MATHEMATICS OF COMPUTATION
VOLUME 63, NUMBER 207
JULY 1994, PAGES 195–213

APPROXIMATION OF ANALYTIC FUNCTIONS: A METHOD OF ENHANCED CONVERGENCE

OSCAR P. BRUNO AND FERNANDO REITICH

ABSTRACT. We deal with a method of enhanced convergence for the approximation of analytic functions. This method introduces conformal transformations in the approximation problems, in order to help extract the values of a given analytic function from its Taylor expansion around a point. An instance of this method, based on the Euler transform, has long been known; recently we introduced more general versions of it in connection with certain problems in wave scattering. In §2 we present a general discussion of this approach.

As is known in the case of the Euler transform, conformal transformations can enlarge the region of convergence of power series and can enhance substantially the convergence rates inside the circles of convergence. We show that conformal maps can also produce a rather dramatic improvement in the conditioning of Padé approximation. This improvement, which we discuss theoretically for Stieltjes-type functions, is most notorious in cases of very poorly conditioned Padé problems. In many instances, an application of enhanced convergence in conjunction with Padé approximation leads to results which are many orders of magnitude more accurate than those obtained by either classical Padé approximants or the summation of a truncated enhanced series.

1. INTRODUCTION

Perturbation methods and series expansions lie at the heart of most mathematical discussions of problems in science and engineering. Linear partial and ordinary differential equations amount, in many cases, to first-order perturbation theory applied to basic principles of physics. Perturbation theory of higher order, on the other hand, has led to an understanding of phenomena that cannot be accounted for accurately by low-order expansions [22, 16, 1, 17, 4, 21, 7]. Yet, high-order perturbation series are often regarded critically. Convergence results for the classical approximation methods are not always available, and numerical ill-conditioning is always a concern. Thus, new summation methods and further understanding of classical methods are necessary.

Here we deal with a method of enhanced convergence for the approximation of analytic functions. This method introduces conformal transformations in the approximation problems in order to help extract the values of a given analytic function from its Taylor expansion around a point. An instance of this method, based on the Euler transform, has long been known; recently we introduced more

Received by the editor August 5, 1992 and, in revised form, August 30, 1993.
1991 *Mathematics Subject Classification.* Primary 65B10, 41A21, 41A25.
Key words and phrases. Power series, enhanced convergence, Padé approximation, conditioning.

0025-5718/94 $1.00 + $.25 per page

general versions of it in connection with certain problems in wave scattering [7, 6]. In §2 we present a general discussion of this approach.

As in known in the case of the Euler transform [23, 14, 15], conformal transformations can extend the region of convergence of power series and can enhance substantially the convergence rates inside the circles of convergence. In §§3 and 4 we show that they can also produce a rather dramatic improvement in the conditioning of Padé approximation. This improvement, which we discuss theoretically for Stieltjes-type functions, is most notorious in cases of very poorly conditioned Padé problems. In many instances, an application of enhanced convergence in conjunction with Padé approximation leads to results which are many orders of magnitude more accurate than those obtained by either classical Padé approximants or the summation of a truncated enhanced series (see §4).

Consider the Taylor series of a function $f(z)$ around $z = 0$. Clearly, poor convergence and lack of convergence of the Taylor series at a point z_0 are related to the arrangement of the singularities of f and the point z_0 relative to $z = 0$. The method of enhanced convergence uses conformal maps to manipulate the complex z-plane so as to produce an arrangement of z_0 and the singularities of f which is favorable for the summation of the series. In addition to the sum of the truncated enhanced series, Padé approximants of the enhanced function *with denominators of low degree* can be used at any point at which the conformal transformation has produced a convergent series, and can yield much better approximations than the enhanced series itself with a negligible additional computational cost (see §4).

A different phenomenon occurs in connection with Padé approximation with denominators of high degree: the conditioning of the Padé problem of the function f in the new variables improves very substantially. In other words, enhanced convergence acts as a preconditioner for Padé approximation. For example, a Fortran double-precision calculation of the function $f(z) = \log(1 + z)$ via regular Padé approximation, of any order, will not yield, at $z = 20$, more than the first four correct digits of $\log(21)$. After composition with an appropriate conformal map, 13 correct digits can be obtained. In fact, diagonal enhanced approximants of orders 50 already yield 9 correct digits, while the number of correct decimals is 13 for approximations of orders 120 to 180. For $z = 200$, an enhanced Padé approximation can produce up to 6 correct figures, while only one correct digit can be obtained through direct Padé calculation.

Since Padé approximants are connected with J-continued fractions and the latter, if written in partial fraction form, with Gaussian quadrature, the conditioning problem for Padé approximants is closely related to the conditioning problem for Gaussian quadrature rules[1]. The possible implications of the methods in this paper on the conditioning problem for Gaussian quadrature rules [11, 12] remain to be explored.

A comment is in order with regard to the calculation of the coefficients of the power series of a function f in the new variables. These coefficients can be produced either (1) by certain linear operations on the coefficients of the series of $f(z)$, or (2) by some alternative direct calculation of the coefficients

[1]We thank Professor W. Gautschi for pointing out this connection to us.

of the composite function. The information contained in the enhanced coefficients calculated by the first method will be limited by that contained in the coefficients of the series of $f(z)$, even though the Taylor coefficients of the true enhanced series encode, in certain cases, more information than those of the regular series. Indeed, as shown in [15, 24], the linear algebra that produces the enhanced coefficients from the coefficients of the direct series is ill-conditioned in some situations. Therefore, the second approach is to be preferred, if the corresponding accurate calculation is possible.

2. Enhanced series

Let Ω be a connected domain in the complex plane, $0 \in \Omega$, let $f(z)$ be an analytic function defined in Ω, and let D denote the circle of convergence of the Taylor series of f about $z = 0$. The divergence of the Taylor series of f at a point z_0 outside D is related to the presence of singularities of f on the boundary of D. We shall show that such singularities are also the cause of numerical ill-conditioning in Padé approximation; see §3. It is clear, then, that it should be useful to deform the domain Ω conformally, keeping $z = 0$ fixed, in such a way that the image of the point z_0 is closer to the origin than the image of any of the singularities. This simple observation is the basis of the method under investigation.

The implementation of this procedure relies on some a priori knowledge of the domain of analyticity of the function f. In applications, such information can usually be obtained from physical considerations, Padé approximation [3, §2.2], or even by studying the convergence of several enhanced series [7]. Once this information is available, the rearrangement of the singularities can be performed in many ways; in the following subsection we discuss some natural choices of conformal transformations that have proved to perform well. A few simple examples follow in §2.2. Further examples and applications, together with a discussion of the numerical aspects of the method in high-order applications will be given in §§3 and 4.

We begin our study by considering conformal maps which extend the radius of convergence of the Taylor series of an analytic function.

2.1. **Geometrical considerations.** Let f be an analytic function defined in Ω. We seek a conformal map $\xi = g(z)$ defined in Ω with $g(0) = 0$ and such that the image $\xi_0 = g(z_0)$ of a given point z_0 lies inside the circle of convergence of the Taylor series of $f \circ g^{-1}$ about $\xi = 0$. If such a function g is available, the value $f(z_0)$ can then be approximated by summing the truncated power series of $f \circ g^{-1}(\xi)$ at $\xi = \xi_0$.

Motivated by their geometrical properties as well as by their simplicity, rational fractions of the form

$$g(z) = \sum_{i=1}^{K} \frac{A_i z}{z + B_i} \tag{1}$$

appear as natural choices. We have found [7] that powers of these transformations can also be useful. The particular case $K = 1$ in (1) corresponds to the Euler transformation [23]. Our intuition here is that, if f is conformal, then clearly, an enhancer g that eliminates the singularities of f completely is the function f itself: $g = f$. It therefore seems reasonable to allow for g to

mimic part of the singular behavior of the function f. In this way, some of the singularities of f are mapped away to infinity.

The performance of the method depends in a critical way on the parameters A_i and B_i in (1). If we are interested in the computation of $f(z_0)$ by composition with g^{-1}, we may seek a combination of parameters for which the convergence of the series of $f \circ g^{-1}$ is fastest. Such optimal convergence rates result if the parameters A_i and B_i are chosen in such a way as to minimize the quotient

$$\left|\frac{g(z_0)}{R}\right| = \left|\frac{\xi_0}{R}\right|, \qquad R = \text{radius of convergence of } f \circ g^{-1} \text{ about } \xi = 0, \tag{2}$$

since the error in a truncated expansion of degree n is of order $|\xi_0/R|^{n+1}$. This fact was noted by Scraton [23] in his study of the Euler transform. Note that the parameters can be selected numerically by optimizing the convergence rates even if no information is known about the singularities of the function f.

To illustrate these ideas, let f be an arbitrary function and assume we know its singularities lie in the interval $[-1/a, -1/b]$. For example, we can take f to be a Stieltjes or Hamburger function of the form [3, Chapter 5]

$$f(z) = \int_a^b \frac{\phi(u)\,du}{1+uz} = \sum_{n=0}^{\infty} c_n z^n \tag{3}$$

with $\phi \geq 0$. For the conformal map we shall first use

$$g_1(z) = \frac{Az}{z+B} \qquad (A, B \in \mathbb{R}), \tag{4}$$

namely, the function that produces the Euler transform. The singularities of $f \circ g_1^{-1}$ are delimited by $g_1(-1/a)$ and $g_1(-1/b)$, and, therefore, the radius of convergence of the composite map is the smaller of the absolute values of these two numbers. It follows from (2) that a choice of parameters that gives optimal convergence rates must minimize the expression

$$\max\left(\left|\frac{g_1(z_0)}{g_1(-1/a)}\right|, \left|\frac{g_1(z_0)}{g_1(-1/b)}\right|\right). \tag{5}$$

It is easily seen from (5) that the optimal B does not depend on z_0 and that it is given by

$$B = \frac{2}{a+b}; \tag{6}$$

see [23]. The parameter A cancels in formula (5) and can be normalized to 1.

The next simplest example of conformal maps of the type (1) is

$$g_2(z) = \frac{A_1 z}{z+B_1} + \frac{A_2 z}{z+B_2}. \tag{7}$$

Motivated by (3) and in order to ensure the invertibility of g_2, we assume $A_1, A_2, B_1, B_2 > 0$. The (relevant branch of the) function g_2^{-1} is then given by

$$g_2^{-1}(\xi) = \frac{(B_1+B_2)\xi - (A_1B_2 + A_2B_1) + \sqrt{\Delta}}{2(A_1+A_2-\xi)},$$

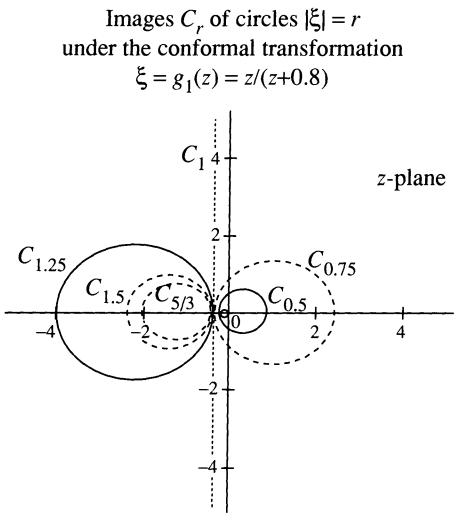

FIGURE 1(a)

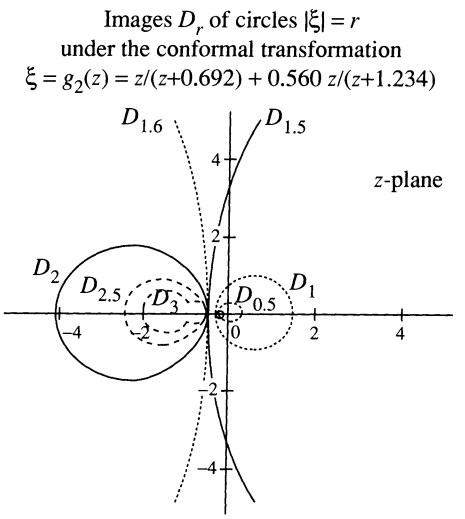

FIGURE 1(b)

where

$$\Delta = (B_1 - B_2)^2\xi^2 + 2(B_1 - B_2)(A_1B_2 - A_2B_1)\xi + (A_1B_2 + A_2B_1)^2.$$

Again, the optimal choice of parameters minimizes the quotient in (2). In this case it is not possible to derive a simple formula such as (6) for the parameters A_i and B_i. Here we need to deal not only with the singularities of f but also with those introduced by g_2^{-1}. It is not difficult to check, however, that the optimal situation is the one in which the parameters minimize the expression

$$\max\left(\left|\frac{g_2(z_0)}{g_2(-1/a)}\right|, \left|\frac{g_2(z_0)}{g_2(-1/b)}\right|, \left|\frac{g_2(z_0)}{r_\Delta}\right|\right), \tag{8}$$

where r_Δ denotes the absolute value of the (complex conjugate) roots of Δ as a function of ξ:

$$r_\Delta = \frac{A_2B_1 + A_1B_2}{B_2 - B_1}.$$

As in (5), we can take one of the parameters A_i in (8), say A_1, to equal 1.

It is reasonable in some cases to take $z_0 = \infty$ in (8) so as to optimize the convergence rates of the approximator in the positive real axis (see §2.2). With this provision (and taking $A_1 = 1$), the parameters A_2, B_1, and B_2 must be chosen so as to minimize

$$\max\left(\left|\frac{1 + A_2}{g_2(-1/a)}\right|, \left|\frac{1 + A_2}{g_2(-1/b)}\right|, \left|\frac{1 + A_2}{r_\Delta}\right|\right). \tag{9}$$

Geometrical insight can be gained by inspection of the effect of the conformal maps described above on circles in the ξ-plane. In Figure 1(a) (resp. 1(b)) we have plotted the images C_r (resp. D_r) in the z-plane of the circles $|\xi| = r$ under the transformation $\xi = g_1(z)$ (resp. $\xi = g_2(z)$). We have chosen the singularity region to be the interval $[-2, -1/2]$, i.e., $a = 1/2$, $b = 2$, which corresponds to the function f in (10) below. Thus, the Taylor series of f around $z = 0$ converges in the circle of radius $\frac{1}{2}$ in the z-plane centered at the origin. From (6) it follows that the parameter B in g_1 is, in this case, equal

to 0.8, while a numerical minimization of (9) yields the values $A_2 = 0.560$, $B_1 = 0.692$, and $B_2 = 1.234$ (cf. (12)) for the conformal map g_2.

Take, for example, the curve $C_{1.25}$ in Figure 1(a). This circle is the image under the map $\xi = g_1(z)$ of the circle $|\xi| = 1.25$. The region $|\xi| < 1.25$ is mapped onto the exterior of $C_{1.25}$, i.e., onto the connected component containing $z = 0$. Thus, since this region does not intersect the interval $[-2, -1/2]$, we see that, for *all* points z outside $C_{1.25}$, the value of $f(z)$ can be obtained by summing the Taylor series of $f \circ g_1^{-1}$ at $\xi = g_1(z)$. Similar considerations hold for all other curves in Figure 1(a) and 1(b). The radius of convergence of $f \circ g_1^{-1}$ is $r = 5/3$ while that of $f \circ g_2^{-1}$ is $r = 3$. We see that the enhanced series will in fact converge to $f(z)$ for all z outside the critical curves $C_{5/3}$ and D_3.

In the following subsection we illustrate the ideas above with a few low-order approximation problems. Higher-order approximation will be dealt with in §§3 and 4.

2.2. **Some simple examples.** Consider first the function

$$f(z) = \sqrt{\frac{1+z/2}{1+2z}} = 1 - \frac{3}{4}z + \frac{39}{32}z^2 - \cdots. \tag{10}$$

A second-order approximation problem for this function is used in [3] to demonstrate some of the outstanding properties of Padé approximants. The $[L/M]$ Padé approximant of a function $f(z) = \sum_{n=0}^{\infty} c_n z^n$ is defined (see [3]) as a rational function

$$[L/M] = \frac{a_0 + a_1 z + \cdots + a_L z^L}{1 + b_1 z + \cdots + b_M z^M}$$

whose Taylor series agrees with that of f up to order $L + M$. A particular $[L/M]$ approximant may fail to exist but, generically, $[L/M]$ Padé approximants exist and are uniquely determined by L, M, and the first $L + M + 1$ coefficients of the Taylor series of f. For convergence studies and numerical calculation of Padé approximants see [3, 5, 8, 13].

The $[1/1]$ Padé approximant of the function f in (10) is given by

$$[1/1] = \frac{1 + \frac{7}{8}z}{1 + \frac{13}{8}z}. \tag{11}$$

Certainly, the information used to construct (11) (namely the first three terms in the Taylor series of f) would permit us to compute the $[2/0]$ and $[0/2]$ approximants also. The choice of the $[1/1]$ approximant may be seen as incorporating certain additional structural information one has about the function f.

Let us now find enhanced series of order 2 for (10). Consider first the conformal map g_1 defined in (4). In this case, (6) yields $B = 0.8$, and taking $A = 1$, we obtain $g_1(z) = z/(z + 0.8)$. Therefore, our approximation reads

$$1 - \frac{0.6z}{z+0.8} + \frac{0.18z^2}{(z+0.8)^2} = \frac{29z^2 + 56z + 32}{2(5z+4)^2}.$$

Another enhanced series can be obtained by using the conformal map g_2 in (7). The expression in (9) can be (numerically) minimized, and the optimal parameters turn out to be

$$A_1 = 1, \quad A_2 = 0.560, \quad B_1 = 0.692, \quad B_2 = 1.234. \tag{12}$$

Our second enhanced series approximation is then given by

$$1 - \frac{0.395z}{z+0.692} - \frac{0.221z}{z+1.234} + 0.069\left(\frac{z}{z+0.692} + \frac{0.560z}{z+1.234}\right)^2.$$

The three approximations lie close to the function f, with errors at $z = \infty$ of 8%, 16%, and 10% for Padé, g_1-enhanced, and g_2-enhanced, respectively. The Padé approximation is slightly more accurate than the other two; this need not be the case, however, as we illustrate with the following example.

Let

$$f(z) = \sqrt{\frac{z}{(z+2)(z+3)} + 1}.$$

The [1/1] Padé approximant of f is given by

$$[1/1] = \frac{1 + \frac{23}{24}z}{1 + \frac{7}{8}z}.$$

To compute the enhanced series corresponding to the map g_1, we find from (6) that $B = 2$, and we take $A = 1$. Thus, the enhanced series is given by

$$1 + \frac{z}{6(z+2)} - \frac{z^2}{8(z+2)^2} = \frac{25z^2 + 104z + 96}{24(z+2)^2}.$$

Analogously, it is found that the parameters corresponding to the conformal map g_2 are, in this case, given by

$$A_2 = 0.578, \quad B_1 = 1.732, \quad B_2 = 3.000,$$

so that the enhanced series is

$$1 + \frac{0.108z}{z+1.732} + \frac{0.063z}{z+3} - 0.051\left(\frac{z}{z+1.732} + \frac{0.578z}{z+3}\right)^2.$$

Again, the three approximations are fairly accurate, taking into account the fact that they have been obtained by using only the first three coefficients of the Taylor expansion. In this case, either of the two enhanced series is a better approximation to the function f than the [1/1] Padé approximant: the errors at $z = \infty$ are of 9.5%, 4.2%, and 4.5% for Padé, g_1-enhanced, and g_2-enhanced, respectively.

It is often the case in applications that a high number of terms in the power series representing the relevant quantities are required in order to reach a reasonable approximation (see, e.g., [22, 16, 1, 17, 4, 21, 7]). A variety of experiments (including those in [7]) have led us to believe that in such cases Padé approximation yields much better results than the summation of the truncated enhanced series. This is so even in cases in which Padé results give limited accuracy due to ill-conditioning. We therefore turn out attention to improving the conditioning of the Padé problem. As we shall show, such improvements can be obtained by means of the conformal changes of variables discussed above. For a study of the conditioning of the value problem for the truncated enhanced series in the particular case of the Euler transformation see [15, 24].

3. High-order approximations: conditioning of Padé and enhanced Padé

In problems in which high-order approximations are necessary, small numerical errors in the Taylor coefficients may lead to poor results for the values of the function. In this section we discuss the effect of these errors in the values of high-order Padé approximants and those of the Padé approximants of the enhanced series—which we will refer to as *enhanced Padé approximants*. To do this, we first introduce, in §3.1, appropriate norms and corresponding condition numbers for the values of the Padé *denominator*. In §3.2 we then restrict ourselves to Stieltjes-type functions and investigate the conditioning of the value problem for the Padé denominator of both direct and enhanced series. We conclude that the conditioning of the Padé denominator problem of an appropriate enhanced series can be substantially better than that of the original series. Examples in §4 show that improvement in the denominator condition number is closely related to a corresponding substantial improvement in the overall quality of the approximations, and that much better numerical results can be expected of enhanced Padé than of Padé approximants for general analytic functions.

Our theoretical discussion in §3.2 concerns the value problem for the Padé denominator. While theoretical studies of the conditioning of the value problem for Padé approximants are not available at present, the generalized belief is that the conditioning of the denominator problem determines that of the whole fraction (see [3]). It must not be understood, however, that the amplification of errors observed in the values of the Padé denominator is to be expected in the whole fraction. Indeed, the conditioning of the Padé fraction is observed to be very substantially better than that of the Padé denominator. The numerical experiments of Luke [20] shed some light on this astonishing property of Padé approximants, which remains, otherwise, not understood. Luke does not study the relative error in the denominator itself, and, indeed, our discussion in this regard appears to be the first one in the literature. At any rate, our present study of the conditioning of the Padé denominator in both the direct and enhanced variables together with the numerical experiments of §4 indicate clearly that there is a close correlation between the conditioning of the denominator problem and that of the whole fraction; or, in other words, improvement in the denominator condition leads to improvement in the condition of the whole fraction, even though a quantitative measure of the former is not necessarily a good quantitative measure of the latter.

3.1. **Condition numbers for the denominator value problem.** The coefficients of the denominator of the $[L/M]$ Padé approximant of the function

$$f(z) = \sum_{n=0}^{\infty} c_n z^n \tag{13}$$

are given by the solution of the linear system of equations

$$\begin{bmatrix} c_{L-M+1} & c_{L-M+2} & \cdots & c_L \\ c_{L-M+2} & c_{L-M+3} & \cdots & c_{L+1} \\ \vdots & \vdots & & \vdots \\ c_L & c_{L+1} & \cdots & c_{L+M-1} \end{bmatrix} \begin{bmatrix} b_M \\ b_{M-1} \\ \vdots \\ b_1 \end{bmatrix} = - \begin{bmatrix} c_{L+1} \\ c_{L+2} \\ \vdots \\ c_{L+M} \end{bmatrix}. \tag{14}$$

The numerical stability of the problem of computing the coefficients b_n from

(14) is governed by the (e.g. l^2) condition number of the matrix C in (14), i.e., by

$$\kappa(C) = \|C\| \, \|C^{-1}\|. \tag{15}$$

Here, $\|C\|$ is the matrix norm associated with the l^2 vector norm

$$\|(b_n)\| = \sqrt{\sum_{n=1}^{M} |b_n|^2}.$$

In other words, errors (δc_n) in the coefficients of (13) are amplified in the calculation of the denominator coefficients (b_n), and result in relative errors which can be estimated by

$$\frac{\|(\delta b_n)\|}{\|(b_n)\|} \leq \kappa(C) \frac{\|(\delta c_n)\|}{\|(c_n)\|}.$$

Our main problem, however, is that of calculating the *values* of the Padé approximant, and the condition number (15) does not provide a measure of the error in these values. In fact, the natural measure for the error in the values of the denominator is

$$\sum_{n=1}^{M} |\delta b_n| \, |z|^n. \tag{16}$$

For convenience we shall use a norm closely related to, but different from, (16), namely

$$\|b\|_z \equiv \sqrt{\sum_{n=1}^{M} |b_n|^2 |z|^{2n}}. \tag{17}$$

To treat the value problem for the Padé denominator, we observe that its coefficients satisfy the following system of equations

$$\begin{bmatrix} c_{L-M+1} z^{L-M+1} & c_{L-M+2} z^{L-M+2} & \cdots & c_L z^L \\ c_{L-M+2} z^{L-M+2} & c_{L-M+3} z^{L-M+3} & \cdots & c_{L+1} z^{L+1} \\ \vdots & \vdots & & \vdots \\ c_L z^L & c_{L+1} z^{L+1} & \cdots & c_{L+M-1} z^{L+M-1} \end{bmatrix} \begin{bmatrix} b_M z^M \\ b_{M-1} z^{M-1} \\ \vdots \\ b_1 z \end{bmatrix} = - \begin{bmatrix} c_{L+1} z^{L+1} \\ c_{L+2} z^{L+2} \\ \vdots \\ c_{L+M} z^{L+M} \end{bmatrix}, \tag{18}$$

as follows from equations (14). We then define the condition number of the denominator value problem as the l^2 condition number $\kappa_v(z)$ of the matrix in (18). This number permits us to bound the error $\varepsilon_{b,z} = \|\delta b\|_z / \|b\|_z$ by the error $\varepsilon_{c,z} = \|\delta c\|_z / \|c\|_z$, i.e., roughly

$$\varepsilon_{b,z} \leq \kappa_v(z) \varepsilon_{c,z}.$$

Notice that this condition number is unchanged if the problem is transformed via $z \to \lambda z$ $(\lambda \in \mathbb{R})$, as expected from dimensional considerations. This is not true, however, of the condition number (15) for the coefficient problem; that is, the number $\kappa(C)$ *does* change if the variable z is transformed homothetically.

3.2. **Conditioning of the denominator value problem of Padé and enhanced Padé approximants.** To gain insight on the effect that a rearrangement of the singularities of the function f can have on the conditioning of the denominator problem (and therefore on the conditioning of the complete Padé fraction, according to the discussion at the beginning of §3), assume that the singularity of f that lies closest to the origin is a simple pole at $z = z_0$. Then in (13), (14) we have $c_n = \text{const} \cdot z_0^{-n} + o(|z_0|^{-n})$ for large n, which explains the ill-conditioning of the matrix for large values of M. A conformal change of variables which equilibrates the influence of the closest singularities is therefore expected to have a beneficial effect on the conditioning of the denominator problem. In this section we provide a quantitative measure of the improvement under the assumption that f is a Stieltjes function with a positive radius of convergence.

In what follows, we shall consider Stieltjes functions [3, Chapter 5], that is, functions which admit an integral representation

$$f(z) = \int_a^b \frac{\phi(u)\,du}{1+zu}. \tag{19}$$

Theorem 1 below permits one to estimate the improvement produced by a conformal transformation of the type (4) in the condition number of the Padé problems for functions of the form $zf(z)$, where f is a Stieltjes function (see Remark 1).

We begin our discussion of the conditioning of the denominator problem with the following lemma, which follows readily from a change of variables.

Lemma 1. *Let f be a Stieltjes function of the form* (19). *Then, for any A and B we have*

$$f(z) = \frac{A}{z+B} \int_{(Ba-1)/A}^{(Bb-1)/A} \frac{\phi((Au+1)/B)\,du}{1+\xi u},$$

where

$$\xi = \frac{Az}{z+B}. \tag{20}$$

In other words, calling

$$e(\xi) = \int_{(Ba-1)/A}^{(Bb-1)/A} \frac{\phi((Au+1)/B)\,du}{1+\xi u}, \tag{21}$$

we have

$$zf(z) = \xi e(\xi). \tag{22}$$

We continue with two lemmas about certain quadratic forms for the vector $x = (x_0, x_1, \ldots, x_n) \in \mathbb{R}^{n+1}$. These quadratic forms are closely related to the Padé approximants of Stieltjes functions, and they are given by integrals such as

$$\begin{aligned} x^t A_\phi^{n,m} x &= \int_a^b (x_0 + x_1 u + \cdots + x_n u^n)^2 u^m \phi(u)\,du \\ &= \sum_{i=0}^{n} \sum_{j=0}^{n} (-1)^{i+j+m} c_{i+j+m} x_i x_j, \end{aligned} \tag{23}$$

where a and b are real numbers, $a < b$, and c_k are the Taylor coefficients of f in (19), i.e.,

$$c_k = (-1)^k \int_a^b u^k \phi(u)\, du.$$

Also, we shall denote

$$A^{n,m} = A_\phi^{n,m} \quad \text{if } \phi(u) \equiv 1.$$

We see that the matrices $A^{n,m}$ are positive definite provided either $0 < a < b$ or m is even; in the latter case we shall write

$$E^{n,l} = A^{n,m}, \quad \text{if } m = 2l. \tag{24}$$

Clearly, then

$$x^t E^{n,l} x = \int_a^b (x_0 u^l + x_1 u^{l+1} + \cdots + x_n u^{l+n})^2\, du.$$

Lemma 2. *Let $m = 2l$. Then, we have*

$$x^t E^{n,l} x = (b-a) y^t H^{n+l} y,$$

where y is related to x via

$$y = D T_a \mathcal{J} x. \tag{25}$$

Here, H^{n+l} denotes the $(n+l+1) \times (n+l+1)$ Hilbert matrix

$$H_{ij}^{n+l} = \frac{1}{i+j-1} \qquad (1 \le i, j \le n+l+1),$$

$D = D_{(b-a)}$ is the $(n+l+1) \times (n+l+1)$ diagonal matrix

$$D_{ii} = (b-a)^{i-1} \qquad (1 \le i \le n+l+1), \tag{26}$$

T_a is the $(n+l+1) \times (n+l+1)$ matrix

$$(T_a)_{ij} = \binom{j-1}{i-1} a^{j-i} \qquad (1 \le i, j \le n+l+1), \tag{27}$$

and $\mathcal{J}$ is the matrix of the inclusion of $\mathbb{R}^{n+1}$ into $\mathbb{R}^{l+n+1}$

$$\mathcal{J} x = (0, \ldots, 0, x_0, \ldots, x_n). \tag{28}$$

Proof. By a change of variables, we obtain

$$\begin{aligned} x^t E^{n,l} x &= \int_0^{b-a} (x_0 (v+a)^l + x_1 (v+a)^{l+1} + \cdots + x_n (v+a)^{l+n})^2\, dv \\ &= \int_0^{b-a} (\overline{x}_0 + \overline{x}_1 v + \cdots + \overline{x}_{l+n} v^{l+n})^2\, dv, \end{aligned}$$

where $\overline{x} = (\overline{x}_0, \ldots, \overline{x}_{l+n})$ is given by $\overline{x} = T_a \mathcal{J} x$ with the matrices T_a and $\mathcal{J}$ defined by (27) and (28), respectively. A further change of variables yields

$$x^t E^{n,l} x = \int_0^1 (y_0 + y_1 u + \cdots + y_{l+n} u^{l+n})^2 (b-a)\, du = (b-a) y^t H^{n+l} y,$$

with $y = (y_0, \ldots, y_{l+n}) = D\overline{x}$, and D given by (26). □

Lemma 3. *Assume the function ϕ is positive and bounded, $0 < C_1 < \phi < C_2 < \infty$. Then, the following inequalities hold:*

- *$a > 0$ and m an arbitrary nonnegative integer: then*

$$K_1 y^t H^n y \le x^t A_\phi^{n,m} x \le K_2 y^t H^n y$$

for certain constants K_1 and K_2. Here, x and y are related through the equation $y = D_{b-a} T_a x$, where D_{b-a} and T_a are the $(n+1)\times(n+1)$ matrices whose entries are given by equations (26) and (27) (with $l = 0$), respectively.

- *$a \in \mathbb{R}$ arbitrary and $m = 2l$ a nonnegative even integer: then*

$$K_1 y^t H^{n+l} y \le x^t A_\phi^{n,m} x \le K_2 y^t H^{n+l} y.$$

Here, x and y are related through equation (25).

Proof. Follows easily from the previous lemma. □

The estimation of the condition number for the denominator value problem will result from the following theorem.

Theorem 1. *Let f and e be defined by (19) and (21), respectively. Let*

$$m = L - (M-1) = 2l \quad \text{and} \quad n = M - 1,$$

and, for a given complex z, define $\tilde{D}_z \in \mathbb{R}^{n+1\times n+1}$ to be a diagonal matrix given by

$$(\tilde{D}_z)_{ii} = z^{i+l-1} \qquad (1 \le i \le n+1).$$

Then the condition numbers $\kappa_v(z)$ and $\kappa_v(\xi)$ for the (denominator) value problem of the $[L/M]$ Padé approximants for the functions $f(z)$ and $e(\xi)$ are given by

$$\kappa_v(z) \simeq \kappa((D_{b-a} T_a \mathcal{F} \tilde{D}_z)^t H^{n+l} (D_{b-a} T_a \mathcal{F} \tilde{D}_z)) \tag{29}$$

and

$$\kappa_v(\xi) \simeq \kappa((D_{\gamma_2} T_{\gamma_1} \mathcal{F} \tilde{D}_\xi)^t H^{n+l} (D_{\gamma_2} T_{\gamma_1} \mathcal{F} \tilde{D}_\xi)), \tag{30}$$

where

$$\gamma_1 = \frac{Ba-1}{A} \quad \text{and} \quad \gamma_2 = \frac{B(b-a)}{A}$$

and $\kappa(C)$ denotes the l^2 condition number of a matrix C.

If $a > 0$, then the estimate (29) sharpens to

$$\kappa_v(z) \simeq \kappa((D_{b-a} T_a \tilde{D}_z)^t H^n (D_{b-a} T_a \tilde{D}_z)).$$

Remark 1. The theorem above permits us to obtain expressions for the $[L+1/M]$ denominator condition number of functions $zf(z)$ and its transform in the ξ-variables $\xi e(\xi)$, see (20), where f is a Stieltjes function (19) and $L-(M-1)$ is even. Indeed, these numbers are identical to the corresponding ones for the $[L/M]$ approximants of the functions $f(z)$ and $e(\xi)$ respectively; the latter are calculated in the theorem.

Proof of Theorem 1. We shall only show how to obtain (29) since, using Lemma 1, we can establish (30) in a similar way. Let S and $S(z)$ denote the matrices in (14) and (18), respectively, and let J be the $(n+1)\times(n+1)$ diagonal matrix with entries $J_{ii} = (-1)^{l+i-1}$. Since

$$\kappa(S(z)) = \kappa(JS(z)J),$$

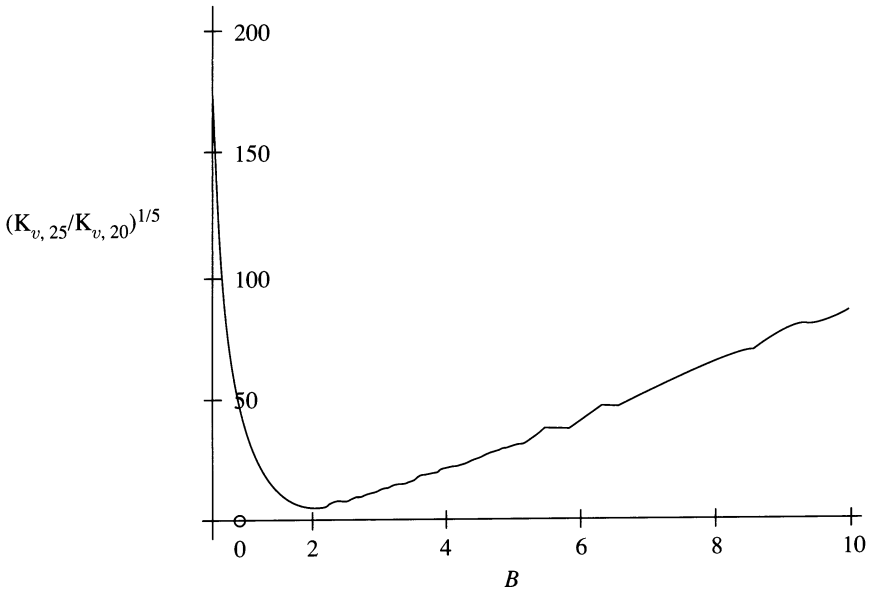

FIGURE 2. Condition number as a function of B, $\xi = z/(z+B)$: $a = 0$, $b = 1$, $m = 0$, $z = 20$

it suffices to estimate the condition number of the matrix $R = JS(z)J$. Now, from the equation

$$R_{(i+1)(j+1)} = (JS(z)J)_{(i+1)(j+1)} = (-1)^{i+j+m} c_{i+j+m} z^{i+j+m} \qquad (0 \le i, j \le n)$$

and (23) we deduce that

$$R = \tilde{D}_z A_\phi^{n,m} \tilde{D}_z.$$

Since for a positive definite and symmetric matrix F we have

$$\|F\| = \max_{\|x\|=1} x^t F x \quad \text{and} \quad \|F^{-1}\| = \frac{1}{\min_{\|x\|=1} x^t F x},$$

we conclude, from Lemma 3, that

$$\kappa(R) = \kappa(\tilde{D}_z A_\phi^{n,m} \tilde{D}_z) \simeq \kappa((D_{b-a} T_a \mathcal{F} \tilde{D}_z)^t H^{n+l} (D_{b-a} T_a \mathcal{F} \tilde{D}_z)). \quad \square$$

The right-hand sides of (29) and (30) can be evaluated using the fact that $T_s^{-1} = T_{-s}$ and the explicit formula for the inverse of the Hilbert matrix (see, e.g., [10])

$$(H^N)_{ij}^{-1} = \frac{(-1)^{i+j}(N+i)!(N+j)!}{(N+1-j)!(N+1-i)!(j-1)!^2(i-1)!^2(i+j-1)}.$$

In Figure 2 we show the dependence of the right-hand side of (30) on the parameter B in (4) in the case $m = 0$, $a = 0$, and $b = 1$ (which yields, in particular, numbers that apply to the function $f(z) = \log(1+z)/z$). In the figure, A was normalized to 1, so that

$$\xi = \frac{z}{z+B},$$

and we have plotted

$$(\kappa_{v,25}/\kappa_{v,20})^{1/5} \quad \text{for } z = 20$$

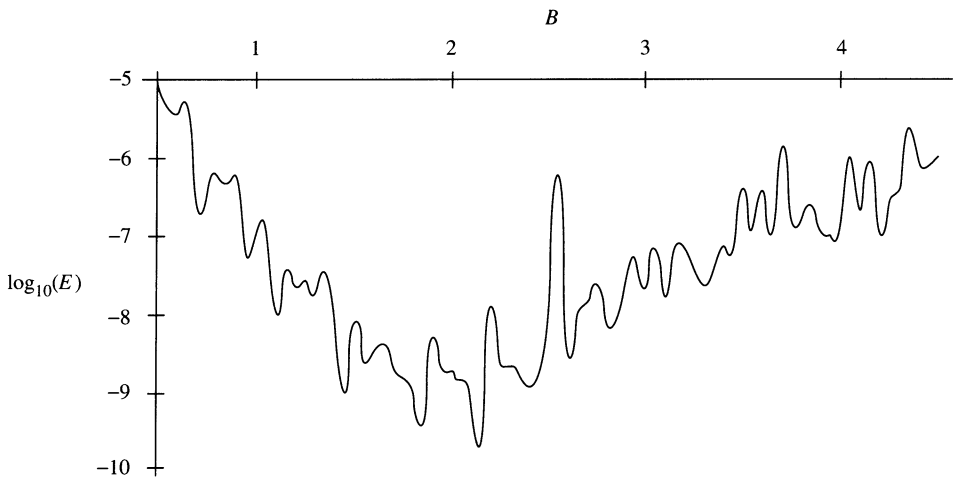

FIGURE 3

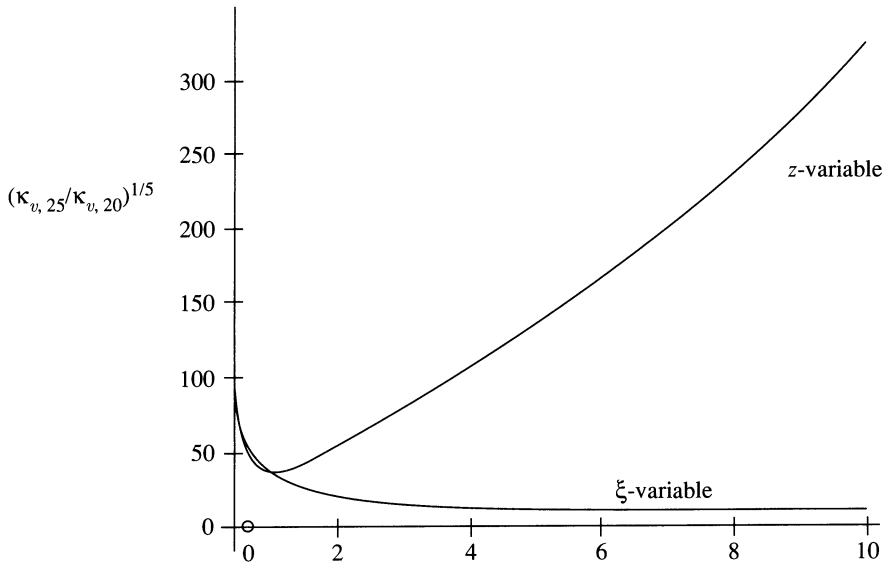

FIGURE 4

as a function of B. Here $\kappa_{v,k}$ is the condition number (30) with $l = 0$ and $n = k$.

In Figure 3 we plot the errors in the [20/21] enhanced Padé approximants at $z = 20$ for $f(z) = \log(1 + z)/z$ *as a function of* B. We observe that, as claimed at the beginning of this section, the condition number in Figure 2 as well as the errors in Figure 3 are smallest for the value of B in (6).

Finally, in Figure 4 we present a plot of the condition numbers as a function of z for the functions f ("z-variable") and e ("ξ-variable") (cf. (21)) again in the case $m = 0$, $a = 0$, and $b = 1$. Here the parameters for the conformal

map are $A = 1$ and $B = 2$, and we see that the conditioning is in fact improved by the change of variables.

4. Examples

In this section we apply the ideas presented in this paper to some elementary analytic functions. These functions have been chosen so as to illustrate the quality of the approximations that can be obtained—by means of a simple change of variables—in problems in which classical approximants have had limited success. As we have said, the key to the most successful approximations is an accurate calculation of the coefficients of the enhanced series. As noted in [15, 24], the calculation of these coefficients by direct composition of power series may produce enhanced coefficients of poor quality. In most of the examples that follow, we will therefore obtain the enhanced coefficients by alternative means. Because of the simplicity of the elementary functions used below, such accurate calculations do not represent a challenge, and they will be described in each case. In more complex applications, an accurate calculation of the enhanced coefficients may not be a simple matter and must be regarded as an integral part of the problem. These questions will, at any rate, be left for future work.

From many numerical experiments, among which the ones in this section were chosen, a clear picture emerges: diagonal or close to diagonal enhanced Padé fractions are probably never worse and can be very substantially better than classical Padé approximants or truncated enhanced series. The degree of improvement of the enhanced Padé method over regular Padé approximants is most notorious in cases of poorly conditioned Padé problems. Enhanced Padé fractions with denominators of low degree can be as accurate or, if the number of coefficients is large enough, slightly more accurate than enhanced diagonal Padé fractions. If a large number of coefficients are available, this option may be attractive since it reduces the ill-conditioning of the problem and, at the same time, it results in a lower computational cost. Summation of the truncated enhanced series, on the other hand, is an alternative to other approximants in low-order problems (see §2.2), but it appears that any of the other proposed methods performs better in problems of higher order.

The computations that follow have been performed in Fortran, and double-precision arithmetic has been used in all cases. Padé approximants have been calculated by means of the approach recommended in [13, 3], that is, via solution of the denominator equations by Gaussian elimination with partial pivoting and iterative refinement [10]. Also, for simplicity, attention is restricted to conformal maps of the type (4). The accuracy of the enhanced approximants is independent of the parameter A in (4), and we have therefore taken $A = 1$. Other conformal maps can, of course, be useful in these and other circumstances.

Our first example is a classical one in approximation theory.

- $f(z) = \log(1 + z)$

In Table 1 we show the values of the $[\frac{N}{2}/\frac{N}{2}]$ Padé and enhanced Padé approximants for the function $f(z) = \log(1 + z)$. Since the singularities of $\log(1 + z)$ lie on the interval $[-\infty, -1]$, we see from (6) that the optimal constant B is $B = 2$.

A technical point here relates to the calculation of the coefficients of the

TABLE 1 $[\frac{N}{2}/\frac{N}{2}]$ approximants for $\log(1+z)$

N	z	$\log(1+z)$	Padé	Enh. Padé
20	20	3.044522437723	3.04399	3.04398878414
40			3.04461	3.044522360574
60			3.04448	3.044522437596
80			3.04418	3.044522437727
100			3.04446	3.044522437722
120			3.04449	3.044522437724
140			3.04450	3.044522437723
160			3.04462	3.044522437723
180			3.04436	3.044522437723
20	200	5.30330	5.04	5.03577
40			5.33	5.28588
60			5.18	5.30093
80			5.09	5.30276
100			5.17	5.30305
120			5.19	5.30324
140			5.20	5.30328
160			5.71	5.30329
180			5.14	5.30330

enhanced series. Because the composite function $f \circ g_1^{-1}$ is given by

$$f \circ g_1^{-1} = \log\left(1 + \frac{2\xi}{1-\xi}\right) = \log(1+\xi) - \log(1-\xi),$$

the enhanced series can simply be obtained as the difference of the series of $\log(1+\xi)$ and $\log(1-\xi)$. A calculation of the enhanced coefficients by composition of the series of f and g_1 results in enhanced approximants of comparable or worse quality than the corresponding Padé fractions.

Table 1 shows that, as noted in the introduction, enhanced Padé approximants produce up to 13 correct digits of $\log(21)$, while ordinary Padé fractions do not produce more than the first four digits.

A point of interest here relates to the fact that, for approximants with denominator and numerator of the same degree, *in exact arithmetic*, and for the conformal map (4) which is being used, the Padé and enhanced Padé calculations coincide. This is a well-known and simple fact, sometimes called the theorem of Baker, Gammel, and Wills [2]; see also Edrei [9]. We conclude that the improvement in the approximation is solely due to a better conditioning for the value problem of enhanced approximants.

• Enhanced series and low-degree denominator enhanced Padé

Let ξ be a point in the disk of convergence of the enhanced series. Since the series in the enhanced variables converges at ξ, its Padé approximants with denominators of low degree usually converge there also. These high-order low-denominator-degree enhanced Padé approximants can produce very good results (with a low computational cost) as we illustrate in Table 2. It is a remarkable fact that a Padé approximant with a denominator of degree as low as 5 can produce such a substantial improvement of the convergence rate of the enhanced series.

In Table 2 we show higher-order approximations for the function $f(z) = \log(1+z)/z$ at $z = 20$, where

$$\log(21)/20 = 0.1522261218861711.$$

TABLE 2. High-order approximants for $\log(1+z)/z$

N	Padé	Enh. Series	Enh. Padé (*low*)	Enh. Padé (*high*)
60	0.1522228	0.15224926416	0.15222613236258	0.15222612189111
100	0.1522238	0.15222642286	0.15222612190656	0.15222612188430
160	0.1522252	0.15222612249	0.15222612188617	0.15222612188623
180	0.1522015	0.15222612197	0.15222612188617	0.15222612188614

TABLE 3. $[\frac{N}{2}/\frac{N}{2}]$ approximants for $\sqrt{\frac{(1+(1+i)z)(1+(1-i)z)}{(1+10(1+i)z)(1+10(1-i)z)}}$

N	z	$f(z)$	Padé	Enh. Padé
20	50	0.100903995976172	0.101691	0.101690955078874
40			0.100852	0.100907485427384
60			0.100830	0.100904026370804
80			0.100863	0.100903996274708
100			0.100859	0.100903996124993
120			0.100979	0.100903995972737
140			0.100961	0.100903995978357
160			0.101015	0.100903995976236
180			0.101039	0.100903995976198
20	500	0.100090040445593	0.100925	0.100925224160906
40			0.100033	0.100093996177684
60			0.100009	0.100090077056242
80			0.100045	0.100090040827442
100			0.100041	0.100090040634719
120			0.100170	0.100090040440984
140			0.100152	0.100090040448810
160			0.100210	0.100090040445689
180			NaN	0.100090040445629

Besides the regular Padé approximants and enhanced series we include enhanced Padé approximants with denominators of degree 5 (*low*) and of degree $N/2+1$ (*high*). Note that, for very large N, approximants of low denominator degree perform better than diagonal ones.

Finally, we present an example of a function whose singularities are not real. Even the simple conformal transformation (4) can provide excellent approximations in such cases.

- $f(z) = \sqrt{\frac{(1+(1+i)z)(1+(1-i)z)}{(1+10(1+i)z)(1+10(1-i)z)}}$

In this case, the coefficients of the enhanced series were calculated as products of series whose coefficients are given by simple formulae. It is easy to check that the optimal value for the parameter is $B = 0.1$. The computer produced NaN ("Not a Number"), an overflow indicator, in the case of the [90/90] direct approximant for $z = 500$.

In Table 3 we show some values of several $[\frac{N}{2}/\frac{N}{2}]$ Padé and enhanced Padé fractions. The qualitative picture remains unchanged.

ACKNOWLEDGMENTS

We thank the reviewer and Professor W. Gautschi for their valuable comments. The first author gratefully acknowledges support from NSF through grant No. DMS-9200002. This work was partially supported by the Army Research

Office and the National Science Foundation through the Center for Nonlinear Analysis.

BIBLIOGRAPHY

1. G. A. Baker, *The theory and application of the Padé approximant method*, Advances in Theoretical Physics, Vol. I (K. A. Brueckner, ed.), Academic Press, New York, 1965.
2. G. A. Baker, J. L. Gammel, and J. G. Wills, *An investigation of applicability of the Padé approximant method*, J. Math. Anal. Appl. **2** (1961), 405–418.
3. G. A. Baker and P. Graves-Morris, *Padé approximants. Part* I: *Basic theory*, Addison-Wesley, Reading, MA, 1981.
4. ——, *Padé approximants. Part* II: *Extensions and applications*, Addison-Wesley, Reading, MA, 1981.
5. C. Brezinski, *Procedures for estimating the error in Padé approximation*, Math. Comp. **53** (1965), 639–648.
6. O. P. Bruno and F. Reitich, *Solution of a boundary value problem for Helmholtz equation via variation of the boundary into the complex domain*, Proc. Roy. Soc. Edinburgh Sect. A **122** (1992), 317–340.
7. ——, *Numerical solution of diffraction problems: a method of variation of boundaries*, J. Opt. Soc. Amer. A **10** (1993), 1168–1175.
8. S. Cabay and D. Choi, *Algebraic computations of scaled Padé fractions*, SIAM J. Comput. **15** (1986), 243–270.
9. A. Edrei, *Sur les déterminants récurrents et les singularités d'une fonction donnée par son développement de Taylor*, Compositio Math. **7** (1939), 20–88.
10. G. E. Forsythe and C. B. Moler, *Computer solution of linear algebraic systems*, Prentice-Hall, Englewood Cliffs, NJ, 1967.
11. W. Gautschi, *Construction of Gauss-Christoffel quadrature formulas*, Math. Comp. **22** (1968), 251–270.
12. ——, *On generating orthogonal polynomials*, SIAM J. Sci. Statist. Comput. **3** (1982), 289–317.
13. P. Graves-Morris, *The numerical calculation of Padé approximants*, Lecture Notes in Math., vol. 765 (L. Wuytack, ed.), Springer-Verlag, Berlin and New York, 1979, pp. 231–245.
14. S. Gustafson, *Convergence accerleration on a general class of power series*, Computing **21** (1978), 53–69.
15. ——, *On stable calculation of linear functionals*, Math. Comp. **33** (1979), 694–704.
16. C. Isenberg, *Moment calculations in lattice dynamics. I. fcc lattice with nearest-neighbor interactions*, Phys. Rev. **132** (1963), 2427–2433.
17. ——, *Expansion of the vibrational spectrum at low frequencies*, Phys. Rev. **150** (1966), 712–719.
18. Y. L. Luke, *Mathematical functions and their approximations*, Academic Press, New York, 1977.
19. ——, *Algorithms for the computation of mathematical functions*, Academic Press, New York, 1977.
20. ——, *Computations of coefficients in the polynomials of Padé approximations by solving systems of linear equations*, J. Comput. Appl. Math. **6** (1980), 213–218.
21. L. R. Mead and N. Papanicolaou, *Maximum entropy in the problem of moments*, J. Math. Phys. **25** (1984), 2404–2417.
22. P. M. Morse and H. Feshbach, *Methods of theoretical physics*, Vol. 2, McGraw-Hill, New York, 1953.
23. R. E. Scraton, *A note on the summation of divergent power series*, Proc. Cambridge Philos. Soc. **66** (1969), 109–114.

24. ———, *The practical use of the Euler transformation*, BIT **29** (1989), 356–360.

25. J. M. Taylor, *The condition of gram matrices and related problems*, Proc. Roy. Soc. Edinburgh **25** (1978), 45–56.

SCHOOL OF MATHEMATICS, GEORGIA INSTITUTE OF TECHNOLOGY, ATLANTA, GEORGIA 30332-0160

E-mail address: bruno@math.gatech.edu

DEPARTMENT OF MATHEMATICS, CARNEGIE MELLON UNIVERSITY, PITTSBURGH, PENNSYLVANIA 15213-3890

E-mail address: reitich@andrew.cmu.edu