. One could go one step further by permitting the denominator to be free also and define the *type* 3 *Laurent*–CF *approximation* r_3^{cf} as the Laurent–Padé approximation of $\tilde{r}$. However, r_3^{cf} might have a pole on ∂D, in contrast to r_1^{cf} and r_2^{cf}. Experiments indicate that type 2 is better than type 1 and type 3, cf. [48]. Finally, *type* 4 is again defined as the best L_2 approximation with fixed denominator q of $\tilde{r}$.

If we assume $f \in \mathscr{F}'_M$ (i.e., real-valued) in the above definitions, then $r_j^{\mathrm{cf}} \in \mathscr{R}'_{m\bar{n}}$ $(j = 1, \ldots, 4)$ is called type j *Fourier*–CF *approximation*. Likewise, $f \in \mathscr{F}''_M$, $g \in \mathscr{F}''_L$ implies that the *Chebyshev*–CF *approximation* r_j^{cf} is in $\mathscr{R}''_{m\bar{n}}$, i.e., r_j^{cf} can be thought of as a real rational function (of $x = (z + 1/z)/2$) regular on I. Unweighted Chebyshev–CF approximations are also treated in [26] (polynomial) and [48] (rational). The CF approximation on the disk (Section 2) might be called *Taylor*–CF *approximation*, cf. the summary in Table 1. Laurent–CF approximation could be generalized by allowing different values $m^\pm$ and $n^\pm$ in the two Problems A that are solved. However, we do not want to proceed in this direction.

One is also tempted to modify the definition of r_j^{cf} slightly in the case where $\bar{m} < m$, i.e., when the numerators of both $\tilde{r}^+$ and $\tilde{r}^-$ have lower degree than m. Replacing m by $\bar{m}$ in the definitions (3.6a) and (3.6b) would lead to $p_1 \in \mathscr{P}_{\bar{m}}$ and to a *modified Laurent*–CF *approximation* $\hat{r}_1^{\mathrm{cf}} \in \mathscr{R}_{\bar{m}\bar{n}}$. So, the numerator degree of r_1^{cf} could be reduced (and thus evaluation of r_1^{cf} be made cheaper) by deleting the terms $e_{-m}^\pm z^{\pm m}, \ldots, e_{-\bar{m}-1}^\pm z^{\pm(\bar{m}+1)}$ in e^T, terms that

TABLE I

Taylor–, Laurent–, Fourier–, and Chebyshev–CF Approximation

	$f \in$	$g \in$	$\tilde{r} \in$	$r^{\mathrm{cf}} \in$
(Taylor–) CF	$\tilde{\mathscr{R}}_{M0}$	$\mathscr{P}_L$, zeros in D	$\tilde{\mathscr{R}}_{mn}$	$\mathscr{R}^0_{mn}$
Laurent–CF	$\mathscr{E}_{\max\{M^+,M^-\}}$	$\mathscr{E}'_L$, >0 on ∂D	$L_\infty(\partial D)$	$\mathscr{E}_{mn}$
Fourier–CF	$\mathscr{T}'_M$	$\mathscr{T}'_L$, >0 on ∂D	$L'_\infty(\partial D)$	$\mathscr{T}'_{mn}$
Chebyshev–CF	$\mathscr{T}''_M$	$\mathscr{T}''_L$, >0 on ∂D	$L''_\infty(\partial D)$	$\mathscr{T}''_{mn}$

are typically very small anyway if f and g are smooth enough so that the CF method works well. Though this modified definition may be suitable in practice, we reject it here since the asymptotic results of [48] would no longer hold. $\hat{r}_2^{\mathrm{cf}}$ could be defined likewise as the $(\bar{m}, \bar{n})$ Laurent–Padé-type approximation of $\tilde{r}$ with fixed denominator q.

In general, it is difficult to find out—both a priori and a posteriori—whether a complex-valued Laurent–CF approximant is close to the best approximation, or even whether its error $\|\sqrt{g}\,(f - r_j^{\mathrm{cf}})\|$ is close to the error $E_{mn}(f;\sqrt{g})$ of the best approximation. Although there is a general inclusion theorem by Meinardus and Schwedt [36, Theorem 85], its applicability in practice is very limited. The situation is different if f is real-valued and approximants are restricted to $\mathscr{T}'_{mn}$ or $\mathscr{T}''_{mn}$, because both for trigonometric rational and for ordinary rational approximation a de la Vallée–Poussin-type inclusion theorem holds [35, 36, Theorem 98]. For CF approximations we can even state a bound that holds globally for all approximations whose "truncation error" $\|\sqrt{g}\,(r_j^{\mathrm{cf}} - \tilde{r})\|$ is known to be small. Note that for $j = 1$

$$\|\sqrt{g}\,(r_1^{\mathrm{cf}} - \tilde{r})\| = \|g^+ e^R/q\|, \tag{3.11}$$

so a bound for this error may be obtained from estimating the factors on the right-hand side. For the Chebyshev–CF approximation of real functions on I with fast converging Chebyshev series this was accomplished in [26] (polynomial) and [48] (rational). Here we state the underlying basic theorem in more general form. Note that we drop the superscript $^+$ whenever there is no chance of confusion. (For example, we replace M^+, ν^+ by M, ν, but we refrain from replacing $f^+, \tilde{r}^+$.) In particular, $\bar{n} = \nu - \gamma = n'$ and $\bar{m} = m - n + \nu - \delta = m'$ now, cf. (2.7). Let $E'_{mn} := E'_{mn}(\tilde{f}; \sqrt{g})$ and $E''_{mn} := E''_{mn}(\tilde{f};\sqrt{g})$ denote the errors of the best approximations of $\tilde{f}$ out of $\mathscr{T}'_{mn}$ and $\mathscr{T}''_{mn}$, respectively, with respect to the weight function $\sqrt{g}$.

THEOREM 3.1. *Let $\tilde{f}\colon \partial D \to \mathbb{R}$, $g \in \mathscr{T}'_L \backslash \mathscr{T}'_{L-1}$ $(L \geqslant 0)$ with $g(z) > 0$ on ∂D, and $m, n \geqslant 0$ be given. Assume $\tilde{f}$ is approximated by $f \in \mathscr{T}'_M \backslash \mathscr{T}'_{M-1}$ such that $\|\sqrt{g}\,(\tilde{f} - f)\| < \varepsilon_1$. Assume the Fourier–*CF *extension $\tilde{r}$ of f and the Fourier–*CF *approximation $r_j^{\mathrm{cf}} \in \mathscr{T}'_{mn'}$ $(j = 1, 2, 3$ or 4, $n' = \nu - \gamma)$ satisfy*

$\|\sqrt{g}(\tilde{r}-r_j^{\text{cf}})\| < \varepsilon_2$, *and* $\varepsilon := \varepsilon_1 + \varepsilon_2 < 2\sigma_n$ *(where* σ_n *is the singular value appearing in the solution of Problem* A). *Then*

$$\big|\, \|\sqrt{g}(\tilde{f}-r_j^{\text{cf}})\| - 2\sigma_n \big| \leqslant \varepsilon, \tag{3.12}$$

$$|E'_{m+\lambda,n'+\lambda} - 2\sigma_n| \leqslant \varepsilon, \tag{3.13}$$

for $0 \leqslant \lambda \leqslant l' := \gamma + \nu + \mu - n$, *and hence*

$$\|\sqrt{g}(\tilde{f}-r_j^{\text{cf}})\| \leqslant E'_{m+\lambda,n'+\lambda} + 2\varepsilon, \qquad 0 \leqslant \lambda \leqslant l'. \tag{3.14}$$

If $\tilde{f}(\bar{z}) = \overline{\tilde{f}(z)}$ $(\forall z \in \partial D)$, $f \in \mathscr{F}''_M$, *and* $g \in \mathscr{F}''_L$, (3.13) *and* (3.14) *hold a fortiori for* $E''_{m+\lambda,n'+\lambda}$.

Proof. According to Theorem 2.1, $g^+(f^+ - \tilde{r}^+)(\partial D) = g^- e^+(\partial D)$ is a circle of radius σ_n around 0 with winding number $L + \iota^+$, where $\iota^+ := \iota(f^+ - r^+) = m - n + \mu + 2\nu + 1$ is the winding number of the unweighted error curve $e^+(\partial D)$. Consequently, there exists a set $A := \{z_k;\, k = 1,\ldots, 2\iota^+\}$ of $2\iota^+$ ordered points $z_k \in \partial D$ such that $e^+(z_K)$ is real and alternately positive and negative. By (3.3) and $e^-(1/z) = \overline{e^+(z)}$,

$$f(z_k) - \tilde{r}(z_k) = 2 \operatorname{Re} e^+(z_K) = 2e^+(z_K), \tag{3.15}$$

so $f - \tilde{r}$ alternates in sign on A. Moreover, since $g(z) = |g^+(z)|^2$ on ∂D,

$$\sqrt{g(z_k)}\, |f(z_k) - \tilde{r}(z_k)| = 2\, |g^+(z_K)\, e^+(z_k)| = 2\sigma_n. \tag{3.16}$$

Of course, $\|\sqrt{g}(f-\tilde{r})\| \leqslant 2\, \|g^+ e^+\| = 2\sigma_n$, hence

$$\|\sqrt{g}(f-\tilde{r})\| = 2\, \|g^+ e^-\| = 2\sigma_n, \tag{3.17}$$

and by (3.15)–(3.17) $\sqrt{g}(f-\tilde{r})$ alternates on A. By assumption,

$$\|\sqrt{g}(\tilde{f}-r_j^{\text{cf}}) - \sqrt{g}(f-\tilde{r})\| \leqslant \varepsilon_1 + \varepsilon_2 = \varepsilon < 2\sigma_n,$$

so (3.17) implies (3.12), and $\sqrt{g}(\tilde{f}-r_j^{\text{cf}})$ alternates in sign on A and deviates there at most by ε from $\pm 2\sigma_n$. Finally, since $2\iota^+ = 2(m + n' + l' + 1)$, the set $\{(z_k, (-1)^k);\ k = 1,\ldots, 2\iota^+\}$ is an H-set (or extremal signature) for $r_j^{\text{cf}} \in \mathscr{R}'_{mn'}$ with respect to $\mathscr{R}'_{m+\lambda,n'+\lambda}$, $0 \leqslant \lambda \leqslant l'$, cf. [35]. From the general inclusion theorem [36, Theorem 85] one can therefore derive $E'_{m+\lambda,n'+\lambda} \geqslant 2\sigma_n - \varepsilon$. On the other hand, $E'_{m+\lambda,n'+\lambda} \leqslant 2\sigma_n + \varepsilon$ by (3.12). ∎

We can interpret (3.14) in terms of the CF table: Consider one of the square blocks of the table. Every pair (m, n) of a column leads to the same r_j^{cf} (if $j = 1$, 2 or 4) since $\tilde{r}$ is the same for the whole block and truncation depends only on m. The numerator degree is typically m (but may be smaller), and the denominator degree is exactly n', which is equal to the value of n in the first row of the block. Inequality (3.14) implies that r_j^{cf} is nearly optimal with respect to spaces $\mathscr{E}'_{mn}$ whose subscripts (m, n) belong to a subsquare whose upper- and right-side edges are part of the boundary of the block while the left-side edge is part of the fixed column. Unfortunately, the pair (m, n) we start with when computing $\tilde{r}$ need not lie in this induced subsquare: $n > n' + l'$ is possible. In other words: Given m and n, Theorem 3.1 does not yet ensure that the computed r_j^{cf} is nearly optimal in $\mathscr{E}'_{mn}$ (even if the assumptions on the truncation errors hold). However, the subsquare coincides with the whole block if we choose $m = m'$ (corresponding to the first column of the block). Of course, we cannot attain that if $m' < 0$.

The situation is different if we consider the *modified* Fourier–CF approximation $\hat{r}_j^{\mathrm{cf}} \in \mathscr{E}'_{m'n'}$ (the conjugate symmetric case of the modified Laurent–CF approximation). The only modification we need in the proof of Theorem 3.1 is that $\iota^+ = m' + n' + l + 1$ iff $l := \gamma + \delta + \mu$ (as in Theorem 2.1).

COROLLARY 3.2. *For the modified Fourier*–CF *approximation* $\hat{r}_j^{\mathrm{cf}} \in \mathscr{E}'_{m'n'} (j = 1$ *or* $2)$ *Theorem* 3.1 *holds with* $E'_{m'+\lambda, n'+\lambda}$, $0 \leqslant \lambda \leqslant l := \gamma + \delta + \mu$ *in* (3.13) *and* (3.14).

Hence, $\hat{r}_j^{\mathrm{cf}}$ is near-best with respect to whole block in the CF table. This discrepancy between r_j^{cf} and $\hat{r}_j^{\mathrm{cf}}$ has little effect in practice, however: Typically $\|r_j^{\mathrm{cf}} - \hat{r}_j^{\mathrm{cf}}\|$ is very small, so that r_j^{cf} is also near-best with respect to the whole block:

COROLLARY 3.3. *Assume* $\|\sqrt{g}(\tilde{f} - f)\| \leqslant \varepsilon_1$, $\|\sqrt{g}(\tilde{r} - \hat{r}_j^{\mathrm{cf}})\| \leqslant \varepsilon_2$, $\|\sqrt{g}(r_j^{\mathrm{cf}} - \hat{r}_j^{\mathrm{cf}})\| < \varepsilon_3$ $(j = 1$ *or* $2)$, *and* $\varepsilon := \varepsilon_1 + \varepsilon_2 + \varepsilon_3 < 2\sigma_n$. *Then*

$$\|\sqrt{g}(\tilde{f} - r_j^{\mathrm{cf}})\| \leqslant E'_{m'+\lambda, n'+\lambda} + 2\varepsilon, \qquad 0 \leqslant \lambda \leqslant l := \gamma + \delta + \mu. \quad (3.18)$$

Of course, in the case of symmetry with respect to the real axis we may also replace E' by E'' in both corollaries.

ACKNOWLEDGMENTS

Much of this work was developed jointly with L. N. Trefethen during the author's valuable and pleasurable stays at the Computer Science Department of Stanford University in 1979/1980 (supported by the Swiss NSF fellowship 82.723.0.79) and in September 1981. The author is also indebted to Professors A. Bultheel, D. Elliott, and G. H. Golub, and Dr. P. van Dooren for pointing out useful references.

References

1. N. I. Achieser, Sur les propriétés asymptotiques de quelques polynomes, *C. R. Acad. Sci. Paris* **191** (1930), 908–910.
2. N. I. Achieser, On the extremal properties of certain fractional functions (Russian), *C. R. Acad. Sci. URSS* **1930** (1930), 495–499.
3. N. I. Achieser, On a minimum problem in the theory of functions and on the number of roots of an algebraic equation which are located in the interior of the unit circle (Russian), *Bull. Acad. Sci. Leningrad* **1930** (1931), 1169–1189.
4. N. I. Achieser, Über ein Tschebyscheffsches Extremumproblem, *Math. Ann.* **104** (1931), 739–744.
5. N. I. Achieser, "Theory of Approximation," Ungar, New York, 1965.
6. V. M. Adamjan, D. Z. Arov, and M. G. Krein, Analytic properties of Schmidt pairs for a Hankel operator and the generalized Schur–Takagi problem, *Math. USSR-Sb.* **15** (1971), 31–73.
7. F. L. Bauer, Ein direktes Iterationsverfahren zur Hurwitz-Zerlegung eines Polynoms, *Arch. Elektrotech. Uebertragung* **9** (1955), 285–290.
8. S. N. Bernstein, "Leçons sur les propriétés extrémales et la meilleure approximation des fonctions analytiques d'une variable réelle," Gauthier–Villars, Paris, 1926.
9. C. Brezinski, "Padé-Type Approximation and General Orthogonal Polynomials," Birkhäuser, Basel, 1980.
10. A. Bultheel and P. Dewilde, "On the Adamjan–Arov–Krein Approximation, Identification and Balanced Realization of a System," Report TW 48, Kath. Univ. Leuven, 1980, and ECCTD Conference, Warsaw, 1980.
11. C. Carathéodory and L. Fejér, Über den Zusammenhang der Extremen von harmonischen Funktionen mit ihren Koeffizienten und über den Picard–Landauschen Satz, *Rend. Circ. Mat. Palermo* **32** (1911), 218–239.
12. D. N. Clark, On the spectra of bounded Hermitian Hankel matrices, *Amer. J. Math.* **90** (1968) 627–656.
13. D. N. Clark, Hankel forms, Toeplitz forms and meromorphic functions, *Trans. Amer. Math. Soc.* **134** (1968), 109–116.
14. C. W. Clenshaw, A comparison of "best" polynomial approximations with truncated Chebyshev series expansions, *J. SIAM Numer. Anal. Ser. B* **1** (1964), 26–37.
15. S. Darlington, Analytical approximations to approximations in the Chebyshev sense, *Bell System Tech. J.* **49** (1970), 1–32.
16. S. W. Ellacott, On the Faber transform and efficient numerical rational approximation, *SIAM J. Numer. Anal.* **20** (1983), 989–1000.
17. S. W. Ellacott and M. H. Gutknecht, The polynomial Carathéodory–Fejér approximation method for Jordan regions, *IMA J. Numer. Anal.* **3** (1983), 207–220.
18. D. Elliott, A direct method for "almost" best uniform approximations, *in* "Error, Approximation, and Accuracy" (F. R. de Hoog and C. L. Jarvis, Eds.), pp. 129–143, Univ. of Queensland Press, St. Lucia, Queensland, 1973.
19. G. H. Elliott, "The Construction of Chebyshev Approximations in the Plane," Ph. D. thesis, University of London, 1978.
20. F. R. Gantmacher, "Theory of Matrices," Chelsea, New York, 1959.
21. Y. V. Genin and S. Kung, A two-variable approach to the model reduction problem with Hankel norm criterion, *IEEE Trans. Circuits and Systems* **CAS-28** (1981), 912–924.
22. W. B. Gragg, Laurent, Fourier, and Chebyshev–Padé tables, *in* "Padé and Rational Approximation" (E. B. Saff and R. S. Varga, Eds.), pp. 61–72, Academic Press, New York, 1977.

23. W. B. Gragg and G. D. Johnson, The Laurent–Padé table, *in* "Information Processing 74," pp. 632–637, North-Holland, Amsterdam, 1974.
24. M. H. Gutknecht, J. O. Smith, and L. N. Trefethen, The Carathéodory–Fejér method for recursive digital filter design, *IEEE Trans. Acoust. Speech Signal Process.* **ASSP–31** (1983), 1417–1426.
25. M. H. Gutknecht and L. N. Trefethen, Recursive digital filter design by the Carathéodory–Fejér method, Numer. Anal. Proj. Manuscript NA-80-01, 1980.
26. M. H. Gutknecht and L. N. Trefethen, Real polynomial Chebyshev approximation by the Carathéodory–Fejér method, *SIAM J. Numer. Anal.* **19** (1982), 358–371.
27. M. Hollenhorst, "Nichtlineare Verfahren bei der Polynomapproximation," Diss. Universität Erlangen-Nürnberg, 1976.
28. S. Kung, Optimal Hankel-norm model reduction—Scalar systems, *in* "Proceedings, 1980 Joint Automatic Control Conference, San Francisco," Inst. Elec. Electronics Eng., New York, 1980.
29. B. Lam, "Some Exact and Asymptotic Results for Best Uniform Approximation," Ph. D. thesis, University of Tasmania, Australia, 1972.
30. B. Lam, A note on best uniform rational approximation, *SIAM J. Numer. Anal.* **13** (1976), 962–965.
31. B. Lam, Some exact and asymptotic results for best uniform approximations, unpublished manuscript.
32. B. Lam and D. Elliott, On a conjecture of C. W. Clenshaw, *SIAM J. Numer. Anal.* **9** (1972), 44–52.
33. B. Lam and D. Elliott, An estimate for $E_n(f)$ for large n, *SIAM J. Numer. Anal.* **10** (1973), 1091–1102.
34. E. Landau, Ueber einen Bieberbachschen Satz, *Rend. Circ. Mat. Palermo* **46** (1922), 456–462.
35. H. L. Loeb, Approximation by generalized rationals, *SIAM J. Numer. Anal.* **3** (1966), 34–55.
36. G. Meinardus, "Approximation of Functions: Theory and Numerical Methods," Springer-Verlag, Berlin/New York, 1967.
37. J. Rissanen, Algorithms for triangular decomposition of block Hankel and Toeplitz matrices with application to factoring positive matrices, *Math. Comp.* **27** (1973), 147–154.
38. W. Rudin, "Real and Complex Analysis," 2nd ed., McGraw–Hill, New York, 1974.
39. I. Schur, Ueber Potenzreihen, die im Innern des Einheitskreises beschränkt sind, *Crelle's J.* (*J. Reine Angew. Math.*) **147** (1917), 205–232; **148** (1918), 122–145.
40. L. M. Silverman and M. Bettayeb, Optimal approximation of linear systems, *in* "Proceedings, 1980 Joint Automatic Control Conference, San Francisco," Inst. Elec. Electronics Eng., New York, 1980.
41. T. Takagi, On an algebraic problem related to an analytic theorem of Carathéodory and Fejér and on an allied theorem of Landau, *Japan J. Math.* **1** (1924), 83–93.
42. T. Takagi, Remarks on an algebraic problem, *Japan J. Math.* **2** (1925), 13–17.
43. A. Talbot, On a class of Tchebysheffian approximation problems solvable algebraically, *Proc. Cambridge Philos. Soc.* **58** (1962), 244–267.
44. A. Talbot, The uniform approximation of polynomials by polynomials of lower degree, *J. Approx. Theory* **17** (1976), 254–279.
45. R. C. Thompson, Principal submatrices IX. Interlacing inequalities for singular values of submatrices, *Linear Algebra Appl.* **5** (1972), 1–12.
46. L. N. Trefethen, Near-circularity of the error curve in complex Chebyshev approximation, *J. Approx. Theory* **31** (1981), 344–367.
47. L. N. Trefethen, Rational Chebyshev approximation on the unit disk, *Numer. Math.* **37** (1981), 297–320.

48. L. N. Trefethen and M. H. Gutknecht, The Carathéodory–Fejér method for real rational approximation, *SIAM J. Numer. Anal.* **20** (1983), 420–436.
49. Z. Vostrý, New algorithms for polynomial spectral factorization with quadratic convergence, *Kybernetika* **11** (1975), 415–422.
50. G. Wilson, Factorization of the covariance generating function of a pure moving average process, *SIAM J. Numer. Anal.* **6** (1969) 1–7.