SQUARE BLOCKS AND EQUIOSCILLATION IN THE PADÉ, WALSH, AND CF TABLES

Lloyd N. Trefethen*
Courant Institute of Mathematical Sciences
New York University
New York, NY 10012

Abstract. It is well known that degeneracies in the form of repeated entries always occupy square blocks in the Padé table, and likewise in the Walsh table of real rational Chebyshev approximants on an interval. The same is true in complex CF (Carathéodory-Fejér) approximation on a circle. We show that these block structure results have a common origin in the existence of equioscillation-type characterization theorems for each of these three approximation problems. Consideration of position within a block is then shown to be a fruitful guide to various questions whose answers are affected by degeneracy.

0. Introduction

Consider the following three problems in rational approximation. In each case m and n are nonnegative integers, except that m may be negative in the CF case.

CHEBYSHEV ("T"). Let f be real and continuous on $I = [-1,1]$, and let R^r_{mn} be the set of rational functions of type (m,n) with real coefficients. Problem: find $r^* \in R^r_{mn}$ such that

(1.T) $$\|f-r^*\|_I \le \|f-r\|_I \quad \forall r \in R^r_{mn},$$

where $\|\cdot\|_I$ is the supremum norm on I.

PADÉ ("P"). Let f be a complex formal power series in z, and let R_{mn} be the set of rational functions of type (m,n) with complex coefficients. Problem: find $r^P \in R_{mn}$ such that

(1.P) $$(f-r^P)(z) = O((f-r)(z)) \quad \text{as} \quad z \to 0 \quad \forall r \in R_{mn}.$$

*Supported by an NSF Postdoctoral Fellowship and by the U.S. Dept. of Energy under contract DE-AC02-76-ER03077-V.

Reprinted from Rational Approximation and Interpolation, edited by P.R. Graves-Morris, E.B. Saff, and R.S. Varga, Springer-Verlag Lecture Notes in Mathematics, v. 1105, 1984.

CF ("K"). Let f be a continuous function on $S = \{z: |z|=1\}$ whose Fourier series converges absolutely. Let $\tilde{R}_{mn}$ be the set of "extended rational" functions representable in the form

$$\tilde{r}(z) = \frac{\tilde{p}(z)}{q(z)} = \sum_{k=-\infty}^{m} a_k z^k \Big/ \sum_{k=0}^{n} b_k z^k , \tag{2}$$

where q has all of its zeros in $|z| > 1$, and the series for $\tilde{p}$ converges there and is bounded except possibly near $z = \infty$. Problem: find $\tilde{r}^* \in \tilde{R}_{mn}$ such that

$$\|f-\tilde{r}^*\|_S \le \|f-\tilde{r}\|_S \quad \forall \tilde{r} \in \tilde{R}_{mn} , \tag{1.K}$$

where $\|\cdot\|_S$ is the L^∞ norm on S.

See [12] for information on Chebyshev approximation, [1,6] for Padé, and [7,14,15] for CF. (The CF approximant defined above is actually the "extended CF approximant"; in practice it would be projected onto a function $r^{cf} \in R_{mn}$ to yield a near-best Chebyshev approximant on a disk.) All three problems have unique solutions, and these can be constructed numerically: r^p by solving a finite Hankel system of linear equations, r^* by a procedure such as the Remes algorithm, and $\tilde{r}^*$ via a singular value decomposition of an infinite Hankel matrix of Laurent series coefficients. We will not go into this.

The Padé table is the array obtained by arranging the various approximants r^p for a given f in sequence in the lower-right quadrant of the plane, with m as the column index and n as the row index. (Sometimes these indices are reversed [6].) The corresponding array of Chebyshev approximants r^* is called the Walsh table. The (extended) CF table is the analogous array of CF approximants $\tilde{r}^*$, except that since $m < 0$ is permitted, it fills the entire lower half plane instead of a quadrant.

The first purpose of this paper is to publicize the connection between equioscillation theorems and block structure in approximation tables (Secs. 1,2). We will show by examples that the equioscillation/square blocks point of view is of mnemonic and heuristic value in investigating problems whose answers are affected by degeneracy (Secs. 2,3).

The second purpose is to elucidate the close analogy between the Chebyshev, Padé, and CF problems. Many aspects of these problems depend only on the superficial structure imposed by an equioscillation theorem, not on the details of the type of approximation. Pursuing this analogy gives insight into how things are different in related problems where there is no equioscillation principle (Sec. 4). It also raises an interesting question of whether the Walsh, Padé, and CF tables all admit the same set of possible block configurations (Sec. 5).

1. Equioscillation characterizations

The word "equioscillation" comes from the Chebyshev problem, where it describes an error curve $(f-r^*)(I)$ that oscillates sufficiently many times between positive and negative extrema with equal magnitude. In CF approximation, the analogous object is an error curve $(f-\tilde{r}^*)(S)$ that is a perfect circle of sufficiently large winding number. In Padé approximation, it is an error function $f-r^P$ that is zero to sufficiently high order at the origin. One can think of this as a circular error curve condition too, for as $\varepsilon \to 0$, $(f-r^P)(\varepsilon S)$ approaches a circle with winding number equal to the degree of the first nonzero coefficient in $f-r^P$.

The question of how great an equioscillation number is "sufficient" depends on the <u>defect</u> δ. Given r ($\tilde{r}$), let it be expressed as a quotient p/q ($\tilde{p}/q$) in lowest terms, i.e. in which the numerator and denominator have no common zeros. Let $\mu \le m$ and $\nu \le n$ be the exact degrees of p ($\tilde{p}$) and q, with $\mu = -\infty$ if $p \equiv 0$ ($\tilde{p} \equiv 0$). Then δ is defined by

(3) $$\delta = \min\{m-\mu, n-\nu\}.$$

<u>THEOREM 1T</u>. *If $r \in R^r_{mn}$ has defect δ, then $r = r^*(f)$ if and only if the error curve $(f-r)(I)$ oscillates between $\pm \|f-r\|_I$ on some sequence of points $-1 \le x_0 < \dots < x_N \le 1$ with $N \ge m+n+1-\delta$.*

<u>THEOREM 1P</u>. *If $r \in R_{mn}$ has defect δ, then $r = r^P(f)$ if and only if*

(4) $$(f-r)(z) = O(z^N) \quad \text{as } z \to 0 \text{ with } N \ge m+n+1-\delta.$$

<u>THEOREM 1K</u>. *If $\tilde{r} \in \tilde{R}_{mn}$ has defect δ, then $\tilde{r} = \tilde{r}^*(f)$ if and only if $\tilde{r}$ is continuous in $|z| \ge 1$ and the error curve $(f-\tilde{r})(S)$ is a circle of winding number $N \ge m+n+1-\delta$ in the positive sense.*

<u>Remark</u>. In each case we will assume N is chosen as large as possible (possibly ∞), and refer to it as "the equioscillation number".

<u>Proofs</u>. These assertions have two halves, namely "equioscillation implies best" and "best implies equioscillation". The result in the first direction is easily obtained by counting zeros. For example if $r \in R_{mn}$ satisfies (4), then by (1.P) one has also $f-r^P = O(z^N)$.

Therefore $r-r^P$ has at least $N \geq m+n+1-\delta$ zeros at the origin. But since $r-r^P \in R_{m+n-\delta,2n-\delta}$, this implies $r = r^P$. (This argument also shows that r^P is unique.) Analogous proofs work for Chebyshev and CF.

Showing "best implies equioscillation" is less trivial, but the threefold analogy can be maintained by arguing in the following way. Suppose r ($\tilde{r}$) does <u>not</u> equioscillate sufficiently many times. Then it can be perturbed slightly to a new function $r' = r+\Delta r$ ($\tilde{r}'$) which is a better approximant. The method of constructing this perturbation depends on which approximation problem is being considered. See [12] for the Chebyshev case (quite straightforward) and [8] for CF (trickier). For Padé approximation one could also write down a perturbation argument, but it is unnecessary since the problem of determining coefficients is actually linear. Therefore it is as well to obtain Thm. 1P as a corollary of the usual Padé table derivation via linear algebra. See [1] or [6].

2. Square blocks

The following arguments run the same way for Chebyshev, Padé, or CF; we consider the Padé case for definiteness. Given f, suppose a function r of exact type (μ,ν) with $\mu > -\infty$ happens to satisfy

$$(5) \qquad (f-r)(z) = O(z^N), \neq O(z^{N+1}) \quad \text{as} \quad z \to 0 \quad \text{with} \quad N = \mu+\nu+1+\Delta$$

for some $\Delta \geq 0$ (possibly ∞, in which case the $\neq O(z^{N+1})$ clause is dropped). For which (m,n), if any, is r the Padé approximant r^P ? The answer is, for precisely those (m,n) that lie in any of the positions of the following $(\Delta+1) \times (\Delta+1)$ block:

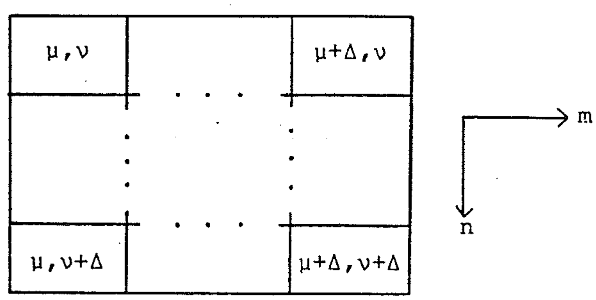


Figure 1

To verify this claim, combine (4) and (5) to obtain the following condition: $r = r^P$ if and only if $m \geq \mu$, $n \geq \nu$, and

$$\mu+\nu+\Delta \geq m+n-\delta ,$$

or by (3),

$$\mu+\nu+\Delta \geq m+n-\min\{m-\mu, n-\nu\} .$$

If $m-\mu \leq n-\nu$, this becomes $\nu+\Delta \geq n$, which gives the lower-left half of the square block of Fig. 1. The alternative $n-\nu \leq m-\mu$ leads to $\mu+\Delta \geq m$, which gives the upper-right half of the block.

This argument carries over directly to Chebyshev and CF approximation. We can summarize the situation for all three problems as follows: if $f-r$ equioscillates the "normal" number of times $N = \mu+\nu+1$, then r is the desired approximant in the (μ,ν) position but nowhere else. With each "extra" oscillation, the size of the block in which r is the approximant increases by 1.

In (5) we have excluded the possibility $\mu = -\infty$, i.e. $r \equiv 0$ ($\tilde{r} \equiv 0$). Here (3) gives $\delta = n$, and so 0 is the desired approximant if and only if f itself equioscillates with $N \geq m+1$. That is, the zero function fills all columns of the table with $m \leq N-1$, if any. This is the only situation in which non-square blocks occur.

Here is a general block structure statement.

THEOREM 2. *The Walsh, Padé, and CF tables all break down into precisely square blocks containing identical entries. (One of these may be infinite in extent, if* f *can be approximated exactly for large enough* (m,n)*.) The only exception is that if an entry* $r \equiv 0$ ($\tilde{r} \equiv 0$) *appears in the table, then it fills all of the columns to the left of some fixed index* $m = N$.

To make these conclusions vivid, Fig. 2 shows how m, n, δ, and N are distributed within a square block.

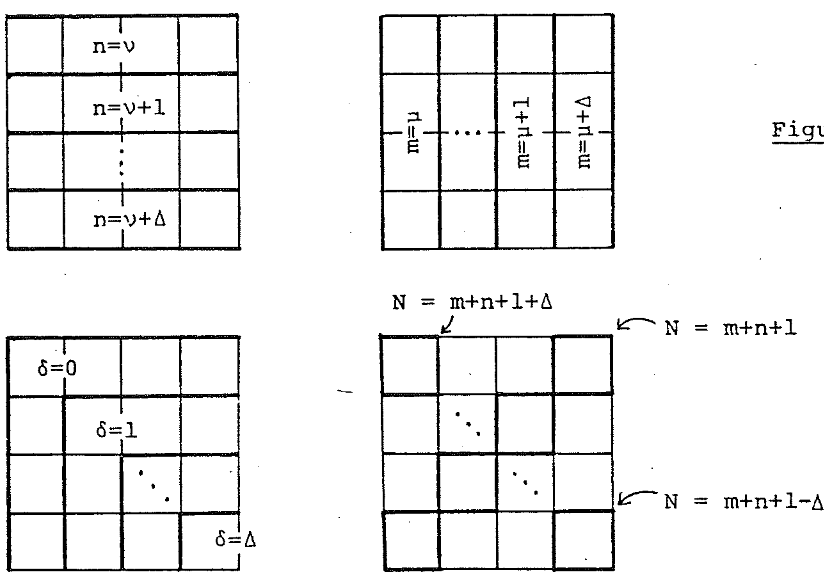


Figure 2

Here is a summary of the notation we have introduced:

f - function to be approximated

m,n - nonnegative integers (m may be negative in CF case)

$R_{mn}^{r}, R_{mn}, \tilde{R}_{mn}$ - spaces of real, complex, extended rational functions

$r^*, r^p, \tilde{r}^*$ - Chebyshev, Padé, CF approximants of type (m,n)

$\mu \le m, \nu \le n$ - exact degrees of approximant

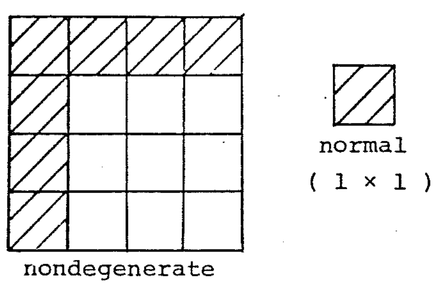
$\delta = \min\{m-\mu, n-\nu\}$ - defect

$\Delta+1$ - dimension of square block

$N = \mu+\nu+1+\Delta$ - equioscillation number

3. Continuity of the Chebyshev, Padé, and CF operators

Complications often arise when one deals with degenerate approximants. For this reason we say that a Walsh, Padé, or CF table is <u>normal</u> if no entry appears twice, that is, if every block has size 1×1 . The word "normal" has also been applied to individual entries in a table, but unfortunately its uses in the Chebyshev and Padé literature have been inconsistent. Perhaps the following problem-independent definitions make the most sense: an approximant is <u>nondegenerate</u> if $m = \mu$ or $n = \nu$ (i.e. if $\delta = 0$), and <u>normal</u> if $m = \mu$, $n = \nu$, and $\Delta = 0$. (This use of "normal" follows the Padé convention [6]; in Chebyshev approximation "normal" has meant what we call "nondegenerate" [12,17].) Figure 3 illustrates the two definitions.



<u>Figure 3</u>

Various approximation results can be stated in terms of hypotheses on position within the square block. For example: (1) Walsh showed in 1974 that if r^p is nondegenerate, then $r_{\varepsilon I}^* \to r^p$ as $\varepsilon \to 0$, where $r_{\varepsilon I}^*$ is the Chebyshev approximant on $[-\varepsilon,\varepsilon]$; on the other hand if r^p is degenerate, this need not hold [16]. (2) The analogous result for CF-approximation appears to be that $r_{\varepsilon S}^* \to r^p$ as $\varepsilon \to 0$ is guaranteed if r^p lies in the upper-right or lower-left corner of its block (Gutknecht and Trefethen, forthcoming). (3) A theorem of Ruttan states that if the Chebyshev approximant r^* of a real function f lies in the strict lower-right subtriangle of its square block, then f can be better approximated in R_{mn} than R^r_{mn} [13]. (4) In the corresponding strict lower-right subtriangle of a square block in the Padé table, the so-called Padé equations are inconsistent, and according to the "Baker definition" of the Padé approximant (different from ours), r^p does not exist [1].

We will now describe a particularly appealing application of block structure arguments. Let the Chebyshev, Padé, and CF approximation operators be defined by

$$T: f \mapsto r^* , \qquad P: f \mapsto r^p , \qquad K: f \mapsto \tilde{r}^* .$$

The question is, when are these operators continuous? We omit details concerning the precise definitions of continuity -- see [8,16].

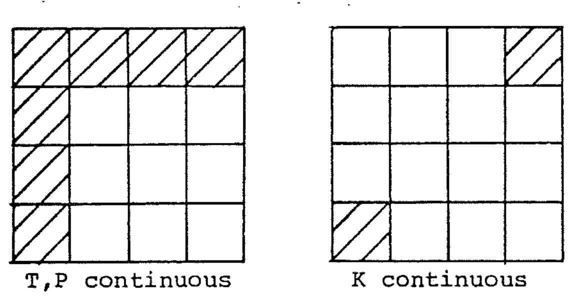
For T and P the answers turn out to be the same, and were obtained by Werner [12,17] and by Werner and Wuytack [16,18], respectively:

<u>THEOREMS 3T, 3P</u>. <u>The operator</u> T (P) <u>is continuous at</u> f <u>if and only if</u> $T(f)$ ($P(f)$) <u>is nondegenerate</u>.

In contrast, the forthcoming paper [8] will establish the following result for K :

<u>THEOREM 3K</u>. <u>The operator</u> K <u>is continuous at</u> f <u>if and only if</u> $K(f)$ <u>is nondegenerate and in addition has equioscillation number exactly</u> $N = m+n+1$.

Figure 4 summarizes these results:



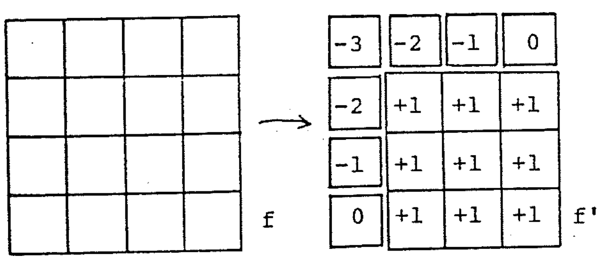
<u>Figure 4</u>

The interesting thing is that despite the apparent discrepancy between Thms. 3T,3P and Thm. 3K, all of these results actually have a single explanation in terms of block structure. The underlying principle is this: small perturbations can break square blocks, but not make them (cf. [10]).

<u>THEOREM 4. Suppose</u> (m_1,n_1) <u>and</u> (m_2,n_2) <u>lie in distinct square blocks of the Walsh, Padé, or CF table for the function</u> f . <u>Then the same is true for all sufficiently small perturbations</u> $f' = f + \Delta f$.

<u>Proof</u>. Without loss of generality we can assume $m_1 \le m_2$ and $n_1 \le n_2$. (Otherwise, an easy argument based on block structure shows we can replace either (m_1,n_1) or (m_2,n_2) by $(\min\{m_1,m_2\},\min\{n_1, n_2\})$.) In the Chebyshev case, the hypotheses imply $\|f-r_2^*\|_I < \|f-r_1^*\|_I$, and by the definition of r^* , the inequality will persist under perturbations with $\|\Delta f\|_I < \frac{1}{2}(\|f-r_1^*\|_I - \|f-r_2^*\|_I)$. The same argument (with $\|\cdot\|_S$) works in the CF case. For Padé approximation an analogous proof can also be constructed, or one can appeal to known results about the linear algebra of the Padé table and use the fact that a small perturbation of a nonsingular matrix is nonsingular.

Theorem 4 now suggests the following idea for the proofs of Thms. 3T,3P,3K, though it is quite disguised in the original papers: assuming that perturbations of f can be constructed to fracture a square block in any desired way, what does this imply about discontinuity? As it happens, the following single fracture pattern is the only one needed, and it turns out that a perturbation can always be found that accomplishes it. Of course, constructing this perturbation requires consideration of problem-dependent details.



<u>Figure 5</u>

From Fig. 2 we know that in any square block, the equioscillation number N is determined by the position of the main cross-diagonal. Therefore under the perturbation of Fig. 5, the value of N in each position must change by exactly the quantity indicated. In particular, it increases in the lower-right subblock, stays constant in two corners, and decreases in the remaining positions on the upper-left border.

Now from Thms. 1, it can be seen that in all three cases T, P, K, an increase in N cannot be obtained through an arbitrarily small perturbation. For example, the winding number of a circular error curve with positive radius ρ cannot increase except by the addition of a function of norm at least ρ . Therefore the construction of Fig. 5 proves discontinuity of T , P , and K whenever the approximant to f is degenerate.

On the other hand, Thms. 1 also imply that N can decrease under a small perturbation in cases T and P , but not in case K . For example, a point x_j in the alternant set for the error curve $(f-r^*)$ (I) may cease to be extremal in response to arbitrarily small perturbations of f -- whereas a circular CF error curve cannot decrease smoothly to another circle of lower winding number. Therefore Fig. 5 proves discontinuity of K , but not of T or P , in the upper-left border positions away from the corners.

These arguments establish discontinuity in all of the unshaded positions of Fig. 4. A little experimentation quickly shows that no alternative fragmentations of the block produce discontinuity in any further positions; in fact the shaded positions of Fig. 4 can be characterized as those locations at which N cannot increase through an arbitrarily small perturbation of f (cases T , P), and those where it cannot change at all (case K). To complete the proofs of Thms. 3, all that remains is to verify continuity in these shaded positions. This is another problem-dependent argument, which we omit.

4. Related problems

The Chebyshev, Padé, and CF problems are unusual in obeying equioscillation principles. Many related problems have no such simple characterization theorems, and as a result, do not possess square block structure.

Within the realm of Chebyshev approximation, the conspicuous context where equioscillation fails is in complex Chebyshev approximation. On either a disk or an interval, complex best approximations have no simple characterization, need not be unique, and need not lie in square blocks [7,16]. A circular error curve of high enough winding number is here still sufficient for best approximation, but no longer necessary.

Both equioscillation and square blocks also vanish in general if one approximates in other norms, such as L^2 .

The equioscillation principle fails in Padé approximation as soon as one generalizes from interpolation just at the origin to any kind of multipoint (or "Newton-Padé") scheme. In fact in problems of this kind the interpolation table itself requires careful definition, since the approximation obtained depends not just on the number of interpolation conditions specified, but also on the order in which they are taken [3, 4].

The analogous multipoint version of CF approximation is the Pick-Nevanlinna problem, with its extension to the rational case due to Achieser. The block structure (not square) in the "PNA table" is currently being investigated by Gutknecht.

All of these remarks show that the ideas of this paper are far from universal in application. However, they are not exhausted, either. On one hand, there are probably other interesting problems besides the three we have mentioned in which one gets equioscillation and square blocks. (This appears to be at least nearly true in so-called Chebyshev-Padé approximation; see [5] and the paper by Bultheel in this volume.) If so, their analysis will be aided by recognition of the recurring patterns of reasoning illustrated here. On the other hand, the strength of the analogies between various problems with equioscillation theorems suggests that useful connections between more complicated problems are also worth looking for. The multipoint Padé/PNA analogy mentioned above is a step in this direction.

5. Which block patterns are possible?

We know that the Walsh, Padé, and CF tables break into square blocks, but what about the converse? Can an arbitrary tiling of a quadrant by squares of various sizes, say, be realized as the block pattern of the Padé table of some formal power series f ? It seems this question has not appeared in print before, but apparently it was asked and answered several years ago by A. Magnus [11] (and possibly others).

The answer is no, and a simple example proves it. Suppose that the top two rows of the Padé (or Walsh or CF) table of f are known to break precisely into a chain of 2×2 blocks. Then it is easily seen that f is even, and this implies that the rest of its table also divides into 2×2 blocks, or larger. Therefore any finer patterns of subdivision are impossible. Of course one can readily generalize this example in various directions.

Since arbitrary block patterns are not permitted, one is naturally led to the following

PROBLEM. *Characterize all patterns of square blocks that can occur in the Walsh, Padé, and CF tables.*

One reason this problem is of interest is that its solution would help one judge the extent of validity of the theme of this paper --

that block patterns are mainly a function of superficial structure, not of problem-dependent details. What is the most one could hope for in this direction? One possibility is the following, offered here not with conviction but as a stimulus to further thought.

CONJECTURE. The sets of possible block patterns in the Walsh and Padé tables are identical.

Acknowledgments

I am happy to thank Martin Gutknecht, William Gragg, William Jones, and Arne Magnus for many useful comments.

References

[1] G. Baker and P. Graves-Morris, Padé Approximants (2 vols.), Encyc. of Math. v. 13 and 14, Addison-Wesley, 1981.

[2] A. Bultheel, paper in this volume.

[3] G. Claessens, On the structure of the Newton-Padé table, J. Approx. Theory 22 (1978), 304-319.

[4] M. A. Gallucci and W. B. Jones, Rational approximations corresponding to Newton series, J. Approx. Theory 17 (1976), 366-392.

[5] K. O. Geddes, Block structure in the Chebyshev-Padé table, SIAM J. Numer. Anal. 18 (1981), 844-861.

[6] W. Gragg, The Padé table and its relation to certain algorithms of numerical analysis, SIAM Review 14 (1972), 1-62.

[7] M. H. Gutknecht, On complex rational approximation II, in Computational Aspects of Complex Analysis, H. Werner et al. (eds.), D. Reidel, Dordrecht/Boston/Lancaster, 1983.

[8] M. H. Gutknecht, E. Hayashi, and L. N. Trefethen, The CF table, in preparation.

[9] M. H. Gutknecht and L. N. Trefethen, Nonuniqueness of best rational Chebyshev approximations on the unit disk, J. Approx. Theory 39 (1983), 275-288.

[10] A. Magnus, The connection between P-fractions and associated fractions, Proc. Amer. Math. Soc. 25 (1970), 676-679.

[11] A. Magnus, private communication, 1983.

[12] G. Meinardus, Approximation of Functions: Theory and Numerical Methods, Springer, 1967.

[13] A. Ruttan, The length of the alternation set as a factor in determining when a best real rational approximation is also a best complex rational approximation, J. Approx. Theory 31 (1981), 230-243.

[14] L. N. Trefethen, Rational Chebyshev approximation on the unit disk, Numer. Math. 37 (1981), 297-320.

[15] L. N. Trefethen, Chebyshev approximation on the unit disk, in Computational Aspects of Complex Analysis, H. Werner et al. (eds.), D. Reidel, Dordrecht/Boston/Lancaster, 1983.

[16] L. N. Trefethen and M. H. Gutknecht, On convergence and degeneracy in rational Padé and Chebyshev approximation, SIAM J. Math. Anal., to appear.

[17] H. Werner, On the rational Tschebyscheff operator, Math. Zeit. 86 (1964), 317-326.

[18] H. Werner and L. Wuytack, On the continuity of the Padé operator, SIAM J. Numer. Anal. 20 (1983), 1273-1280.