BULL. AUSTRAL. MATH. SOC. 41A30
VOL. 10 (1974), 263-270.

Padé approximation and orthogonal polynomials

G.D. Allen, C.K. Chui, W.R. Madych, F.J. Narcowich, and P.W. Smith

By using a variational method, we study the structure of the Padé table for a formal power series. For series of Stieltjes, this method is employed to study the relations of the Padé approximants with orthogonal polynomials and gaussian quadrature formulas. Hence, we can study convergence, precise locations of poles and zeros, monotonicity, and so on, of these approximants. Our methods have nothing to do with determinant theory and the theory of continued fractions which were used extensively in the past.

1. Introduction

Consider the formal power series

(1) $$f(z) = a_0 + a_1 z + a_2 z^2 + \dots .$$

The $[N, M]$ Padé approximant of f is defined as the unique rational function $P_{N,M}(z)/Q_{N,M}(z)$ where $P_{N,M}$ and $Q_{N,M}$ are polynomials of degree no greater than M and N respectively and satisfy

$$f(z)Q_{N,M}(z) - P_{N,M}(z) = O\left(z^{N+M+1}\right)$$

as $z \to 0$; (*cf.* Wall [19]). The collection of all Padé approximants form a doubly infinite array of rational functions:

Received 22 November 1973.

$$\begin{matrix} [0,0] & [0,1] & [0,2] & \cdots \\ [1,0] & [1,1] & [1,2] & \cdots \\ \vdots & \vdots & \vdots & \cdots \end{matrix}$$

These rational functions were first studied by Frobenius [5], who systematically investigated their properties and relations. It was Padé [12] who, in his thesis, investigated when approximants of continued fractions of a function f fit into the above table, now called the Padé table. Active interest in this subject seems to have ceased about 1900. In about 1960, theoretical physicists and chemists discovered that this approximation method assisted with high speed computing machines proved to be very useful in solving problems in scattering theory, turbulence, field theory, and a host of other applications. The modern fundamental paper in this subject was written in 1965 by Baker [2]. However, the known methods he uses (such as determinant theory and continued fractions) are very cumbersome to mathematicians.

It is the intent of this paper to streamline some of these ideas and in particular to render the subject of series of Stieltjes elementary. Namely, we do not use determinant theory and continued fractions but instead use the simple concept of orthogonality.

Series of Stieltjes are examples of (1) where the series comes from the formal expansion of the integral

$$\int_0^\infty \frac{1}{1-tz}\, d\phi(t) \ , \tag{2}$$

where $d\phi$ is a positive measure with infinite support and the coefficients a_n are the n-th moments with respect to $d\phi$. That is,

$$a_n = \int_0^\infty t^n d\phi(t) \ . \tag{3}$$

The subject of these series was developed in conjunction with the theory of continued fractions by Stieltjes [15, 16]. It is with integrals of the type (2) and their continued fractions that the subject of orthogonal polynomials was originated. (For reference, see Szegö [17].) Some of the first papers were due to Tchebyshev [18], Heine [9], and Markoff [10].

That is, associated with respect to the measure $d\phi$ there is an orthonormal set of polynomials, with which the continued fraction of (2) is computed. The connections between these polynomials and the classical moment problem were developed by Hamburger [6], [7], [8], and Riesz [13], [14]. Much of this can be found in the recent book of Akhieser [1].

2. Main results

We relate the $[N, M]$ Padé approximant to the solution of the so-called Schwinger variational principle (*cf.* [11]) as follows: associated with the formal power series (1), consider the formal series

$$(4) \qquad F(c_0, c_1, \ldots) = \sum_{k=0}^{\infty} 2c_k a_k - \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} c_k c_j (a_{k+j} - z a_{k+j+1}) .$$

Formally we compute the variation of F and set it equal to zero. We observe the solution, that is, the stationary point, of this form is at

$$(c_0, c_1, \ldots) = (1, z, z^2, \ldots)$$

and so, formally, $F(1, z, z^2, \ldots) = f(z)$. If we "approximate" (4) by its partial sums

$$(5) \qquad F_N(c_0, \ldots, c_{N-1}) = \sum_{k=0}^{N-1} 2c_k a_k - \sum_{k=0}^{N-1} \sum_{j=0}^{N-1} c_k c_j (a_{k+j} - z a_{k+j+1})$$

when $N \geq 1$, it is clear that $(c_0, \ldots, c_{N-1})$ is a solution of the system

$$(6) \qquad \frac{\partial}{\partial c_j} F_N(c_0, \ldots, c_{N-1}) = 0 ,$$

$j = 0, \ldots, N-1$, if the column n-vector $\hat{c}_{0,N} = (c_0, \ldots, c_{N-1})^T$ satisfies the following matrix equation for each z :

$$(7) \qquad K_{0,N} \hat{c}_{0,N} = \hat{a}_{0,N} .$$

Here, and throughout the rest of the paper, we use the notation $\hat{a}_{M,N} = (a_M, \ldots, a_{M+N-1})^T$ and

$$(8) \qquad K_{M,N} = \begin{bmatrix} a_M - a_{M+1}z & a_{M+1} - a_{M+2}z & \cdots & a_{M+N-1} - a_{M+N}z \\ a_{M+1} - a_{M+2}z & a_{M+2} - a_{M+3}z & \cdots & a_{M+N} - a_{M+N+1}z \\ & & \cdots & \\ a_{M+N-1} - a_{M+N}z & a_{M+N} - a_{M+N+1}z & \cdots & a_{M+2N-2} - a_{M-2N-1}z \end{bmatrix},$$

where $M = 0, \pm 1, \ldots$ and we set $a_j = 0$ for $j < 0$. We obtain the following

LEMMA 1. *There exists a* $\hat{c}_{0,N}$ *satisfying* (7).

We remark that $\hat{c}_{0,N}$ may not be unique but the matrix product $\hat{a}^T_{0,N}\hat{c}_{0,N}$ is unique, independent of solution $\hat{c}_{0,N}$ of (7), and in fact it is the $[N, N-1]$ Padé approximant of the formal series f. We include this in the following more general theorem.

THEOREM 1. *The Padé approximants of the formal power series are uniquely given by*

$$(9) \qquad [N, M](f)(z) = \sum_{j=0}^{M-N} a_j z^j + z^{M-N+1} a^T_{M-N+1,N} K^g_{M-N+1,N} a_{M-N+1,N}$$

where $K^g_{M-N+1,N}$ *denotes a generalized inverse of the matrix* $K_{M-N+1,N}$

Note that $K^g_{M-N+1,N}\hat{a}_{M-N+1,N}$ is the analogue of $\hat{c}_{0,N}$ for $M = N - 1$. When $M - N < 0$, the sum in (9) is considered to be zero. For generalized inverses of matrices, see [4]. In the case when $K_{M-N+1,N}$ is non-singular, this formula reduces to the Baker-Nuttall Compact Formula [3]. Using the above formula (9) we have the Padé table explicitly for any formal power series (1).

We now restrict our attention to series of Stieltjes defined by (2) and (3). From equation (7), we arrive at the following system of equations:

$$\int_0^\infty P_N(t, z) t^k d\phi(t) = 0 ,$$

where $k = 0, \ldots, N-1$, where $P_N(t, z)$ is given by

$$P_N(t,\ z) = c_0 - 1 + \sum_{j=1}^{N-1} \left(c_j - c_{j-1}z\right)t^j - c_{N-1}zt^N .$$

Let $L_k(t)$ be the orthonormal polynomials with positive leading coefficients with respect to $d\phi$. Then P_N is a constant multiple of L_N. From this we arrive at the following

THEOREM 2. *The* $[N,\ N-1]$ *Padé approximant of the series of Stieltjes given by* (1), (2), *and* (3) *is*

(10) $$[N,\ N-1](f)(z) = \frac{1}{L_N\left(z^{-1}\right)} \int_0^\infty \frac{L_N\left(z^{-1}\right) - L_N(t)}{1-zt}\, d\phi(t) ,$$

and the error $f(z) - [N,\ N-1](f)(z)$ *is*

(11) $$\frac{1}{L_N\left(z^{-1}\right)} \int_0^\infty \frac{L_N(t)}{1-zt}\, d\phi(t) .$$

COROLLARY 1. *For* $j \geq -1$,

(12) $$[N,\ N+j](f)(z) = \sum_{k=0}^{j} a_k z^k + \frac{1}{L_{j,N}\left(z^{-1}\right)} \int_0^\infty \frac{L_{j,N}\left(z^{-1}\right) - L_{j,N}(t)}{1-zt}\, d\phi_j(t)$$

where $\{L_{j,N}\}$ *is the orthonormal set of polynomials with positive coefficients with respect to the measure* $t^{j+1}d\phi(t)$.

We note that $L_{-1,N} = L_N$, and when $j = -1$ the sum in (12) is interpreted as zero.

COROLLARY 2. *For* $j < -1$,

(13) $$[N,\ N+j](f)(z) = z^{j+1}[N,\ N-1](F)(z) ,$$

where $F(z) = z^{-(j+1)}f(z)$.

Thus, the entire Padé table for series of Stieltjes can be determined explicitly in terms of orthogonal polynomials. Consequently, we know the precise location (and hence interlacing properties) of the poles and zeros of the Padé approximants. The formula (10) can be modified as a gaussian quadrature formula as in the following

COROLLARY 3. *For* $j \geq -1$, *the* $[N, N+j]$ *Padé approximants are given by*

$$(14) \qquad [N, N+j](f)(z) = \sum_{k=0}^{j} a_k z^k + z^{j+1} \sum_{k=1}^{N} \frac{\alpha_{j,N,k}}{1-zx_{j,N,k}} ,$$

where $x_{j,N,k}$, $k = 1, \ldots, N$, *are the zeros of the orthonormal polynomial* $L_{j,N}$ *and* $\alpha_{j,N,k}$ *are the corresponding Cotes-Christoffel numbers.*

Since the Cotes-Christoffel numbers $\alpha_{j,N,k}$ are positive it is easy to see that for each fixed $j \geq -1$ the sequence $\{[N, N+j](f)\}$, $N = 1, 2, \ldots$, is a normal family in the domain $D_j = C - R_j$, where C denotes the complex plane and R_j denotes the reciprocal of the support of the measure $t^{j+1}d\phi(t)$. In the theorem to follow, we impose a mild condition on the measures $t^{j+1}d\phi(t)$, namely, $a_{j,m} = O\left((2m+1)!\, R^{2m}\right)$, $R > 0$, and

$$a_{j,m} = \int_0^\infty t^{m+j+1} d\phi(t) .$$

Of course, $a_{-1,m} = a_m$.

THEOREM 3. *For each* $j \geq -1$, *the sequence* $\{[N, N+j](f)\}$ *converges uniformly on each compact subset of* D_j *to the function* f *in* (2).

In Baker [2], it is only shown that the limit exists, and we have shown that indeed the limit is the function f . It might be of interest to point out that our method gives elementary proofs to the following inequalities. For each $x \leq 0$,

$$(15) \qquad [N, N-1](f)(x) < f(x) < [N, N](f)(x)$$

and

$$(16) \qquad [1, 0](f)(x) < [2, 1](f)(x) < \ldots < [N, N-1](f)(x) < \ldots .$$

Note that the inequalities are strict, compared with the results in Baker [2].

In conclusion, we remark that there are many directions of

https://doi.org/10.1017/S0004972700040879 Published online by Cambridge University Press

generalization of these techniques. The details, related results, and generalizations will appear elsewhere.

References

[1] N.I. Akhiezer, *The classical moment problem and some related questions in analysis* (translated by N. Kemmer. Hafner, New York; Oliver & Boyd, Edinburgh and London, 1965).

[2] George A. Baker, Jr, "The theory and application of the Pade approximant method", *Adv. Theor. Phys.* 1 (1965), 1-58.

[3] George A. Baker, Jr and John L. Gammel (editors), *The Padé approximant in theoretical physics* (Mathematics in Science and Engineering, 71. Academic Press, New York and London, 1970).

[4] Thomas L. Boullion, Patrick L. Odell, *Generalized inverse matrices* (Wiley-Interscience, New York, London, Sydney, Toronto, 1971).

[5] G. Frobenius, "Ueber Relationen zwischen den Näherungsbrüchen von Potenzreihen", *J. reine angew. Math.* 90 (1881), 1-17. (See also: Ferdinand Georg Frobenius, *Gesammelte Abhandlungen*, Band II, 47-63 (Springer-Verlag, Berlin, Heidelberg, New York, 1968).)

[6] Hans Hamburger, "Beiträge zur Konvergenztheorie der Stieltjesschen Kettenbrüche", *Math. Z.* 4 (1919), 186-222.

[7] Hans Hamburger, "Über eine Erweiterung des Stieltjesschen Momentenproblems. I", *Math. Ann.* 81 (1920), 235-319.

[8] Hans Hamburger, "Über eine Erweiterung des Stieltjesschen Momentenproblems. II, III", *Math. Ann.* 82 (1921), 120-164, 168-187.

[9] E. Heine, *Handbuch der Kugelfunctionen:* Erster Band, *Theorie der Kugelfunctionen und der verwandten Functionen*; Zweiter Band, *Anwendungen der Kugelfunctionen und der verwandten Functionen* (zweite umbearbeitete und vermehrte Auflage. G. Reimer, Berlin, 1878, 1881).

[10] A.A. Markoff, *Differenzenrechnung* (B.G. Teubner, Leipzig, 1896).

[11] J. Nuttall, "The connection of Padé approximants with stationary variational principles and the convergence of certain Padé approximants", *The Padé approximant in theoretical physics*, 219-230 (Mathematics in Science and Engineering, 71. Academic Press, New York and London, 1970); see [3].

[12] H. Padé, "Sur la représentation approchée d'une fonction par der fractions rationelles", Thèse, *Ann. de l'École Norm.* (3) 9 (1892), Suppl. 3-93.

[13] M. Riesz, "Sur le problème des moments. I, II", *Ark. f. Mat., Astron. Fys.* 16 (1922): Nr. 12, 23 pp; Nr. 19, 21 pp.

[14] M. Riesz, "Sur le problème des moments, III", *Ark. f. Mat., Astron. Fys.* 17 (1923), Nr. 16.

[15] T.-J. Stieltjes, "Recherches sur les fractions continues", *Ann. Fac. Sci., Toulouse* 8 (1894), 1-122.

[16] T.-J. Stieltjes, "Recherches sur les fractions continues", *Ann. Fac. Sci., Toulouse* 9 (1895), 1-47.

[17] Gabor Szegö, *Orthogonal polynomials* (Amer. Math. Soc. Colloquium Publications, 23. Amer. Math. Soc., New York City, 1939; revised edition, 1959).

[18] P.L. Tchebyshev, *Oeuvres*. Tomes I, II (Commissionaires Acad. Impér. Sci., St Petersburg, 1899-1907; reprinted Chelsea Publishing Co., New York, 1962).

[19] H.S. Wall, *Analytic theory of continued fractions* (Van Nostrand, New York, Toronto, London, 1948).

Department of Mathematics,
Texas A&M University,
College of Science,
Texas,
USA.