Padé Approximants

Maciej Pindor

Instytut Fizyki Teoretycznej,
Uniwersytet Warszawski ul.Hoża 69,
00-681 Warszawa, Poland.

1 Introduction

The frequent situation one encounters in applied science is the following: the information we need is contained in values, or some features of the analytical structure, of some function of which we have a knowledge only in the form of its power expansion in a vicinity of some point. Favourably, it is the Taylor expansion with some finite radius of convergence, but it may also be an asymptotic expansion. Let us concentrate on the first case, some remarks concerning the second one will be given later, if time allows.

If the information we need concerns points within the circle of convergence of the Taylor series, then the problem is (almost) trivial. If it concerns points outside the circle, then the problem becomes that of analytic continuation. Unfortunately, the method of direct rearrangements of the series, used in theoretical considerations on the analytic continuation, is practically useless here. The method of the "practical analytic continuation" which I shall discuss is called the method of "Padé Approximation". There exist ample monographs on Padé Approximants [6], [2], [3] and my purpose here is to present you a subjective glimpse of the subject.

Actually, the method is based on the very direct idea of using rational functions instead of polynomials to approximate the function of interest. They are practically as easy to calculate as polynomials, but when we recall that the truncated Laurent expansions is just a rational function, we can expect that they could provide reasonable approximations of functions also in a vicinity of the poles of the latter, not only in circles of analyticity. Therefore the concept, born already in XIXth century, was to substitute partial sums of the Taylor series, by rational functions having the corresponding partial sums of their own Taylor series identical to that former one. To formulate it precisely, let us assume we have a function $f(z)$ with its Taylor expansion

$$f(z) = \sum_{i=0}^{\infty} f_i z^i \text{ for } |z| < R \; . \qquad (1)$$

J.-D. Fournier et al. (Eds.): Harm. Analysis and Ratio. Approx., LNCIS 327, pp. 59–69, 2006.

Having the partial sum of the above series up to the power M, we seek a rational function $r_{m,n}(z)$ which will have first $M+1$ terms of its Taylor expansion identical to that of $f(z)$ what I shall represent by

$$r_{m,n}(z) - f(z) = O(z^{M+1}) \ . \tag{2}$$

Unfortunately this problem seems to be badly defined – there are probably many rational functions that can satisfy this condition: possibly all such that $m + n = M$. In other words, assuming for the moment that all such rational functions can be found, to the infinite *sequence* of partial sums of the Taylor series (1) there correspond an infinite *table* (or a double sequence) of *rational approximants* defined by (2). As we are after the analytic continuation of $f(z)$, we expect that some sequence of rational approximants defined this way would converge, in some sense, to $f(z)$ outside the convergence circle of (1). But which one? Is it a case of advantageous flexibility, or that of “embarras du choix”? I shall argue in a moment that it is this first one!

2 The Padé Table

Let us, however, discuss first the problem of existence of rational functions defined by (2). If we denote the numerator of $r_{m,n}(z)$ by $P_m(z)$ and its denominator by $Q_n(z)$ and $r_{m,n}(z)$ by $[m/n]_f(z)$ then (2) becomes

$$[m/n]_f(z) - f(z) = \frac{P_m(z)}{Q_n(z)} - f(z) = O(z^{m+n+1}) \ . \tag{3}$$

Let me make here an obvious remark that P_m depends also on n and Q_n depends on m and they should be denoted, e.g., $P_m^{[m/n]}$, but for hygienic reasons I shall almost everywhere skip this additional index. Finding coefficients of P_m and Q_n by the expansion of $[m/n]_f(z)$ and then comparing the two series, would be a horror, but the problem can immediately be reduced to the linear one:

$$P_m(z) - Q_n(z)f(z) = O(z^{m+n+1}) \ . \tag{4}$$

This is how Frobenius [5] defined “Näruhngsbrüchen” already in 1881 and therefore (4) is called the Frobenius definition. One can immediately see that it leads to a system of linear equations for coefficients of $Q_n(z)$ and formulae expressing coefficients of P_m by those of Q_n. Denoting the former by $\{p_i\}_0^m$ and the later by $\{q_i\}_0^n$ we have (assuming that $f_i \equiv 0$ for $i < 0$)

$$\begin{pmatrix} f_{m+1} & f_m & \cdots & f_{m-n+2} & f_{m-n+1} \\ f_{m+2} & f_{m+1} & \cdots & f_{m-n+3} & f_{m-n+2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ f_{m+n} & f_{m+n-1} & \cdots & f_{m+1} & f_m \end{pmatrix} \begin{pmatrix} q_0 \\ q_1 \\ \cdots \\ q_n \end{pmatrix} = 0 \tag{5}$$

and

$$
\begin{aligned}
p_0 &= f_0 q_0 \\
p_1 &= f_1 q_0 + f_0 q_1 \\
p_2 &= f_2 q_0 + f_1 q_1 + f_0 q_2 \\
\ldots&\ldots \\
p_m &= \sum_{i=0}^{\min(m,n)} f_{m-i} q_i \,.
\end{aligned}
\tag{6}
$$

The system (5) is an homogeneous one but it has n equations for $n+1$ unknowns. The reason is that $[m/n]_f$ has $m+n+1$ free coefficients, but we have written equations for $m+n+2$ ones. The result is that the system (5) has always at least one nontrivial solution. One could think that we can take then an arbitrary value for one of the coefficients q_i and next solve (5) for the remaining coefficients of Q_n. However, it may happen that the determinant of this linear system vanishes and either we have again an infinite number of solutions, or no solution at all. The problem can also be stated in this way: although the rational approximant defined by (4) always exists, it may happen that it does not satisfy (3). The first study of the table of all approximants $[m/n]_f$ has been done by Henry Padé [8] in his PhD dissertation and it is why now they are called *Padé Approximants* and the table is called *Padé Table*. The result was that there were square areas of the Padé Table where all entries were identical rational functions of degrees equal to those of its upper left corner and they all fulfill (4). However, only approximants on the antidiagonal of the square and to the left (up) to it fulfill also (3). According to one of the contemporary definitions introduced by Baker [1], we take $Q_n(0) \neq 0$ (e.g. 1, i.e. $q_0 = 1$) which is possible only when (3) is satisfied, and say that only in this case "Padé Approximants exist". See Fig. 1.

I shall present here only the very brief discussion of situations leading to an appearance of blocks in the Padé table. Of course their existence is due to special relations between coefficients of the Taylor series – e.g. vanishing of some coefficients or possibility of representing higher coefficients by algebraic functions of lower ones.

The first situation is exemplified by the series containing only even powers of the variable

$$f(z) = \frac{\log(1+z^2)}{z^2} = 1 - \frac{z^2}{2} + \frac{z^4}{3} - \frac{z^6}{4} + \frac{z^8}{5} + \cdots$$

For this series we have

$$[2/2]_f = \frac{1 + \frac{z^2}{6}}{1 + \frac{2z^2}{3}} = 1 - \frac{z^2}{2} + \frac{z^4}{3} - \frac{2z^6}{9} + \cdots$$

Obviously, $[2/2]_f$ is simultaneously $[3/2]_f$ and $[2/3]_f$ because its Taylor series matches that of $f(z)$ up to z^5. On the other hand, there is no rational function

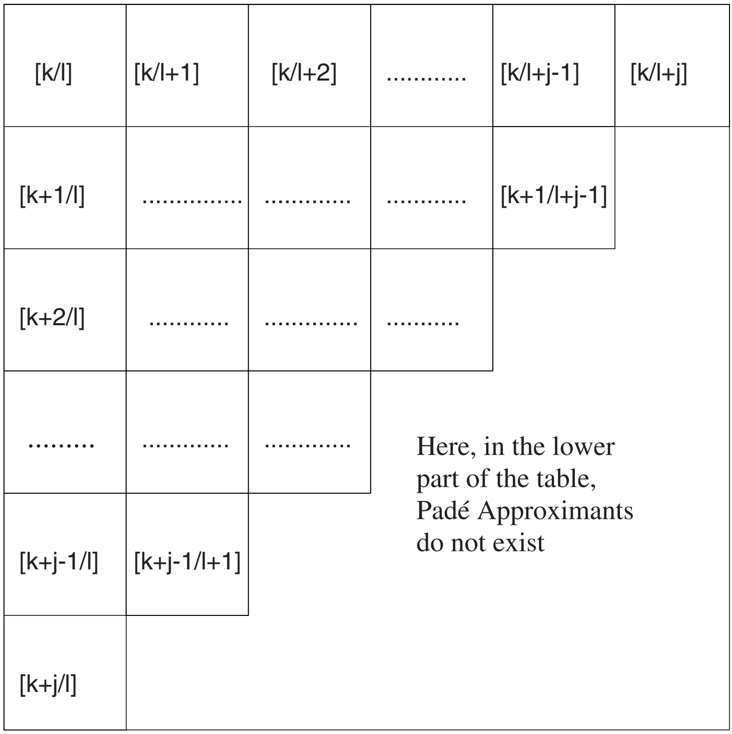

Fig. 1. A block of the size $j+1$ in the Padé Table. All Padé Approximants on the positions indicated by their symbols, or by dots, exist and are identical to $[k/l]$, therefore they are rational functions of degrees k and l in the numerator and the denominator, however they fulfill equation (3) with m and n corresponding to their positions in the Padé Table.

of degrees of the numerator and of the denominator both ≤ 3 that would match the series for $f(z)$ up to z^6. $[4/2]_f$ satisfies this condition, but then it is identical with $[5/3]_f$ and $[4/3]_f$

$$[4/2] = \frac{1 + \frac{z^2}{4} - \frac{z^4}{24}}{1 + \frac{3z^2}{4}} = 1 - \frac{z^2}{2} + \frac{z^4}{3} - \frac{z^6}{4} + \frac{3z^8}{16} + \cdots$$

In this case the whole Padé Table consists of blocks of the size 2.

The second situation appears typically when $f(z)$ is a rational function itself. In this case there is one infinite block with the left upper corner at the entry corresponding to the exact degrees of the numerator and the denominator of this function. Obviously all the Padé Approximants with degrees of numerators and denominators larger or equal to these of the function, are equal to this function, because it matches it own Taylor expansion to any order!

3 Convergence

Rational functions are meromorphic, and therefore the first speculation that comes to the mind (at least mine) is that Padé approximants should be well suited to approximate just the former ones.

This speculation appears to absolutely correct, because there holds the de Montessus theorem ([3] p. 246):

Theorem 1 (de Montessus, 1902). *Let $f(z)$ be a function meromorphic in the disk $|z| < R$ with m poles at distinct points $z_1, z_2, ..., z_m$ with*

$$0 < |z_1| \leq |z_2| \leq \cdots \leq |z_m| < R \, .$$

Let the pole at z_k have multiplicity μ_k and let the total multiplicity $\sum_{k=1}^{m} \mu_k = M$ precisely. Then

$$f(z) = \lim_{L \to \infty} [L/M]$$

uniformly on any compact subset of

$$\mathcal{D} = \{z, \, |z| \leq R, \, z \neq z_k, \; k = 1, 2, \ldots, m\} \, .$$

One could be very enthusiastic about this theorem, considering that it "solves" completely the problem of analytic continuation inside a disc of meromorphy. There is however a practical obstacle in applying the theorem: generally, we cannot say what M we should use. We cannot expect anything particularly interesting if M is too small (e.g. smaller than the multiplicity of the nearest singularity), but when it is too large, the uniform convergence can be expected only for subsequences on rows in the Padé Table. This is well illustrated by [4]:

Theorem 2 (Beardon, 1968). *Let $f(z)$ be analytic in $|z| \leq R$. Then an infinite subsequence of $[L/1]$ Padé approximants converges to $f(z)$ uniformly in $|z| \leq R$.*

which casts into doubt whether the sequence $[L/1]$ must converge even in a disc of analyticity of the function! Although the theorem does not exclude that the subsequence could be the complete sequence, many counterexamples were constructed to show that the above theorem is the optimal result. Maybe the best known is the one due to Perron [9] – he has constructed the series representing an entire function, but such that poles of $[L/1]$ were dense in the plane.

On the other hand such ugly phenomena do not appear in "practice" – e.g. for $f(z) = \mathrm{e}^z$ poles of $[L/a]$ lie at $L+1$, while these of $[L/2]$ at $L+1 \pm \mathrm{i}\sqrt{L+1}$ and both rows (and also all the other ones) of the Padé table converge to $f(z)$ on any compact subset of the complex plane containing the origin.

Happily, problems caused by the stray poles are not as acute as one could think, as explained by the following theorem ([3] p. 264)

Theorem 3. *Let $f(z)$ be analytic at the origin and also in a given disk $|z| \leq R$ except for m poles counting multiplicity. Consider a row of Padé table $[L/M]$ of $f(z)$ with M fixed, $M \geq m$, and $L \to \infty$. Suppose that arbitrarily small, positive ε and δ are given. Then L_0 exists such that $|f(z) - [L/M]| < \varepsilon$ for any $L > L_0$ and for all $|z| \leq R$ except for $z \in \mathcal{E}_L$ where $\mathcal{E}_L$ is a set of points in the z-plane of measure less than δ.*

This type of convergence is known as the convergence in measure and seems to be used in this context first by Nuttal [7]. It means that we cannot guarantee convergence at any given point in the z-plane, but it assures us that the area where our Padé approximants do not approximate $f(z)$ arbitrarily well can be made as small as we wish.

It is important to understand that the theorem says nothing about where this set $\mathcal{E}_L$ is, and the practice shows that undesired poles are accompanied by undesired zeros and form so called *defects* which spoil convergence in smaller and smaller neighborhoods, but shift unpredictably from order to order.

But what about functions with more rich analytical structure – essential singularities and branch points?

The amazing (at least for me) fact is that if we are content with convergence in measure (or even stronger convergence in capacity) also such functions can be approximated by Padé approximants, if we consider sequences with growing degrees of the numerator and of the denominator. The fundamental theorem on convergence of Padé approximants for functions with essential singularities is due to Pommerenke [10]

Theorem 4 (Pommerenke, 1973). *Let $f(z)$ be a function which is analytic at the origin and analytic in the entire z-plane except for a countable number of isolated poles and essential singularities. Suppose $\varepsilon > 0$ and $\delta > 0$ are given. Then M_0 exists such that any $[L/M]$ Padé approximant of the ray sequence $(L/M = \lambda;\ \lambda \neq 0,\ \lambda \neq \infty)$ satisfies*

$$|f(z) - [L/M]_f(z)| < \varepsilon$$

for any $M \geq M_0$, on any compact set of the z-plane except for a set $\mathcal{E}_L$ of capacity less than δ.

As you see, the essential notion here is that of capacity. It is also known as Chebishev constant, or transfinite diameter. I do not have time here to define it, as it is a difficult concept concerning geometry of the complex plane. Anyway to understand practical implications of the theorem above and the ones to follow, it is sufficient to know that the capacity is a function on sets in the complex plane such that it vanishes for countable sets of points, but is different from zero on line segments, e.g. for a section of a straight line it equals to one fourth of its length. For a circle it is the same as for the disk inside the circle and equals to their radius. Actually it is proportional to the electrostatic capacity in the plane electrostatics.

If we want to approximate functions having branchpoints the first question that comes to mind is how can rational functions approximate a function in a vicinity of its branchpoint? The astonishing answer is – they can do it very well, simulating a cut as a line of coalescence of infinite number of zeros and poles! This answer may seem puzzling for you – which cut? There seem to be the enormous arbitrariness in joining branchpoints by cuts, and why should Padé approximants choose just this set of cuts and not another, or why should all Padé approximants choose the same cuts? The answer to these questions lies in the interesting fact that although "all cuts are equal", but some of them "are more equal than others".

This fact is established by the following theorem ($\hat{\mathbb{C}}$ denotes here the extended complex plane) [11]

Theorem 5 (Stahl, 1985). *Let f be given by an analytic function element in a neighborhood of infinity. There uniquely exists a compact set $\mathcal{K}_0 \subseteq \mathbb{C}$ such that*

(i) $\mathcal{D}_0 := \hat{\mathbb{C}} \backslash \mathcal{K}_0$ *is a domain in which $f(z)$ has a single-valued analytic continuation,*
(ii) $\mathrm{cap}(\mathcal{K}_0) = \inf \mathrm{cap}(\mathcal{K})$, *where the infimum extends over all compact sets $\mathcal{K} \subseteq \mathbb{C}$ satisfying (i),*
(iii) $\mathcal{K}_0 \subseteq \mathcal{K}$ *for all compact sets $\mathcal{K} \subseteq \mathbb{C}$ satisfying (i) and (ii).*

The set $\mathcal{K}_0$ is called *minimal set* (for single-valued analytical continuation of $f(z)$) and the domain $\mathcal{D}_0 \subseteq \hat{\mathbb{C}}$ – *extremal domain*.

The following theorem, due to H. Stahl [11], refers to, so called, close-to-diagonal sequences of Padé approximants. By the latter one means the sequence $[m/n]$ such that $\lim_{m+n\to\infty} m/n = 1$.

Theorem 6 (Stahl, 1985). *Let the function $f(z)$ be defined by*

$$f(z) = \sum_{j=0}^{\infty} f_j z^{-j}$$

and have all its singularities in a compact set $E \subseteq \hat{\mathbb{C}}$ of capacity zero. Then any close to diagonal sequence of Padé approximants $[m/n](z)$ to the function $f(z)$ converges in capacity to $f(z)$ in the extremal domain $\mathcal{D}_0$.

In simple words, the theorem says that close-to-diagonal sequences of Padé approximants converge "practically", for a very wide class of functions, everywhere, except on a set of "optimal" cuts. However, we must keep in the mind that it is not the uniform convergence, therefore when applying Padé approximants, we must be careful and compare few different approximants from a close-to-diagonal sequence.

4 Examples

Let us see some examples how Padé approximants work for different types of functions. In illustrations below, I shall devote more attention to demonstrating that Padé approximants "discover" correctly singularities and zeros than to approximating values of functions, though I shall not forget about the latter.

Let $f(z) = \tanh(z)/z + 1/[2(1+z)]$. This function has an infinite number of poles uniformly distributed on the imaginary axis at $z = (2k+1)\pi/2$ $k = 0, \pm 1, \pm 2, \ldots$ and the pole at $z = -1$. It has also infinite number of zeros, the ones closest to origin are: $z = -2.06727$, $-.491559 \pm 2.93395$i, $-.535753 \pm 6.17741$i, $-.545977 \pm 12.5134$i and so on. I have added the geometric series mainly to have a function with a series containing all powers of z, not the one with even powers only. A small curiosity is that there is a block in the Padé table of this function – the one consisting of [0/1], [0/2], [1/1], [1/2].

As in any circle centred at the origin there is an odd number of zeros and poles, we consider the sequence $[M/3]$. In the tables below I shall compare positions of zeros and poles of the approximants in this sequence.

P.A.	zeros	poles
[3/3]	-2.06806, -1.02990 ± 3.17939i	-1.00065, $-.002435 \pm 1.58229$i
[4/3]	-1.98353, $-.963506 \pm 2.89352$i 25.5413	$-.999348$, $-.006303 \pm 1.57462$i
[5/3]	-2.06711, $-.645195 \pm 2.98225$i $-6.10166, 8.37705$	$-.999974$, $-.000237 \pm 1.57193$
[6/3]	-2.08494, $-.632818 \pm 2.91477$i $-4.85867, 10.7332 \pm 4.29698$i	-1.00003, $-.000584 \pm 1.57118$i
[7/3]	-2.06730, $-.547353 \pm 2.93852$i -4.73415 ± 2.76689i, 5.76997 ± 3.47249i	-1.00000, $-.000025 \pm 1.57091$i
[8/3]	-2.06422, $-.540386 \pm 2.91715$i -4.13620 ± 2.49272i, 11.6809 5.90481 ± 4.63395i	$-.999999$, $-.000062 \pm 1.57084$i

We clearly see that first three poles and first three zeros of $[M/3]$ converge to corresponding zeros and poles of $f(z)$ as expected from the de Montessus theorem. We could have also checked that values of $[M/3]$ converge to values of $f(z)$ in the circle of the radius smaller than $3\pi/2$ – the distance of the next pair of poles. There appeared also "stray zeros" but they were outside this circle.

Our function has infinite number of poles, so let us see how "diagonal" Padé approximants work here.

P.A.	zeros	poles
[4/4]	$-\,2.02230,\ -.906076 \pm 3.08279\mathrm{i}$ 2.07416	$-.999772,\ -.001036 \pm 1.57569\mathrm{i}$ 2.08395
[5/5]	$-2.06727,\ -.499934 \pm 2.93370\mathrm{i}$ $-2.74928 \pm 8.40343\mathrm{i}$	$-1.00000,\ -2 \cdot 10^{-6} \pm 1.57081\mathrm{i}$ $-.003750 \pm 5.06207\mathrm{i}$
[6/6]	$-2.06716,\ -.497714 \pm 2.93312\mathrm{i}$ $2.03302,\ -2.66063 \pm 8.25624\mathrm{i}$	$-1.00000,\ -1 \cdot 10^{-6} \pm 1.57080\mathrm{i}$ $2.03304,\ -.003394 \pm 5.02527\mathrm{i}$

If we remember that [7/3] and [5/5] both use the same number of the coefficients (11), we can conclude that the diagonal Padé approximants approximate our function better than approximants with a prescribed degree of the denominator. We could say, there is a price to pay: [4/4] – using 9 coefficients like [5/3] – has an unwanted pole at 2.08395. We see however that it is accompanied by a zero at 2.07416 and can (correctly) guess that values of [4/4] deviate considerably from those of $f(x)$ only close to the pair, which is called *the defect.* The analogous *defect* appears in [6/6], but the pair is much more "tight" here and we can (correctly) guess that it spoils the approximative quality of [6/6] in even smaller area close to the defect. This is just how convergence in measure (and in capacity) manifests itself.

We can also see on Fig. 2 how the behavior of some Padé approximants, mentioned above, compares with the behavior of $f(x)$ on the interval $[-6, -1.5]$ i.e. "behind" the singularity at $x = -1$.

If you are curious what happens when $f(z)$ has a multiple pole – let me tell you that in that case Padé approximants have as many single poles as is a multiplicity of that pole and they all converge to this one when order the of the approximation increases.

Finally, let me say that I would be glad if you have read the message: diagonal Padé approximants are beautiful – do not be discouraged by their defects – others can also have defects, but none are as useful.

You should not, however, think that diagonal Padé approximants are always the best ones – there are some situations when *paradiagonal* sequences of Padé approximants, i.e. sequences $[m+k/m]$ with k constant, are optimal. It can happen if we have some information on the behavior of the function at infinity. Obviously $[m+k/m](x)$ behaves like x^k for $x \to \infty$. If our function behaves at infinity in a similar way, such sequences of Padé approximants can converge faster. This is well exemplified by a study of Padé approximants for $f(x) = \sqrt{x+1}\sqrt{2x+1} + 2/(1-x)$. It has zeros at 1.60415 and -1.39193, a pole at $x = 1$ and two branch points at $x = -1/2$ and $x = -1$. Look at zeros and poles of [4/3] and [4/4], remembering that [4/4] uses one coefficient the series more.

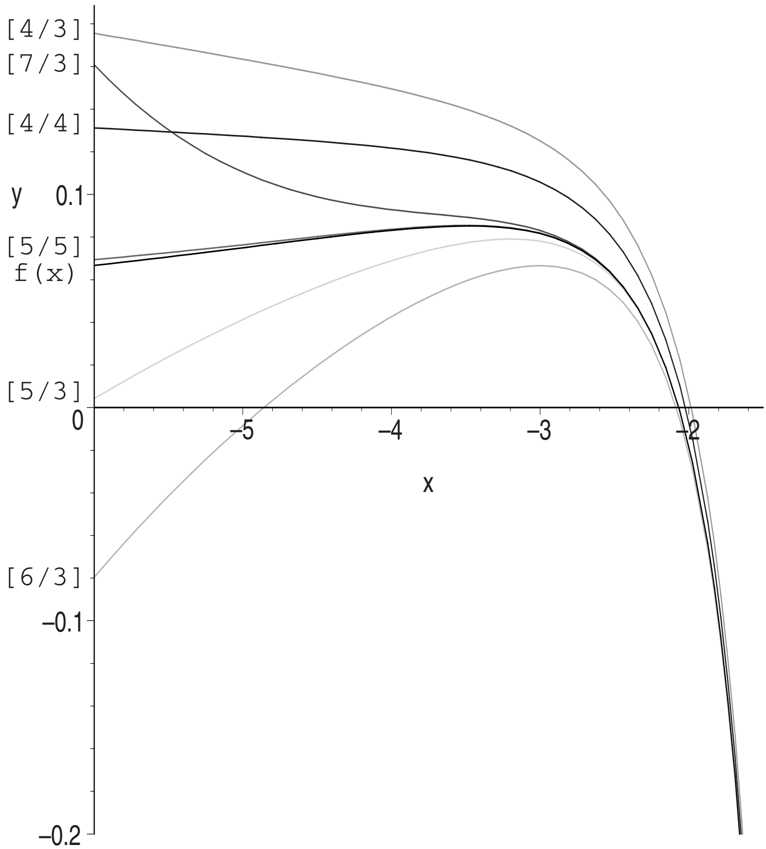

Fig. 2. Values of different PA to $f(x) = \tanh(x) + 1/(1+x)/2$

P.A.	zeros	poles
[4/3]	$-1.38833, -.754876,$ $-.556928, 1.60403$	$-.782628, -.564096$ $.999985$
[4/4]	$-1.38548, -.739811,$ $-.552442, 1.60441$	$-2499.85, -.767038$ $-.558807, 1.00002$

Positions of zeros and of the pole are clearly better reproduced by [4/3] than [4/4]. Moreover, when $x \to \infty$ [4/3](x) behaves like $1.4146x$ ($\sqrt{2} \approx 1.4142$). Additionally we see that the cut $(-1, -1/2)$ is simulated by a line of interlacing zeros and poles – the line of minimal capacity connecting branch-points.

5 Calculation of Padé approximants

In practical applications there appears a problem of how to calculate the given Padé approximants. In principle one should avoid solving a system of linear equations, because it is the process very sensitive both to errors of data and to precision of calculations. Forty and thirty years ago much activity was devoted to finding different algorithms of recursive calculation of Padé approximants. It is well documented in [3] ch. 2.4. However you can see that the system of equations for coefficients of the denominator is the one with the Toeplitz matrix and for such systems there exist relatively fast and reliable routines in all numerical programs libraries. With the speed of computers now in use, quadruple precision as a standard option in all modern Fortran compilers and also multiprecision libraries spreading around, I think that finding Padé approximants this way is in practice the most convenient solution. This is, e.g., the method used for calculation of Padé approximants in the symbolic algebra system Maple.

References

1. G.A. Baker, Jr. Existence and Convergence of Subsequences of Padé Approximants. *J. Math. Anal. Appl.*, 43:498–528, 1973.
2. G.A. Baker, Jr. *Essentials of Padé Approximants.* Academic Press, 1975.
3. G.A. Baker, Jr., P. Graves-Morris. *Padé Approximants*, volume 13 and 14 of *Encyclopedia of Mathematics and Applications.* Addison-Wesley, 1981.
4. A.F. Beardon. The convergence of Padé Approximants. *J. Math. Anal. Appl.*, 21:344–346, 1968.
5. G. Frobenius. Ueber Relationen zwischen den Näherungsbrüchen von potenzreihen. *J für Reine und Angewandte Math.*, 90:1–17, 1881.
6. J. Gilewicz. *Approximants de Padé.* Number 667 in Springer Lecture Notes in Mathematics. Springer Verlag, 1978.
7. J. Nuttal. Convergence of Padé approximants of meromorphic functions. *J. Math. Anal. Appl.*, 31:147–153, 1970.
8. H. Padé. Sur la représentation approchée d'une fonction par des fractions rationelles. *Ann. de l'Ecole Normale*, 9(3ieme série, Suppl. 3-93).
9. O. Perron. *Die Lehre von den Kettenbrüchen.* B.G. Tuebner, 1957. Chapter 4.
10. Ch. Pommerenke. Padé approximants and convergence in capacity. *J. Math. Anal. Appl.*, 31:775–780, 1973.
11. H. Stahl. Three different approaches to a proof of convergence for Padé approximants. In *Rational Approximation and its Applications in Mathematics and Physics*, number 1237 in Lecture Notes in Mathematics. Springer Verlag, 1987.