An Introduction to the Convergence Theory of Padé Approximants

E. B. SAFF
University of South Florida, Tampa, U.S.A.

1 Introduction

Our purpose is to introduce the reader to the convergence theory of interpolating rational functions known as *Padé Approximants.* Though the subject originated around the turn of the century, it has only been in the last 15 years that significant applications to physics (particularly critical point phenomena) has rekindled interest in this topic. We make no attempt in this brief note to give a comprehensive survey, but we do hope to provide the flavour of old and new results. The (hopefully) interested reader can consult the more extensive bibliographies contained in the references for further details.

2 Definitions and Notation

Let π_m denote the collection of all polynomials of degree at most m. A rational function $R(z)$ is said to be of *type* (n, ν), where n, ν are non-negative integers, if

$$R \in \pi_{n,\nu} \equiv \{p/q : p \in \pi_n, q \in \pi_\nu, q \not\equiv 0\}.$$

Given (n, ν) and a formal power series

$$f(z) = \sum_{k=0}^{\infty} a_k z^k, \tag{2.1}$$

we select polynomials $P_{n\nu} \in \pi_n$, $Q_{n\nu} \in \pi_\nu (Q_{n\nu} \not\equiv 0)$ such that

$$Q_{n\nu}(z)f(z) - P_{n\nu}(z) = \sum_{k=n+\nu+1}^{\infty} C_k z^k \equiv \mathcal{O}(z^{n+\nu+1}). \tag{2.2}$$

If f is $(n + \nu)$-times differentiable at $z = 0$, then (2.2) can be equivalently written as

$$(Q_{n\nu} f - P_{n\nu})^{(k)}|_{z=0} = 0, \quad k = 0, 1, \ldots, n + \nu. \tag{2.3}$$

The equations (2.2) and (2.3) are known as the *Padé equations.* Since (2.2) represents $n + \nu + 1$ equations in $n + \nu + 2$ unknowns (the coefficients of $P_{n\nu}$ and $Q_{n\nu}$), it follows that this homogeneous system has a non-trivial solution (necessarily with $Q_{n\nu} \not\equiv 0$). Although $P_{n\nu}$ and $Q_{n\nu}$ are not unique, the ratio $P_{n\nu}/Q_{n\nu}$ (in lowest terms) does, however, determine a unique rational function of type (n, ν).

Definition 2.1. The *Padé approximant* (PA) of type (n, ν) to f is given by $R_{n\nu} = P_{n\nu}/Q_{n\nu}$, where $P_{n\nu} \in \pi_n$ and $Q_{n\nu}(\not\equiv 0) \in \pi_\nu$ are polynomials which satisfy (2.2).

In particular, for $\nu = 0$ the PA reduces to a section of the power series, i.e.,

$$R_{n0}(z) = \sum_{k=0}^{n} a_k z^k.$$

We caution the reader that (2.2) does not in general imply that

$$f(z) - R_{n\nu}(z) = \mathcal{O}(z^{n+\nu+1}).$$

Indeed if $Q_{n\nu}(0) = 0$ (so that $P_{n\nu}(0) = 0$) the contact at the origin for the PA may be diminished. However, in the class $\pi_{n\nu}$, the PA $R_{n\nu}$ is that unique rational function with maximum possible contact with f at the origin.

The Padé numerators and denominators are rich in algebraic properties such as the 3-term recurrence relations found by Frobenius. For a detailed discussion of these classical properties as well as more recent ones we refer the reader to [2, **10**, **15**]. Here we pause only to give a determinant representation for $Q_{n\nu}(z)$ which illustrates the important role played by Hankel determinants:

$$(2.4) \qquad Q_{n\nu}(z) = \begin{vmatrix} 1 & z & \cdots & z^{\nu} \\ a_{n+1} & a_n & \cdots & a_{n-\nu+1} \\ a_{n+2} & a_{n+1} & \cdots & a_{n-\nu+2} \\ \vdots & \vdots & & \vdots \\ a_{n+\nu} & a_{n+\nu-1} & \cdots & a_n \end{vmatrix},$$

provided this determinant does not vanish identically.

The PAs for (2.1) are typically displayed in a doubly infinite array known as the *Padé Table:*

$$\begin{array}{|llll} \hline R_{00} & R_{10} & R_{20} & \cdots \\ R_{01} & R_{11} & R_{21} & \cdots \\ R_{02} & R_{12} & R_{22} & \cdots \\ \vdots & \vdots & \vdots & \end{array}$$

Here the first row represents the polynomial sections, the 2nd row consists of PAs with at most one pole, the third row PAs with at most two poles, etc. Of special interest are the diagonal files of the table, for these represent continued fraction expansions. Indeed if

$$(2.5) \qquad f(z) = d_0 + \cfrac{d_1 z}{1 + \cfrac{d_2 z}{1 + \cdots}} = \sum_{k=0}^{\infty} a_k z^k,$$

the successive truncations d_0, $d_0 + d_1 z$, $d_0 + d_1 z/(1 + d_2 z), \ldots$ are rational functions of the respective types

$$(0, 0), (1, 0), (1, 1), (2, 1), \ldots, (n, n), (n + 1, n), \ldots.$$

It is easy to see by induction that for each n, the coefficients a_n in (2.5) depend only on $d_0, d_1, \ldots, d_n$ and hence the Maclaurin series of the truncations agrees with that of f to the maximum number of terms. Consequently these truncations are PAs which form the staircase entries of the table, i.e., the entries along the main diagonal and first super diagonal.

3 Convergence Theorems

The convergence question for PAs can be stated as follows. Given the power series (2.1), what can be said about the convergence of sequences extracted from the table, such as rows, columns, diagonals, and, in general, any sequence $R_{n\nu}$ for which $n + \nu \to \infty$?

One of the earliest results concerning row convergence for meromorphic functions $f(z)$ is due to Montessus de Ballore [5]. His proof relied mainly on Hadamard's theory for locating polar singularities–particularly the properties of Hankel determinants. In 1970 this author gave an elementary proof [17] of the Montessus theorem, which applies as well to general interpolating schemes. Here we present an even more abbreviated version of the proof which H. Wallin attributes to H. Shapiro.

THEOREM 3.1. *Let f be meromorphic with precisely ν poles (counting multiplicities) in the disk D: $|z| < r$, with f analytic at $z = 0$. Then, as $n \to \infty$, the sequence of PAs $\{R_{n\nu}\}_{n=0}^{\infty}$ converges to $f(z)$ uniformly on any compact set in $D - \{\nu$ poles of $f\}$. Furthermore, as $n \to \infty$, the poles of the $R_{n\nu}$ tend, respectively, to the ν poles of f in D.*

Proof. Set $g(z) \equiv \prod_{k=1}^{\nu} (z - \alpha_k)$, where $\alpha_1, \alpha_2, \ldots, \alpha_\nu$ are the ν poles of f in D. In the Padé equations

$$(3.1) \qquad Q_{n\nu} f - P_{n\nu} = \mathcal{O}(z^{n+\nu+1}), \quad n = 0, 1, \ldots,$$

we normalize each $Q_{n\nu}$ so that its coefficient of largest modulus is 1 (in the case when there is more than one largest coefficient we choose the one with smallest subscript). Multiplying (3.1) by $g(z)$ we obtain

$$(3.2) \qquad Q_{n\nu}gf - gP_{n\nu} = \mathcal{O}(z^{n+\nu+1}), \quad n = 0, 1, \ldots .$$

Note that $h_n \equiv Q_{n\nu}gf$ is analytic in D and that $gP_{n\nu} \in \pi_{n+\nu}$. Hence $gP_{n\nu}$ is the $(n+\nu)$th partial sum of the Maclaurin series for $h_n(z)$ and we can write, by Hermite's formula,

$$(3.3) \quad h_n(z) - g(z)P_{n\nu}(z) = \frac{1}{2\pi i}\int_C \frac{z^{n+\nu+1}h_n(t)\,\mathrm{d}t}{t^{n+\nu+1}(t-z)}, \; z \text{ inside } C,$$

where C is any circle centred at the origin and contained in D. As the functions $h_n(t)$ are uniformly bounded on C, it follows from (3.3) that

$$(3.4) \qquad h_n(z) - g(z)P_{n\nu}(z) \to 0,$$

uniformly (in fact, geometrically) on any compact set in D.

Now let Q_∞ be any limit function of the normal family $\{Q_{n\nu}\}_{n=0}^{\infty}$, so that Q_∞ is necessarily a polynomial of degree at most ν. From (3.4) we see that $Q_\infty gf$ is the uniform limit of analytic functions which vanish at each of the points α_i. Consequently, as the analytic function gf is non-zero in these points, we have $g|Q_\infty$. Thus from the form of Q_∞ it follows that $Q_\infty(z) \equiv \lambda g(z)$ for some constant $\lambda \neq 0$. As Q_∞ is an arbitrary limit function, the zeros of the sequence $\{Q_{n\nu}\}_{n=0}^{\infty}$ therefore tend to the ν poles of f in D, and from (3.4) we deduce that $R_{n\nu}(z) \to f(z)$ locally uniformly in $D - \{\nu$ poles of $f\}$. ∎

Another early result, due to H. Padé **[15]**, concerns the Padé table for the exponential function.

THEOREM 3.2. *For any sequence* (n_i, ν_i), $i = 1, 2, \ldots$, *with* $n_i + \nu_i \to \infty$, *the poles of the PAs to* e^z *all tend to* ∞, *and*

$$\lim_{i\to\infty} R_{n_i\nu_i}(z) = \mathrm{e}^z$$

uniformly on any compact set in $\mathbb{C}$.

While this result is perhaps more an example than a theorem, the PAs to e^z have several important applications. The stability properties of certain numerical schemes for solving differential equations can be rephrased in terms of these approximants, and the Padé numerators are related in a simple way to Bessel polynomials [12] which have applications to number theory.

The Padé table for e^z has a particularly elegant behaviour with regard to the distribution of poles and zeros of the PAs. In 1924 Szegö characterized the asymptotic behaviour of the zeros of the partial sums

$$R_{n0}(z) = \sum_{0}^{n} z^k/k!$$

Specifically, he showed that, as $n \to \infty$, the zeros of the normalized polynomials $R_{n0}(nz)$ accumulate at every point on the level curve $|ze^{1-z}| = 1$, $|z| \leqslant 1$. In [19] a generalization of this result is given for sequences of PAs to e^z for which the ratio ν/n ($\nu = \nu(n)$) tends to a finite limit σ as $n \to \infty$. In such a case the zeros as well as the poles of the normalized PAs $R_{n\nu}((n+\nu)z)$ all accumulate on an explicitly given level curve $\Gamma_\sigma : |\omega_\sigma(z)| = 1$, $|z| \leqslant 1$, with zeros tending to a subarc (with endpoints on $|z| = 1$) of Γ_σ and poles accumulating on the complementary subarc.

An important extension of Theorem 3.2 to the class of functions generated by totally positive sequences was obtained by Arms and Edrei [1]. Such functions were shown by Edrei to be of the form

$$(3.5) \qquad f(z) = a_0 e^{\gamma z} \prod_{j=1}^{\infty} \left(\frac{1 + \alpha_j z}{1 - \beta_j z} \right), \quad a_0, \gamma, \alpha_j, \beta_j \geqslant 0,$$

$$\sum (\alpha_j + \beta_j) < \infty.$$

For these functions we have

THEOREM 3.3. *Let $f(z)$ be of the form (3.5). If (n_i, ν_i), $i = 1, 2, \ldots,$ is such that $\nu_i/n_i \to \sigma$, and the Padé denominators are normalized so that $Q_{n\nu}(0) = 1$, then*

$$(3.6) \qquad \lim_{i \to \infty} P_{n_i, \nu_i}(z) = a_0 \exp\,[\gamma z/(1 + \sigma)] \prod_{j=1}^{\infty} (1 + \alpha_j z),$$

$$\lim_{i\to\infty} Q_{n_i, \nu_i}(z) = \exp\,[-\gamma\sigma z/(1+\sigma)] \prod_{j=1}^{\infty} (1-\beta_j z);$$

the convergence being uniform on any compact subset of $\mathbb{C}$.

In the proof of Theorem 3.3 the positivity of the Hankel determinants plays a crucial role.

A second class of functions for which convergence properties of PAs are known is the series of Stieltjes (cf. [15]).

Definition 3.1. A power series $f(z) = \sum_{j=0}^{\infty} c_j(-z)^j$ is said to be a series of Stieltjes if

$$c_j = \int_0^{\infty} t^j \mathrm{d}\phi(t), \quad j = 0, 1, 2. \ldots, \tag{3.7}$$

where ϕ is bounded, non-decreasing, left continuous, and assumes infinitely many values.

For this class we have the following result which essentially follows from Stieltjes' work on the moment problem (cf. [2, 15, 20]).

THEOREM 3.4. *For* $n \geqslant 1$, $\mu \geqslant -1$, *each PA* $R_{n+\mu,n}$ *for a series of Stieltjes has all its poles on the negative real axis. Furthermore, each diagonal file* $\{R_{n+\mu,n}\}_{n=1}^{\infty}$, $\mu \geqslant -1$, *converges uniformly on compact sets of the slit plane* $\mathbb{C} - (-\infty, 0]$.

Notice that Theorem 3.4 does not assert that the diagonal files all have the same limit. This will be the case only when the Stieltjes moment problem is *determinant,* i.e., when ϕ is the unique distribution whose moments generate the sequence c_j. The proof of Theorem 3.4 again relies on the special sign pattern for the Hankel determinants.

4 All Is Not Roses

While convergence results exist for PAs to other special functions, (e.g., [2, 4, 8]) there are "nice" functions $f(z)$ for which the poles of the PAs misbehave and destroy convergence. To see this, consider a sequence of non-zero coefficients a_n for which

$$(4.1) \qquad 0 = \lim_{n \to \infty} |a_n|^{1/n} < \limsup_{n \to \infty} |a_{n+1}/a_n|,$$

so that $f(z) = \sum_0^\infty a_n z^n$ is entire. As shown by Perron [15], it is not difficult to construct such a sequence for which $\{a_n/a_{n+1}\}_{n=0}^\infty$ has limit points that are dense in the plane. But, by (2.4), a_n/a_{n+1} is the zero of the Padé denominator $Q_{n1}(z)$, and so the second row of the Padé table for this entire function has poles everywhere dense in the plane. Even more startling is the following result due to H. Wallin [21].

THEOREM 4.1. *There exists an entire function f such that the sequence of diagonal PAs $\{R_{nn}(z)\}_{n=0}^\infty$ for f is unbounded at every point in the plane except $z = 0$.*

In the light of this result, two possible directions suggest themselves:

(1) consider a weaker form of convergence, such as convergence in measure or in capacity; or

(2) try to extract *subsequences* which do have the desired uniform convergence properties.

In the first direction there is a result due to Pommerenke [16] who generalized earlier work of Nuttall. Roughly speaking it asserts that near-diagonal PAs will be inaccurate approximations to f only on sets of small capacity (transfinite diameter). For the precise statement of this result it is more convenient to deal with PAs which interpolate an analytic function at ∞ instead of at $z = 0$. Simply replacing z by $1/z$ in the discussion of Section 2 defines these approximants, which we denote by $\tilde{R}_{n\nu}$.

THEOREM 4.2. *Let f be analytic on a domain containing ∞ in the extended complex plane whose complement is a set of zero capacity. Let $r > 1$ and $\lambda > 1$. Then for each $\epsilon > 0$ and $\eta > 0$, there exists an n_0 such that $|\tilde{R}_{n\nu}(z) - f(z)| < \epsilon^n$ for all n, ν for which $n > n_0$ and $1/\lambda \leq n/\nu \leq \lambda$ and for all z such that $|z| \leq r$ and $z \notin E_{n\nu}$, where* cap $E_{n\nu} < \eta$.

In the case when bounds are known for the function $f(z)$, FitzGerald [9] showed how Theorem 4.2 can be used to generate "truncated" PAs whose areal means converge normally to $f(z)$.

In the second direction, Baker and Graves–Morris [3] (extending earlier work of Beardon) have proved that if $f(z)$ is analytic in a neighbourhood of the origin, then it is possible to select subsequences of the 2nd and 3rd rows of the Padé table which will converge to $f(z)$. Recently Edrei [7] has, for the case of entire functions of finite order, extended these results by proving the existence of uniformly convergent subsequences for certain other rows (depending on the order of f) of the table. There still remains open, however, the following.

Conjecture (Baker, Gammel, Wills). Let f be analytic on $|z| \leqslant 1$ except for m poles (different from $z = 0$) in $D: |z| < 1$ and except for $z = 1$ at which continuity holds only when points in $|z| \leqslant 1$ are considered. Then there exists a subsequence of the diagonal PAs $\{R_{nn}\}$ to f which converges uniformly to f on compact sets of $D - \{m \text{ poles of } f\}$.

5 Converse Theorems

Recently A. A. Gončar has suggested the study of converse theorems of the following form: given a function f known merely to be analytic at $z = 0$, and given information about the behaviour of the poles of a sequence of its PAs, what can be said about analytic continuations of f and the location of their singularities? Such results would generalize a theorem of Fabry [6, p. 377] who proved that if

$$\lim_{n \to \infty} a_n / a_{n+1} = z_0$$

exists, then z_0 is a singular point of $f(z)$. A converse to the Theorem 3.1 was announced by Kovačeva in [14]. She proved:

THEOREM 5.1. *If f is analytic at $z = 0$ and the PAs in a fixed row, say $\{R_{n,\nu}\}_{n=0}^{\infty}$, have poles which tend at geometric rates to ν finite points $\alpha_1, \alpha_2, \ldots, \alpha_\nu$ (not necessarily distinct, but different from $z = 0$), then there exists a disk $D_r: |z| < r$ containing the α_i in which f is analytic except for actual poles at the α_i.*

In a similar spirit we state the following recent result:

THEOREM 5.2 (Karlsson, Saff [13]). *If $f(z)$ is analytic at $z = 0$ and the poles of the diagonal sequence of PAs $\{R_{nn}\}_{n=0}^{\infty}$ for f all tend to infinity, then f can be extended to an entire function.*

References

1. R. J. Arms and A. Edrei, The Padé tables and continued fractions generated by totally positive sequences, *in* "Mathematical Essays dedicated to A. J. Macintyre", pp. 1–21. Ohio University Press, Athens, Ohio, 1970.
2. G. A. Baker, Jr, "Essentials of Padé Approximants". Academic Press, New York, 1975.
3. G. A. Baker, Jr and P. Graves–Morris, Convergence of rows of the Padé table, *J. Math. Anal. Appl.* **57** (1977) 323–339.
4. M. G. de Bruin, Some classes of Padé tables whose upper halves are normal, *Nieuw Archief voor Wiskunde* **25** (1977) 148–160.
5. R. de Montessus de Ballore, Sur les fractions continues algébriques, *Bull. Soc. Math. France* **30** (1902) 28–36.
6. P. Dienes, "The Taylor Series". Dover, New York, 1957.
7. A. Edrei, The Padé tables of entire functions, *J. Approx. Theory,* **28** (1980) 54–82.
8. A. Edrei, The Padé table of functions having a finite number of essential singularities, *Pacific J. Math.* **56** (1975) 429–453.
9. C. FitzGerald, Confirming the accuracy of Padé table approximants, *in* "Padé and rational approximation: Theory and applications". (E. B. Saff and R. S. Varga, eds) pp. 51–60. Academic Press, New York, 1976.
10. W. B. Gragg, The Padé table and its relation to certain algorithms of numerical analysis, *SIAM Rev.* **14** (1972) 1–62.
11. P. Graves-Morris, "Padé approximants". Institute of Physics, London, 1973 (Lectures delivered at a summer school at the University of Kent, UK, 1972).
12. E. Grosswald, "Bessel polynomials", Springer Lecture Notes in Mathematics No. 698. Springer-Verlag, Berlin, 1978.
13. J. Karlsson and E. B. Saff, Singularities of analytic functions determined by the behaviour of poles of interpolating rational functions (to appear).
14. R. K. Kovačeva, On rational approximations to meromorphic functions, *Dokl. Akad. Nauk. SSSR* **241** (1978) 540–543.
15. O. Perron, "Die Lehre von den Kettenbrüchen". Chelsea, New York, 1957.
16. Ch. Pommerenke, Padé approximants and convergence in capacity, *J. Math. Anal. Appl.* **41** (1973) 775–780.
17. E. B. Saff, An extension of Montessus de Ballore's Theorem on the convergence of interpolating rational functions, *J. Approx. Theory* **6** (1972) 63–67.
18. E. B. Saff and R. S. Varga, "Padé and rational approximation: Theory and applications". Academic Press, New York, 1977.
19. E. B. Saff and R. S. Varga, On the zeros and poles of Padé approximants to e^z III, *Numer. Math.* **30** (1979) 241–266.
20. H. S. Wall, "Analytic theory of continued fractions". Van Nostrand-Reinhold, Princeton, NJ, 1948.
21. H. Wallin, On the convergence theory of Padé approximants *in* "Linear operators and approximation" (Proceedings of the Conference Oberwolfach, 1971), International Series Numer. Mathematics, Vol. 20, pp. 461–469. Birkhauser, Basel, 1972.