ELSEVIER
Available at www.ElsevierMathematics.com
POWERED BY SCIENCE @ DIRECT®
Journal of Computational and Applied Mathematics 161 (2003) 371–391
JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS
www.elsevier.com/locate/cam

Some structural properties of two counter-examples to the Baker–Gammel–Wills conjecture

George A. Baker Jr.

Theoretical Division, Los Alamos National Laboratory, University of California, Los Alamos, NM 87545, USA

Received 13 November 2002; received in revised form 28 May 2003

Abstract

I improve the counter-example of Lubinsky, and show that the counter-example of Buslaev is also relevant to the original form of the Baker–Gammel–Wills conjecture. I notice that these counter-examples have only a single spurious pole and that a patchwork of just two subsequences of diagonal Padé approximants provides uniform convergence in compact subsets of $|z| < 1$. I find that both counter-examples can be characterized by the observation that they are associated with bounded J-matrices. I prove a number of results for the convergence of diagonal Padé approximants to functions which have bounded J-matrices.

MSC: 41A21; 49A15; 30B70; 46c07

Keywords: Padé approximants; Baker–Gammel–Wills conjecture; Spurious poles

1. Introduction and Summary

The study of Padé approximants [6] remained at a rather low level of activity until the second half of the last century. A Padé approximant to a function $f(z)$ which is defined by a power series at the origin is defined by the equations,

$$Q_M(z)f(z) - P_L(z) = o(z^{L+M}), \quad Q_M(0) = 1, \tag{1.1}$$

where $Q_M(z)$ and $P_L(z)$ are polynomials of degrees at most M and L, respectively. The notation for such a Padé approximant is

$$[L/M] = \frac{P_L(z)}{Q_M(z)}. \tag{1.2}$$

E-mail address: gbj@viking.lanl.gov (G.A. Baker, Jr.).

0377-0427/$ - see front matter
doi:10.1016/j.cam.2003.05.005

I think that the principal reason for the low level of activity was that the computation, aside from those results which could be obtained analytically, involved the numerical solution of sets of frequently ill-conditioned, linear equations, a rather tedious task. With the advent of even primitive digital computers, the tedium was handled by machine computation and enough approximants could be computed to begin a numerical exploration of their convergence properties. Much to the surprise of the early explorers, the convergence was very much better than was expected, based on experience with Taylor series. It was found that the diagonal sequences $M \approx L$ tended to be the best converged, for the number of terms of the series used. However, it was soon noticed that in many cases there were "defects" in the approximants. By a defect, is meant that there are a pole and a zero close together. They only effect the value of the approximant over a very small region of the complex plane. In the cases studied, these defects only occur occasionally and these features lead Baker, Gammel, and Wills to propose the conjecture [5].

Conjecture 1.1. *If $P(z)$ is a power series representing a function which is regular for $|z| \leqslant 1$, except for m poles within this circle and except for $z = +1$, at which point the function is assumed continuous when only points $|z| \leqslant 1$ are considered, then at least a subsequence of the $[N/N]$ Padé approximants converge uniformly to the function (as N tends to infinity) in the domain formed by removing the interiors of small circles with centers at these poles.*

Over time, many different versions of this conjecture were proposed and studied. I quote a second version given by Baker [2].

Conjecture 1.2. *If $P(z)$ is a power series which is meromorphic in $|z| < 1$ and continuous on the sphere [See Definition (3.1) below] in $|z| \leqslant 1$, then at least a subsequence of the $[M/M]$ Padé approximants is equicontinuous on the sphere in $|z| \leqslant 1$.*

By Theorem 3.3, this conjecture implies that at least a subsequence of the $[M/M]$ Padé approximants converge uniformly on the sphere to $f(z)$.

A weaker version of this conjecture was proposed by Stahl [26].

Conjecture 1.3. *Let the function f be algebraic and meromorphic in the unit disc $\mathscr{D}$. Then there exists an infinite subsequence $N \in \mathscr{N}$ such that*

$$[n/n](z) \to f(z) \quad \text{as } n \to \infty, \ n \in \mathscr{N} \tag{1.3}$$

holds true locally uniformly for $z \in \mathscr{D} \setminus \{\text{poles of } f\}$.

From the point of view of workers who are trying to evaluate function values by means of Padé approximants, the sum and substance of these conjectures has been to interpret them to mean, "just disregard the approximants with defects and use the rest of them and you will be OK."

After 40 years of study by a number of workers, Lubinsky [19] produced a counter-example to Conjecture 1.2. Shortly thereafter, and apparently motivated by the work of Lubinsky, Buslaev [10,11] produced an algebraic counter-example to Conjectures 1.2 and 1.3. As I explain in

Section 2, his counter-example can straightforwardly be converted into a counter-example to Conjecture 1.1 as well.

In this paper I examine the structural properties of the counter examples. Both of them are constructed from functions which are particularly simply described by continued fractions. In particular they can both be cast into the form of associated continued fractions [18]. These continued fractions can be derived from Wall's J-matrix formulation [28], and both correspond to the J-matrices being bounded operators.

The counter-examples also have the property that there exist two fixed subsequences of the diagonal Padé approximant such that for any point z in $|z| \leqslant 1$, not a singular point, that one or the other of these two sequences converges strongly to the function value $f(z)$. In this paper I raise the question as to how general this type of property is.

In the Section 2, I discuss Buslaev's [10] counter-example. I locate all the branch points and poles of his function. I find that the contradictions to the conjectures occur due to the presence of a single rotating defect. I remind the reader of Nuttall's conjecture on the possible limitation of the number of spurious poles, and Gončar's theorem on the convergence of diagonal Padé approximants in regions where there are only a finite number of spurious poles.

In Section 3, I discuss Lubinsky's [19] counter-example. By discussing it in the context of the convergence on the Riemann sphere, I am able to show that a contradiction to Conjecture 1.2 occurs at a distance from the origin of less than $\frac{1}{3}$ rather than his original result of a distance of less than 0.46.

In Section 4, I discuss the Padé approximation to functions which correspond to bounded J-matrices. I show that the whole sequence converges if the J-matrix is a compact operator. If the J-matrix, which is always a symmetric, tri-diagonal infinite matrix, has a subsequence of off diagonal elements which tend to zero, then there exists a subsequence of diagonal Padé approximants which converge except at singular points. If no such subsequence of off diagonal elements exits, I prove, subject to certain assumptions on the number of poles and zeros of the approximants, uniform convergence of a patchwork of a finite number of subsequences of Padé approximants. A transfer-matrix type recursion relation is given for the Padé polynomials in terms of the coefficients of the associated continued fraction.

2. Buslaev's counter-example

Inspired by the work of Lubinsky [19], which I shall discuss in the next section, Buslaev [10] discovered an algebraic counter-example to the Padé (Baker–Gammel–Wills) conjecture [5]. His example is given by the periodic continued fraction (an associated continued fraction),

$$f(z) = \frac{z/3}{1-\omega^2 z} {}_{+} \frac{\omega z^2/9}{1-\omega z} {}_{+} \frac{\omega^2 z^2/9}{1-z} {}_{+} \frac{z^2/9}{1-\omega^2 z} {}_{+} \cdots, \tag{2.1}$$

where $\omega = -\frac{1}{2} + \sqrt{3}\mathrm{i}/2$ is a cube root of unity, and $\omega^2 = \omega^* = \omega^{-1}$. I denote the complex conjugate of ω by ω^*. By use of the standard methods [18] he is able to sum this continued fraction to give

$$f(z) = \frac{-27 + 6z^2 + 3(9+\omega)z^3 + \sqrt{81[3-(3+\omega)z^3]^2 + 4z^6}}{2z[9 + 9z + (9+\omega)z^2]}. \tag{2.2}$$

The analytical properties of this function are as follows. As there is a square root of a sixth-order polynomial, there are six square root type branch points. They are

$$z^3 = \frac{3}{3+\omega \pm 2i/9},$$

$$z_1 \approx 1.0227140 - 0.14084275i, \quad z_2 = \omega z_1, \quad z_3 = \omega^2 z_1,$$

$$z_4 \approx 1.0476385 - 0.088224985i, \quad z_5 = \omega z_4, \; z_6 = \omega^2 z_4. \tag{2.3}$$

The absolute values of these branch points are

$$|z_1| = |z_2| = |z_3| \approx 1.0323665, \quad |z_4| = |z_5| = |z_6| \approx 1.0513468, \tag{2.4}$$

so all the branch points are outside the unit circle. Of further interest are the zeros of the quadratic denominator polynomial. They are

$$r_1 \approx -0.55280095 - 0.82813567i, \quad |r_1| \approx 0.99568951,$$

$$r_2 \approx -0.49514431 + 0.93490595i, \quad |r_2| \approx 1.0579306. \tag{2.5}$$

We see that r_1 does lie in the unit circle, but direct evaluation of the numerator in (2.2) shows that this is an ordinary point and not a pole. However, r_2 is a pole, and $\omega r_2 \approx -0.56208014-0.89626050i$ is a zero, owing to cancellation in the numerator of (2.1). $\omega^2 r_2$ turns out to be an ordinary point. Thus, on the first Riemann sheet, $f(z)$ is holomorphic for $|z| \leqslant 1$.

On the second Riemann sheet, $f(z)$ analytically continues to

$$\tilde{f}(z) = \frac{-27+6z^2+3(9+\omega)z^3 - \sqrt{81[3-(3+\omega)z^3]^2+4z^6}}{2z[9+9z+(9+\omega)z^2]}. \tag{2.6}$$

In this case, there is the relation,

$$\tilde{f}(z) \equiv -\frac{9-9(1+\omega)z-(1-8\omega)z^2}{[9+9z+(9+\omega)z^2]f(z)}, \tag{2.7}$$

which can be derived by multiplying (2.2) by (2.6). It is important to notice that r_1 is a pole of $f(z)$ inside $|z| \leqslant 1$ on the second sheet.

The main result of Buslaev is that

$$[3n+j-1/3n+j-1](\omega^j r_1) = \tilde{f}(\omega^j r_1) \neq f(\omega^j r_1) \tag{2.8}$$

for all $n \geqslant 1$ and $j = 0$, 1, 2. This result provides a direct counter-example to Conjecture 1.2 and as $f(z)$ is algebraic, it also provides a counter-example to Conjecture 1.3. It is well known [6] that the diagonal Padé approximants are invariant under linear fractional transformations of the argument which leave the origin fixed. That is to say that the $[n/n](w)$ Padé approximant to $g[aw/(1+bw)]$ is the Padé approximant to $g(z)$ evaluated at $z = aw/(1+bw)$. In order to investigate Conjecture 1.1, I need $z_1 \mapsto 1$ in such a way that the maps of $z_2, \ldots, z_6$ all lie out side the unit circle while $r_1, \omega r_1$ and $\omega^2 r_1$ remain inside the unit circle. By the numerical results quoted above, the mapping $w = 1.001(z/z_1)/[1+0.001(z/z_1)]$ will accomplish this transformation. Thus I will satisfy the conditions of Conjecture 1.1 and hence I have derived from Buslaev's counter-example a counter-example to Conjecture 1.1 as well!

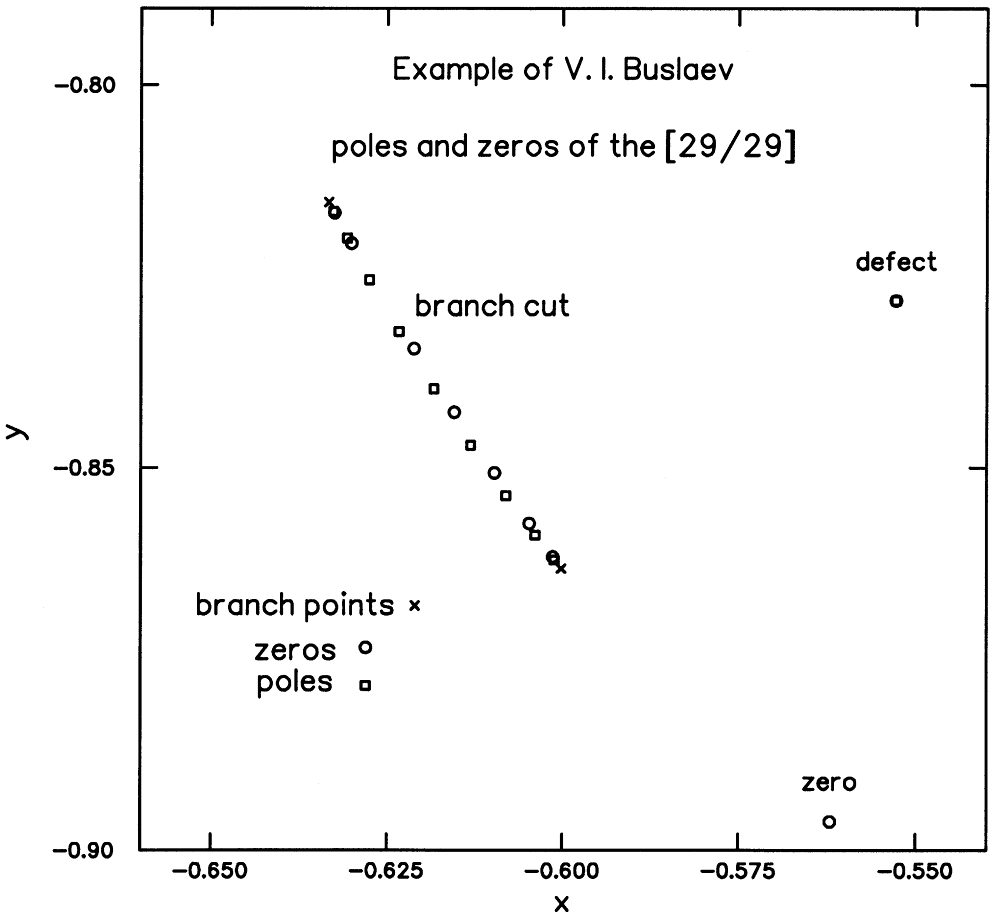

Fig. 1. A plot of the poles and zeros of the [29/29] Padé approximant to $f(z)$ from (2.2) in the region of one of the three branch cuts. The point labeled "defect" has both a pole and a zero which are at the same location to within plotting accuracy. This location corresponds to the point at which Buslaev found that there is convergence to the wrong value for the $[3n-1/3n-1]$ sequence. It is to be noticed that most of the poles and zeros lie neatly along the branch cut predicted by Stahl's convergence theorem.

The function $f(z)$ belongs to the class of functions to which Stahl's [27] convergence theorem applies. For this case, if I take the set $\mathscr{S}$ consisting of the three straight lines, $\overline{z_1^{-1}z_4^{-1}}$, $\overline{z_2^{-1}z_5^{-1}}$, $\overline{z_1^{-1}z_4^{-1}}$, in the $w = 1/z$ plane, then the diagonal Padé approximants converge uniformly on compact subsets of $\bar{\mathbb{C}}\backslash\mathscr{S}$ except for a set of capacity 0. Hence it must be that these "bad" points lie in defects. A defect necessarily has a very close pole-zero pair which causes a large fluctuation from the value of the approximated function. The occurrence of a pole is required as there is a theorem [6] which shows that if the approximants are uniformly bounded, then the Padé approximants must converge. By the just quoted theorem of Stahl this disruption must tend to a set of capacity 0 as $n \to \infty$. Thus there must also be a zero to cancel the effects of the pole and they must get arbitrarily close to each other as $n \to \infty$. I illustrate the behavior of a typical Padé approximant to $f(z)$ in Fig. 1.

There are some points worth mentioning. First, there is never more than one spurious pole. This result is in line with Nuttall's conjecture ([21, p. 305] as phrased by Stahl [26]) which states that

Conjecture 2.1. *Let the function f be algebraic and have no branch point at the origin. Then there exists a finite upper bound for the number of spurious poles* (*in the sense of total order*) *which each* $[n/n]$ (*diagonal*) *Padé approximant may have*, $n = 1, 2, \ldots$.

Stahl [26] adds the remark that in most cases the upper bound is equal to the genus of the Riemann surface of the function. In the case of Buslaev's example, the genus of the surface is 2 so this conjecture and Stahl's remark are valid in this case. The second point is that the defect moves around among three locations. Thus, by patching them together I find that at most two subsequences are sufficient to provide convergence uniformly in any compact subset of the unit disk for Buslaev's example.

In addition I mention a theorem of Gončar [15], which relates to the case where the number of poles is uniformly bounded for the whole sequence of diagonal Padé approximants.

Theorem 2.2. *Let $\mathscr{D}$ be a domain which satisfies the condition $\partial\mathscr{D} \subset \partial\tilde{\mathscr{D}}$ where $\tilde{\mathscr{D}}$ is the complement in the extended complex plane of the convex hull of $\mathscr{D}$, and ∂ denotes the boundary. Further let $\mathscr{E}$ be a relatively closed subset of $\mathscr{D}$ of capacity* 0. *Let $\Omega = \mathscr{D} \setminus \mathscr{E}$ be a domain containing the origin as an element, and suppose that the number of poles of the* $[n/n]$ *Padé approximant in any compact subset $\mathscr{K} \subset \Omega$ is uniformly bounded in n by $\kappa(\mathscr{K}) < \infty$. Further, if the* $[n/n]$ *Padé approximants converge to a holomorphic function in a neighborhood of the origin, then the diagonal Padé approximants converge uniformly on compact subsets to a function which is meromorphic in Ω and holomorphic at the origin, except for a possible set of Hausdorf measure* 0. *If $\kappa(\mathscr{K}) = 0$, then the convergence is uniform on compact subsets of Ω.*

3. Lubinsky's counter-example

Lubinsky [19] was the first to produce a counter-example to the Padé (Baker–Gammel–Wills) Conjecture 1.2 [2]. He has considered a special case of the Rogers–Ramanujan continued fraction. It is defined by the continued fraction (a regular C-fraction),

$$H_q(z) = 1 + \frac{c_1 z}{1} + \frac{c_2 z}{1} + \frac{c_3 z}{1} + \cdots, \tag{3.1}$$

where $c_j = q^j$. This function may be re-expressed in terms of the Rogers–Ramanujan function,

$$G_q(z) = \sum_{j=0}^{\infty} \frac{q^{j^2}}{(1-q)(1-q^2)\cdots(1-q^j)} z^j. \tag{3.2}$$

The formula is

$$H_q(z) = \frac{G_q(z)}{G_q(qz)}. \tag{3.3}$$

Since I am concerned with diagonal sequences of Padé approximants, it is worthwhile to convert (3.1) into the associated continued fraction. As the general relation is

$$F(z) = b_0 + \frac{a_1 z}{1} {}_{+} \frac{a_2 z}{1} {}_{+} \frac{a_3 z}{1} {}_{+} \cdots$$

$$= b_0 + \frac{a_1 z}{1 + a_2 z} {}_{-} \frac{a_2 a_3 z^2}{1 + (a_3 + a_4)z} {}_{-} \frac{a_4 a_5 z^2}{1 + (a_5 + a_6)z} {}_{-} \cdots, \tag{3.4}$$

we may rewrite (3.1) as

$$H_q(z) = 1 + \frac{qz}{1 + q^2 z} {}_{-} \frac{q^5 z^2}{1 + (1 + q^{-1})q^4 z} {}_{-} \frac{q^9 z^2}{1 + (1 + q^{-1})q^6 z} {}_{-} \cdots$$

$${}_{-} \frac{q^{4p+1} z^2}{1 + (1 + q^{-1})q^{2p+2} z} {}_{-} \cdots . \tag{3.5}$$

The specific example chosen by Lubinsky has the $|q| = 1$ where q is given by

$$q = \exp(2\pi \mathrm{i} \tau), \quad \tau = \frac{2}{99 + \sqrt{5}}. \tag{3.6}$$

The diagonal Padé approximants are invariant under linear fractional transformations of the value [6], that is to say, the diagonal Padé approximant to the linear fractional transformation of a function is the linear fractional transformation of the Padé approximant. It is worthwhile at this point to discuss the Riemann sphere. The idea is to construct a sphere whose equator is the unit circle of the complex plane. Then every point on the plane can be mapped in a one-to-one manner onto the sphere by connecting that point by a line to the north (upper) pole of the sphere. That way the origin is mapped into the south pole, the unit circle into itself, and the point at infinity into the north pole. The distance between two points on the sphere is just the length of the chord. In terms of the original points in the plane, the chordal metric is

$$D^2(z_1, z_2) \equiv \frac{4|z_1 - z_2|^2}{|1 + z_1^* z_2|^2 + |z_1 - z_2|^2} = \frac{4|z_1^{-1} - z_2^{-1}|^2}{|1 + (z_1^*)^{-1} z_2^{-1}|^2 + |z_1^{-1} - z_2^{-1}|^2}. \tag{3.7}$$

The maximum distance (between $z_1 = 0$ and $z_2 = \infty$, for example) is 2. The group of linear fractional transformations is equivalent to the group [2] of 2×2 matrices where the composition of two linear fractional transformations corresponds to standard matrix multiplication.

$$z' = \frac{Bz + A}{Dz + C} \quad \Leftrightarrow \quad \begin{pmatrix} B & A \\ D & C \end{pmatrix}. \tag{3.8}$$

Since the multiplication of A, B, C, D by a constant factor leaves z' unchanged, and since when the determinant of the matrix equals zero, $z' =$ a constant is an uninteresting case, I consider only transformations for which $BC - AD = 1$. These transformations can be decomposed into scale changes and unitary transformations (and compositions thereof). Any unitary transformation corresponds to a rotation of the sphere and vice versa. The relevant matrix is

$$\begin{pmatrix} b & a \\ -a^* & b* \end{pmatrix} \quad |a|^2 + |b|^2 = 1. \tag{3.9}$$

Each Padé approximant is clearly a meromorphic function in the whole complex plane (including the point at infinity). It is therefore continuous in terms of the chordal metric (3.7). An important concept is that of equicontinuity.

Definition 3.1. A sequence of functions $\{f_n(z)\}$ all defined in a closed region $\mathscr{R}$ is equicontinuous on the sphere on $\mathscr{R}$, if for each z_0 in $\mathscr{R}$ and each $\varepsilon > 0$ there exists a δ depending on z_0 and ε, but independent of n, such that

$$D(f_n(z), f_n(z_0)) < \varepsilon, \quad |z - z_0| < \delta, \tag{3.10}$$

where z is in $\mathscr{R}$ and D is as given in (3.7).

The Riemann sphere is a compact set so for each particular value of z there is always a subsequence Padé approximants which tends to a limit. However more than this is required to establish convergence to a function. The following relations between equicontinuity and convergence have been proven [6].

Theorem 3.2. *If $P_n(z)$ is any sequence of meromorphic functions which converges uniformly on the sphere in some closed region $\mathscr{R}$ to some limit $f(z)$, then the limit is a meromorphic function in the interior of $\mathscr{R}$ and continuous on the sphere in $\mathscr{R}$, and $\{P_n(z)\}$ is uniformly equicontinuous on the sphere in $\mathscr{R}$.*

Theorem 3.3. *If $P_n(z)$ is any infinite sequence of meromorphic functions which is uniformly equicontinuous on the sphere over a closed region $\mathscr{R}$, then at least a subsequence of the $P_n(z)$ converges uniformly on the sphere to a limit $f(z)$ continuous on the sphere in $\mathscr{R}$ and meromorphic in the interior of $\mathscr{R}$.*

Lubinsky [19] has proven,

Theorem 3.4. *Let q be given by (3.6). Then $H_q(z)$ is meromorphic in the unit ball and analytic at 0. There does not exist any subsequence of $\{[n/n]\}_{n=1}^{\infty}$ that converges uniformly in all compact subsets of*

$$\mathscr{A} := \{z \mid |z| < \Theta\}, \quad \Theta = 0.46, \tag{3.11}$$

omitting poles of $H_q(z)$.

There are several parts to the proof of this theorem. If g_j are the coefficients of the Maclaurin series expansion of a regular C-fraction and we define

$$\Delta_n = \det(g_{1+i+j})_{i,j=0}^{n-1}, \tag{3.12}$$

then, as is well known [28],

$$\Delta_n = c_1^n \prod_{j=1}^{n-1} (c_{2j}c_{2j+1})^{n-j}. \tag{3.13}$$

Since, for $H_q(z)$ $|c_j| = 1$ $\forall j$, we have at once that $|\Delta_n| = 1$ $\forall n$. There is a result of Polya ([22] or more accessible [14]) which gives an upper bound on the radius of meromorphy, σ. It is

$$\limsup_{n\to\infty} |\Delta_n|^{1/n^2} \leqslant \sigma^{-1}, \tag{3.14}$$

which gives us directly that $\sigma \leqslant 1$. Worpitzky's theorem [18], applied to (3.1) gives us a lower bound so that we have $\frac{1}{4} \leqslant \sigma \leqslant 1$. By an examination of the radius of convergence of the series of $G_q(z)$ (3.2) through the use a remarkable identity of Hardy and Littlewood [16] and the Cauchy–Hadamard formula for the radius of convergence, Lubinsky [19] has improved the lower bound on σ to unity for his choice of q. Thus follows the conclusion that $H_q(z)$ is meromorphic in $|z| < 1$. He has also proved that $H_q(z)$ has a natural boundary on the unit circle.

Having established that $H_q(z)$ is meromorphic in the unit circle, and thus a potential candidate for a counter-example to the Padé conjecture, we next need to show that there does not exist a subsequence of the $[n/n]$ Padé approximants which is free of any defects inside the unit circle. The cornerstone of this effort is Hirschhorn's [17] explicit result for the numerator and the denominator polynomials $P_n(z)$ and $Q_n(z)$, respectively, of the $[n/n]$ Padé approximants. They are

$$P_n(z) = \sum_{k=0}^{n} z^k q^{k^2} \begin{bmatrix} 2n+1-k \\ k \end{bmatrix},$$

$$Q_n(z) = \sum_{k=0}^{n} z^k q^{k(k+1)} \begin{bmatrix} 2n-k \\ k \end{bmatrix} \tag{3.15}$$

for irrational τ, where,

$$\begin{bmatrix} \alpha \\ l \end{bmatrix} = \frac{(1-q^\alpha)(1-q^{\alpha-1})\cdots(1-q^{\alpha-l+1})}{(1-q)(1-q^2)\cdots(1-q^l)}, \quad l \geqslant 0,\ \alpha \in \mathbb{C}. \tag{3.16}$$

First Lubinsky proved

Theorem 3.5. *Let q and τ be given by (3.6). There must exist a set $\mathscr{T}$ of integers such that for any $|\beta| = 1$,*

$$\lim_{n\to\infty, n\in\mathscr{T}} q^{2n} = \beta. \tag{3.17}$$

Then uniformly on compact subsets of $\{z \mid |z| < 1\}$

$$\lim_{\substack{n\to\infty \\ n\in\mathscr{T}}} P_n(z) = G_q[(\beta q z)^*]^* G_q(z), \tag{3.18}$$

$$\lim_{\substack{n\to\infty \\ n\in\mathscr{T}}} Q_n(z) = G_q[(\beta q z)^*]^* G_q(qz), \tag{3.19}$$

and

$$\lim_{\substack{n\to\infty \\ n\in\mathscr{T}}} \frac{H_q(z)-(P_n(z)/Q_n(z))}{z^{2n+1}q^{(n+1)(2n+1)}} = \frac{G_q(\beta q^2 z)}{G_q(qz)^2 G_q[(\beta qz)^*]^*}, \tag{3.20}$$

uniformly in compact subsets of $\{z \mid |z|<1\}$ *omitting zeros of* $G_q[(\beta qz)^*]^*$ *and* $G_q(qz)$.

From this theorem, we find immediately that

Corollary 3.6. *For* τ *irrational and* $\mathscr{T}$ *and* β *as given in Theorem* 3.5

$$\lim_{\substack{n\to\infty \\ n\in\mathscr{T}}} D(H_q(z),[n/n](z)) = 0 \tag{3.21}$$

uniformly on compact subsets of $\{z \mid |z|<1\} \setminus \{z \mid G_q[(\beta qz)^*]=0\}$.

That is to say, we obtain convergence on the sphere away from the defects made apparent by Lubinsky's results (3.18) and (3.19). Corollary 3.6, by using convergence on the sphere, gives us, for this case, a slight improvement over Gončar's general Theorem 2.2. That theorem applies here if we choose $D=\{z \mid |z|<1\}$ and proves the convergence of Lubinsky's counter-example, outside an exceptional set. Stahl's [27] theorem does not apply because of the natural boundary on the unit circle.

Lubinsky [19] further points out that the extra pole-zero pairs in the limit of large n are just rotations and reflections of the poles of $H_q(z)$. The poles and zeros, by (3.18) and (3.19), with β replaced by q^{2n} as is allowed in the large n limit, tend as $n\to\infty$ to the same location. Since $D(\infty,0)=2$ it is not possible to give an $\varepsilon>0$ and a $\delta>0$ at such a point. It cannot be that the pole and zero cancel identically because $H_q(z)$ is a regular C-fraction. This fact can be seen from the two-term Padé approximant identities [6],

$$Q_M^{(0)}(z)P_{M+1}^{(0)}(z) - Q_{M+1}^{(0)}(z)P_M^{(0)}(z) = z^{2M+1}\prod_{j=1}^{2M+1} c_j \neq 0, \tag{3.22}$$

where the superscript (0) is the degree of the numerator polynomial minus the degree of the denominator polynomial. The normalization here is defined by $Q_M^{(0)}(0)=1.0$. We have, by the converse of Theorem 3.2, that convergence must fail at such a point. Thus we have,

Corollary 3.7. *The entire sequence* $[n/n]$, *in the limit of large* n, *fails to be equicontinuous in any closed neighborhood of a point where* $G_q[(q^{2n+1}z)^*]=0$. *Hence there is no uniform convergence in any compact subset containing one of these points. In particular, we can exclude uniform convergence in compact subsets of* $\{z \mid |z|<r\}$ *where* r *is any number greater than the smallest* $|\hat{z}|$ *where* $G_q(\hat{z})=0$.

There is one special case of this corollary which is worthy of mention. It may be that, for a suitable choice of β, $(\beta qz)^*=qz$, where z is the location of the pole which is closest to the origin. In this case the spurious pole would converge to the same location as a real pole. However, the accompanying spurious zero, by (3.22), can neither be at the same point as the true pole of $H_q(z)$ nor at the same point as the spurious pole. In this case, it is the spurious zero, tending to the same location as two poles that causes the loss of equi-continuity, and hence the failure of uniform convergence on the sphere. The situation may be clearer if one thinks of $1/H_q(z)$ where there is a pole (and a zero) approaching an ordinary zero.

We will see directly that this corollary improves the value of Θ in Theorem 3.4 when we use convergence on the sphere and do not omit the poles of $H_q(z)$.

The final step in the proof of Theorem 3.4, is to demonstrate the existence of a $\hat{z}$ such that $|\hat{z}| < 1$. Since we have already seen that the radius of convergence of $G_q(z)$ is unity for our choice of q, we may proceed as Lubinsky did and find $\hat{z}$ numerically. The numerical results which I report were computed using the Brent multi-precision arithmetic package [9]. I have solved for the roots using Laguerre's method as expounded in Press et al. [24]. I have retained a minimum of 58 decimal places. In searching for the roots, I use a more stringent criterion than that given in that book. I should obtain an accuracy of about 12 decimal places, according to the authors. I find $\hat{z}_1 \approx -0.3305614689253 - 0.130714239903i$ which yields $|\hat{z}_1| \approx 0.33239517265$.

Hence by Corollary 3.7 we can improve Theorem 3.4 by the choice $\Theta = \frac{1}{3}$ instead of the original choice $\Theta = 0.46$.

There are, in fact, an infinite number of $\hat{z}_k$ interior to the unit circle, but only a finite number in $|z| \leqslant r < 1$. For example $\hat{z}_2 \approx -0.2837288258318 - 0.2929172277400i$ with $|\hat{z}_2| \approx 0.4078020952801$. It is interesting to see the structure of the poles and defects of the diagonal Padé approximants. I illustrate this behavior in Figs. 2–4. It appears that each $[n/n]$ approximant produces poles which are just rotations by a factor of q^{-2} of the poles found in the $[n-1/n-1]$ on the circles $|z| = |\hat{z}_i|$. This observation is in accord with Theorem 3.4.

Hence we can conclude that in addition to the subsequence $n \in \mathscr{T}$, that if we also use the subsequence $n+1$, $n \in \mathscr{T}$, then, as the "bad" points are rotated relative to the other sequence, we can patch the two subsequences together appropriately and thereby obtain uniform convergence in compact subsets of the unit disk, $|z| < 1$.

I remark that there are also a number of poles outside the unit circle. In the approximants I have computed so far I see five such poles which are circling at radii of about 1.27, 2.12, 4.18, 16.8, and 105. It may be that the Padé approximants will converge outside the natural boundary on the unit circle [13].

4. Bounded J-matrix

I observe that both Buslaev's [10] and Lubinsky's [19] counter-examples can be given as associated continued fractions of the general form

$$F(z) = B_0 + A_0 z \left(\frac{1}{1+B_1 z} \, _{-} \, \frac{A_1^2 z^2}{1+B_2 z} \, _{-} \, \frac{A_2^2 z^2}{1+B_3 z} \, _{-} \, \cdots \right). \tag{4.1}$$

It is of interest to investigate what can be said about this class of functions. I can associate Wall's J-matrix with this continued fraction. First let us introduce the elements $\boldsymbol{e}_j$ where $j = 1, 2, 3, \ldots$

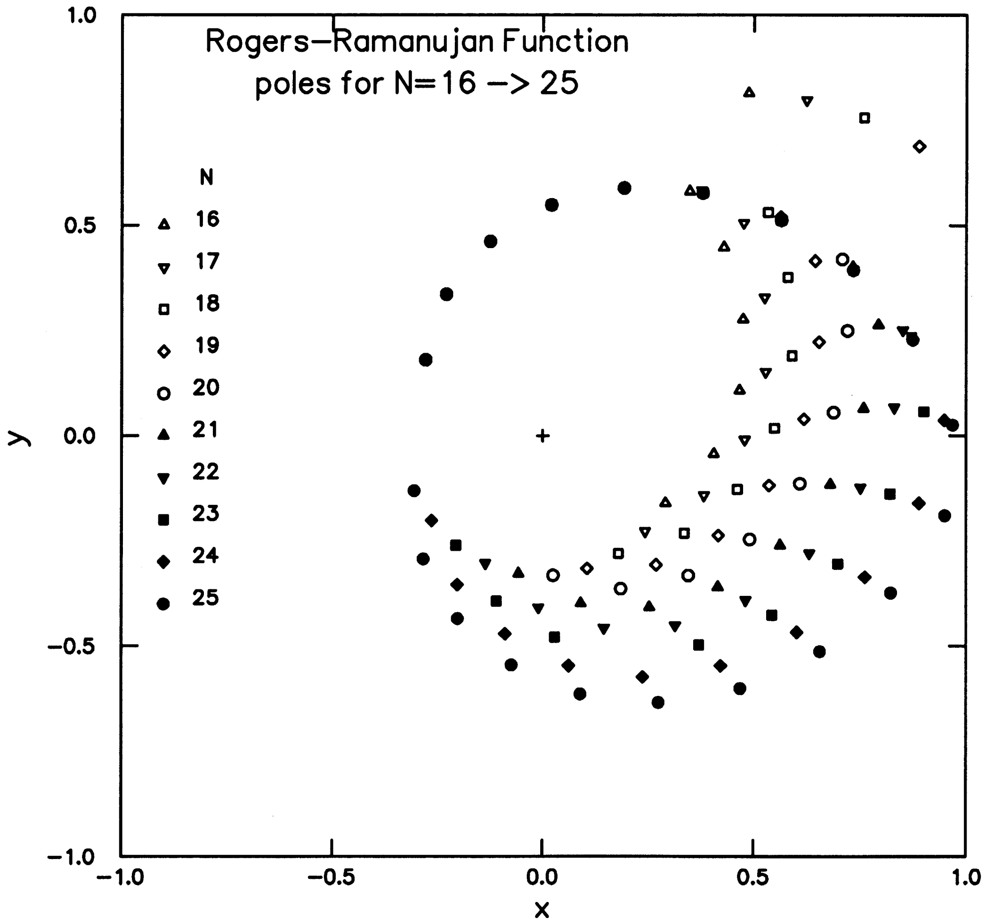

Fig. 2. A plot of the poles of the $[N/N]$ Padé approximants for $N = 16, \ldots, 25$ to the Rogers–Ramanujan function for the value of q given in (3.6). The poles with positive imaginary parts are real poles and those with negative imaginary parts are defects. The eye can easily pick out the path of convergence as N increases. The first nine real poles are converged by $N = 25$ to within graphical accuracy.

and $(e_i, e_j) = 1$ if $i = j$, and zero otherwise. In terms of this representation the J-matrix is

$$\boldsymbol{J} = \begin{pmatrix} B_1 & -A_1 & 0 & 0 & 0 & \cdots \\ -A_1 & B_2 & -A_2 & 0 & 0 & \cdots \\ 0 & -A_2 & B_3 & -A_3 & 0 & \cdots \\ 0 & 0 & -A_3 & B_4 & -A_4 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{4.2}$$

By its structure, $\boldsymbol{J} = \boldsymbol{J}_1 + \mathrm{i}\boldsymbol{J}_2$ where $\boldsymbol{J}_1$ and $\boldsymbol{J}_2$ are real and self-adjoint. It is known [28,3] that if the element $\boldsymbol{f}$ satisfies the equation

$$\boldsymbol{f} = \boldsymbol{e}_1 - z\boldsymbol{J}\boldsymbol{f}, \tag{4.3}$$

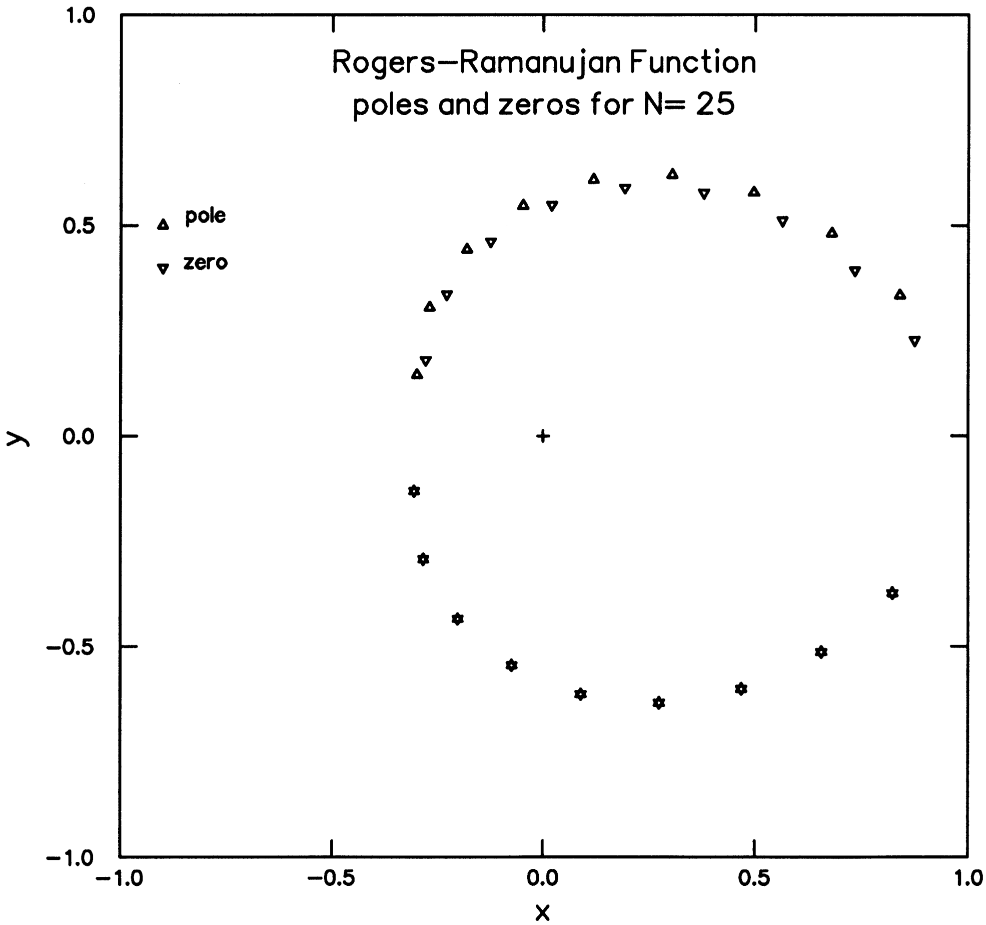

Fig. 3. A plot of the poles and the zeros of the [25/25] Padé approximant to the Rogers–Ramanujan function for the value of q given in (3.6). As the poles are denoted by triangles and the zeros by up-side down triangles, the defects appear as six pointed stars. Only the first nine sets of poles and zeros and the first nine defects are plotted.

and if z is not a singular point of (4.3), then the value of the continued fraction $F(z)$ can be expressed as

$$F(z) = B_0 + A_0 z(\boldsymbol{e}_1, (\boldsymbol{I} + z\boldsymbol{J})^{-1}\boldsymbol{e}_1). \tag{4.4}$$

I now wish to consider the space spanned by the elements, $\boldsymbol{J}^{j-1}\boldsymbol{e}_1$. In general this leads to non-orthogonal projection operators [3]. However, owing to the special circumstance that $\boldsymbol{J}$ is tri-diagonal, I can show by induction that the projection operator on subspace spanned by $\boldsymbol{J}^{j-1}\boldsymbol{e}_1$ for $j = 1, 2, \ldots N$ is just the $N \times N$ diagonal unit matrix. For example $\boldsymbol{J}\boldsymbol{e}_1 = B_1\boldsymbol{e}_1 - A_1\boldsymbol{e}_2$. The only unit vector in the space spanned by $\boldsymbol{e}_1, \boldsymbol{J}\boldsymbol{e}_1$ which is orthogonal to $\boldsymbol{e}_1$ is just $\boldsymbol{e}_2$. This argument carries through by induction. I define the orthonormal projection operator $\boldsymbol{P}_N$ to project the first N states. If the element $\boldsymbol{f}_N$ satisfies the equation,

$$\boldsymbol{f}_N = \boldsymbol{e}_1 - z\boldsymbol{P}_N\boldsymbol{J}\boldsymbol{P}_N\boldsymbol{f}_N, \tag{4.5}$$

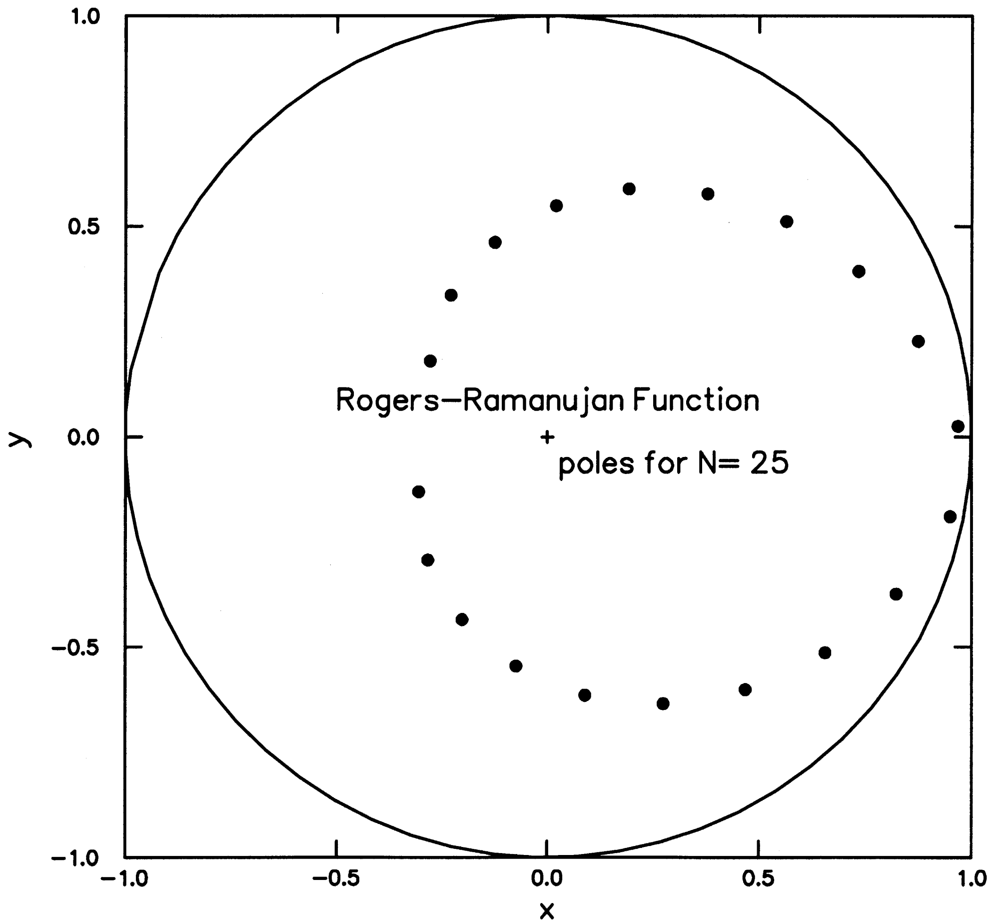

Fig. 4. A plot of the poles of the [25/25] Padé approximant to the Rogers–Ramanujan function for the value of q given in (3.6). This figure illustrates the fact that the poles form a clockwise spiral moving outward towards the unit circle, with only a finite number inside each disk $|z| \leqslant r < 1$. The ninth pole is at a distance of about 0.9036 from the origin. The poles which are part of the defects form a counter-clockwise spiral which is a reflected rotated version of the real poles.

then

$$[N/N]_F(z) = B_0 + A_0 z(\boldsymbol{e}_1, \boldsymbol{f}_N), \tag{4.6}$$

provided z is not a singular point of (4.5).

Next I observe that in both counter-examples, that the corresponding J-matrix is bounded in the sense that [20]

$$\max_i \sum_j |J_{i,j}| \leqslant B < \infty, \qquad \max_j \sum_i |J_{i,j}| \leqslant B < \infty. \tag{4.7}$$

For Buslaev's example, we may choose $B = \frac{5}{3}$ by (2.1), and for Lubinsky's example we may choose $B = 2 + |1 + q^{-1}|$ by (3.5). These results mean that there are no singular points in either Eqs. (4.3) or (4.5) for $|z| < 0.6$ in Buslaev's example and for $|z| < 1/(2 + |1 + q^{-1}|)$ in Lubinsky's example. This latter result was also obtained by Lubinsky [19] by "standard methods." Wall's theorem ([28, 26.3]) proves convergence of the entire diagonal sequence for $|z| < 1/B$. As both of these examples are bounded, we shall confine our attention in this section to the case of bounded J-matrices. Note is taken that if in (4.1) any of the $A_j = 0$, then the J-matrix splits into (at least) two square blocks and this corresponds to the truncation of the associated continued fraction. Hence we have only a rational fraction and convergence questions do not arise. (It may happen that the truncated continued fraction is not equivalent to the power series. This case corresponds to a non-normal Padé table and is not suitable for the present formalism. I will not discuss it here.)

First let us consider the case where $\boldsymbol{J}$ is a compact operator, that is to say, for every sequence of elements $\boldsymbol{c}_m$ with the property $\|\boldsymbol{c}_m\| \equiv (\boldsymbol{c}_m, \boldsymbol{c}_m)^{1/2} \leqslant C$, then the sequence of elements $\boldsymbol{J}\boldsymbol{c}_m$ contains a convergent subsequence $\boldsymbol{g}_m$ which converges in the sense $\|\boldsymbol{g}_j - \boldsymbol{g}_k\| \to 0$ as j and k go to infinity. From the general theory [25] of (4.3) for a compact operator, we know that there are only a finite, or denumerably infinite number of singular values of z for which (4.3) fails to have a unique, bounded solution. These singular values have no limit point in the finite z plane.

It is clear that neither of the counter-examples have compact J-matrices, but if $|q| < 1$ in Lubinsky's example, we would have such an example. The convergence in this case is, of course, covered by a known theorem [18, Theorem 4.55] as in form (3.1) the $\lim_{n\to\infty} c_n = 0$. However, there are cases where such is not the case. For example, suppose the sequence of the c_i, $i = 1, 2, 3, \ldots$, is $1,\ 1,\ 1,\ -1,\ \frac{1}{4},\ -\frac{1}{4},\ 2,\ -2,\ \frac{1}{9},\ -\frac{1}{9},\ 3,\ -3,\ \ldots$. This sequence leads to the associated continued fraction,

$$\frac{z}{1+z}\,_{-}\,\frac{z^2}{1}\,_{+}\,\frac{z^2/4}{1}\,_{+}\,\frac{z^2/2}{1}\,_{+}\,\frac{2z^2/9}{1}\,_{+}\,\frac{z^2/3}{1}\,_{+}\,\cdots, \tag{4.8}$$

For this continued fraction, $\limsup_{n\to\infty} c_n = \infty$, however the limit as $n \to \infty$ of the A_n of (4.1) is zero and all the B_n for $n > 1$ are also zero. From these results, it is straightforward to show that in this case that the J-matrix is a compact operator. Thus the case where $\lim_{n\to\infty} c_n = 0$ is a special case of the following theorem for a compact J-matrix.

Theorem 4.1. *If the J-matrix (4.2) is a compact operator, then the $[N/N]$ Padé approximants converge to $F(z)$ (4.4), provided z is not a singular point of the equation (4.3).*

Proof. Subtract (4.5) from (4.3) and rearrange the result.

$$\boldsymbol{f} - \boldsymbol{f}_N = -z\boldsymbol{J}(\boldsymbol{f} - \boldsymbol{f}_N) - z(\boldsymbol{I} - \boldsymbol{P}_N)\boldsymbol{J}\boldsymbol{f}_N - z\boldsymbol{P}_N\boldsymbol{J}(\boldsymbol{I} - \boldsymbol{P}_N)\boldsymbol{f}_N. \tag{4.9}$$

The last term on the right hand side is identically zero. If $\|\boldsymbol{f}_N\|$ is uniformly bounded for all N, then the next to the last term on the right hand side tends to zero as $N \to \infty$ because $\boldsymbol{J}$ is compact. To see this result, consider the sequence $\boldsymbol{c}_m = \boldsymbol{e}_m$. Then $\boldsymbol{g}_m = \boldsymbol{J}\boldsymbol{c}_m = -A_{m-1}\boldsymbol{e}_{m-1} + B_m\boldsymbol{e}_m - A_m\boldsymbol{e}_{m+1}$. If one assumes that there is a subsequence for which the A's and B's do not tend to zero, then the definition of a compact operator implies there exists a sub-subsequence which converges. However the convergence of the $\boldsymbol{g}_m$'s and perpendicularity of the $\boldsymbol{e}_m$'s imply that those A's and B's tend to zero, which is a contradiction. Thus as $\boldsymbol{J}$ is a tri-diagonal, compact operator, the A_n's and B_n's tend to zero as $n \to \infty$. Therefore $\lim_{N\to\infty} \|(\boldsymbol{I} - \boldsymbol{P}_N)\boldsymbol{J}\| = 0$. Since z is assumed not to be a singular point of (4.3), by the uniqueness of the solution of the equation

$$\boldsymbol{d} = 0 - z\boldsymbol{J}\boldsymbol{d}, \tag{4.10}$$

we can conclude that

$$\begin{aligned} \lim_{N\to\infty} \|\boldsymbol{f} - \boldsymbol{f}_N\| = 0 \quad &\Rightarrow \quad \lim_{N\to\infty} (\boldsymbol{e}_1, \boldsymbol{f}_N) = (\boldsymbol{e}_1, \boldsymbol{f}) \\ &\Rightarrow \quad \lim_{N\to\infty} [N/N]_F(z) = F(z) \end{aligned} \tag{4.11}$$

by (4.6).

Suppose there is a subsequence of the N's for which $\|\boldsymbol{f}_N\|$ is not uniformly bounded. Then for this sequence define $\boldsymbol{d}_N = \boldsymbol{f}_N / \|\boldsymbol{f}_N\|$. Then $\|\boldsymbol{d}_N\| \equiv 1$. Eq. (4.5) becomes

$$\boldsymbol{d}_N = \frac{\boldsymbol{e}_1}{\|\boldsymbol{f}_N\|} - z\boldsymbol{J}\boldsymbol{d}_N + z(\boldsymbol{I} - \boldsymbol{P}_N)\boldsymbol{J}\boldsymbol{d}_N. \tag{4.12}$$

As $\|\boldsymbol{f}_N\|$ goes to infinity, the first term on the right-hand side of (4.12) goes to zero. Since $\boldsymbol{J}$ is compact, the last term on the right-hand side also goes to zero and hence the equation reduces to (4.10) in the limit as $N \to \infty$. This situation implies $\lim_{N\to\infty} \boldsymbol{d}_N = \boldsymbol{0}$ which is a contradiction as $\|\boldsymbol{d}_N\| = 1$. Therefore there do not exist infinite subsequences of divergent norm when z is not a singular point of (4.3) and the conclusion of the theorem follows by the preceding arguments. □

Another case is the one where although $\boldsymbol{J}$ is bounded but not compact, and there is a subsequence of the A_n of (4.1) which tends to zero as $n \to \infty$. I summarize the results in the following theorem.

Theorem 4.2. *Let the J-matrix (4.2) be bounded and let there be a subsequence $\mathscr{T}$ of N's such that $\lim_{N\to\infty, N\in\mathscr{T}} A_N = 0$. Then there is at least a subsequence of $[N/N]$ Padé approximants which converges to $F(z)$ (4.4) provided z is not a singular point of Eq. (4.3).*

Proof. The proof of this theorem is very similar to that of Theorem 4.1. First assume that the sequence of $\boldsymbol{f}_N$ is uniformly bounded. I now need to estimate the size of the middle term on the right-hand side in (4.9). We may compute it directly as

$$-z(\boldsymbol{I} - \boldsymbol{P}_N)\boldsymbol{J}\boldsymbol{f}_N = zA_N\boldsymbol{e}_{N+1}(\boldsymbol{e}_N, \boldsymbol{f}_N). \tag{4.13}$$

Since $\|\boldsymbol{f}_N\|$ is uniformly bounded, the same is true of $|(\boldsymbol{e}_N, \boldsymbol{f}_N)|$. As A_N, $N \in \mathscr{T}$ goes to zero, Eq. (4.9) again reduces to (4.10). Thus as z is not a singular point of (4.3), by the uniqueness of the solution of that equation, we can again conclude convergence in this case.

To establish the boundedness just used, I note that in this case, Eq. (4.12) becomes

$$\boldsymbol{d}_N = \frac{\boldsymbol{e}_1 - zA_N\boldsymbol{e}_{N+1}(\boldsymbol{e}_N,\boldsymbol{f}_N)}{\|\boldsymbol{f}_N\|} - z\boldsymbol{J}\boldsymbol{d}_N \tag{4.14}$$

which again reduces to (4.10) for $N \in \mathscr{T}$ and leads to a contradiction to the existence of an infinite sequence with a divergent norm. Hence we conclude that the sequence $[N/N]$ of Padé approximants where $N \in \mathscr{T}$ converges to $F(z)$ provided that z is not a singular point of Eq. (4.3). □

The alternative case to Theorem 4.2 is the case where there is no subsequence for which $|A_N|$ tends to zero as $N \to \infty$. The following results focus on the observations that in the two counter-examples, there are a finite number of defects (spurious pole-zero pairs) and as they move around, we can patch two subsequences together to obtain uniform convergence. Nuttall's conjecture is also a motivating factor for these results, as it also pertains to limiting the number of defects as well.

Lemma 4.3. *Let the J-matrix (4.2) be bounded and further assume that $|A_n| \geqslant b > 0, \forall n$. It cannot be that if $z_{M,j}$ is a pole of the $[M/M]$ to $F(z)$ of (4.1), that it is also a pole of the $[M+1/M+1]$. Likewise, it cannot be that if $\hat{z}_{M,j}$ is a zero of the $[M/M]$ that it is also a zero of the $[M+1/M+1]$.*

Proof. Let us re-express (3.22) in terms of the coefficients of the associated continued fraction. It is

$$Q_M^{(0)}(z)P_{M+1}^{(0)}(z) - Q_{M+1}^{(0)}(z)P_M^{(0)}(z) = z^{2M+1}a_1\prod_{j=1}^{M} A_j^2, \tag{4.15}$$

Since the right-hand side of (4.15) is just $z^{2M+1}a_1 \prod_{j=1}^{M} A_j^2$ with a magnitude $\geqslant z^{2M+1}b^{2M} \neq 0$, the movement of the poles and zeros follows directly. □

In assessing the magnitudes of the polynomials in (4.15) it is worthwhile to note the following recursion relation,

$$\begin{pmatrix} S_{N+2}^{(0)}(z) \\ S_{N+1}^{(0)}(z) \end{pmatrix} = \begin{pmatrix} (1+B_{N+1}z)(1+B_{N+2}z) - A_{N+1}^2z^2 & -(1+B_{N+2}z)A_N^2z^2 \\ 1+B_{N+1}z & -A_N^2z^2 \end{pmatrix} \begin{pmatrix} S_N^{(0)}(z) \\ S_{N-1}^{(0)}(z) \end{pmatrix}, \tag{4.16}$$

which can be derived in a straightforward manner from the standard recursion relations for the approximants to continued fractions. The notation is as in (4.1) and $S_N^{(0)}(z) = rQ_N^{(0)}(z) + sP_N^{(0)}(z)$ with r and s arbitrary. The equation for the eigenvalues of the transfer matrix in (4.16) is

$$\lambda^2 - [(1+B_{N+1}z)(1+B_{N+2}z) - z^2(A_N^2 + A_{N+1}^2)]\lambda + A_n^2A_{N+1}^2z^4 = 0. \tag{4.17}$$

Note that the absolute value of the geometric mean of the solutions of (4.17) is just $|A_NA_{N+1}z^2|$ so that a characteristic size for $S_N(z)$ might be taken to be something like $z^N \prod_{j=1}^{N} A_j$, which if used in the left-hand side of (4.15) gives a result which is not dissimilar in size to the right-hand side. As an example, when z is small, $\lambda_+ = 1 + O(z)$ and $\lambda_- = O(z^4)$.

Although not mentioned explicitly in the Gončar's Theorem 2.2, it is evident from his proof that the exceptional set of possible non-convergence is associated with neighborhoods of the poles of the diagonal approximants. By applying his theorem to the reciprocal function (remember that the diagonal Padé approximant to the reciprocal function is the reciprocal of the Padé approximant) we obtain convergence on the sphere at the actual poles of the function. Since the Hausdorf measure of the exceptional set is zero, there must be a zero which tends to each spurious pole and so the intersection of the exceptional set for the function and for its reciprocal will contain (shrinking in size as $N \to \infty$) neighborhoods of all the spurious poles, and nothing else. Thus given a compact set $\mathscr{K}$ there will be at most only a finite number $\kappa(\mathscr{K})$ of spurious poles under the assumptions of Gončar's Theorem 2.2. In spite of Lemma 4.3, in the case of a real pole the motion of the corresponding pole in successive approximants is very small and the poles of the approximants tend to limit points.

This situation however is different in the case of the spurious poles in the counter-examples. They move significantly from one diagonal approximant to the next. This feature has been commonly observed in a wide variety of other cases as well. Suppose we have an isolated, spurious pole. If we evaluate (4.15) at $z_{M,j}$, a spurious zero of $Q_M^{(0)}(z)$, we get

$$Q_{M+1}^{(0)}(z_{M,j})P_M^{(0)}(z_{M,j}) = -z_{M,j}^{2M+1}\prod_{j=1}^{M} A_j^2. \tag{4.18}$$

Its magnitude is $\geqslant z_{M,j}^{2M+1}b^{2M} \neq 0$. Now as we have just seen, since $P_M^{(0)}(z)$ must have a zero near by, both $P_M^{(0)}(z_{M,j})$ and the residue at a spurious pole are very small. Eq. (4.18) thus implies that $Q_{M+1}^{(0)}(z_{M,j})$ is not at all small, as the right-hand side is bounded from below, and hence $Q_{M+1}^{(0)}(z)$ does not have a nearby zero. The comparison "very small" is relative to its size for a real pole. In the case of a real pole, the zero of $Q_{M+1}^{(0)}(z)$ does not move much. It is this ratio of sizes which is what forces the zero of $Q_{M+1}^{(0)}(z)$ away from $z_{M,j}$.

The above argument indicates the plausibility of the idea that for the entire sequence of diagonal Padé approximants isolated, spurious poles do not have a single limit point, but is not yet a proof.

Theorem 4.4. *Let the J-matrix (4.2) be bounded and further assume that $|A_n| \geqslant b > 0$, $\forall n$. Let $\check{z}$ be any point in an open, simply connected region $\mathscr{R}$ of the extended complex plane which is bounded by a simple closed curve. Let $\mathscr{R}$ contain the origin and be in the domain of meromorphy of $F(z)$ (4.4). Further let $\check{z}$ not be a limit point of spurious poles of the entire sequence of the $[k/k]$ Padé approximants. Let the number of poles and zeros $n_k(d)$ of $[k/k]$ more distant on the sphere than any $d > 0$ from the boundary of $\mathscr{R}$ satisfy*

$$\lim_{k\to\infty} n_k(d)/k = 0. \tag{4.19}$$

Then there exists at least a subsequence of $[k/k](\check{z})$ Padé approximants which converge on the sphere to $F(\check{z})$.

Note, condition (4.19) insures that $\mathscr{R}$ avoids the branch cuts, natural boundariers and essential singularities of $F(z)$. One of the other consequences of condition (4.19) is to allow the construction of a path in order to prove analytic continuation from the origin to $\check{z}$ so as to insure that the convergence is in fact to $F(\check{z})$ and not, for example, to $F(\check{z})$ on a different Riemann sheet.

Proof. As the J-matrix is infinite and bounded, all the $[k/k]$ Padé approximants exist. If $\mathscr{R}$ is not already that size we may enlarge it to include all $|z| < 1/B$. Select $\mathscr{T}$, a closed set, consisting of $|z| \leqslant \beta < 1/B$ and the point $\check{z}$. In this set there are no poles nor limit points of poles as the J-matrix is bounded by B and by hypothesis $\check{z}$ is not a limit point of poles of the entire $[k/k]$ sequence. Select d so that no point in $\mathscr{T}$ is closer than d to the boundary of $\mathscr{R}$. If, as $k \to \infty$ the neighborhood of $\check{z}$ is free of spurious poles, then by Baker's theorem [2, Theorem 12.5]; [4] the entire sequence converges uniformly on the sphere in $\mathscr{T}$. If on the other hand there are spurious poles in some of the diagonal Padé approximants at or near $\check{z}$, we are guaranteed by the assumptions of the theorem that there exists a subsequence for which there is a neighborhood of $\check{z}$ which is free from spurious poles and again by Baker's theorem we get uniform convergence on the sphere in $\mathscr{T}$. □

Theorem 4.5. *Let the J-matrix (4.2) be bounded and further assume that $|A_n| \geqslant b > 0$, $\forall n$. Let $\mathscr{K}$ be any simply connected, compact set containing the origin as an element but not containing any element of a set $\mathscr{E}$ of capacity zero. Let the number of poles of $[n/n]$ in $\mathscr{K}$ be bounded uniformly in n by $\kappa(\mathscr{K}) < \infty$. Further assume that, for the entire sequence of $[n/n]$ Padé approximants, none of the spurious poles have only a single limit point in $\mathscr{K}$. Then by use of at most $\kappa(\mathscr{K}) + 1$ subsequences, we can obtain uniform convergence in $\mathscr{K}$.*

Note that the restriction on the number of poles means that the set $\mathscr{K}$ avoids branch cuts and natural boundaries of $F(z)$. Essential singularities can be incorporated in the set $\mathscr{E}$ as was done by Pommerenke [23].

Proof. First, by Gončar's theorem 2.2, there is convergence everywhere in $\mathscr{K}$ except for a set of Hausdorf measure 0 in the neighborhoods of the spurious poles. If there are no spurious poles, the entire sequence of $[n/n]$ Padé approximants converges in $\mathscr{K}$ by Theorem 2.2. I confine the rest of the proof to the case where there are spurious poles. More precisely, it is the union of $\kappa(\mathscr{K})$, ε/n^{γ} neighborhoods for all $n > N_0$ for appropriately chosen ε, γ, and N_0 as detailed in Gončar's proof. Since the number of poles in $\mathscr{K}$ is bounded, the number of spurious poles is bounded a fortiori. Suppose the number is unity. Since $\mathscr{K}$ is compact, there exists a subsequence for which this spurious pole tends to a limit. By the assumption on the lack of a limit point for the whole sequence, there must exist another subsequence which converges in the neighborhood of that limit point. Thus by patching together the two subsequences, we may obtain uniform convergence in $\mathscr{K}$. Since we are dealing with a bounded J-matrix, any infinite subsequence converges in $|z| < 1/B$. Suppose that there are two spurious poles. First we find a subsequence which has a limit point of one of the spurious poles. From this infinite subsequence we may choose a sub-subsequence for which the second pole also converges to a limit point. We may now choose another subsequence for which the value converges in a neighborhood of at the first limit point of spurious poles. It may happen (plausibly does) that it also converges at the second limit point of poles. If not we may select a third subsequence which does converge there. Thus by patching the three subsequences together, we obtain uniform convergence in $\mathscr{K}$. By continuing the argument in the same way we obtain by induction that there exist at most $\kappa(\mathscr{K}) + 1$ subsequences which can be patched together to provide uniform convergence in $\mathscr{K}$. □

I remark that both Buslaev's and Lubinsky's counter-examples satisfy the conditions of this theorem.

The above results point to two open questions:

(1) Can it be proven for the class of functions which correspond to bounded J-matrices with $|A_N| \geqslant b > 0 \ \forall N$, that there are only a finite number of spurious poles in any compact set inside the domain of meromorphy, excluding branch cuts, natural boundariers and essential singularities?
(2) Can it be shown for the same class of functions that in the $[n+1/n+1]$ approximants there are no spurious poles in the neighborhoods of the isolated, spurious poles of the $[n/n]$ approximants?

One of the referees has brought to my attention that there is an alternate approach to the material of this section [1,7,12], which is reviewed in [8]. This approach applies spectral analysis to the issues. Some of the fundamental ideas are: (i) To use in the resolvent set of the matrix the exponential decay, with the distance away from the principal diagonal, of the matrix elements of the inverse of a tridiagonal matrix to bound the product of the Padé denominator polynomials times the residues of the linear Padé equations. (ii) To analyze the behavior of $Q_M^{(0)}(z)/Q_{M+1}^{(0)}(z) = (\boldsymbol{e}_{M+1}, (\boldsymbol{I} + z\boldsymbol{P}_{M+1}\boldsymbol{J}\boldsymbol{P}_{M+1})^{-1}\boldsymbol{e}_{M+1})$ by means of its chordal derivative with respect to z. The results of this section and the answers to the open questions have been partly obtain by that approach.

The author is happy to acknowledge helpful correspondence with Professors D.S. Lubinsky, V.I. Buslaev, and J.L. Gammel.

References

[1] A.I. Aptekarev, V. Kaliaguine, W. Van Assche, Criterion for the resolvent set of nonsymmetric tridiagonal operators, Proc. Amer. Math. Soc. 123 (1995) 2423–2430.
[2] G.A. Baker Jr., Essentials of Padé Approximants, Academic Press, New York, 1975.
[3] G.A. Baker Jr., Convergence of Padé approximants using the solution of linear functional equations, J. Math. Phys. 16 (1975) 813–822 This reference has the following unfortunate typographic errors. Eq. (5.11) should read $\liminf_{N\to\infty} \|\mathscr{P}_N \mathbf{A}\mathbf{P}_N \mathbf{A}\mathbf{e}_N\| = 0$ and Eq. (5.15) should read $\lim_{N\to\infty} \|(\mathbf{I} - \mathscr{P}'_N)\mathscr{P}_N\| = 0$.
[4] G.A. Baker Jr., A theorem on convergence of Padé approximants, Stud. Appl. Math. 55 (1976) 107–117.
[5] G.A. Baker Jr., J.L. Gammel, J.G. Wills, An investigation of the applicability of the Padé approximant method, J. Math. Anal. Appl. 2 (1961) 405–418.
[6] G.A. Baker Jr., P.R. Graves–Morris, Padé approximants, in: G.-C. Rota (Ed.), Encyclopedia of Mathematics and its Applications, Vol. 59, 2nd Edition, Cambridge University Press, New York, 1996.
[7] B. Beckermann, On the convergence of bounded J-fractions on the resolvent set of the corresponding second order difference operator, J. Approx. Theory 99 (1999) 369–408.
[8] B. Beckermann, Complex Jacobi matrices, J. Comput. Appl. Math. 127 (2001) 17–65.
[9] R.P. Brent, Algorithm 524: MP a fortran multiple-precision arithmetic package [A1], ACM Trans. Math. Software 4 (1978) 71–81, and accompanying software on a computer diskette.
[10] V.I. Buslaev, Simple counterexample to the Baker–Gammel–Wills conjecture, East J. Approx. 7 (2001) 515–517.
[11] V.I. Buslaev, On the Baker–Gammel–Wills conjecture in the theory of Padé approximants, Sbornik Math. 193 (2002) 811–829.
[12] S. Demko, W.F. Moss, P.W. Smith, Decay rates for inverses of band matrices, Math. Comp. 43 (1984) 491–499.
[13] J.L. Gammel, J. Nuttall, Convergence of Padé approximants to quasianalytic functions beyond natural boundaries, J. Math. Anal. Appl. 43 (1973) 694–696.
[14] G.M. Goluzin, Geometric Theory of Functions of a Complex Variable, American Mathematical Society, Providence, RI, 1969.

[15] A.A. Gončar, On uniform convergence of diagonal Padé approximants, Math. USSR Sbornik 46 (1983) 539–559.
[16] G.H. Hardy, J.E. Littlewood, Notes on the theory of series (XXIV): a curious power series, Proc. Cambridge Philos. Soc. 42 (1946) 85–90.
[17] M.D. Hirschhorn, Partitions and Ramanujan's continued fraction, Duke Math. J. 39 (1972) 789–791.
[18] W.B. Jones, W.J. Thron, Continued fractions, analytic theory and applications, in: G.-C. Rota (Ed.), Encyclopedia of Mathematics and its Applications, Vol. 11, Addison-Weslely, Reading, MA, 1980.
[19] D.S. Lubinsky, Rogers–Ramanujan and the Baker–Gammel–Wills (Padé) conjecture, Ann. Math. 157 (2003) 847–889.
[20] M. Marcus, Basic Theorems in Matrix Theory, Appl. Math. Ser., Vol. 57, Theorem 3.3, National Bureau of Standards, Washington, DC, 1960.
[21] J. Nuttall, Asymptotics of diagonal Hermite–Padé polynomials, J. Approx. Theory 42 (1984) 299–386.
[22] G. Polya, Über gewisse notwendige Determinantkriterien für die Forsetzbarkeit einer Potenzreihe, Math. Annalen 99 (1928) 687–706.
[23] C. Pommerenke, Padé approximants and convergence in capacity, J. Math. Anal. Appl. 41 (1973) 775–780.
[24] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, Numerical Recipes in Fortran, Cambridge University Press, New York, 1992.
[25] F. Riesz, B. Sz.-Nagy, Functional Analysis, Ungar, New York, 1955 (translation by L.F. Boron).
[26] H. Stahl, The convergence of diagonal Padé approximants and the Padé conjecture, J. Comput. Appl. Math. 86 (1997) 287–296.
[27] H. Stahl, The convergence of Padé approximants to functions with branch points, J. Approx. Theory 91 (1997) 139–204.
[28] H.S. Wall, Analytic Theory of Continued Fractions, Van Nostrand, Princeton, NJ, 1948.