Applied Numerical Mathematics 157 (2020) 291–306

Contents lists available at ScienceDirect

Applied Numerical Mathematics

www.elsevier.com/locate/apnum

Robust formula for N-point Padé approximant calculation based on Wynn identity

Todor M. Mishonov [a,b,*], Albert M. Varonov [a,b]

[a] Georgi Nadjakov Institute of Solid State Physics, Bulgarian Academy of Sciences, 72 Tzarigradsko Chaussee Blvd., BG-1784 Sofia, Bulgaria
[b] Faculty of Physics, St. Clement of Ohrid University at Sofia, 5 James Bourchier Blvd., BG-1164 Sofia, Bulgaria

ARTICLE INFO

Article history:
Received 22 March 2019
Received in revised form 3 June 2020
Accepted 14 June 2020
Available online 22 June 2020

Keywords:
Padé approximants
Wynn identity
Multipoint Padé approximant
N-point Padé approximant
Error criterion
Cauchy-Jacobi problem
Numerical extrapolation

ABSTRACT

The performed numerical analysis reveals that Wynn's identity for the compass $1/(N-C)+1/(S-C)=1/(W-C)+1/(E-C)=1/\eta$ (here C stands for centre, the other letters correspond to the four directions of the compass) gives the long sought criterion, the minimal $|\eta|$ or p-criterion, for the choice of the optimal Padé approximant. The work of this method is illustrated by calculation of multipoint Padé approximation by a new formula for calculation of this best rational approximation. The calculation of the optimal Padé approximant by this criterion is demonstrated in calculation of series summation – frequently encountered problem in theoretical physics. This study originates from a magneto-hydrodynamic problem of heating of solar corona by Alfvén waves, where the present method is used for a predictor in solution of differential equations. In such a way, an efficient and valuable control mechanism for N-point Padé approximant calculation is proposed. We believe that the suggested method and criterion can be useful for many applied problems in numerous areas not only in physics but in any scientific application where differential equations are solved. The obtained new solution of the Cauchy-Jacobi problem is illustrated by a Fortran program. The algorithm is generalized for the case of the first K-derivatives at N-nodal points. The numerical analysis of probability distribution function of the errors will review, which formula for extrapolation of functions deserves to be implemented in the applied mathematics software like Mathematica and Maple.

1. Introduction and motivation

As the analytical structures take central place in theoretical and mathematical physics, the Padé approximants of functions become one of the most important problems of applied mathematics. For a general introduction in the problem of Padé approximants, see the well-known monographs Refs. [5,10,8]. Since the problem was first systematically studied by Frobenius [7,13] and the Padé table introduced by Padé [7,27] in the 19th century, the contemporary literature is enormous. We mention only two well-known problems, which can be of broad interest for physicists: 1) series summation and 2) extrapolation of functions related to predictor-corrector methods for solution of ordinary differential equations. Searching for numerical recipes in this field, a beginner can read that the downside of the Padé approximation is that it is uncontrolled and the dangers of extrapolation cannot be overemphasized [28]. And after that our beginner will most probably

* Corresponding author.
E-mail addresses: mishonov@bgphysics.eu (T.M. Mishonov), varonov@issp.bas.bg (A.M. Varonov).

https://doi.org/10.1016/j.apnum.2020.06.007

start searching for a commercial software, where the appropriate analytical formulae are programmed. By the very famous law: whatever can go wrong, will go wrong, in very interesting for the physics problems, the commercial software stops working and the research would continue only in co-authorship with the source code authors.

In 1947 Hannes Alvfén [3] suggested that the solar corona heating (the corona is 2 orders of magnitude hotter than the photosphere) is due to the absorption of magneto-hydrodynamic (now called Alfvén) waves. The programming of the magneto-hydrodynamic equations is relatively easy, but after 30-time temperature increase all available for us commercial computation software stops working. Facing such a problem, we had to develop an own method for a frequently encountered numerical problem (four first order linear differential equations coupled to four first order non-linear differential equations) [25,26]. Our research goal was to develop this method based on calculation of Padé approximants with a mechanism to select the most accurate approximant. This mechanism is precisely the main development of the current research and allows the usage of the described method for solution of differential equations. That is why we suppose that many colleagues are in our situation and we share know-how in calculation of Padé approximants.

Before starting let us recall the definition of multipoint Padé approximation [5]. A rational function $f(z)$ which fits given function values y_i at various points x_i, $i = 1, \ldots, N$, i.e. $f(x_i) = y_i$ is called multipoint Padé approximant. The associate problem of interpolation by rational functions is called Cauchy [12]-Jacobi [18] problem [4]. If the Cauchy-Jacobi problem is solvable then the multipoint Pade approximant is the solution [29]. Multipoint Padé approximants are also called: rational interpolants, N-point Padé approximants, Newton-Padé approximants, etc., depending on the context. Addressing to experts in the present paper we present a general robust solution of the Cauchy [12]-Jacobi [18] problem which can attract interest and become a convenient tool in the applied numerical analysis.

2. Announcement of the results

For a practical implementation of an analytical result, we have to use a computer with finite accuracy of digital representation of real numbers. The Padé approximants give fast convergence and actually the best rational approximation but they are sensitive to the noise of the discrete representation of real numbers. Roughly speaking, each value representing a real number suffers from error of discretization as it is multiplied by a random factor $1 + \varepsilon_{\text{mach}}(\text{rnd} - 1/2)$, where $\varepsilon_{\text{mach}}$ is the machine epsilon (the biggest positive value for which $1 + \varepsilon_{\text{mach}} = 1$), and "rnd" is a programming operator generating a random number homogeneously distributed between 0 and 1. As formulae for Padé approximants picturally speaking are very sensitive to the random noise of truncation, it is necessary to study the influence of the machine error $\varepsilon_{\text{mach}}$ on the final result. In such a way beyond the analytical problem we have also a statistical one which already belongs to another branch of mathematics. Probably for some special cases it is possible to calculate the probability distribution functions (PDF) of the errors of the final result of calculation of Padé approximants, but it deserves to start with descriptive statistics of some well-known examples, which in our opinion illustrate how to calculate and choose the optimal Padé approximant in physical applications. Here the determination of the optimal Padé approximant is crucial, as our beginner has already learned from the numerical recipes book. To experts we propose a robust method which at least empirically works well.

Let $r_{l,m}(z)$ be different Padé approximants of the complex function $f(z)$

$$f(z) \approx r_{l,m}(z) = \frac{\sum_{i=0}^{l} a_i z^i}{\sum_{k=0}^{m} b_k z^k}, \qquad b_0 = 1, \tag{1}$$

where l is the maximal power of the numerator and m is the maximal power in the polynomial in the denominator. The first index in the notations of our paper $r_{l,m} = [l/m]$ is for the power of the numerator l and the second one m is for the power of the denominator. It is common practice to represent different approximants $r_{l,m}$ in a table with called a Padé table [5, Chap. 1, Sec. 1.1, p. 7, Tables 1 and 2]. For these approximants in 1966 Wynn discovered the remarkable relation [31, Eqs. (15) and (16)]

$$\frac{1}{\eta_{lm}} \equiv \frac{1}{r_{l,m+1} - r_{l,m}} + \frac{1}{r_{l,m-1} - r_{l,m}} = \frac{1}{r_{l+1,m} - r_{l,m}} + \frac{1}{r_{l-1,m} - r_{l,m}}, \tag{2}$$

later on baptized by Gragg [17, Theorem 5.5] as *missing identity of Frobenius* [13]. According the Arnold's principle: if a notion bears a personal name, then this name is not the name of the discoverer; while mathematicians such as Bernoulli, Cauchy, Jacobi, and Frobenius developed many of the ideas above long before Hermite or Padé [22]. The criterion for choosing an optimal Padé approximant is based on this Wynn relation which can be rewritten also as

$$\frac{1}{\eta} \equiv \frac{1}{S - C} + \frac{1}{N - C} = \frac{1}{E - C} + \frac{1}{W - C}, \tag{3}$$

$$C \equiv r_{l,m}, \quad N \equiv r_{l,m-1}, \quad E \equiv r_{l+1,m}, \quad W \equiv r_{l-1,m}, \quad S \equiv r_{l,m+1}. \tag{4}$$

If C is in the Centre, the notations of the other elements N, E, W, S follows the direction of the compass: North, East, West, and South, also found in the original paper of Wynn [31] (i.e. we use Central NEWS algorithm). Probably this representation

was an idea of Gragg, who has also described it [17], because of the footnote in Wynn's paper on page 266 (if that is the case, Gragg was a referee of Wynn's paper). The notations originate from the comparison of the Padé table with a geographical map: index l represents a horizontal position and index m a vertical one. For convergent Padé approximants

$$f(z) = \lim_{l,m\to\infty} r_{l,m}(z) \tag{5}$$

in the Padé table C, N, E, W, $S \to f(z)$, $1/\eta \to \infty$, and

$$\lim_{l,m\to\infty} \eta_{lm} = 0. \tag{6}$$

In numerical implementations this limit leads to minimal-$|\eta|$ criterion of the CNEWS algorithm.

Let us repeat in other words. Often the difference between sequential Padé approximants gives a reasonable evaluation of the error, see Ref. [5]. In case of convergence, we have vanishing differences between the values of different cells of the Padé table $(r_{l,m+1} - r_{l,m}) \to 0$, for $l, m \to \infty$ and $l/m = \text{const}$. In this case also $\eta_{lm} \to 0$ and the minimal value of η_{lm} gives a reasonable criterion for the minimal error and the choice of the optimal Padé approximant. In many examples of calculation of Padé approximants we arrive at the conclusion that the long sought criterion for the choice of the optimal Padé approximant is simply a search for the minimal value $|\eta_{l,m}|$, i.e.

$$\mathrm{p} \equiv \eta_{\min} \equiv |\eta_{L,M}| = \min_{l,m}|\eta_{l,m}| \tag{7}$$

and $f(z) \approx r_{L,M}$. The letter p (with spelling wynn, also spelled wen, pynn, or pen) is a letter of the Old English alphabet, where it is used to represent the sound /w/, is used in honor of one of the creators of numerical analysis Peter Wynn (mathematician, born 1932) with main achievements concern approximation theory, in particular the theory of Padé approximants and its application in numerical methods for improving the rate of convergence of sequences of real numbers [9]. We consider that the Wynn identity gives the long sought criterion for the choice of the optimal Padé approximant. It is technologically and aesthetically attracting that the Wynn identity gives simultaneously 1) explicit method for the calculation of Padé approximants and 2) method for the evaluation of the error; $\lim \eta_{\min} = 0$ is actually a criterion for convergence of Padé approximation if we consider real numbers. Our criterion is an alternative of the singular value decomposition (SVD) method described in great detail in Refs. [16,6], see also Refs. [19,14,15,21,32]. Our minimal η or p criterion and the tolerance level of SVD are perhaps different implementations of one and the same idea. For both methods the rounding error is something external for the theory of Padé approximation. There are no theoretical justifications which method is better. The numerical experiment and the descriptive statistics are just the first steps in the practical realizations on the robust Padé approximants.

The purpose of the present work is to demonstrate how this criterion works and how it can be used in applications. In the next section we illustrate in detail this criterion in the problem of calculation of limits of numerical series.

3. The new algorithm for determination of the optimal Padé approximant

Let us have a numerical sequence $\{S_0, S_1, S_2, \ldots, S_N\}$ and we need to calculate the limit $\tilde{S} = \lim_{l\to\infty} S_l$. The well-known method is to initialize $r_{l,0} = S_l$, for $l = 0, \ldots, N$, and also $r_{-1,l} = \infty$ for $l = 0, \ldots, N+1$. In other words we treat every term of the initial series as polynomial approximation of some function

$$r_{l,0} = \sum_{i=0}^{l} a_i z^i = S_l. \tag{8}$$

The method can work even if this sequence is convergent, and those are the most interesting cases for applications. Then we calculate in the "south direction" the corresponding Padé approximants

$$S = C + \cfrac{1}{\cfrac{1}{E-C} + \cfrac{1}{W-C} - \cfrac{1}{N-C}} \tag{9}$$

or

$$r_{l,m+1} = r_{l,m} + \cfrac{1}{\cfrac{1}{r_{l+1,m} - r_{l,m}} + \cfrac{1}{r_{l-1,m} - r_{l,m}} - \cfrac{1}{r_{l,m-1} - r_{l,m}}} \tag{10}$$

and simultaneously calculate the empirical error

$$\eta_{lm} = \cfrac{1}{\cfrac{1}{r_{l+1,m} - r_{l,m}} + \cfrac{1}{r_{l-1,m} - r_{l,m}}} \equiv \eta = \cfrac{1}{\cfrac{1}{E-C} + \cfrac{1}{W-C}}, \tag{11}$$

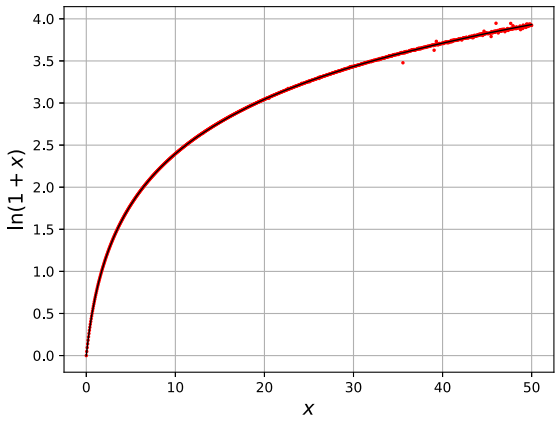

Fig. 1. The logarithmic function (line) and the series summation with $\eta_{\min}$ criterion for the optimal choice (dots). The calculated values are *evaporated* from the analytical curve, but $\eta_{\min}$ criterion gives reliable warning depicted in Fig. 2.

with special cases

$$\eta = 0, \qquad \text{if} \quad E = C \quad \text{or} \quad W = C. \tag{12}$$

The special case $1/(N-C) = 1/(W-C) + 1/(W-C)$ of zero denominator in Eq. (9) means that the reciprocal value of the quantity we calculate $1/S \approx 0$ is within the used accuracy. The minimal absolute value $\mathrm{p} \equiv \eta_{\min}$ is our criterion for the determination of the optimal Padé approximant. According to the best we know, this minimal-$|\eta|$ criterion has never been implemented in the numerical recipes so far (authors will very much appreciate any information about this). In the beginning when $N = \infty$, the Wynn identity Eq. (9) gives the well-known Δ-process by Aitken $S \approx C + \eta$, which is exact for the geometrical progression series.

For the programming task, we have to calculate all the values, for which division is possible. In the next section we thoroughly describe 2 technical examples.

4. Two easy pieces

4.1. Calculation of divergent series $\ln(1+x)$

Perhaps the most famous example summation of divergent series by Padé approximants is the calculation of the divergent Taylor series $\ln(1+x)$ beyond the radius of convergence $|x| < 1$

$$\ln(1+x) = \tilde{S} \equiv \lim_{n\to\infty} S_n, \qquad S_n(x) \equiv \sum_{k=1}^{n} (-1)^{(k+1)} \frac{x^k}{k}. \tag{13}$$

The application of Eq. (10) and Eq. (11) gives the optimal Padé approximant for every positive x shown in Fig. 1. Even for $x \approx 20$ for the calculation of the series $\ln(1+x)$ a pixel accuracy is present. In many articles results of numerical calculations are illustrated by the figures presented by computer graphics with million pixels. For such a figures pixel accuracy means maximal error in which the approximation is undistinguished from the exact result when graphically presented on screen. For many applied studies this is an acceptable beginning. For larger values of x the dots representing the calculation *evaporate* from the line representing the exact value. The most important detail is how and when the method stops working and when the resources of the numerical accuracy are exhausted. The real $\varepsilon_{\text{real}}$ and empirical $\varepsilon_{\text{emp}} \equiv \eta_{\min}$ of the calculation are shown in Fig. 2. The empirical error $\varepsilon_{\text{emp}} \equiv \eta_{\min}$ shows saturation and is a reliable indicator for the calculation accuracy. Statistical analysis of the correlation between both errors from Fig. 2 shows very strong dependence between them in Fig. 3. Now we can safely tell how accurate our Padé approximation is and even how far out in x it can be extended. This dependence reveals that the long sought criterion for empirical evaluation of the accuracy of the Padé approximants calculated by ε-algorithm has already been found.

4.2. Extrapolation of a sine arch from a preceding one

The second technical example we present is the problem of extrapolation of function which we illustrate in the case of the sine function. We take N equidistant interpolation points from one arch of the $\sin(x)$ function and extrapolate the next arch and even beyond. We use the well-known Aitken interpolation method [2,11,1,20] to order the interpolation points within the arch and the point to be extrapolated. In more details: we calculate $y_i = \sin(x_i)$ for all N points $i = 1, \ldots, N$. The digital noise is minimized if the points are ordered with the proximity with the point of interpolation, i.e.

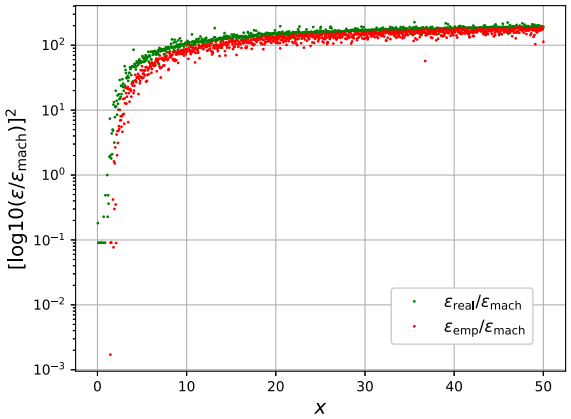

Fig. 2. Square of decimal logarithms of empirical $\varepsilon_{\mathrm{emp}} \equiv \eta_{\mathrm{min}}$ and real $\varepsilon_{\mathrm{real}}$ error versus the argument of the function for the calculation of the $\ln(1+x)$ series in Fig. 1. Close to the convergence radius the errors are small and almost linearly increase, then the errors reach saturation when the numerical resources of the fixed accuracy are exhausted.

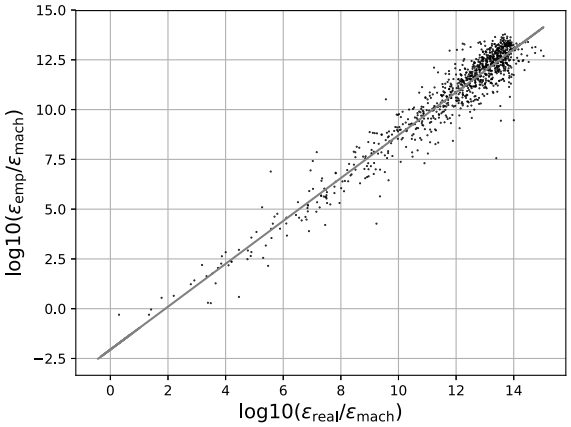

Fig. 3. The logarithm of the empirical $\eta_{\mathrm{min}} \equiv \varepsilon_{\mathrm{emp}}$ versus logarithm of the real $\varepsilon_{\mathrm{real}}$ error for the calculation of the $\ln(1+x)$ series in Fig. 1. The high correlation coefficient 0.961 of the linear regression reveals that the long sought criterion for empirical evaluation of the accuracy of the Padé approximants calculated by ε-algorithm has already been found.

$$|x - x_1| < |x - x_2| < |x - x_2| < \cdots < |x - x_N|. \tag{14}$$

Then using sequentially l interpolation points in the Aitken scheme we calculate the corresponding polynomial approximation S_l for $l = 1, \ldots, N$. Next we apply the CNEWS method and use the minimal-$|\eta|$ for the best Padé approximant. In this manner we have a numerical sequence that we give to the ε-algorithm to calculate its limit. The described calculation for the $\sin(x)$ function is shown in Fig. 4. The preceding arch is in the interval $(-\pi, 0)$ and contains 21 interpolation points. Using these points, we extrapolate one arch with 2000 points in the interval $(0, \pi)$ and continue the extrapolation in an attempt to obtain a second arch with the same number of points in the next interval $(\pi, 2\pi)$. We repeat the procedure: every extrapolated point is calculated independently using the Aitken method for interpolation followed by calculation of Padé approximants by the Wynn identity and the optimal Padé approximant is chosen by the advocated in the present paper η_{min} criterion. The deviation of the points from the real function represented with the line in Fig. 4 shows the limit of applicability of the Aitken-Wynn extrapolation algorithm. Detailed error estimates of the extrapolation are shown in Fig. 5. The η_{min} criterion shown in Fig. 5 gives the reliable order estimation of the error. The similar behaviour of both errors shows that our criterion is a reliable method for error estimation. Furthermore, a statistical analysis of the correlation between the real $\varepsilon_{\mathrm{real}}$ and the empirical error $\varepsilon_{\mathrm{emp}}$ shows very strong dependence in Fig. 6.

The literature on the problem of extrapolation is enormous, however we have not found a discussion of the problem in mathematical statistics. If we know the values of the function with some accuracy, how to choose the best Padé approximant among the many, some of which have serious noise of rounding and others are even *uncontrolled* [28]. There is a huge literature about the formulae, but a convenient criterion for the choice of optimal Padé approximant is not discussed. We give our illustrative example not as an exercise in programming, but as an illustration that we have a good working criterion for many practical cases. As a rule, articles on Padé approximation are illustrated by examples describing how some method is good working. But the instructive analysis reveals where the method stops working and most importantly

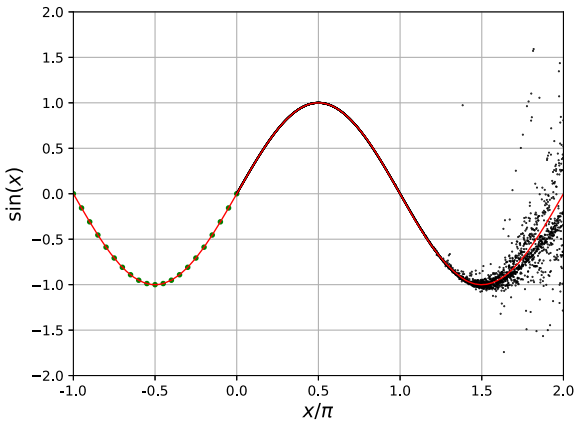

Fig. 4. Extrapolation of the function $\sin(x)$ (small dots) from 21 interpolation points (larger dots) in the interval $[-\pi, 0]$ compared with the real function (line). In the interval $(0, \pi)$ the function is reliably extrapolated and the limit of the numerical implementation of the Aitken-Wynn extrapolation is clearly shown – a *gas* of extrapolated points *evaporated* from the analytical function.

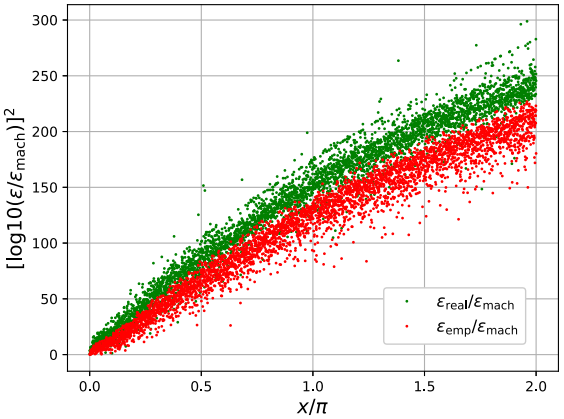

Fig. 5. Squared logarithm of the error estimates ε of the $\sin(x)$ extrapolation shown in Fig. 4. The error $\varepsilon_{\text{real}}$ is the real error of the extrapolation and the error $\varepsilon_{\text{emp}} \equiv \eta_{\min}$. The similar behaviour of both errors shows that our criterion is a reliable method for error estimation, which is evident in Fig. 6. The important problem in front of the applied mathematics is to research real extrapolation error beyond the extrapolation interval.

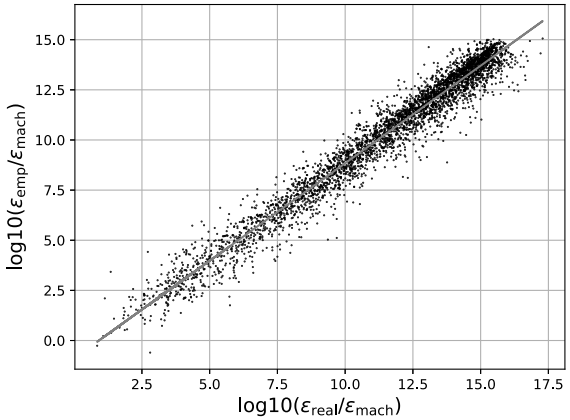

Fig. 6. Logarithmic dependence of $\eta_{\min} \equiv \varepsilon_{\text{emp}}$ as a function of the real error $\varepsilon_{\text{real}}$ for the extrapolation shown in Fig. 4. The slope of the line of the linear regression is 0.973, the intercept is -0.887 with a correlation coefficient 0.979. This strong correlation sets in the agenda the problem of the statistical properties of the Padé approximants and the investigation of the corresponding probability distribution functions (PDF).

whether any criterion warns that calculation resource has been exhausted. From a practical point of view, exhaustion of resources for extrapolation is demonstrated by *evaporation* of the extrapolated points from the analytical curve, and one of the achievements of the present work is that order of the magnitude of this deviation is reliably indicated by our criterion based on the Wynn identity.

5. Discussion and conclusions

In spite that the literature for Padé approximants is enormous we were unable to find alternative formulas, methods, algorithms and programs to compare the work of our method. One can put in the agenda for the development of the numerical analysis a simple test: which method for extrapolation of functions from tabulated values in N points gives the best extrapolation of a function far from the interpolation interval. This juxtaposition will give the final verdict which method is appropriate to be implemented in commercial software like Mathematica and Maple.

The new result of the implementation of calculation of Padé approximants and their application is the modulus minimization $\eta_{\min}$ of Wynn's η_{lm} as reliable empirical criterion of the error. The preformed analysis of several simple examples reveals that for practical implementation of Padé approximants we can use the empirical $\eta_{\min}$ error extracted from the Wynn's identity. The comparison of the descriptive statistics data for the PDF of errors of calculation of Padé approximants by different criteria will give the answer what general recommendation as a numerical recipe has to be given to users not willing to understand how.

In short, the practical implementation of Padé approximants can reach one order of magnitude more applications in theoretical physics and applied mathematics. More than half a century after its discovery, the ε-algorithm has not been included for calculation of divergent series with convergent Padé approximants and for extrapolation of functions in commercial software. Now the time for this inclusion has come, the herein implemented control mechanism has rendered this mission possible.

Last but not least, the suggested criterion Eq. (11) is applicable in solution of differential equations, numerically performed analytical continuation, perturbation, series summation and other analogous problems in theoretical physics.

Let us summarize our results: 1) we have suggested a method for empirical choice of optimal Padé approximant which works for all examples we have met in the literature and seems to have universal applicability, 2) we derived a new formula for the N-point Padé approximant, for a function tabulated in N-points. This formula based on Aitken interpolation formula and Wynn identity is maximally robust with respect to truncation errors in the numerical calculations and can be used even for extrapolation. We have developed this method in order to obtain a predictor method for predictor-corrector methods for solving of one irrelevant for the present paper magneto-hydrodynamic problem, the programs where our formulae are implemented are given as appendices in the arXiv version of the present paper [24] and early and more detailed version has already been published in conference proceedings [23]. The figures in the paper are calculated by those programs. Reproducing the figures is the criterion that our formula can be used directly in numerical calculations. Up to now a non-specialist knows that extrapolation and summation of divergent series is a forbidden procedure used only by elite mathematicians. We have made a popularization addressed to user of applied mathematics working for the industry. 3) The described method can be used as a predictor method for solving ordinary differential equations $\mathrm{d}y/\mathrm{d}x = \mathcal{F}(x, y)$. The described N-point Padé approximation can be used as a predictor using former calculated points and their first derivatives as interpolation points to calculate the predicted value of the function y_{pred} at the new point x_{new} of the argument. Then we calculate the derivative at the new point $\mathcal{F}(x_{\mathrm{new}}, y_{\mathrm{pred}})$. We consider this possibility as a new method for solution of ordinary differential equations. The method can be additionally improved if in the $(x\text{-}y)$ plane we perform a rotation, and the new abscissa of the interpolation to be taken along the local tangent of the curve. We started with the problem of calculation of the temperature profile of the solar corona and velocity of solar wind [26] but this method can be applied for many other problems for which standard software is not (good) working. 4) If we know the first K derivatives of the function using of Taylor expansion we can calculate

$$\tilde{y}_i = y_i + (x - x_i)\, y_i' + \frac{1}{2}(x - x_i)^2 y_i'' + \cdots + \frac{1}{K!}(x - x_i)^K\, y_i^{(K)}. \tag{15}$$

Then we can use the Taylor approximants $\tilde{y}_i$ as interpolation points x_i in the already described method. At the basic points the interpolated rational function will have exact coincidence not only for the function but also for its first K derivatives. The main idea of the algorithm is the same: using a node point x_i we calculate the interpolated value $\tilde{y}(x)$ at the argument of interpolation, and then use this calculated value $\tilde{y}(x; x_i)$ at this point as a new value in the node point $y_i = f(x_i) \leftarrow \tilde{y}(x; x_i)$. If we need a formal proof of convergence it is necessary in the right side of Eq. (15) to put insert the evaluation of error $h^{(K+1)} M^{(K+1)}/(K+1)!$, where h is the maximal distance between interpolating an node argument and $M^{(K+1)}$ is the supremum of the modulus of the $(K+1)$ derivative. This estimation of the error reveal why the analytical approximation is not applicable to function $|x|$ and the trajectory of a Brownial particle.

Finally we have to remark. There is no general criterion for the convergence of the Padé approximation, and even to choose the optimal approximant when the method is convergent, that is why there are no theorems which could be cited. As all known existing methods are given without any empirical prescription for the choice of the optimal approximant, the considered minimal-$|\eta|$ criterion of the CNEWS method can not be compared with other methods. The area of the

applicability of the suggested method is however well defined. For many applied problems it is known that solution exists, while for the analytical problems the solution is analytical. In many problems in physics, for example, there are alternative methods to check whether the sought solution is numerically obtained in spite that there will be no alternative methods. Contra-examples of mathematical series for which the Padé method does not work can be easily constructed. We are unaware of a real problem for which the solution is experimentally measured but the series can not be summarized by the exciting combinations of Euler-McLaurin summation and Padé approximation. That is why we believe that the suggested method, criterion and software can be useful for many applied problems in numerous areas not only in physics but in any scientific application where differential equations are solved. While other solutions of the Cauchy-Jacobi problem give calculation of the coefficients of the Padé approximants, we give a fast method for the calculation of its numerical value.

Acknowledgements

The authors are thankful to Evgeni Penev for the collaboration in early stages in the present research, and for writing the ε-algorithm in Fortran-90, Michail Mishonov gave a version in Java and Nedeltcho Zahariev and Zlatan Dimitrov in C-language. The authors are also grateful to Claude Brezinski for pointing out and recommendation of Refs. [16,6], to Ognyan Kounchev for Refs. [19,14,15,21,32], to Emil Horozov, Radostina Kamburova, Dantchi Koulova, Peter Kenderov and Hassan Chamati for considerations, comments and stimulating discussions. This publication is supported by the Union of the Physicists of Bulgaria, special acknowledgements to the president acad. Alexander Petrov. This work is partially supported by National program "Young scientists and postdoctoral researchers" approved by DCM 577, 17.08.2018. Last but not least, authors of the manuscript are grateful to the anonymous reviewers for their careful and thorough reading of our manuscript, and its constructive criticism.

Appendix A. Wynn CNEWS-algorithm source code in Fortran-90

This section includes the Wynn CNEWS algorithm written in Fortran-90 subroutine and used for the calculations presented in the figures. This Fortran program is our understanding of the Cauchy [12]-Jacobi [18] problem an if the time of floating point operation $\tau_{\text{FPO}} \to 0$ and also the machine epsilon $\varepsilon_{\text{mach}} \to 0$ we obtain in the God computer the analytical result. The program can be converted in a constructive proof.

In the foolproof programming it is necessary to point out what to do if any denominator is zero, for example the special cases marked in Eq. (12). Here we have to make a social remark related to teaching. Equality of some real numbers presented by double precision looks like event with a negligible probability. However especially in our case it is not true. The program for calculation of limit of e sequence is working like a microscope magnifying the real axes around the searched value. Very often the empirical error p is comparable with machine error $\varepsilon_{\text{mach}}$ and the dimensionless ratio $N_{\mathrm{p}} \equiv \mathrm{p}/\varepsilon_{\text{mach}} \gg 1$ can be not very big number, say $N_{\mathrm{p}} \simeq 10^5$. In this case the probability $P \simeq 1/N_{\mathrm{p}}$ to use special case avoiding division by zero is not very rare event $P \simeq 10^{-5}$ and even a personal computer can detect equality of real numbers ones per hour. The following Fortran listing contains a lot of analogous details. We note that the Padé index convention in the following Fortran program is transposed to that in the current paper.

```
Subroutine LimesCross(N, S, rLimes, i_Pade, k_Pade, i_err, error, Nopt)
          implicit integer (i–n)
          implicit real(8) (a–h), real(8) (o–z)
          dimension s(0:N), a(0:N+2), b(0:N+1), c(0:N+1)
          Zero=0.d0
          One =1.d0
          rLimes=s(0)
          i_Pade=0
          k_Pade=0
          rNorm=Zero

          if (N<2) then
                  i_err=1 ! N>=2
                  Nopt=2
                  rLimes=s(N)
                  goto 137
          endif

          do i=0,N
                  b(i) = s(i)
                  c(i) = Zero
                  rNorm=rNorm+dabs(s(i))
          end do
```

```
      if (rNorm==0) then
              i_err=2 ! S_i=0

              goto 137
      endif

      error= dabs(s(1)-s(0))
      do i=2, N
      eloc= dabs(s(i)-s(i-1))
      if(eloc<error) then
              error = eloc
              k_Pade=0
              i_Pade=i
              rLimes=s(i)
              i_err=3
              Nopt=i_Pade + k_Pade+1
      endif
      end do

      Nb=N-2 ! N boundary
      k=0
123   continue

      !       beginning of the calculation of Pade table P_{k_Pade,i_Pade}
      !       k_Pade= power of the denominator
      !       i_Pade= power of the numerator
      !       Table ordered as a matrix numbering

      do i=0, Nb
              Center= b(i+1)
              if(k>=1) then
                      rNorth= a(i+2)
                      rNC= rNorth-Center
                      if(rNC==Zero) then
                              rLimes= Center
                              k_Pade=k
                              i_Pade=k_Pade+i
                              Nopt=i_Pade + k_Pade +1
                              error=zero
                              i_err=4
                              goto 137
                      endif
              endif
              West= b(i)
      WC= West-Center
      if(WC==Zero) then
          rLimes= Center
          k_Pade=k
          i_Pade=k_Pade+i +1
          Nopt=i_Pade + k_Pade +1
          error=zero
          i_err=5
          goto 137
      endif
      East= b(i+2)
      EC= East-Center
      if(EC==Zero)  then
          rLimes= Center
          k_Pade=k
          i_Pade=k_Pade+i+2
          Nopt=i_Pade + k_Pade +1
```

```
                error=zero
                i_err=6
                goto 137

            endif
!                       The new criterion Wynn–Frobenius Eq. (5) from the present paper
                        Wynn=One/WC + One/EC
                        if(Wynn.NE.Zero) eloc=One/dabs(Wynn) ! in search of local minimum
                        denom=Wynn
                        if(k>=1) denom=denom–One/rNC
                        if(denom==Zero) then
                                i_err=7 ! 1/rLimes=0, i.e. divergent result
                                Nopt=i_Pade + k_Pade +1
                                goto 137
                        endif
                        South= Center+One/denom
                        C(i)=South
                        SC=South–Center
!                       Optimal Pade approximant
            if(eloc<error) then
                error= eloc
                                i_err=0
                k_Pade=k+1
                i_Pade=k_Pade+i
                Nopt=i_Pade + k_Pade +1
                rLimes= South ! = P(k_Pade,i_Pade)
            endif
        end do ! i cycle
        k=k+1
!       Change of the succesions
        do i=0, Nb
                        a(i)=b(i)
                        b(i)=c(i)
            enddo
!           Reduction of the boundary
        Nb=Nb–2
        if(Nb>=0) goto 123
!       Exit
137     continue
        end Subroutine LimesCross
```

Appendix B. Aitken algorithm for polynomial interpolation

In order to explain the novelty of the present achievement in the beginning we will recall the well known Aitken algorithm. Later on using this basis we will explain the general polynomial exponential. We follow the well-known reference book by Bronstein and Semendyayev [11]. If we need to calculate value of the function $\varphi_n(x)$ at some fixed argument x, one can use following schematics ("crest by crest") which is especially convenient for computer programming

$$
\begin{array}{lllll}
x_0 - x & f_0 & & & \\
x_1 - x & f_1 & (f_0, f_1) & & \\
x_2 - x & f_2 & (f_0, f_2) & (f_0, f_1, f_2) & \\
\quad \ldots\ldots\ldots\ldots\ldots & & & & \\
x_n - x & f_n & (f_0, f_n) & (f_0, f_1, f_n) & \ldots \quad (f_0, f_1, \ldots, f_n).
\end{array}
$$

Every symbol $(f_0, f_1, \ldots, f_k)$ denotes the value in the point x of the interpolating polynomial function, build on nodes $x_0, x_1, \ldots, x_k$, i.e. $f_0 = f(x_0)$, $f_1 = f(x_1)$, …, $f_n = f(x_n)$. Those values calculate column by column, in the following manner. Numbers of the column (f_0, f_k) are obtained by

$$(f_0, f_k) = \frac{(x_0 - x) f_k - (x_k - x) f_0}{(x_0 - x) - (x_k - x)}. \tag{B.1}$$

Every following column is calculating by the former one at the same schematics, for example

$$(f_0, f_1, f_k) = \frac{(x_1 - x)(f_0, f_k) - (x_k - x)(f_0, f_1)}{(x_1 - x) - (x_k - x)}. \tag{B.2}$$

The rounding error is minimized if we reorder the points as $|x_0 - x| < |x_1 - x| < |x_2 - x| < \dots |x_n - x|$. In such a manner we obtain the series

$$S_0 \equiv f_0, \quad S_1 \equiv (f_0, f_1), \quad S_2 \equiv (f_0, f_1, f_2), \quad \dots, \quad S_n \equiv (f_0, f_1, f_2, \dots, f_n). \tag{B.3}$$

In order to obtain an approximation of limit of this sequence

$$\tilde{S} = \lim_{n\to\infty} S_n \tag{B.4}$$

we use the Wynn algorithm CNEWS algorithm with minimal-$|\eta|$ criterion, which works even if the sequence $(S_0, S_1, S_2, \dots)$ is divergent. It is only necessary the problem which we solve to be analytical one born from a real physical problem. One can easily check that if $x = x_0$ then $S_0 = S_1 = S_2 = \dots = S_n = f_0$ and the limit has been reached, i.e. the function passes through the node point x_0, and consequently through every nodal point.

Let us repeat the Aitken formula as an algorithm. Having fixed node point x_0 we can calculate the linear approximation using a second node point $\tilde{x}$

$$\varphi(x; x_0, \tilde{x}) = f(\tilde{x}) + \frac{f_0 - f(\tilde{x})}{x_0 - \tilde{x}}(x - \tilde{x}) = \frac{(x_0 - \tilde{x}) f(\tilde{x}) - (\tilde{x} - x) f_0}{(x_0 - x) - (\tilde{x} - x)}, \tag{B.5}$$

where node arguments x_0 and $\tilde{x}$ are parameters of the linear function. In such a way omitting in the notations the argument x we can calculate the sequence

$$(f_0, f_1) = \varphi(x; x_0, x_1), \quad (f_0, f_2) = \varphi(x; x_0, x_2), \quad (f_0, f_3) = \varphi(x; x_0, x_3) \quad , \dots, \quad (f_0, f_n) = \varphi(x; x_0, x_n). \tag{B.6}$$

In order to rewrite this conveniently for programming algorithm, we can use the same notations for the changing new values

$$f_1 \leftarrow (f_0, f_1), \quad f_2 \leftarrow (f_0, f_2), \quad f_3 \leftarrow (f_0, f_3), \quad \dots, \quad f_k \leftarrow (f_0, f_n). \tag{B.7}$$

Then for the next column of the Aitken table according Eq. (B.2) using so re-denoted values we calculate

$$f_k \leftarrow \frac{(x_1 - x) f_k - (x_k - x) f_1}{(x_1 - x) - (x_k - x)}, \qquad \text{for} \quad k = 2, 3, \dots, n. \tag{B.8}$$

The formula for a general column l is the same

$$f_k \leftarrow \frac{(x_l - x) f_k - (x_k - x) f_l}{(x_l - x) - (x_k - x)} = \frac{\begin{vmatrix} (x_l - x) & f_l \\ (x_k - x) & f_k \end{vmatrix}}{x_l - x_k}, \qquad \text{for} \qquad k = l + 1, \dots, n; \tag{B.9}$$

i.e. linearly interpolated value of the function at point x is ascribed as the new function value at the node point x_k. In other words, the polynomial interpolation is reduced to sequential linear interpolations This procedure repeats until $l = n - 1$. In such a way we program the series identically by the so defined polynomial interpolations $(f_0, f_1, f_2, \dots, f_n) = (S_0, S_1, S_2, \dots, S_n)$ using sequential $(x_0), (x_0, x_1), (x_0, x_1, x_2), \dots, (x_0, x_1, x_2, \dots, x_n)$ set of nodes – this is the Aitken algorithm. According to the general theory of Padé approximants, the Wynn CNEWS algorithm converts the series of polynomial approximation into Padé ones, and for $N = n - 1$ we obtain the N-point Padé approximant. The calculation of the coefficients of the polynomials is a completely different mathematical problem. In the present paper we analyze the calculation of the approximant at some fixed point x, but not the coefficients in the polynomials in the numerator and denominator.

At solution of differential equations the node points can be ordered $x_n < x_{n-1} < \dots < x_2 < x_1 < x_0 < x$. In this case the numerical noise of rounding at calculations of differences can be minimized if we use the differences

$$x = x_0 + h, \quad x_1 = x_0 + h_1, \quad x_2 = x_1 + h_2, \quad x_3 = x_2 + h_3, \quad \dots, \quad x_n = x_{n-1} + h_n. \tag{B.10}$$

In this case is convenient to use the representation

$$x_k - x_l = h_k + h_{k-1} + \dots + h_{l+1} = \sum_{i=l+1}^{k} h_i, \quad \text{for} \quad k > l, \qquad \text{and} \quad x_k - x = (x_k - x_0) + (-h), \tag{B.11}$$

where all terms have one and the same sign. For the function values one can perform an analogous representation.

If the first K points are going to coincide $x_0 \to x_1, \ldots, \to x_K$ the corresponding polynomial approximant is just the Taylor series

$$\phi(x; x_0) = f_0 + f_0'(x - x_0) + \frac{1}{2} f_0''(x - x_0)^2 + \cdots + \frac{1}{K!} f_0^{(K)} (x - x_0)^K, \tag{B.12}$$

where f_0', f_0'', …, $f_0^{(K)}$ are the first K derivatives of the approximated function, and x_0 is the parameter of this series expansion. Re-denoting

$$f_0 = \phi(x; x_0), \quad f_1 = \phi(x; x_1), \quad f_2 = \phi(x; x_2) \quad , \ldots, \quad f_n = \phi(x; x_n), \tag{B.13}$$

we can apply the regular Aitken procedure using the new Taylor expanded functional values. In such a way, we have finally solved the generalized Cauchy [12]-Jacobi [18] problem: to obtain a rational approximation of a function passing through the N-points $x_1, x_2, \ldots, x_N$ with fixed values of its first derivatives f_i', f_i'', …, $f_i^{(K)}$ at every at those points $i = 1, \ldots, N$. We give only an idea for proof and a Fortran program to illustrate our understanding.

We will use this achievement for many applied problems. For example, for solutions of ordinary differential equations

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \mathcal{F}(x, y). \tag{B.14}$$

Using the former N-points with first derivative Taylor approximation $K = 1$, we can calculate the functional value $y(x)$ using the explained method of Padé extrapolation and then calculate the first derivative at this point explicitly $f'(x) = \mathcal{F}(x, y(x))$. Renumbering later $x_{n+1} \leftarrow x_n$, $x_n \leftarrow x_{n-1}$,…$x_2 \leftarrow x_1$, $x_1 = x_0$ and $x_0 \leftarrow x$ we can continue the procedure. The number of former points is determined by the indices of the optimal Padé approximants according to the p-(minimal-$|\eta|$) criterion. For the step h for the next point $x = x_0 + h$ one can apply many different criteria. To our knowledge such combination of methods for solution of ordinary differential equation without fixed order is a new one. We consider the content of this appendix is implicitly given in the body of the article and we give this clarification only for the beginner users.

Appendix C. Final formula for the Cauchy-Jacobi-Taylor problem

In many articles on applied mathematics and especially in physics, the accent falls how the result can be derived, reproduced, explained or programmed. For programming is very useful to use one and the same memory cell writing sequentially different values in it. On the other side, it is convenient to give a final set of formulae even if the price is the introduction of additional notations. In this section we make an extract of formulae in order to present in short the final solution of the Cauchy [12]-Jacobi [18]-Taylor problem using the Wynn CNEWS algorithm with p-criterion of the error.

Let us have explicit an analytical expression for a function $\Phi(x)$, and its first K derivatives $\Phi'(x)$, $\Phi''(x)$, $\Phi'''(x)$, …, and $\Phi^{(K)}(x)$. We can calculate the values of these functions

$$\Phi_i = \Phi(x_i), \qquad \Phi_i' = \Phi'(x_i), \qquad \Phi_i'' = \Phi''(x_i), \qquad \Phi_i''' = \Phi_i'''(x_i), \qquad \ldots, \qquad \Phi_i^{(K)} = \Phi^{(K)}(x_i) \tag{C.1}$$

at $N = n + 1$ nodal points x_0, x_1, x_2,…, x_n. In short, we have $(K + 2)N$ real parameters $f_i^{(k)}$ and x_i for $i = 0, \ldots, n$ and $k = 0, \ldots, K$. The problem is to derive a rational function $f(x)$, which at every nodal point has fixed functional values for its derivatives.

According Eq. (B.12) our first step is to calculate values of the Taylor expansion

$$f_i = \Phi_i + (x - x_i)\Phi_i' + \frac{(x - x_i)^2}{2}\Phi_i'' + \cdots + \frac{(x - x_i)^K}{K!}\Phi_i^{(K)} = \sum_{k=0}^{K} \frac{\Phi_i^{(k)}}{k!}(x - x_i)^k, \qquad \text{for} \quad i = 0, \ldots, n, \tag{C.2}$$

and to use it as values f_i of a new function $f(x)$ at nodal points x_i.

The problem of Cauchy-Jacobi (C-J) [12,18] is to derive a rational function of the argument x passing exactly at the points (x_i, f_i). The rational approximation has $2N$ real parameters and this is our beginning.

The second step is to calculate the corresponding triangular Aitken table of the polynomial approximations according to the Aitken algorithm described in Sec. B. For the left column of functional values we have $a_{i,0} = f_i$ for $i = 0, \ldots, n$. Then we have to apply a double cycle filling of the table according to the Aitken algorithm Then we have to calculate sequentially l-th column of the Aitken table for $l = 1, 2, \ldots, n - 1$ and for every fixed column number l we have to calculate the values in every row

$$a_{k,l+1} = \frac{\begin{vmatrix} (x_l - x) & a_{l,l} \\ (x_k - x) & a_{k,l} \end{vmatrix}}{x_l - x_k}, \qquad \text{for} \qquad k = l + 1, \ldots, n. \tag{C.3}$$

In such a way we obtain the sequence $S_0 = f_0 = a_{0,0}$, $S_1 = (f_1, f_0) = a_{1,1}$, $S_2 = (f_2, f_1, f_0) = a_{2,2}$, …, $S_n = (f_n, f_{n-1}, \ldots, f_2, f_1, f_0) = a_{n,n}$, which limit we will search using the Wynn CNEWS algorithm. Here are given some often used notations.

The third and final step is to calculate the Padé table. The (-1)-th row is the initialization

$$r_{i,-1} = \infty, \qquad \text{for} \quad i = 0,\ 1,\ 2\ \ldots,\ n,\ n+1. \tag{C.4}$$

Let us recall that $1/\infty = 0$. The main detail is the initialization of the 0-th row

$$r_{i,0} = a_{i,i}, \qquad \text{for} \quad i = 0,\ 1,\ 2,\ \ldots,\ n. \tag{C.5}$$

Having two sequential rows in the Padé table we can calculate the elements in third one using the CNEWS Wynn algorithm. For simplicity we can make the local notation

$$\begin{pmatrix} \ldots & N & \ldots \\ W & C & E \\ \ldots & S & \ldots \end{pmatrix} \equiv \begin{pmatrix} \ldots & r_{l,m-1} & \ldots \\ r_{l-1,m} & r_{l,m} & r_{l+1,m} \\ \ldots & r_{l,m+1} & \ldots \end{pmatrix} \tag{C.6}$$

in order to use Gragg notations in the Wynn identity Eq. (3). Then we calculate

$$\eta_{l,m} = \begin{cases} \dfrac{1}{\dfrac{1}{E-C} + \dfrac{1}{W-C}}, & \text{for} \quad E \neq C \quad \text{and} \quad W \neq C, \\ 0, & \text{for} \quad E = C \quad \text{or} \quad W = C. \end{cases} \tag{C.7}$$

Then we calculate "South" element (using $\infty = 1/0$, if some of denominators is zero)

$$D \equiv \frac{1}{E-C} + \frac{1}{W-C} - \frac{1}{N-C}, \qquad S = \begin{cases} C + 1/D, & \text{for} \quad D \neq 0 \\ \infty, & \text{for} \quad D = 0 \\ C, & \text{for} \quad D = \infty \end{cases}, \qquad r_{l,m+1} = S. \tag{C.8}$$

During these calculations we look for the minimal $\eta_{l,m}$ element

$$\mathrm{p} = |\eta_{L,M}| = \min_{l,m} |\eta_{l,m}|. \tag{C.9}$$

The indices L, M mark the first time when this minimal value p is reached. We recommend that $r_{L,M}$ is the optimal Padé approximant and p is the empirical error. For calculation of $r_{L,M}$ are actually used only the first $L + M + 1$ node points (x_0, f_0), (x_1, f_1), $\ldots$, (x_{L+M}, f_{L+M}). The calculated area in the Padé table has V-shape. First we calculate the first row $r_{1,1}$, $r_{2,1}$, $r_{3,1}, \ldots, r_{n-2,1}$, $r_{n-1,1}$, then the second one $r_{2,2}$, $r_{3,2}$, $r_{4,2}, \ldots, r_{n-2,2}$. The process finishes with the calculation of $r_{k,k}$ in $n = 2k$ is even, or with $r_{k,k}$, $r_{k+1,k}$ if $n = 2k + 1$ is odd. The total number of interpolation node points x_0, x_1, $\ldots$ x_n is $N = n + 1$. One can easily check the special case: if $E = C$ or $W = C$ then $\eta = 0$, $D = 0$ and $S = C$, i.e. the limit of the sequence $\tilde{S} = C$ is this central element. This means that the calculated Padé approximant passes through the node points.

Calculated according the derived formulae rational function of x

$$r_{l,m}(x;\ \mathrm{p}, (x_0, f_0),\ (x_1, f_1),\ (x_2, f_2),\ (x_{l+m}, f_{l+m})) \tag{C.10}$$

passes through all node $N = l + m + 1$ points $\{(x_0, f_0),\ (x_1, f_1),\ (x_2, f_2),\ (x_{l+m}, f_{l+m})\}$ which are $2(l + m + 1)$ parameters of the sought N-point Padé approximant. Moreover, as a result of the performed numerical experiments we consider that the p-criterion gives the best empirical evaluation of the searched result $r_{L,M} \pm \mathrm{p}$. This p-criterion can be used for practical calculations of the limit of many sequences S_0, S_1, $\ldots$, S_n obtained from some analytical procedure. For complex numbers the formulae are the same. Actually it is extremely difficult to find some real physical problem derived by an analytical procedure for which the calculated in such a way limit does not converge to the experimental result. The price of every exception is 10$, write us now!

In the epoch of Cauchy and Jacobi (C-J) [12,18] the problem of finding of rational approximation of a tabulated function was a problem of pure mathematics. The appearance of computers made possible different solutions of the C-J problem to be used in many applied tasks and at contemporary fragmentation of mathematics, the solution of the C-J problem administratively belongs to the applied numerical analysis. It will be very instructive to compare present set of formulae for the solution of C-J problem by calculation of N-point Padé approximant by Wynn CNEWS algorithm with alternative solutions implemented in commercial software as, for example, Mathematica and Maple products.

Studies what to do when we have a zero denominator and boundaries of the cycles only fogs the clear idea. We repeat the obtained chain of formulae which have self-verifying condition of applicability. Often the diagonal Padé approximants $r_{1,1}$, $r_{2,2}$, $r_{3,3}$, $\ldots$ have best convergence and using them we repeat our prescription in short.

Numerical recipe: Using tabulated values of an analytical function $\Phi(x)$ and its K first derivatives $\Phi^{(k)}(x)$ at N node points x_i calculate when possible

$$n = N-1, \qquad k = n/2, \qquad a_{i,0} = f_i = \sum_{k=0}^{K} \frac{\Phi_i^{(k)}}{k!}(x - x_i)^k, \quad i = 0, \ldots, n, \qquad a_{k,l+1} = \frac{(x_l - x)\, a_{k,l} - (x_k - x)\, a_{k,l}}{x_l - x_k},$$

$$r_{i,0} = S_i \equiv a_{i,i}, \qquad r_{l,m+1} = r_{l,m} + \frac{1}{\dfrac{1}{r_{l+1,m} - r_{l,m}} + \dfrac{1}{r_{l-1,m} - r_{l,m}} - \dfrac{1}{r_{l,m-1} - r_{l,m}}}, \tag{C.11}$$

$$\eta_{lm} = \frac{1}{\dfrac{1}{r_{l+1,m} - r_{l,m}} + \dfrac{1}{r_{l-1,m} - r_{l,m}}}, \qquad \mathrm{p} = \min_{l,m} |\eta_{l,m}|, \qquad f_{\mathrm{CJT}}(x) \equiv \tilde{S} \approx r_{k,k}, \qquad \Phi(x) \equiv \tilde{S} \approx r_{k,k} \pm \mathrm{p},$$

and you obtain a good rational appoximation of the tabulated in N point function $\Phi(x)$. We are unaware of an alternative explicit solution of the Cauchy-Jacobi-Taylor (CJT) problem. Our main conclusion is that obtained by the Wynn identity value p gives the long sought criterion for the choice of the optimal Padé approximant. If the approximated function is actually a ratio of two polynomial, then $\mathrm{p} \simeq \varepsilon_{\mathrm{mach}}$. The original C-J problem corresponds to the special case of $K = 0$. For approximate limit of a numerical sequence $\tilde{S} = \lim_{n\to\infty} S_n$ we use CNEWS Wynn algorithm with p-criterion for the choice of optimal Padé approximant.

In conclusion we consider that p-criterion for the choice of optimal Padé approximant is the crucial missing link p's central-news algorithm to find many applications in physics: analytical continuation, series summation and solving differential equations.

Appendix D. Drawing a circle $x^2 + y^2 = 1$. Etude for a student

D.1. Homework task

Let solve the Cauchy problem, and to find the solution of the ordinary differential equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \mathcal{F}(x, y), \qquad \mathcal{F}(x, y) = -\frac{x}{y}, \qquad y(0) = -1. \tag{D.1}$$

D.2. Teachers psychology

This is a trivial problem with separation of variables. Differentiation the equation of the circle $x^2 + y^2 = r^2 = \mathrm{const}$ we obtain $x\mathrm{d}x = -y\mathrm{d}y$, whence $\mathcal{F}(x, y) = -x/y$. Simplest test of accuracy of the used numerical method is to look at $1 - \sqrt{x^2 + y^2}$ versus x.

D.3. Solution

Let have recorded sequential "time" steps of the trajectory (x_0, y_0), (x_1, y_1), (x_2, y_2), …, (x_t, y_t) where index t is integer. In the last point we calculate the derivative $\mathcal{F}_t = -x_t/y_t = \tan(\theta_t)$, where $\theta_t = \arctan(y_t, x_t)$ is the angle of the slope of the curve the solution of the differential equation. This last point (x_t, y_t) we use as the centre of a new coordinate system, with the x' axis chosen along the tangent of the trajectory. In such a way we obtain the trajectory in the new system

$$\begin{pmatrix} x'_i \\ y'_i \end{pmatrix} = \begin{pmatrix} c_t & s_t \\ -s_t & c_t \end{pmatrix} \begin{pmatrix} x_{t-i} - x_t \\ y_{t-i} - x_t \end{pmatrix}, \qquad c_t = \cos(\theta_t), \quad s_t = \sin(\theta_t), \qquad \text{for} \quad i = 0,\ 1,\ 2,\ \ldots,\ n_t. \tag{D.2}$$

Here the integer number n_t describes how many former points we take into account. Analogously one can find the angles between the tangents in moment t and the present one

$$f'_i = \tan(\theta_{t-i} - \theta_t), \qquad \text{for} \quad i = 0,\ 1,\ 2,\ \ldots,\ n_t. \tag{D.3}$$

Having argument, function and its derivative in $N_t = n_t + 1$ points $(x'_i,\ y'_i,\ f'_i)$ for $i = 0,\ 1,\ \ldots,\ n_t$ we can use the described procedure of extrapolation and to calculate the function $\Delta y'_{t+1}$ at argument $\Delta x'_{t+1} = h_t$. The method gives the empirical error p_{t+1} at the calculation of the new point, and the used optimal powers L_{t+1} and M_{t+1} in the numerator and denominator of the Padé ratio. This means that actually values of only $N_{t+1} = L_{t+1} + M_{t+1} + 1$ point is used and returning in time is necessary only for $n_t = N_{t+1} - 1$ points back.

The differences of the coordinates have to be calculated in the initial system

$$\begin{pmatrix} \Delta x_{t+1} \\ \Delta y_{t+1} \end{pmatrix} = \begin{pmatrix} c_t & -s_t \\ s_t & c_t \end{pmatrix} \begin{pmatrix} \Delta x'_{t+1} \\ \Delta y'_{t+1} \end{pmatrix}. \tag{D.4}$$

Now we have to check the accuracy. If the empirical error $\mathrm{p}_{t+1} > \varepsilon_{\max}$ is larger than the maximal acceptable error $\varepsilon_{\max}$ we have to use halve step $h_t /= 2.0$ and to try again to calculate the next time step in the trajectory.

If the empirical error $\mathrm{p}_{t+1} < \varepsilon_{\max}$ is smaller than the maximal accepted error the new point is recorded and we can calculate the next point of the curve

$$x_{t+1} = x_t + \Delta x_{t+1}, \qquad y_{t+1} = y_t + \Delta y_{t+1}. \tag{D.5}$$

For the next step we can try to use slightly bigger step

$$h_{t+1} = 1.1 \left(\frac{\varepsilon_{\mathrm{opt}}}{\mathrm{p}} \right)^{1/n_{t+1}} h_t, \tag{D.6}$$

where optimal error $\varepsilon_{\mathrm{opt}} = \varepsilon_{\max}/10.0$ we keep to be, say, one order of magnitude less than maximal acceptable error $\varepsilon_{\max}$. Moreover for the next point we will try to use one node point more $n_{t+1} = n_t + 1$. Here we wish to stress out that our numerical method for solution of ordinary differential equations is "disordered". The number of the node point in the optimal Padé approximant is also adaptive as the step of the integration h_t.

In the beginning we have to start with a very small step, for example the square of the machine epsilon $h_0 = \sqrt{\varepsilon_{\mathrm{mach}}}$. In this case the procedure of extrapolation is trivial – we start with the method of Euler

$$y_1 = y_0 + \mathcal{F}(x_0, y_0) h_0, \qquad x_1 = x_0 + h_0. \tag{D.7}$$

There are no former points to return at $n_0 = 0$. The second is another Euler steps with $h_1 = h_0$

$$y_2 = y_1 + \mathcal{F}(x_1, y_1) h_1, \qquad x_2 = x_1 + h_1. \tag{D.8}$$

Having already 3 points we can for first time apply the CNEWS p's algorithm $W = y_0$, $C = y_1$, $E = y_2$ and we can calculate according to Eq. (9) the improved value

$$\tilde{y}_2 = r_{1,1} = y_2 + \frac{1}{\dfrac{1}{y_2 - y_1} + \dfrac{1}{y_0 - y_1} - \dfrac{1}{\infty - y_1}}, \qquad \frac{1}{\infty - y_1} = 0. \tag{D.9}$$

This is the well-known Δ-process by Aitken, with empirical accuracy at calculation of second "time" step $t = 2$

$$\eta_{1,1} = \frac{1}{\dfrac{1}{y_2 - y_1} + \dfrac{1}{y_0 - y_1}}, \qquad \mathrm{p}_{t=2} = |\eta_{1,1}|, \qquad L_2 = 1, \quad M_2 = 1, \quad n_2 = L_2 + M_2 = 2, \quad N_2 = n_2 + 1. \tag{D.10}$$

This empirical error p_2 has to be compared with the maximal one $\varepsilon_{\max}$ Now after this initialization we can continue with the calculation of the third time step $n_3 = n_2 + 1 = 3$. In other words for calculation of the third point y_3 we can return $n_3 = 3$ points back and for extrapolation we go to the beginning y_2, y_1, y_0. After some steps returning back in "time" n_t is limited by the numerical noise of rounding errors.

We use this method in a magnetohydrodynamic problem [30] and long time ago it was used in one-dimensional three-body problem. If the considered method is useful for many physical problem following Gragg style it can be called missing Brezinski method.

The conventional methods are applicable for solving the differential equation (D.1). But standard numerical software can be used if we rewrite it as a time dependent system with real time τ

$$\frac{\mathrm{d}x}{\mathrm{d}\tau} = -y, \qquad x(0) = 0, \tag{D.11}$$

$$\frac{\mathrm{d}y}{\mathrm{d}\tau} = +x, \qquad y(0) = -1, \tag{D.12}$$

and to check the solution $x = \cos(\tau - \pi/2)$ and $y = \sin(\tau - \pi/2)$ with almost machine error.

A small variation of the same excise is to start with equation of the unit circle $|y| = \sqrt{1 - x^2}$ for which the differentiation gives $\mathrm{d}y/\mathrm{d}x = -x\,\mathrm{sgn}(y)/\sqrt{1 - x^2}$. Having the best rational approximation for the Cauchy [12]-Jacobi [18]-Taylor problem (interpolation with fixed values of the derivatives), we can use it as predictor to calculate the new (x, y)-point and then to calculate the derivative in the new point $\mathcal{F}(x, y)$ exactly. In such a way we have obtained the best rational approximation for solving Cauchy problem when the solution $y(x)$ is an analytical function. In some cases extrapolation of a curve along the tangent is very useful especially when derivative is divergent. In any cases having a good rational approximation we have to try tho minimize the noise related to errors of numerical truncation avoiding calculations of differences when it is possible.

References

[1] M. Abramowitz, I.A. Stegun, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables (10th pr.), United States Department of Commerce, National Bureau of Standards, Washington D.C., 1972, Subsec. 25.2. Interpolation.
[2] A.C. Aitken, On Interpolation by Iteration of Proportional Parts, without the Use of Differences, Proc. Edinburgh Math. Soc. 3 (1) (1932) 56–76.
[3] H. Alfvén, Granulation, magneto-hydrodynamic waves, and the heating of the solar corona, MNRAS 107 (1947) 211.
[4] George A. Baker Jr., Essentials of Padé Approximants, Academic Press, New York, 1975, p. 100, Sec. 8.
[5] G.A. Baker Jr., P. Graves-Morris, Padé Approximants, 2nd ed., Cambridge Univ. Press, Cambridge, 1996, Sec. 7.1. Multipoint Padé approximation.
[6] B. Beckermann, A.C. Matos, Algebraic properties of robust Padé approximants, arXiv:1310.2438v2 [math.NA].
[7] C. Brezinski, History of Continued Fractions and Padé Approximants, Springer-Verlag, Berlin, 1991, Sec. 5.2.5. Padé approximants.
[8] C. Brezinski, Extrapolation algorithms and Padé approximations: a historical survey, Applied Numerical Mathematics 20 (1996) 299–318.
[9] C. Brezinski, Some pioneers of extrapolation methods, in: A. Bultheel, R. Cools (Eds.), The Birth of Numerical Analysis, World Scientific, Singapore, 2010.
[10] C. Brezinski, M. Redivo Zaglia, Extrapolation Methods. Theory and Practice, North-Holland, 1991.
[11] I.N. Bronshtein, K.A. Semendyayev, Handbook of mathematics for engineers and students of technical universities, 5th ed., Gosudarstvennoe Izdatel'stvo Fiziko-Matematicheskoy Literatury, Moscow, 1955 (in Russian), Sec. 6. Experimental data processing.
[12] A.L. Cauchy, Analyse Algèbrique. Cours d'Analyse de l'Ecole royale polytechnique, L'Imprimerie Royale, Debure fréres, Libraires du Roi et de la Bibliothéque du Roi, 1821, Note V., Sur la Formule de Lagrange relative à l'Interpolation., 525–529 (in French);
R.E. Bradley, C.E. Sandifer, Cauchy's Cours d'analyse: An Annotated Translation, in: J.Z. Buchwald (Ed.), Sources and Studies in the History of Mathematics and Physical Sciences, Springer, New York, 2009.
[13] G. Frobenius, Über Relationen zwischen den Näherungsbrüchen von Potenzreihen, J. Reine Angew. Mathem. 90 (1881) 1–17 (in German).
[14] A.A. Gonchar, G. López Lagomasino, On Markov's theorem for multipoint Padé approximants, Mat. Sb. (N. S.) 105(147) (4) (1978) 512–524.
[15] A.A. Gonchar, E.A. Rakhmanov, Equilibrium distributions and degree of rational approximation of analytic functions, Math. USSR Sbornik 62 (2) (1989) 305–348.
[16] P. Gonnet, S. Güttel, L.N. Trefethen, Robust Padé Approximation via SVD, SIAM Review 55 (1) (2013) 101–117.
[17] W.B. Gragg, The Padé Table and Its Relation to Certain Algorithms of Numerical Analysis, SIAM Rev. 14 (1) (1972) 1–62.
[18] C.G.J. Jacobi, Über die Darstellung einer Reihe gegebner Werthe durch eine gebrochne rationale Function, J. Reine Angew. Math. (Crelle) 30 (1846) 127–156.
[19] J. Karlsson, Rational Interpolation and Best Rational Approximation, J. of Math. Anal. and Appl. 53 (1976) 38–52.
[20] G.A. Korn, T.M. Korn, Mathematical Handbook for Scientists and Engineers: Definitions, Theorems, and Formulas for Reference and Review, Revised Edition, McGraw-Hill, 1968, Sec. 20.5. Approximation of Functions by Interpolation.
[21] F.E. Levis, C.N. Rodriguez, The Best Multipoint Padé Approximant, in: Numerical Functional Analysis and Optimization, Taylor & Francis, 2018.
[22] D.S. Lubinsky, The Spurious Side of Diagonal Multipoint Pade Approximants, in: M. Abell, E. Iacob, A. Stokolos, S. Taylor, S. Tikhonov, J. Zhu (Eds.), Topics in Classical and Modern Analysis, Birkhäuser, Cham, 2019, p. 245.
[23] T.M. Mishonov, A.M. Varonov, On the practical realization of ε-algorithm for calculation of limits of numerical sequences and N-point Padé approximations, in: T.M. Mishonov, A.M. Varonov (Eds.), 10th Jubilee International Balkan Physical Union Conference Proceedings, in: AIP Conference Proceedings, vol. 2075, 2019, p. 110004.
[24] T.M. Mishonov, A.M. Varonov, The Wynn identity as the long sought criterion for the choice of the optimal Padé approximant, arXiv:1901.06014v1 [cs.NA].
[25] T.M. Mishonov, A.M. Varonov, N.I. Zahariev, et al., Magnetohydrodynamic calculation of the temperature and wind velocity profile of the solar transition region. Preliminary results, in: V. Vassilev, S. Nikolov, M. Datcheva, Y. Ivanova (Eds.), NCTAM 2017–13th Bulgarian National Congress on Theoretical and Applied Mechanics 2017, in: MATEC Web Conf., vol. 145, 2018, p. 03009, https://www.matec-conferences.org/articles/matecconf/abs/2018/04/contents/contents.html.
[26] T.M. Mishonov, Y.G. Maneva, M.V. Stoev, A.M. Varonov, Heating of the solar corona by Alfvén waves – magnetohydrodynamic calculation of height temperature profile of the transition region, Magnetohydrodynamics 55 (3) (2019) 261–284.
[27] H. Padé, Sur la représentation approchée d'une fonction par des fractions rationelles, Ann. Sci. Ec. Norm. Super. 9 (1892) 1–93 (in French).
[28] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, Numerical Recipes in C: The Art of Scientific Computing, 2nd ed., Cambridge University Press, Cambridge, 1992.
[29] H. Stahl, Existence and Uniqueness of Rational Interpolants with Free and Prescribed Poles, in: E.B. Saff (Ed.), Approximation Theory, Proc. of a Seminar, Tampa, 1985-1986, Springer-Verlag, Berlin, 1987, p. 206.
[30] A.M. Varonov, The Theory of Heating of the Solar Corona and Launching of the Solar Wind by Alfvén Waves, https://arxiv.org/abs/1903.07688.
[31] P. Wynn, Upon Systems of Recursions which Obtain Among the Quotients of the Padé Table, Numerische Mathematik 8 (1966) 264–269.
[32] M.L. Yattselev, Symmetric contours and convergent interpolation, J. of Appr. Theory 225 (2018) 76–105.