ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 4, Number 2, Spring 1974

CERTAIN INVARIANCE AND CONVERGENCE PROPERTIES OF THE PADÉ APPROXIMANT*

GEORGE A. BAKER, JR.†

We define the Padé approximant, invented by Jacobi [12], to the formal power series $f(x)$, by the equations

$$[L/M] = \frac{P_L(x)}{Q_M(x)} \tag{1}$$

$$Q_M(x)f(x) - P_L(x) = O(x^{M+L+1}) \tag{2}$$

$$Q_M(0) = 1 \tag{3}$$

where P_L and Q_M are polynomials of degree L and M respectively. This definition differs from the classical one of Frobenius [9] and Padé [14], in the use of (3). Under our definition the Padé approximants do not always exist, but an infinite number on each row, column, and diagonal of the Padé table always do exist [4].

The diagonal, $L = M$, Padé approximants satisfy the following invariance theorem,

THEOREM (INVARIANCE). *If $P_M(x)/Q_M(x)$ is the $[M/M]$ Padé approximant to $f(x)$, and $C + Df(0) \neq 0$, then*

$$\frac{A + B\left[P_M\left(\frac{\alpha y}{1+\beta y}\right) \Big/ Q_M\left(\frac{\alpha y}{1+\beta y}\right)\right]}{C + D\left[P_M\left(\frac{\alpha y}{1+\beta y}\right) \Big/ Q_M\left(\frac{\alpha y}{1+\beta y}\right)\right]} \tag{4}$$

is the $[M/M]$ Padé approximant to

$$\{A + Bf[\alpha y/(1+\beta y)]\}/\{C + Df[\alpha y/(1+\beta y)]\}. \tag{5}$$

The proof of this theorem is easily constructed by multiplying numerator and denominator by $(1+\beta y)^M Q_M(\alpha y/(1+\beta y))$. This operation reduces form (4) to the ratio of two polynomials of degree M, and thus the invariance theorem can be made to follow from the uniqueness theorem:

Received by the editors February 8, 1973.

*Work supported in part by the U.S.A.E.C. and in part by ARPA through the Materials Science Center of Cornell University.

†On sabbatical leave from Brookhaven to Cornell 1971–1972.

THEOREM (UNIQUENESS). *When it exists, the* $[L/M]$ *Padé approximant to any formal power series is unique.*

The proof of this theorem consists of assuming that there are two $[L/M]$ Padé approximants and showing, by means of (2), that they are the same.

Next we will consider the geometrical significance of the invariance theorem. We seek a measure of the "distance" between two Padé's which is likewise invariant. Since the Padé approximant is invariant under the linear fractional group, let us remind ourselves of some of its properties. If

$$w = T(z) = \frac{Bz + A}{Dz + C}, \tag{6}$$

then to insure that (6) is not degenerate so that all z map into a single w, we specify that (B, A) not be simply proportional to (D, C). This condition is most conveniently imposed as

$$BC - AD = 1 \tag{7}$$

The law of composition of two successive transformations is given by

$$\begin{gathered} w = T_2(T_1(z)) = T_3(z) \\ \begin{pmatrix} B_3 & A_3 \\ D_3 & C_3 \end{pmatrix} = \begin{pmatrix} B_2 & A_2 \\ D_2 & C_2 \end{pmatrix} \begin{pmatrix} B_1 & A_1 \\ D_1 & C_1 \end{pmatrix} \end{gathered} \tag{8}$$

where ordinary matrix multiplication is implied, and (7) implies that we have the subgroup of two-by-two matrices with unit determinant. Now any complex matrix can be factored in the form

$$T = U_1 D U_2 \tag{9}$$

where U_1 and U_2 are unitary and D is diagonal. However, refering to (6) and (8), we see that a diagonal matrix corresponds to

$$w = \frac{Bz + 0}{0 + C} = \frac{B}{C} z \tag{10}$$

or an uninteresting multiplication by a constant factor. Thus from our present point of view it is not unreasonable to confine our attention to the unitary subgroup of 2×2 matrices, i.e.

$$T = \begin{pmatrix} b & a \\ -a^* & b^* \end{pmatrix}. \tag{11}$$

Let us introduce Riemann's spherical representation of the complex numbers. This representation follows by imagining a unit sphere

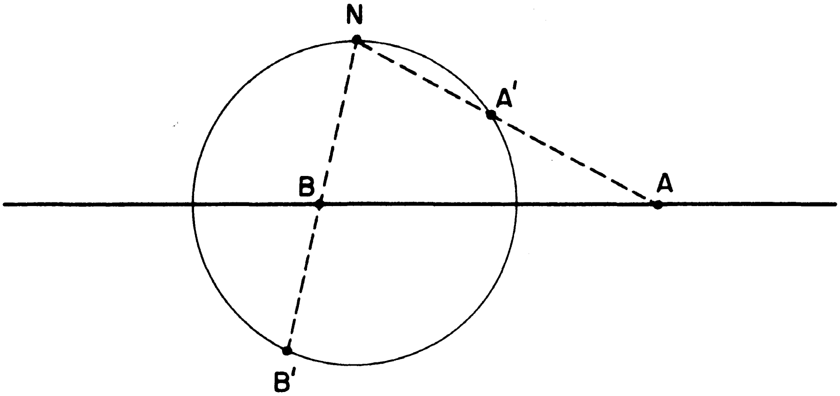

FIGURE 1. A cross-section of the Riemann sphere and the complex plane. The points A' and B' are projections on the sphere of points A and B in the plane using vertex N.

centered at the origin with its equatorial plane the intersection with the complex plane. If we now project a line from the north pole of the sphere to any point in the complex plane, we intersect the sphere in one and only one point. The north pole corresponds to the point at infinity. This intersection (see Fig. 1) is Riemann's spherical representation. It is not hard to show that 2×2 unitary transformations (11) correspond exactly to *rotations* of the Riemann sphere. Thus any measurement on the sphere will be invariant under the unitary group of transformations and a satisfactory distance between two Padé approximants. We pick the chord length, which is easily calculated as

$$D^2(z,w) = \frac{4|z - w|^2}{|1 + z^*w|^2 + |z - w|^2} \tag{12}$$

and introduce the

DEFINITION. A sequence of complex numbers w_n is said to *converge on the sphere*, if for any $\epsilon > 0$, there exists an N such that

$$D^2(w_n, w_m) \leqq \epsilon^2 \tag{13}$$

for all $n,m \geqq N$.

This definition allows sequences which tend to infinity to be treated on the same basis as sequences tending to any other limit point. For, we can rewrite (12) as

$$D^2(z,w) = \frac{4|z^{-1} - w^{-1}|^2}{|1 + z^{-1*}w^{-1}|^2 + |z^{-1} - w^{-1}|^2} \tag{14}$$

which converts the problem of $z \to \infty$ into $z^{-1} \to 0$. We can now treat convergence of Padé approximants at a pole in a much more convenient manner.

Next we will review, consolidate and extend the known theorems on the convergence of vertical (degree of the denominator fixed) sequences in the Padé table. These are due to Montesuss [**13**], Wilson [**18**], [**19**], and Baker [4], [5]. They treat "smooth" series and we will describe them generally rather than state them in detail. The underlying feature is, if there is an unambiguous location for each pole to converge to, then the Padé approximants converge on the sphere, inside the radius of convergence of the Taylor series to

$$B(z) = f(z) \prod_{i=1}^{n} (1 - z/z_i)^{m_i} \tag{15}$$

where the z_i are the n locations of convergence of the poles and m_i are the multiplicities of the converging poles. The theorems then make these statements precise. The simplest case is where there is a finite number of poles inside a circle of radius R. With no assumption on the behavior outside or on $|z| = R$, one can prove that the $[L/M]$, $M = \Sigma\, m_i$ converges on the sphere uniformly in $|z| \leqq \rho < R$ for any such ρ. When the number of poles in the Padé approximants exceeds the number of simple poles closer to the origin than the closest non-polar singularity, the theorems involve "smoothness" assumptions on the coefficients. These conditions are of the sort obeyed by the series expansions for algebraic, or logarithmic type singularities. Specifically, we assume, over a finite range of n near L after the pole contributions are removed, that

$$b_n \approx \Gamma^{-n}\beta(L) \left[\sum_{j=0}^{2\mu-1} \alpha_j(L) \left(\frac{n}{L} - 1\right)^j + 0\left(\left(\frac{n}{L} - 1\right)^{2\mu}\right) \right] \tag{16}$$

for the $[L/M + \mu]$ Padé approximants and that

$$\det \begin{vmatrix} \alpha_0 & \cdots & (\mu-1)!\alpha_{\mu-1} \\ \vdots & & \vdots \\ (\mu-1)!\alpha_{\mu-1} & \cdots & (2\mu-2)!\alpha_{2\mu-2} \end{vmatrix} \neq 0 \tag{17}$$

in the limit as $L \to \infty$. Finally, there are results for smooth series for entire functions where the form of (16) also includes a factor of $(n!)^{-\theta}$, $\theta > 0$. In this case, condition (17) is not required. In these cases the "extra" poles converge to $z = \Gamma$ for form (16) or to (ΓL^{θ}) for the case of a "smooth" entire function.

In the case of several competing non-polar, singularities the "extra" poles are distributed in such a way as to represent form (16) with minimum error as L goes to infinity using the rule that 1 pole can represent a constant, 2 poles a quadratic polynomial in $(n/L - 1)$, 3 poles a quartic, and so on. When this distribution is unique, that vertical sequence of Padé approximants converges. Otherwise, examples show that the whole sequence need not.

Since they satisfy an invariance property, both in their argument and in their value, one would expect that the diagonal or near diagonal $[L/L + J]$ sequences would be more powerful methods of summing power series than horizontal or vertical sequences. Indeed many examples show that this expectation is so. It is however very hard to find useful general criteria satisfied by either the functions themselves or their series coefficients to imply the convergence of diagonal sequences of Padé approximants. In the case where the continued fraction representation of the functions is known, as for the continued fraction of Gauss, then as continued fractions are just limits of stairstep sequences in the Padé table, the known theorems (dealt with in another lecture), involving the coefficients of the continued fractions are very satisfactory. Otherwise the theorems involve conditions on the Padé approximants themselves (see, for example, the reviews [1], [3].) These conditions have to do mainly with the location of the poles ([7], [17], [3]), or the stronger condition of boundedness [1]. As a sample of this type of theorem, we quote a special case of Chisholm's theorem

THEOREM (CHISHOLM). *Let $f(z)$ be analytic in $|z| \leqq R$ and $f(0) \neq 0$). Let $[L/M]$ be an infinite sequence of Padé approximants to $f(z)$ with $L \to \infty$ and $M \to \infty$ in any way, such that they contain no poles in $|z| \leqq R$. Then the sequence $[L/M]$ converges uniformly to $f(z)$ in the region*

$$|z| \leqq (\sqrt{2} - 1)R - \epsilon \tag{18}$$

for any $\epsilon > 0$.

The proof of this theorem involves constructing bounds for the numerator of $[L/M] - f(z)$. There are a number of general theorems on convergence in measure, but, as they are being treated in a separate lecture, I will not discuss them here.

For one class of function very complete results are available. Those are the series of Stieltjes. Let

$$f(z) = \sum_{j=0}^{\infty} f_j(-z)^j = \int_0^{\infty} \frac{d\phi(u)}{1+uz} \tag{19}$$

where $d\phi \geqq 0$. It is necessary and sufficient for $f(z)$ to be of form (19) that

$$C(L/M) = det \begin{vmatrix} f_{L-M+1} & \cdot & \cdot & \cdot & f_L \\ \vdots & & \cdot & & \vdots \\ & & & \cdot & \\ f_L & \cdot & \cdot & \cdot & f_{L+M-1} \end{vmatrix} \geqq 0 \tag{20}$$

for all $L \geqq M-1$. We will hear from other speakers of some of the examples of functions of this structure which appear in physics.

For series of Stieltjes one can prove that all the poles lie on the negative real axis and have positive residue. This result serves effectively to control the location and strength of the singularities. For z real and positive, the Padé approximants obey $(J \geqq -1)$

$$\begin{aligned} &(-1)^{1+J}\{[M+1+J/M+1] - [M+J/M]\} \geqq 0, \\ &(-1)^{1+J}\{[M+J/M] - [M+J+1/M-1]\} \geqq 0, \\ &\qquad [M/M] \geqq f(z) \geqq [M-1/M], \\ &\qquad [M/M]' \geqq f'(z) \geqq [M-1/M]'. \end{aligned} \tag{21}$$

For z complex, one can give lens shaped error inclusion regions for the Padé approximants. [Pfluger and Henrici [15], Gargantini and Henrici [11], Common [8], Baker [2].] If the radius of convergence is known all these estimates can be sharpened up.

Even if the radius of convergence is zero, it can be proved that the diagonal sequences of Padé approximants sum series of Stieltjes. If the series diverges less rapidly than about $(f_P) \approx (2P)!$ so that the determinate case holds, then every diagonal sequence tends to the same limit.

In order to fill partially the gap between what can be proved about diagonal sequences and what seems to hold true for examples, Baker, Gammel and Wills [6] proposed the following conjecture (slightly modernized):

PADE CONJECTURE. *If $P(z)$ is a power series which is regular for $|z| \leqq 1$, except for m poles within this circle and except for $z = +1$,* *at which point the function is assumed continuous when only points* $|z| \leqq 1$ *are considered, then at least a subsequence of the* $[M/M]$ *Padé approximants converges on the sphere, uniformly for* $|z| \leqq 1$.

The effect of this conjecture can be greatly extended through the use of the invariance theorem. There are examples ([10], [16]) which show that the whole sequence of Padé approximants need not converge in the pointwise sense.

Finally I will quote two theorems of a different character. These theorems concern the existence of convergent, vertical subsequences of Padé approximants [4]. They represent, I think, just a beginning for theorems of this character.

THEOREM. *Let* $f(z)$ *be an entire function, then there exists an infinite subsequence of* $[L/1]$ *Padé approximants which converges uniformly in any closed bounded region of the complex plane to the function defined by the power series.*

The proof follows simply from a few observations. First,

$$[L/1] = \sum_{j=0}^{L-1} f_j z^j + \frac{f_L z^L}{1 - f_{L+1}/f_L z} \quad . \tag{22}$$

Since

$$f_j = f_0 \cdot \prod_{k=1}^{j} (f_k/f_{k-1}), \tag{23}$$

and f_j goes to zero faster than any geometric progression, as $f(z)$ is an entire function, one can prove that there must exist an infinite sequence of ratios which go to zero. Thus the pole term, by the convergence of $f(z)$ can be made negligible.

THEOREM. *Let* $f(z)$ *be an entire function, which satisfies*

$$f(z) = \sum_{j=0}^{\infty} f_j z^j, \quad |f_j| \leqq K/[(j)!]^{\theta} \tag{24}$$

for some $K > 0$, $\theta > 1$. *Then there exists an infinite subsequence of* $[L/2]$ *Padé approximants which converge uniformly in any closed bounded region of the complex plane to the entire function defined by the power series.*

The proof of this theorem is more difficult. It is based on the fact that the absolute value of the determinant of the coefficients $C(L/2)$ of the Padé equations going uniformly to zero with L faster than f_L^2/L is inconsistent with restriction (24).

These theorems indicate, I think, that while convergence in measure is the best that can be proved for the sequence as a whole, there is much more to be said about subsequences. The theorems on "smooth" series are reasonably complete for vertical and, by duality, horizonal sequences, the only major gap is for additional subsequence type theorems for the cases where the sequences as a whole do not converge. These theorems seem likely to be provable, although there may be a fair number of cases to consider.

As far as the Padé conjecture is concerned, all the possible counter examples I know of have been checked and found not to be. Therefore, the Padé conjecture remains today, unproved and uncontradicted.

References

1. G. A. Baker, Jr., *The Theory and Application of the Padé Approximant Method,* Advances Theoret. Phys. 1 (1965) 1-58 (ed. K. A. Brueckner) (Academic Press, New York).
2. ——, *Best Error Bounds for Padé Approximants to Convergent Series of Stieltjes,* J. Math. Phys. **10** (1969), 814-820.
3. ——, *The Padé Approximants Method and Some Related Generalizations* in *The Padé Approximant in Theoretical Physics,* (ed. G. A. Baker, Jr. and J. L. Gammel), Academic Press, New York, 1970.
4. ——, *The Existence and Convergence of Subsequences of Padé Approximants,* J. Math. Anal. and Appl. 43 (1973), 498-528.
5. ——, *Essentials of Padé Approximants,* In production, Academic Press, (1974).
6. G. A. Baker, Jr., J. L. Gammel and J. G. Wills, *An Investigation of the Applicability of the Padé Approximant Method,* J. Math. Analysis and Application **2** (1961), 405-418.
7. J. S. R. Chisholm, *Approximation by Sequences of Padé Approximants in Regions of Meromorphy,* J. Math. Phys. **7** (1966), 39-44.
8. A. K. Common, *Padé Approximants and Bounds to Series of Stieltjes,* J. Math. Phys. **9** (1968), 32-38.
9. G. Frobenius, *Uber Relationen zwischen den Näherungsbrüchen von Potenzreihen,* Jour. für Math. (Crelle) **90** (1881) 1-17.
10. J. L. Gammel, Private communication, (1970).
11. J. Gargantini and P. Henrici, Math. Comput. **21**, (1967) 18-29.
12. C. G. J. Jacobi, *Über die Darstellung einer Reihe gegebner Werthe durch einer gebrochene rationale Function,* Jour. für Reine u. Angewandte Math. (Crelle) **30** (1846) 127-156.
13. R. Montessus, *Sur les Fractions Continues Algébriques,* Bull. Soc. Math. France **30** (1902) 28-36.
14. H. Padé, *Sur la Représentation approchée d'une Fonction par des Fractions Rationnelles,* Ann. Sci. École Norm. Sup. (3) **9** (1892), 1-93.
15. A. Pfluger and P. Henrici in *Contemporary Problems in the Theory of Analytic Functions,* Int. Conf. Theory of Analytic Functions, Erevan 1965, (ed. M. A. Lavrentiev) Izdat. "Nauka" Moscow, 1966.
16. H. Wallin, *The Convergence of Padé Approximants and the Size of the Power Series Coefficients,* To be published, Applicable Analysis, 1972.

17. J. L. Walsh, *Padé Approximants as Limit of Rational Functions of Best Approximation,* J. Math. Mech. **13** (1964) 305-312.
18. R. Wilson, *Divergent Continued Fractions and Polar Singularities,* Proc. London Math. Soc. **26** (1927) 159–168; II, **27** (1928) 497–512; III, **28** (1928) 128–144.
19. R. Wilson, *Divergent Continued Fractions and Non-Polar Singularities,* Proc. London Math. Soc. **30** (1930), 38–57.

CORNELL UNIVERSITY, ITHACA, NEW YORK 14850
BROOKHAVEN NATIONAL LABORATORY, UPTON, NEW YORK 11973