UNIVERSITÄT REGENSBURG
Institut für Physikalische und Theoretische Chemie

D-93040 Regensburg
Universitätsstraße 31
GERMANY

PD Dr. Ernst Joachim Weniger

30 October 2007

Dr. Saul Teukolsky
Hans A. Bethe Professor of Physics and Astrophysics
608 Space Sciences Building
Cornell University
Ithaca NY 14853
USA

Dear Dr. Teukolsky:

When I obtained my copy of the most recent (3rd) edition of your book *Numerical Recipes* [68], I noticed with pleasure that nonlinear sequence transformations are now also treated.

When writing a book like *Numerical Recipes*, which tries to incorporate virtually all computationally useful mathematical techniques in a singe volume, there is an obvious conflict of breadth vs. depth. Thus, the authors of such a book necessarily have to compromise, and according to experience it is normally not possible to satisfy everybody with a compromise.

However, I think that your book is an excellent compromise, and that it is extremely useful for a very wide audience of computationally oriented scientists and engineers. This was already true for the FORTRAN version of the 1st edition [67], which I had bought many years ago. I am positive that this new edition will also have the same favorable reception and the same impact on the whole scientific community as the previous editions. Congratulations!

In the following text I will make some remarks on sequence transformations in general and on the way that they are presented in your book. I do hope that these remarks will be of interest for you, and they may even be helpful when you and your co-authors will write the next edition of your book.

As we all know, in applied mathematics or in the mathematical treatment of scientific or engineering problems, slowly convergent or even divergent sequences and series abound. Accordingly, convergence acceleration and summation techniques are at least potentially extremely useful in a large variety of different contexts, and I can probably claim with some justification that anybody involved in computational work should have at least some basic knowledge about the power, but also about the shortcomings and limitations of these techniques.

Unfortunately, in the university curricula these techniques are treated at best at a very superficial level, and more often not at all. Accordingly, university graduates tend to be fairly ignorant about these things.

This was also true in my case. During my undergraduate studies, I had never heard anything about these things, and even during the work for my PhD thesis [80], where I had to struggle with the efficient and reliable evaluation of (complicated) series expansions for so-called molecular multicenter integrals of exponentially decaying basis functions, I was completely ignorant about Padé approximants or other nonlinear sequence transformation[1]. During my PhD thesis, I was only aware of *linear* and *regular* sequence transformations as they were described in the book by Knopp [56], which is undeniably a very useful classic but now completely outdated from a computational point of view.

My ignorance only changed when I did postdoctoral work at the Department of Applied Mathematics of the University of Waterloo in Waterloo, Ontario, Canada, where I – inspired by Jiří Čížek – applied Padé approximants and continued fractions for the summation of divergent power series. Obviously, my stay in Waterloo had a huge impact on my later research.

It is thus my basic assumption that, when writing about convergence acceleration and summation techniques, one always has to take into account that the hypothetical typical reader is fairly ignorant at least about the subtleties of these techniques. Therefore, one has to proceed with great caution and give the uninitiated reader a helping hand. This is also the reason why I like for instance your Chapter 5.12 on Padé approximants very much: You provide a lot of useful information by means of simple examples and do not try to show your mathematical sophistication and/or cleverness by over-flooding the uninitiated reader with difficult mathematical details.

My assessment about the general knowledge of hypothetical typical readers may sound pessimistic, but I actually think that the situation is changing for the better. For example, the treatment of nonlinear sequence transformation in the most recent edition of your book is in my opinion an encouraging sign. Moreover, I recently found on the Internet that the Institute of Computer Science at the University of Wrocław in Poland plans to offer a graduate course with the title "Convergence Acceleration Methods". This course is a part of the so-called "Studies in English" for graduate students. For me, this is a good sign, and I do hope that other universities will follow this example.

When writing about sequence transformations and related topics, one should also take into account that the potential readers tend to be highly heterogeneous. Essentially, there are two radically different prototypes who have very different attitudes and preferences: Firstly, there are researchers – mainly mathematicians – who essentially want to work *on* sequence transformations: They are predominantly interested in the mathematical properties of sequence transformations or in the derivation of error estimates and convergence proofs, but they usually do at best token applications of sequence transformations. Secondly, there are many others – mainly scientists and engineers – who only want to work *with* sequence transformations: They want to use these techniques as computational tools (and often as black boxes) in order to solve practical problems, but they usually do not care too much about theoretical aspects and mathematical subtleties or even convergence proofs that tend to be either unrealistic or not suited for practical applications.

Of course, this pragmatic attitude of scientists and engineers can easily lead to problems. Too often, they simply lack a sufficiently broad background knowledge about the techniques they apply. For

[1] Actually, my ignorance about non-standard numerical techniques was at that time so bad that *divergent series* were for me essentially some kind of mathematical pornography. Fortunately, this has changed not only in my case, and the usefulness of divergent series for computational purposes is widely accepted. For example, I recently came across a PhD thesis by Meurer [65] discussing the summation of divergent series in the context of mechanical engineering. I am too ignorant in mechanical engineering to decide whether this is a good thesis or not, but I nevertheless think that it is noteworthy that divergent series are now also used in mechanical engineering.

example, some years ago an article by Leung and Murakowski [58] was published in Journal of Mathematical Physics in which a seemingly new generalization of Padé approximants was introduced that utilizes information from a weak coupling and a strong coupling expansion. However, these "new" rational approximants are noting but the well known two-point Padé approximants, which are for instance described in a very detailed way in Chapter 7.1 of the book by Baker and Graves-Morris [7] and which I had also applied some time ago for the summation of divergent perturbation expansions [34] (compare also the Acknowledgment in [59]). For me, the shocking thing was not the lack of knowledge of the authors – this can happen – but rather the ignorance of the referees of Journal of Mathematical Physics, which is usually considered to be a highly respected journal[2].

The deep division among those who are interested in sequence transformations and related topics is also obvious in the literature on sequence transformations. So far, books *by* mathematicians written *for* mathematicians clearly dominate (see for example the books by Brezinski [14, 15, 16, 17, 18], Brezinski and Redivo Zaglia [26], Cuyt and Wuytack [36], Delahaye [37], Liem, Lü, and Shih [61], Marchuk and Shaidurov [63], Sidi [75], Walz [79], and Wimp [102]).

Unfortunately, the monographs listed above cannot be digested easily by computationally oriented scientists or engineers who just want to use convergence acceleration and summation techniques as computational tools in order to get their work done.

The problems with these mathematically oriented books become particularly obvious in the case of Sidi's book on sequence transformations [75], which is the most recent monograph on this topic. Avram Sidi is a very good mathematician, and I highly appreciate and respect some of his work *on* sequence transformations. Nobody can deny that Sidi's book contains a wealth of information, and it is very useful for anybody interested in the mathematical properties of sequence transformations. Essentially, this book of more than 500 pages is a book written by an expert in this field for the few other ones who can also claim to be experts in this field. However, Sidi's style makes this book hard to read even for specialists. Moreover, Sidi's choice of topics – and in some cases also the deliberate omission of certain topics – makes this book highly subjective. Therefore, non-specialist readers may get a distorted view about the state of the art and the contributions of other researchers (compare also my book reviews [31, 32]). Frankly speaking, I cannot imagine that anybody, who only wants to work *with* sequence transformations, would bother to study Sidi's impressive but difficult book seriously. Personally, I think that Sidi's book is a missed chance of further popularizing sequence transformations among non-specialists, which I deplore very much. Sidi undoubtedly invested an enormous amount of time and effort.

Nevertheless, I again have the impression that the situation is slowly improving, and that there are now some books – or at least parts of books – that provide an easily digestible introduction to sequence transformations from the perspective of those who predominantly want to apply sequence transformations as computational tools. I consider your book to be just another example that supports my claim.

For example, there is a book by Bornemann, Laurie, Wagon, and Waldvogel [12], which was recently translated to German [13]. The topic of this book is extreme digit hunting in the context of

[2] Apart from this incident, I have a very good opinion of Journal of Mathematical Physics: I published several articles in Journal of Mathematical Physics and plan to submit further manuscripts in the future.

some challenging problems of numerical analysis. For this extreme digit hunting, the authors also use sequence transformations, whose basic theory is described compactly in their Appendix A.

Of course, this Appendix A is much too short to provide a reasonably complete and balanced presentation of sequence transformation, but I think that a novice can benefit considerably from reading this Appendix A. In particular, I like the extremely pragmatic approach of the authors of this book, which is very uncommon among mathematicians. Probably, this is a consequence of the fact that the authors of this book are not primarily interested in the mathematical theory of sequence transformations: They only wanted to apply sequence transformations as computational tools in order to get more precise results at acceptable computational costs.

For example, in [12, p. 230] one finds the following very instructive remark:

> *More than any other branch of numerical analysis, convergence acceleration is an experimental science. The researcher applies the algorithm and looks at the results to assess their worth.*

At least conceptually, this is very similar to a remark which I had made about Levin-type transformations at roughly the same time [95, p. 1241]:

> *As discussed in more details in the following article [94][3], our current level of theoretical understanding does not permit to predict which one of the numerous variants of* $\mathscr{G}_k^{(n)}(q_m, s_n, \omega_n)$ *will give best results for a given convergence acceleration or summation problem. So, if we for example use one of the numerous Levin-type transformation for the summation of a divergent perturbation expansion, we are essentially conducting a numerical experiment. As every good experimentalist knows, a single experiment is only rarely able to provide a definite answer. Normally, a whole set of related experiments is needed to obtain convincing evidence. Of course, this applies also to our numerical experiments. Therefore, we should not insist with a quasi-religious zeal on using only a single (Levin-type) transformation which we for some reason may prefer. Instead, it is usually a much better idea to compare the performance of several different transformations.*

Although I referred in this remark explicitly to Levin-type transformations, it of course also applies to other sequence transformations as well. Moreover, it completely agrees with the following remark from the book by Bornemann, Laurie, Wagon, and Waldvogel [12, p. 250] that should be heeded not only by novices:

> *Whenever possible, use more than one extrapolation method.*

You also mention this shortly in the 1st paragraph of [68, p. 217]. Nevertheless, it is my conviction that the experimental nature of work *with* sequence transformations cannot be overemphasized.

Another instructive remark from the book by Bornemann, Laurie, Wagon, and Waldvogel, which I like very much, is the following one [12, p. 225]:

[3] My article [94] is not yet finished and I am still struggling with some open problems.

The question whether a series converges is largely irrelevant when the reason for using a series is to approximate its sum numerically.

It is my conviction that factorially divergent asymptotic series expansions for special functions can indeed be very useful computationally also for at most moderately large arguments. Levin-type transformations[4] and to a lesser degree also Wynn's epsilon algorithm [104] can sum divergent series of that kind quite effectively (see for example [87, 97]).

Then there is a recent book by Kythe and Schäferkotter [57]. It discusses the use of sequence transformations to speed up the convergence of quadrature schemes beyond Romberg, and it also mentions some of Sidi's highly consequential work on this topic. Generally speaking, the combination of quadrature rules with extrapolation methods currently seems to be a "hot" topic. Thus, the contents of this book may be of interest for you, in particular if you should intend to write some time in the future another edition of your book, which I would welcome very much since computationally oriented mathematics is progressing rapidly. Extrapolation techniques in numerical quadrature are also discussed, albeit in less detail, in the already somewhat older book by Evans [40].

In the next few days a new book by Gil, Segura, and Temme [44] on the evaluation of special functions should come out. It discusses in addition to various other computational techniques also Padé approximants, continued fractions, and sequence transformations.

I would also like to make some comments on what you write in your book [68] about sequence transformations.

In the text following Eq. (5.3.8) on p. 211, you write:

Sometimes convergence acceleration is helpful only once the terms start decreasing.

I had addressed this question as well as related problems with what I call *irregular input data* in my article [92]. In this article, I studied the impact of irregular input data on the performance of sequence transformations. My main model system was the Gaussian hypergeometric series

$$ {}_2F_1(a,b;c;z) \; = \; \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{(c)_n} \, \frac{z^n}{n!} \tag{1} $$

with a negative third parameter $c < 0$ but $-c \notin \mathbb{N}_0$. The terms of such a hypergeometric series initially look like the terms of a mildly divergent series. Only for higher indices do the terms settle down and eventually vanish. The results in [92] show that sequence transformations fail horribly and produce nonsensical results if the leading irregular terms of such a ${}_2F_1$ are used as input data. It is, however, important to note that different sequence transformations produce different nonsense. By comparing the disagreeing results of several different sequence transformations, it is immediately obvious that something is wrong. If we only use a single transformation, this is by no means immediately obvious.

Thus, the numerical results presented in [92] provide strong support for my claim – or also the claim of Bornemann, Laurie, Wagon, and Waldvogel [12, p. 250] – that it is advisable to use more

[4] A fairly complete list of so-called *scalar* Levin-type transformations can be found in an article by Homeier [52].

than a single transformations. Of course, the *agreement* of the results of several different sequence transformation does not prove that these results are correct. Nevertheless, such an agreement helps to gain confidence in the validity of the numerical results. Moreover, the *disagreement* of the results of different transformations as seen in [92] is extremely valuable because it indicates very strongly that one cannot trust these numerical results and that something has gone wrong.

In the 2nd paragraph on p. 212, you recommend using instead of the mathematically simpler *linear* and *regular* sequence transformations the mathematically more complicated, but also more powerful *nonlinear* sequence transformations. I very much welcome this. There are still too many mathematical authors who endorse linear transformations. For example, in Zayed's relatively recent book [106, Chapter 1.11.1], the classic summability methods associated with the names of Cesàro, Abel, and Riesz are reviewed, but the (much) more powerful nonlinear summation methods as for example Padé approximants are completely ignored.

In the text following Eq. (5.3.11) on p. 212, you write about the iterations of Aitken's Δ^2 process [2]:

> *(In practice, this iteration will only rarely do much for you after the first stage.)*

Here, I disagree. The iteration of Aitken's Δ^2 process as described in [81, Eq. (5.1-15)] actually produces a fairly powerful sequence transformation (see for example [81, Table 13-1 on p. 328] or the discussion in [81, Section 15.2]).

However, I also think that Wynn's closely related epsilon algorithm [104], which you discuss on pp. 212 - 213, is usually (much) more stable. Moreover, in the context of series expansions for special functions, Levin-type transformations usually give clearly better results. Therefore, it makes sense not to emphasize Aitken's iterated Δ^2 process as a particularly useful numerical tool in your book, which obviously has serious space constraints.

It may be of interest that Wynn's epsilon algorithm [104] is not restricted to so-called *scalar* input data that are either real or complex numbers. Numerous generalizations of the epsilon algorithm to other types of input data are discussed in a review by Graves-Morris, Roberts, and Salam [50].

In the 3rd paragraph on p. 214, you make the following remark about Levin's sequence transformation [60]:

> *The Levin transformation is probably the best single sequence acceleration method currently known.*

My own work both *on* and *with* sequence transformations should show that I also do believe that Levin's idea of introducing explicit truncation error estimates into the transformation process was a very good and highly consequential idea, and that Levin-type transformations are beyond doubt at least *potentially* more powerful than other transformations that cannot benefit from the input of additional information contained in explicit remainder estimates. Nevertheless, I fear that your remark can easily mislead novices. They might come to the *wrong* conclusion that it would suffice to use *exclusively* Levin's sequence transformation in their numerical work.

I suspect that your verdict was strongly influenced by the extensive numerical studies performed by Smith and Ford [76, 77]. I do not question the correctness of the investigations of Smith and

Ford, who deserve praise because they succeeded in raising public awareness of the usefulness of nonlinear sequence transformations as computational tools. However, I think that it is dangerous to rely too much on statistical arguments which – as I am willing to admit – clearly favor Levin's transformation and other Levin-type transformation. Let me try to explain my point of view by two examples.

Quite a few years ago, I looked for test systems for my various sequence transformation programs. In this context, I came across the following power series expansion for the digamma function [1, Eq. (6.3.14)]:

$$\psi(1+z) \; = \; -\gamma + z \sum_{\nu=0}^{\infty} \zeta(\nu+2)\,(-z)^{\nu} \, . \qquad (2)$$

Since this power series is strictly alternating for $z > 0$, I expected in agreement with the observations of Smith and Ford [76, 77] that certain variants of Levin's sequence transformation should be able to produce very good results for this power series. To my dismay, it turned out that Levin's sequence transformation and also other Levin-type transformations were not particularly powerful in the case of the power series (2) and produced results that were clearly inferior to those obtained by Wynn's epsilon algorithm [104].

I understood these observations only much later when I prepared my article [93], in which I analyzed the index dependence of the partial sums and the truncation errors of the power series (2) for $\psi(z)$ in more detail: The partial sums of the power series (2) possess truncation errors for which Wynn's epsilon algorithm seems to be more or less optimal.

The power series (2) for $\psi(z)$ is relatively simple. Consequently, it is not too difficult to analyze the index dependence of the truncation errors of this series. In the case of more complicated series expansions and in particular if the series terms are determined numerically, such an analysis is either very difficult or not possible at all. Thus, we would not be able to explain why for instance Levin-type transformations fail to accomplish something substantial for a power series with a similar behavior. If, however, we compare the performance of several different sequence transformations, we would at least notice that only Levin-type transformations do not accomplish much, whereas other transformations like the epsilon algorithm may produce (much) better results.

Another example, which shows that the observations of Smith and Ford [76, 77] can be badly misleading in special cases, is the factorially divergent Rayleigh-Schrödinger perturbation series for the ground state energy of the quartic anharmonic oscillator.

In my article [82], I applied Wynn's epsilon algorithm [104], Levin's d transformation (compare [81, Eq. (7.3-9)]), and the d variant of the so-called $\mathscr{S}$ transformation (compare [81, Eq. (8.4-4)]) to this divergent perturbation expansion. In these calculations, I used the first 22 perturbation series coefficients, and I did everything in FORTRAN 77 on a Cyber 180-995 E with a precision of approximately 29 decimal digits.

As shown in [82, Table II], the d variant of $\mathscr{S}$ gave best results, followed by the d variant of Levin's transformation, and Wynn's epsilon algorithm, whose convergence is guaranteed because of the Stieltjes nature of this perturbation series, was least effective. Qualitatively, these results were confirmed in [82, Table I], where I applied the same sequence transformations to a hypergeometric model series ${}_2F_0$ – essentially the asymptotic series for a special complementary error function erfc – which possesses the same rate of divergence as the perturbation series for the quartic anharmonic oscillator.

On the basis of these observation, it seemed to be perfectly legitimate and logical to conclude that the d variant of $\mathscr{S}$ sums the factorially divergent perturbation expansion for the ground state energy of the quartic anharmonic oscillator most effectively to its exact result, that the d variant of Levin's transformation also accomplishes this, albeit less efficiently, and that Wynn's epsilon clearly ranks last and is for this problem the least efficient transformation.

Unfortunately, these conclusions – although perfectly plausible on the basis of my numerical results – were premature and based on incomplete evidence. During a stay in Waterloo in 1991, I repeated my previous calculations [82] using now 200 perturbation series coefficients calculated exactly with the help of Maple's rational arithmetic[5], and I did the summation calculations in Maple with a precision of up to 1000 decimal digits [99].

The results obtained in this way showed that the d variant of $\mathscr{S}$ was clearly the most effective transformation for this and – as it became obvious later on – also for other related perturbation problems. However, the results in [99] also showed unambiguously that the d variant of Levin's transformation – although apparently convergent in the case of small transformation orders – diverged for higher transformation orders. In contrast, the summation results obtained by the d variant of $\mathscr{S}$ seemed to converge also for very high transformation orders[6]. Wynn's epsilon algorithm was clearly less efficient, but no divergence problems were observed.

The divergence of Levin's transformation was also confirmed in [84, Table 2], where the summations were performed with a Levin-type transformation that – depending on the value of a continuous parameter – interpolates between Levin's transformation and the $\mathscr{S}$ transformation.

On the basis of our current level of understanding, no completely satisfactory explanation of the divergence of Levin's d transformation in the case of the anharmonic oscillators is known. Personally, I believe that this divergence my be due to subdominant contributions, but this is just unproven speculation. This is a difficult topic, and subdominant contributions in numerically determined data that also diverge factorially or faster with increasing index, are highly elusive objects. If you happen to read German: A sufficiently detailed discussion of the divergence of Levin's transformation can be found in Chapter 10.6 of my habilitation thesis [85, pp. 211 - 216]. A similar divergence of Levin's transformation was observed by Čížek, Zamastil, and Skála [35, p. 965] in the case of the hydrogen atom in an external magnetic field.

In numerous convergence acceleration and summation problems, it has been observed by various authors (including myself) that suitable variants of Levin's transformation often produce truly remarkable convergence acceleration and summation results. Therefore, it would not at all be justified to issue a general warning that Levin's transformation should not be used since it can lead to divergent results. Here, one should take into account that nonlinear sequence transformations are in general *nonregular*. Accordingly, the convergence of the transformed sequence is not guaranteed, let alone to the correct limit. Thus, I can only say that one should not trust Levin's transformation blindly. However, this applies also to any other nonlinear and nonregular sequence transformation.

If we use sequence transformations for the acceleration of convergence of series expansions of special functions, we are in a relatively fortunate situation. We have explicit analytic expressions

[5]This at that time fairly challenging and very time-consuming calculation was only possible because we had exclusive access to new and powerful Unix workstation that later served several research groups.

[6]Nevertheless, I would not be surprised if similar types of divergence could also occur in summation calculations involving $\mathscr{S}$ with very high transformation orders.

for the series terms, and if we are willing to invest enough time and effort, we may even be able to construct sufficiently simple approximations to the truncation errors that hold in the limit of large indices. Accordingly, we have a relatively good chance that we understand what we are doing when applying a sequence transformation. Under these fortunate circumstances, one can even hope for rigorous convergence proofs (although they certainly would not be trivial).

If, however, we are dealing with series expansions whose terms are determined numerically (as it is the case with quantum mechanical perturbation expansions), we are normally fairly ignorant from a theoretical point of view. Consequently, the use of a sequence transformation is essentially a numerical experiment whose intricacies are at best partly understood. In such a case, it is undeniably helpful to use more than a single sequence transformation.

In [81, Section 3.2 on pp. 211 - 212], I formulated a general construction principle for Levin-type transformations based on annihilation operators, which is also described in your Webnote No. 6 "Derivation of the Levin Transformation" and which is based on the assumption that the truncation errors $\{r_n\}_{n=0}^{\infty}$ of of a model sequence $\{s_n\}_{n=0}^{\infty}$ can be partitioned into the product of a *remainder estimate* ω_n and a *correction term* z_n according to

$$r_n \,=\, s_n - s \,=\, \omega_n z_n \,, \qquad n \in \mathbb{N}_0 \,. \tag{3}$$

From a purely formal point of view, such a partition is always possible and thus a triviality. However, in Levin-type transformations it is assumed that the remainder estimates $\{\omega_n\}_{n=0}^{\infty}$ are explicitly known, which obviously makes (3) not so trivial. The correction terms $\{x_n\}_{n=0}^{\infty}$, which typically contain free parameters, should then be chosen in such a way that the products $\omega_n z_n$ provide sufficiently accurate and rapidly convergent approximations to the actual remainders r_n.

The principal advantage of this approach is that only the correction terms $\{z_n\}_{n=0}^{\infty}$ have to be determined, but not the remainders $\{r_n\}_{n=0}^{\infty}$. If good remainder estimates can be found, the determination of z_n and the subsequent elimination of $\omega_n z_n$ from s_n often leads to significantly better results than the construction and subsequent elimination of other approximations to r_n.

The model sequence (3) has – as also explained in your Webnote No. 6 "Derivation of the Levin Transformation" – another indisputable advantage: There exists a systematic approach for the construction of a sequence transformation which is exact for this model sequence. It is only necessary that a *linear* operator $\hat{T}$ can be found which annihilates for all $n \in \mathbb{N}_0$ the correction term z_n according to $\hat{T}(z_n) = 0$. Then, a sequence transformation, which is exact for the model sequence (3), is given by the following ratio [81, Eq. (3.2-11)]:

$$\mathscr{T}(s_n, \omega_n) \,=\, \frac{\hat{T}(s_n/\omega_n)}{\hat{T}(1/\omega_n)} \,=\, s \,. \tag{4}$$

I introduced the construction of sequence transformations via annihilation operators in [81, Section 3.2] in connection with the rederivation of Levin's transformation [60], which is also presented in your Webnote No. 6 "Derivation of the Levin Transformation", and the construction of some other, closely related Levin-type sequence transformations as for example the $\mathscr{S}$ transformation [81, Sections 7 - 9].

Later, this annihilation operator approach was discussed in books by Brezinski [20] and Brezinski and Redivo Zaglia [26] and in articles by Brezinski [19, 21, 22, 23], Brezinski and Matos [25], Brezinski and Redivo Zaglia [27, 28, 29], Brezinski and Salam [30], Homeier [52], Matos [64], and myself [84, 95] and extended to other sequence transformations. The fact, that my annihilation operator approach can also be used for the derivation of numerous other sequence transformations, highlights, in my opinion, the rather obvious fact that annihilation of the truncation errors is *the* central step of convergence acceleration and summation via sequence transformations.

Pragmatism dictates that the correction terms $\{z_n\}_{n=0}^{\infty}$ should be chosen in such a way that the corresponding annihilation operator $\hat{T}$ satisfying $\hat{T}(z_n)$ possesses a manageable complexity and leads to a convenient expression for the sequence transformation and/or to a simple recursive scheme. However, the condition $\hat{T}(z_n) = 0$ does not fix the choice of the remainder estimates $\{\omega_n\}_{n=0}^{\infty}$, and we have at least in principle quite a lot of freedom.

The choice of both $\{\omega_n\}_{n=0}^{\infty}$ and $\{z_n\}_{n=0}^{\infty}$ determines a Levin-type sequence transformation of the type of (4). However, for a given input sequence $\{s_n\}_{n=0}^{\infty}$, the choice of a suitable sequence $\{\omega_n\}_{n=0}^{\infty}$ of remainder estimates seems to be crucial for the success or failure of a transformation process involving Levin-type transformations: The remainder estimates $\{\omega_n\}_{n=0}^{\infty}$ should correctly describe the characteristic features of the input data $\{s_n\}_{n=0}^{\infty}$, whereas the correction terms $\{z_n\}_{n=0}^{\infty}$ should be fairly unspecific and smooth functions of the index n.

As already mentioned above, the explicit incorporation of the information contained in the remainder estimates makes these transformations at least *potentially* more powerful than other transformations such as for example Wynn's epsilon algorithm. However, the explicit incorporation of this information is also the major weakness of Levin-type transformations. If it is possible to find remainder estimates $\{\omega_n\}_{n=0}^{\infty}$ such that the products $\omega_n z_n$ provide *good* approximations to the actual remainders $r_n = s_n - s$, then we can expect (very) good transformation results. If, however, we cannot find for a given sequence $\{s_n\}_{n=0}^{\infty}$ good remainder estimates $\{\omega_n\}_{n=0}^{\infty}$, then we incorporate *non-existing* or *explicitly wrong* information into the transformation process. In such a case, Levin-type transformations will most likely produce (very) bad or even nonsensical results[7].

In [81, Eq. (7.3-1)], I had emphasized that the remainder estimates $\{\omega_n\}_{n=0}^{\infty}$ should be chosen in such a way that ω_n is proportional to the *dominant* term of an asymptotic expansion of the actual remainder r_n (compare also [95, Eq. (4.5)]):

$$r_n \,=\, s_n - s \,=\, \omega_n \left[c + \mathrm{O}\bigl(1/n\bigr) \right] \,, \qquad c \neq 0 \,, \quad n \to \infty \,. \tag{5}$$

In your book [68, Eq. (5.3.14)], you also state this asymptotic condition, which seems to be well suited to rationalize and motivate the choice of suitable remainder estimates.

Unfortunately, I am no longer fully convinced that (5) is really appropriate: It may well be an over-simplification. Using techniques described in [96], I constructed *improved* asymptotic estimates for

[7] In my own research, I encountered situations of that kind quite frequently. According to my experience, this usually does not happen when trying to accelerate the convergence of a series expansion for a special function. If, however, the series under consideration possesses a very difficult structure such as multiple inner sums or if the series terms were determined by complicated numerical processes, then I frequently observed that Levin-type transformations are less effective than for instance Wynn's epsilon algorithm, which pursues a much less ambitious transformation strategy. The best interpretation that I have to offer for this phenomenon is that series terms produced by complicated processes are somehow "polluted", not only by rounding errors, but more seriously by subdominant contributions. Thus, simple model sequences of the type of $s_n = s + \omega_n z_n$ may be overly simplistic, and the resulting Levin-type transformations may turn out to be unsuited for difficult convergence acceleration and summation problems of the kind mentioned above.

the truncation errors of some series expansions for special functions[8]. When I used these improved asymptotic estimates as remainder estimates in appropriate Levin-type sequence transformations, I observed in some cases that *improved* remainder estimates produced *inferior* transformation results.

I believe that these problems are related to the role of large index asymptotics in the mathematical treatment of sequence transformations. Most sequence transformations are constructed via model sequences, and most model sequences are based on heuristic reasoning inspired by large index asymptotics, whose simplifying power makes it possible to identify and understand underlying hierarchical structures[9].

However, there is a major problem: When using a such a sequence transformation, we implicitly use the simplifying power of large index asymptotics in order to avoid the asymptotic regime of large indices. This is an intrinsic contradiction which occasionally can lead to serious numerical or convergence problems. It is by no means obvious that asymptotic expressions allow a reasonably accurate description of the function it represents also in the nonasymptotic regime, i.e., for small indices. If subdominant contributions play a dominant role in the nonasymptotic regime, problems are extremely likely.

Since Levin-type transformations pursue a much more ambitious transformation strategy than other sequence transformations that do not use explicit remainder estimates, it makes sense to assume that they are more strongly affected by problems with subdominant contributions than other, less ambitious sequence transformations.

If the asymptotic condition (5) should indeed turn out to be an oversimplification, then we have to formulate an alternative criterion which the remainder estimates $\{\omega_n\}_{n=0}^{\infty}$ have to satisfy. Maybe, we can only demand that the remainder estimates $\{\omega_n\}_{n=0}^{\infty}$ should be chosen in such a way that the ratios $[s_n - s]/\omega_n$ are annihilated *effectively* already for *small* indices n by those annihilation operators $\hat{T}$ that satisfy $\hat{T}(z_n) = 0$. Obviously, such a weak and somewhat vague statement would not be as nice as the asymptotic condition (5).

In [68, Eq. (5.3.19)], you discuss the remainder estimates that give rise to Levin's u, t, d, and v transformation, respectively. It may be of interest for you that the asymptotic structure of sequences that motivate these remainder estimates were discussed in a relatively detailed way in [95, Section IV].

In [68, Eq. (5.3.19)], you call Levin's d transformation a modified t transformation. This is not correct. I showed in [95, Eqs. (4.41) - (4.53)] that all t-variants of the transformations considered in

[8] Here, *improved* means that the asymptotic expansion of $r_n = s_n - s$ as $n \to \infty$ in the sense of Poincaré does not differ from ω_n by a term proportional to $1/n$, but by a term proportional to a higher power of $1/n$.

[9] For example, in the summer of 2002 you exchanged on the Numerical Recipes Forum several messages with some Jan M. who wanted to evaluate the Weierstraß $\mathscr{P}$ function with the help of sequence transformations. In my opinion, this Jan M. should have tried to analyze the large index asymptotics of the truncation errors of the series expansion he wanted to use. I do not know whether this would have been feasible since the Weierstraß $\mathscr{P}$ function seems to be a fairly complicated beast, but such an asymptotic approximation to the truncation errors would have shown which – if any – of the conventionally used sequence transformations should be best suited for an acceleration of the convergence of this difficult problem. Alternatively, it could have given Jan M. some inspiration on the construction of a new sequence transformation that might be able to do the job.

this article actually do not exist as an independent variant and in fact are d-variants in disguise. The most relevant formula for you is [95, Eq. (4.46)].

In the last paragraph of Chapter 5.3.2 on p. 216, you briefly mention Wynn's rho algorithm [105]. I do not completely agree with your assessment that the rho or related transformation does not offer any advantages over the transformations which you discussed in more detail. According to my experience, the relative power of Wynn's rho algorithm and of related transformations compared to Levin-type transformations depends very much on the problem under consideration. My personal experiences seem to indicate that transformations like Wynn's epsilon and rho algorithm are able to accomplish at least some acceleration of convergence in situations in which the input data are highly "polluted" and thus are essentially indigestible for Levin-type transformations.

According to my own practical experience, the most severe "pollution" occurred in the case of the extrapolation of quantum chemical oligomer calculations to the infinite chain limit of a stereoregular *quasi*-onedimensional polymer as for example polyacetylene [33, 100, 101]. The input data for these extrapolation calculations are produced by commercial or slightly adapted commercial program packages for quantum chemical *ab initio* calculations as for example GAUSSIAN. These packages do a lot of complicated numerics: First, the matrix elements – the so-called molecular multicenter integrals – have to be calculated and then, a self-consistent diagonalization of something resembling an often huge generalized matrix eigenvalue problem has to be done iteratively and self-consistently. All these computational steps require somewhat drastic approximations to become feasible. Accordingly, in the case of such a program package, it is practically impossible to apply the conventional techniques of numerical mathematics like backward error analysis. Thus, program packages of that kind have to be viewed as huge "black boxes" that respond to some input by producing some output of more or less unknown quality.

Nevertheless, it is possible to accomplish even under such difficult circumstances some stabilization of transformation results. However, I consistently observed in the extrapolation calculations mentioned above that for instance Levin's u transformation was not particularly effective, whereas Wynn's rho algorithm was usually doing a (very) good job.

Wynn's rho algorithm, which in its general form is given by the following recursive scheme [105],

$$\rho_{-1}^{(n)} \; = \; 0 \, , \qquad \rho_{0}^{(n)} \; = \; s_n \, , \qquad n \in \mathbb{N}_0 \, , \tag{6a}$$

$$\rho_{k+1}^{(n)} \; = \; \rho_{k-1}^{(n+1)} + \frac{x_{n+k+1} - x_n}{\rho_k^{(n+1)} - \rho_k^{(n)}} \, , \qquad k, n \in \mathbb{N}_0 \, , \tag{6b}$$

is the best known example of a whole group of closely related nonlinear transformations that can all be very useful.

Wynn's rho algorithm is actually a recursive scheme that computes an interpolating rational function

$$\mathscr{S}_{2k}(x) \; = \; \frac{a_0^{(k)} + a_1^{(k)} x + a_2^{(k)} x^2 + \cdots + a_l^{(k)} x^k}{b_0^{(k)} + b_1^{(k)} x + b_2^{(k)} x^2 + \cdots + b_m^{(k)} x^k} \, , \qquad k, l, m \in \mathbb{N}_0 \, , \tag{7}$$

satisfying

$$\mathscr{S}_{2k}(x_{n+j}) \; = \; s_{n+j} \, , \qquad k, n \in \mathbb{N}_0 \, , \quad 0 \le j \le 2k \, , \tag{8}$$

and interpolates the rational function to infinity. Thus, Wynn's rho algorithm (6) tacitly assumes that the interpolation points $\{x_n\}_{n=0}^{\infty}$ are strictly increasing and unbounded, satisfying

$$0 < x_0 < x_1 < \cdots < x_m < x_{m+1} < \cdots \, , \tag{9a}$$

$$\lim_{n \to \infty} x_n \; = \; \infty \, . \tag{9b}$$

By iterating the explicit expression for $\rho_2^{(n)}$ along the lines of Aitken's iterated Δ^2 process, the following close relative of Wynn's rho algorithm can be constructed [83, Eq. (6.3-3)]:

$$\mathscr{W}_0^{(n)} \; = \; s_n \, , \qquad n \in \mathbb{N}_0 \, , \tag{10a}$$

$$\mathscr{W}_{k+1}^{(n)} \; = \; \mathscr{W}_k^{(n+1)} + \frac{(x_{n+2k+2} - x_n) \bigl[\Delta \mathscr{W}_k^{(n+1)} \bigr] \bigl[\Delta \mathscr{W}_k^{(n)} \bigr]}{(x_{n+2k+2} - x_{n+1}) \bigl[\Delta \mathscr{W}_k^{(n)} \bigr] - (x_{n+2k+1} - x_n) \bigl[\Delta \mathscr{W}_k^{(n+1)} \bigr]} \, , \qquad k, n \in \mathbb{N}_0 \, . \tag{10b}$$

This is not the only possibility of iterating $\rho_2^{(n)}$. However, the iterations derived by Bhowmick, Bhattacharya, and Roy [10] are significantly less efficient than $\mathscr{W}_k^{(n)}$, which has similar properties to Wynn's rho algorithm [83]. In my opinion, it is very important to iterate the general form (6) of the rho algorithm. Bhowmick, Bhattacharya, and Roy [10] were misled because they started from the standard form (11) of Wynn's rho algorithm which will be discussed later.

The main practical problem with sequence transformations based upon interpolation theory is that for a given sequence $\{s_n\}_{n=0}^{\infty}$ one has to find suitable interpolation points $\{x_n\}_{n=0}^{\infty}$ that produce good results. In fortunate cases, there may be additional information which answers this question, but in general, this is both a nontrivial as well as a practically very relevant problem.

In the vast majority of all applications, Wynn's rho algorithm (6) and its iteration (10) have been used in combination with the interpolation points $x_n = n + 1$, yielding the standard forms[10] (see for example [81, Eq. (6.2-4)])

$$\rho_{-1}^{(n)} \; = \; 0 \, , \qquad \rho_0^{(n)} \; = \; s_n \, , \qquad n \in \mathbb{N}_0 \, , \tag{11a}$$

$$\rho_{k+1}^{(n)} \; = \; \rho_{k-1}^{(n+1)} + \frac{k+1}{\rho_k^{(n+1)} - \rho_k^{(n)}} \, , \qquad k, n \in \mathbb{N}_0 \, , \tag{11b}$$

and [83, Section 6.3]

$$\mathscr{W}_0^{(n)} \; = \; s_n \, , \qquad n \in \mathbb{N}_0 \, , \tag{12a}$$

$$\mathscr{W}_{k+1}^{(n)} \; = \; \mathscr{W}_k^{(n+1)} - \frac{(2k+2) \bigl[\Delta \mathscr{W}_k^{(n+1)} \bigr] \bigl[\Delta \mathscr{W}_k^{(n)} \bigr]}{(2k+1) \Delta^2 \mathscr{W}_k^{(n)}} \, , \qquad k, n \in \mathbb{N}_0 \, . \tag{12b}$$

However, these standard forms are not suited for all logarithmically convergent sequences of interest. The elements of many practically relevant logarithmically convergent sequences $\{s_n\}_{n=0}^{\infty}$ can at least for large indices n be represented by series expansions of the following kind:

$$s_n \; = \; s + (n + \beta)^{-\alpha} \sum_{j=0}^{\infty} c_j / (n + \beta)^j \, , \qquad n \in \mathbb{N}_0 \, . \tag{13}$$

[10] Actually, in many articles as well as in some books only the standard form of the rho algorithm is presented. In my opinion, this is not good, because we otherwise tend to forget that the rho algorithm can be viewed to be the recursive solution of a rational interpolation problem that performs extrapolation to infinity.

Here, α is a positive decay parameter and β is a positive shift parameter. In [66, Theorem 3.2], it was shown that the standard form (11) of the rho algorithm accelerates the convergence of sequences of the type of (13) if α is a positive integer, but fails if α is nonintegral. In the case of the iteration of Wynn's rho algorithm, no rigorous theoretical result seems to be known but there is considerable empirical evidence that it only accomplishes something if α is a positive integer.

If the decay parameter α of a sequence of the type of (13) is known, then Osada's variant of Wynn's rho algorithm can be used [66, Eq. (3.1)]:

$$\bar{\rho}_{-1}^{(n)} \,=\, 0\,, \qquad \bar{\rho}_{0}^{(n)} \,=\, s_n\,, \qquad n \in \mathbb{N}_0\,, \tag{14a}$$

$$\bar{\rho}_{k+1}^{(n)} \,=\, \bar{\rho}_{k-1}^{(n+1)} + \frac{k+\alpha}{\bar{\rho}_{k}^{(n+1)} - \bar{\rho}_{k}^{(n)}}\,, \qquad k,n \in \mathbb{N}_0\,. \tag{14b}$$

Osada also demonstrated that his variant accelerates the convergence of sequences of the type of (13) for arbitrary $\alpha > 0$, and that the transformation error satisfies the following asymptotic estimate [66, Theorem 4]:

$$\bar{\rho}_{2k}^{(n)} - s \,=\, \mathrm{O}\bigl(n^{-\alpha-2k}\bigr)\,, \qquad n \to \infty\,. \tag{15}$$

Osada's variant of the rho algorithm can be iterated. From (14) we obtain the following expression for $\bar{\rho}_{2}^{(n)}$ in terms of s_n, s_{n+1}, and s_{n+2}:

$$\bar{\rho}_{2}^{(n)} \,=\, s_{n+1} - \frac{(\alpha+1)}{\alpha} \frac{[\Delta s_n][\Delta s_{n+1}]}{[\Delta^2 s_n]}\,, \qquad n \in \mathbb{N}_0\,. \tag{16}$$

If the iteration is done in such a way that α is increased by two with every recursive step, we obtain the following recursive scheme [83, Eq. (2.29)] which was originally derived by Bjørstad, Dahlquist, and Grosse [11, Eq. (2.4)]:

$$\overline{\mathscr{W}}_{0}^{(n)} \,=\, s_n\,, \qquad n \in \mathbb{N}_0\,, \tag{17a}$$

$$\overline{\mathscr{W}}_{k+1}^{(n)} \,=\, \overline{\mathscr{W}}_{k}^{(n+1)} - \frac{(2k+\alpha+1)}{(2k+\alpha)} \frac{\bigl[\Delta \overline{\mathscr{W}}_{k}^{(n+1)}\bigr]\bigl[\Delta \overline{\mathscr{W}}_{k}^{(n)}\bigr]}{\Delta^2 \overline{\mathscr{W}}_{k}^{(n)}}\,, \quad k,n \in \mathbb{N}_0\,. \tag{17b}$$

Bjørstad, Dahlquist, and Grosse showed that $\overline{\mathscr{W}}_{k}^{(n)}$ accelerates the convergence of sequences of the type of (13), and that the transformation error satisfies the following asymptotic estimate [11, Eq. (3.1)]:

$$\overline{\mathscr{W}}_{k}^{(n)} - s \,=\, \mathrm{O}\bigl(n^{-\alpha-2k}\bigr)\,, \qquad n \to \infty\,. \tag{18}$$

The explicit knowledge of the decay parameter α is crucial for an application of the transformations (14) and (17) to a sequence of the type of (13). An approximation to α can be obtained with the help of the following nonlinear transformation, which was first derived in a somewhat disguised form by Drummond [38] and later rederived by Bjørstad, Dahlquist, and Grosse [11]:

$$T_n \,=\, \frac{[\Delta^2 s_n]\,[\Delta^2 s_{n+1}]}{[\Delta s_{n+1}]\,[\Delta^2 s_{n+1}] - [\Delta s_{n+2}]\,[\Delta^2 s_n]} \,-\, 1\,, \qquad n \in \mathbb{N}_0\,. \tag{19}$$

T_n is essentially a weighted Δ^3 method, which implies that it is potentially very unstable. Thus, stability problems are likely to occur if the relative accuracy of the input data is low. Bjørstad, Dahlquist, and Grosse [11, Eq. (4.1)] also showed that

$$\alpha \,=\, T_n + O(1/n^2)\,, \qquad n \to \infty\,, \tag{20}$$

if the elements of a sequence of the type of (13) are used as input data.

The sequence transformations mentioned above are not particularly relevant in the context of special function evaluation – apart from Dirichlet series for zeta functions logarithmic convergence does not occur too frequently in special functions theory which is dominated by power series – but in different contexts logarithmic convergence occurs quite frequently. The main advantage of Wynn's rho algorithm and its close relatives is that they are (much) more robust than Levin-type transformations.

At the bottom of the 1st paragraph of Chapter 5.3.3 on p. 216, you mention divergent series with *monotonic* terms. While the summation of strictly alternating divergent series has reached a relatively high degree of sophistication[11], the summation of divergent monotonic series is a tricky business that can easily lead to *counter-intuitive* results.

As an example, let us consider the geometric series:

$$\sum_{\nu=0}^{\infty} z^{\nu} \; = \; \frac{1}{1-z} \, . \tag{21}$$

As is well known, this series converges in the interior $|z| < 1$ of the unit circle, and diverges for $|z| \geq 1$. A reasonable summation technique should sum the partial sums $\sum_{\nu=0}^{n} z^{\nu} = [1-z^{n+1}]/[1-z]$ of the geometric series to the generalized limit $1/[1-z]$. If, however, z is positive and greater 1, then we get a highly counter-intuitive *negative* summation result. Let us for instance assume $z = 2$. Then, we obtain a result that cannot be motivated with the help of simple plausibility arguments:

$$\sum_{\nu=0}^{\infty} 2^{\nu} \; = \; 1+2+4+8+\ldots \; = \; \frac{1}{1-2} \; = \; -1 \, . \tag{22}$$

Another very serious problem is that the summation of a monotone factorially divergent power series with real terms usually must produce something complex, and to make things worse, the nonzero imaginary part is a nonanalytic contribution that cannot be deduced in a straightforward way from the divergent power series which is usually asymptotic as the effective argument approaches zero.

These complications can be understood comparatively easily on the basis of the exponential integral [1, Eq. (5.1.1)],

$$E_1(z) \; = \; \int_z^{\infty} \frac{\exp(-t) \mathrm{d}t}{t} \, , \tag{23}$$

which possesses the following asymptotic expansion as $z \to \infty$ [1, Eq. (5.1.51)],

$$z \mathrm{e}^z E_1(z) \; \sim \; \sum_{m=0}^{\infty} \frac{(-1)^m m!}{z^m} \; = \; {}_2F_0(1,1;-1/z) \, , \qquad z \to \infty \, , \tag{24}$$

and which can also be expressed as a Stieltjes integral [1, Eq. (5.1.28)]:

$$\mathrm{e}^z E_1(z) \; = \; \int_0^{\infty} \frac{\exp(-t) \mathrm{d}t}{z+t} \; = \; \frac{1}{z} \int_0^{\infty} \frac{\exp(-t) \mathrm{d}t}{1+t/z} \, . \tag{25}$$

[11] Even the so-called *truncation at the minimal term* often suffices to produce results of acceptable accuracy in the case of strictly alternating divergent series. Moreover, techniques like Borel summation, Padé approximants, and numerous sequence transformations are usually able to accomplish *much* more accurate summation results.

For $z > 0$, the Stieltjes integral on the right-hand side, which in this case coincides with the Laplace integral of the Borel summation method, is perfectly well defined and can safely be used for the evaluation of the exponential integral by numerical quadrature. For $z < 0$, there is, however, a pole on the integration contour at $t = -z$. As a remedy, the integration contour has to be augmented by a semicircular indentation around the pole into either the upper or lower part of the complex plane. In the limit of a vanishing radius of the semicircular indentation, the Stieltjes integral has to be reinterpreted as a *Cauchy principal value integral* of the type of the exponential integral Ei according to [1, Eq. (5.1.2)]

$$\mathrm{Ei}(z) \;=\; -\int_{-z}^{\infty} \frac{\exp(-t)\mathrm{d}t}{t} \;=\; \int_{-\infty}^{z} \frac{\exp(t)\mathrm{d}t}{t}\,, \qquad z > 0\,, \tag{26}$$

and the semicircular indentation essentially produces half of the residue of the integrand $\exp(-t)$ resulting from to the Stieltjes measure $\exp(-t)\mathrm{d}t$ at the pole $t = -z$. Thus, we finally obtain [1, Eq. (5.1.7)]:

$$E_1(-z \pm \mathrm{i}0) \;=\; -\mathrm{Ei}(z) \mp \mathrm{i}\pi\,, \qquad z > 0\,. \tag{27}$$

The sign of the imaginary contribution $\mp \mathrm{i}\pi$ depends on the sign of $-z \pm \mathrm{i}0$, i.e., whether the integration contour is indented into the upper or lower part of the complex plane. Accordingly, the exponential integral $E_1(z)$ has a cut along the negative real axis.

If we replace in (25) z by $1/(-z \pm \mathrm{i}0) = -1/z \mp \mathrm{i}0$ with $z > 0$ and combine the resulting expression with (27), we obtain

$$\frac{\exp\bigl(1/(-z \pm \mathrm{i}0)\bigr)}{-z \pm \mathrm{i}0} E_1\bigl(1/(-z \pm \mathrm{i}0)\bigr) \;=\; -\frac{\exp(-1/z)}{x} E_1(-1/z \mp \mathrm{i}0) \tag{28}$$

$$= \int_0^{\infty} \frac{\exp(-t)\mathrm{d}t}{1 + (-z \pm \mathrm{i}0)t} \tag{29}$$

$$= \frac{\exp(-1/z)}{z} \bigl\{\mathrm{Ei}(1/z) \mp \mathrm{i}\pi\bigr\}\,. \tag{30}$$

This expression contains a nonzero imaginary part $\mp\pi \exp(-1/z)/z$ which is nonanalytic as $z \to 0$. However, the asymptotic series

$$\int_0^{\infty} \frac{\exp(-t)\mathrm{d}t}{1 + (-z \pm \mathrm{i}0)t} \sim \sum_{m=0}^{\infty} m!\,z^m \;=\; {}_2F_0(1,1;z)\,, \qquad z \to 0\,, \tag{31}$$

which we obtain by replacing z by $1/(-z \pm \mathrm{i}0)$ in the asymptotic series (24), does not contain any information on the nonanalytic imaginary part. Thus, we have the seemingly paradoxical situation that a completely successful summation process has to yield something complex from a factorially divergent series of positive terms. We also cannot recover the imaginary part by truncating the divergent series appropriately.

Most summation methods cannot recover such a nonanalytic imaginary part without additional information about the function that is to be reconstructed from the divergent series. This is a serious problem in quantum mechanics where the imaginary part of the summation of a divergent monotone series can usually be linked to a resonance width and thus produces vital physical information (compare for instance [43] and references therein).

This lack of information about a possibly nonzero nonanalytic imaginary part is a serious limitation of asymptotic series, which is related to the principal nonuniqueness of asymptotic series. Given a function, its asymptotic series is unique if it exists. The converse is, however, unfortunately not true. Thus, two or more functions that differ by nonanalytic contributions of the type of $\exp(-1/z)$ all have the same asymptotic series as $z \to 0$. The summation of a divergent power series can be considered to be an attempt at reconstructing the function from its divergent asymptotic series. Consequently, this principal nonuniqueness is a serious problem which can only be overcome if some additional information on the function is available.

As remarked above, the reconstruction of a nonzero imaginary part from a factorially divergent monotonic series is a more or less insurmountable problem for most of the commonly used sequence transformations such as Wynn's epsilon algorithm or Levin-type transformations. However, the *quadratic* or more generally *algebraic* approximants introduced by Shafer [74], which generalize Padé approximants, can accomplish this. Application of these approximants, a description of the algorithmic problems, and further references can for instance be found in articles by Baker [5, 6], Feil and Homeier [41], Fernández [42], Goodson [46, 47, 48], Goodson and Sergeev [49], Sergeev and Goodson [73], Loi and McInnes [62], and Tourigny and Drazin [78].

On p. 217, you discuss the evaluation of the oscillatory integral representation (5.3.22) for the modified Bessel function $K_0(1)$. You essentially suggest converting this semi-infinite integral into an alternating series of integrals between neighboring zeros of the Bessel function $J_0(x)$. The convergence of this alternating series is most likely quite slow because of the slow decay of the integrand in (5.3.22), but can be accelerated by a variety of techniques such as Wynn's epsilon algorithm or Levin-type transformations.

This is a standard approach, which should in principle work for any semi-infinite integral involving a J function, but I strongly suspect that this approach is by no means optimal. I have a colleague

Hassan Safouhi
Faculté Saint-Jean
University of Alberta
8406, 91 Street
Edmonton, Alberta T6C 4G9
Canada

who has worked quite a lot on the numerical evaluation of highly oscillatory integrals with the help of extrapolation techniques (see for example [9, 39, 69, 70, 71, 72] and references therein). I will ask Hassan Safouhi whether he would be able to provide a more efficient approach for the integral representation (5.3.22).

In Chapter 5.12 on pp. 245 - 247, you discuss Padé approximants and define them as the ratio of two polynomials whose coefficients are the solution of the system of coupled linear equations (5.12.5) and (5.12.6). This is standard. However, in this context, you might emphasize – just as you did it in Chapter 3.5 on p. 129 – that this approach makes sense only if the coefficients of the Padé are needed. In most practical applications, however, only the *numerical values* of Padé approximants are needed. In that case, it is a much better idea to compute Padé approximants recursively. In my own work, I have always used Wynn's epsilon which you also describe in your book. A review of different computational approaches for Padé approximants can be found in the book by Cuyt and Wuytack [36, Chapter II §3, pp. 76 - 95] and in an older article by Wuytack [103].

In the 1st paragraph on p. 247, you write about Padé approximants:

> Why does this work? Are there not other functions with the same first five terms in their power series but completely different behavior in the range (say) $2 < x < 10$? Indeed there are. Padé approximation has the uncanny knack of picking the function *you had in mind* from among all the possibilities. *Except when it doesn't!* That is the downside of Padé approximation: It is uncontrolled. There is, in general, no way to tell how accurate it is, or how far in x it can usefully be extended. It is a powerful but in the end still mysterious technique.

Of course, your remarks about Padé approximation apply just as well to other sequence transformations which also transform the partial sums of a formal power series to rational functions. I consider your remark to be a confirmation of my view that any nontrivial application of Padé approximants – or also of other sequence transformations – should be viewed as a numerical experiment, and that the experimentalist should proceed with utmost caution. Theoretical error estimates are in most cases of practical relevance useless.

I suspect that you are worried about the following kind of scenario: Let us assume that we have a family of functions $f_j(z)$ all possessing power series about zero:

$$f_j(z) \; = \; \sum_{m=0}^{\infty} c_m^{(j)} \, z^m \, , \qquad j = 0, 1, 2, \ldots \, . \tag{32}$$

Let us now construct a function $F(z)$ according to the following rule:

$$F(z) \; = \; \sum_{m=0}^{M_0} c_m^{(0)} \, z^m + \sum_{m=M_0+1}^{M_1} c_m^{(1)} \, z^m + \ldots + \sum_{m=M_k+1}^{M_{k+1}} c_m^{(k)} \, z^m + \ldots \, . \tag{33}$$

I am convinced that in this way a power series can be constructed that brings Padé approximants – or any other class of rational approximants – down to their knees.

But firstly, such a function F violates some kind of philosophical principle which we tacitly assume when applying a sequence transformation[12]: A sequence transformation accomplishes an acceleration of convergence and a summation in the case of divergence by detecting and utilizing some regularity in the behavior of the data. It obviously makes no sense to apply a sequence transformation to a sequence with a completely random and arbitrary behavior, even if it ultimately converges.

Secondly, it should at least in principle be possible to detect that for instance the series coefficients $c_m^{(k)}$ and $c_m^{(k+1)}$ possess a different index dependence. This can be accomplished with the help of *Padé prediction*, which is an admittedly expensive technique that is best done symbolically and therefore not necessarily practically applicable in all cases of interest.

[12]Padé approximation can be viewed to be simply some special sequence transformation since the partial sums of a power series are transformed to a doubly indexed sequence of rational functions.

Let us assume that the power series

$$f(z) \; = \; \sum_{\nu=0}^{\infty} \gamma_\nu \, z^\nu \tag{34}$$

converges in a neighborhood of zero. A Padé approximant $[l/m]_f(z)$ to $f(z)$ is a rational function

$$[l/m]_f(z) \; = \; \frac{P^{[l/m]}(z)}{Q^{[l/m]}(z)} \; = \; \frac{p_0 + p_1 z + p_2 z^2 + \cdots + p_l z^l}{1 + q_1 z + q_2 z^2 + \cdots + q_m z^m} \, , \tag{35}$$

which satisfies the asymptotic condition

$$f(z) - P^{[l/m]}(z)/Q^{[l/m]}(z) \; = \; \mathrm{O}(z^{l+m+1}) \, , \quad z \to 0 \, . \tag{36}$$

This *accuracy-through-order relationship* implies that $[l/m]_f(z)$ can be written as the partial sum from which it was constructed, plus a term which was generated by the transformation of the partial sum to the rational approximant:

$$[l/m]_f(z) \; = \; \sum_{\nu=0}^{l+m} \gamma_\nu \, z^\nu + z^{l+m+1} \, \mathscr{P}^{[l/m]}(z) \; = \; f_{l+m}(z) + z^{l+m+1} \, \mathscr{P}^{[l/m]}(z) \, . \tag{37}$$

Similarly, the power series for f can be rewritten as follows:

$$f(z) \; = \; \sum_{\nu=0}^{l+m} \gamma_\nu \, z^\nu + z^{l+m+1} \, \mathscr{F}_{l+m+1}(z) \; = \; f_{l+m}(z) + z^{l+m+1} \, \mathscr{F}_{l+m+1}(z) \, . \tag{38}$$

Let us now assume that the Padé approximants $[l/m]_f(z)$ provide better approximations to $f(z)$ than the partial sums $f_{l+m}(z)$ from which they are constructed:

$$|f(z) - [l/m]_f(z)| \; < \; |f(z) - f_{l+m}(z)| \; = \; |z^{l+m+1} \, \mathscr{F}_{l+m+1}(z)| \, . \tag{39}$$

Thus, already relatively small indices l and m of the Padé approximants suffice to produce reasonably accurate approximations,

$$|f(z) - [l/m]_f(z)| \; < \; \varepsilon \, , \tag{40}$$

where ε is a fixed positive real number. This, however, implies that a Padé transformation term $z^{l+m+1} \mathscr{P}^{[l/m]}(z)$ must also provide a sufficiently accurate approximation to the corresponding truncation error $z^{l+m+1} \mathscr{F}_{l+m+1}(z)$ of the power series according to

$$|z^{l+m+1} \, \mathscr{F}_{l+m+1}(z) - z^{l+m+1} \, \mathscr{P}^{[l/m]}(z)| \; < \; \varepsilon \, . \tag{41}$$

By transforming the partial sum $f_{l+m}(z)$ to a Padé approximant $[l/m]_f(z)$, the partial sum is augmented by a term $z^{l+m+1} \mathscr{P}^{[l/m]}(z)$ which tries to simulate the truncation error $z^{l+m+1} \mathscr{F}_{l+m+1}(z)$ of the power series.

Thus, if the construction of the Padé approximant improves convergence, we have instead of the bare partial sum $f_{l+m}(z)$ something which effectively behaves like a partial sum with a larger number of terms. Consequently, the truncation error of such a Padé approximant is smaller than the truncation error of the partial sum from which it was constructed. This implies that the construction of a Padé approximant also corresponds to something like the construction and elimination of approximations to the actual remainders.

It is of course possible that a Padé approximant $[l/m]_f(z)$ provides an inferior approximation to $f(z)$ than the partial sum $f_{l+m}(z)$ from which it was constructed. However, also in this case the difference between $[l/m]_f(z)$ and $f_{l+m}(z)$ is entirely due to the term $z^{l+m+1}\mathscr{P}^{[l/m]}(z)$.

Padé approximants are by construction analytic in a neighborhood of the origin. Consequently, we can do a Taylor expansion of either $[l/m]_f(z)$ or $\mathscr{P}^{[l/m]}(z)$ around $z=0$, yielding

$$[l/m]_f(z) \; = \; \sum_{\nu=0}^{\infty} \gamma_{\nu}^{[l/m]} z^{\nu} \tag{42}$$

or equivalently

$$\mathscr{P}^{[l/m]}(z) \; = \; \sum_{\nu=0}^{\infty} \gamma_{l+m+\nu+1}^{[l/m]} z^{\nu} \, . \tag{43}$$

If $\mathscr{P}^{[l/m]}(z)$ and $\mathscr{F}_{l+m+1}(z)$ were equal – which is the case if if f is a rational function that can be reproduced exactly by the Padé approximant $[l/m]_f(z)$ – we would of course obtain

$$\gamma_{l+m+\nu+1}^{[l/m]} \; = \; \gamma_{l+m+\nu+1} \, , \qquad \nu \in \mathbb{N}_0 \, . \tag{44}$$

In the case of an essentially arbitrary function f, we have no reason to assume that $\mathscr{P}^{[l/m]}(z)$ and $\mathscr{F}_{l+m+1}(z)$ are identical for finite values of l and m. Consequently, (44) will normally not be valid. Nevertheless, the approximate equality of $\mathscr{P}^{[l/m]}(z)$ and $\mathscr{F}_{l+m+1}(z)$ for finite values of l and m implies that at least the leading coefficients $\gamma_{l+m+1}^{[l/m]}$, $\gamma_{l+m+2}^{[l/m]}, \ldots$ of the Taylor expansions of $\mathscr{P}^{[l/m]}(z)$ should provide approximations to the corresponding coefficients γ_{l+m+1}, $\gamma_{l+m+2}, \ldots$ of the power series.

Our conclusions about the role of the Padé transformation term $z^{l+m+1}\mathscr{P}^{[l/m]}(z)$ do not depend on the initial assumption that the power series for f possesses a nonzero but finite radius of convergence $0 < R < \infty$. If the power series for f has a zero radius of convergence, the Padé approximant $[l/m]_f(z)$ can still be partitioned into the partial sum $f_{l+m}(z)$ and the transformation term $z^{l+m+1}\mathscr{P}^{[l/m]}(z)$ according to (37), just as the now formal power series can still be partitioned into the partial sum $f_{l+m}(z)$ and the truncation error $z^{l+m+1}\mathscr{F}_{l+m+1}(z)$ according to (38). The only difference is that a Taylor series for $\mathscr{F}_{l+m+1}(z)$ now does not converge. Nevertheless, $z^{l+m+1}\mathscr{P}^{[l/m]}(z)$ provides an approximation to $z^{l+m+1}\mathscr{F}_{l+m+1}(z)$ which implies that the approximate equality of the leading coefficients $\gamma_{l+m+1}^{[l/m]}$, $\gamma_{l+m+2}^{[l/m]}, \ldots$ of the Taylor expansion (44) with the corresponding exact series coefficients $\gamma_{l+m+1}, \gamma_{l+m+2}, \ldots$ still holds.

Thus, Padé approximants can be used to make predictions for unknown power series coefficients. This fact was apparently first observed and utilized by Gilewicz [45].

In any way, with the help of Padé prediction, it should be possible to construct a Padé approximants to $F(z)$ defined by (33) from the first coefficients $c_m^{(0)}$ with $0 \le m \le M_0$. If the first and second coefficients $c_m^{(0)}$ with $0 \le m \le M_0$ and $c_m^{(1)}$ with $M_0 + 1 \le m \le M_1$, respectively, differ sufficiently, then we should be able to see it from the first prediction to $c_{M_0+1}^{(1)}$ made by this Padé approximants constructed exclusively from the series terms $c_m^{(0)}$ with $0 \le m \le M_0$.

In [90] I had derived formulas which permit a comparatively convenient *recursive* computation of the first prediction made by those Padé approximants that can be computed by Wynn's epsilon algorithm. In [8]), it was shown that in this way very accurate predictions are possible. Prediction techniques were also used in the articles [54, 55, 89]. These articles also contain numerous other references.

In the context of practical applications, it is extremely important to get at least some idea that explains not only the power of Padé approximation, but also its shortcomings.

In my opinion, one should always take into account that ordinary Padé approximants – or also other rational approximants obtained for instance from Levin-type transformations – are *local* approximants since they only use information about the function, which they try to represent, from a single point. Many people tend to forget that, since in the vast majority of applications, Padé approximants provide *much* better local approximations than the partial sums of the power series from which they are constructed, and in fortunate cases, Padé approximants constructed from a power series expansion about zero even provide meaningful information about the behavior of the function under consideration at infinity[13]. However, reliable approximation results over a larger range of arguments usually require that we use *two-point* or more generally *multi-point* Padé approximants, which use input from two or more points and which are for instance described in [7, Chapter 7.1]. According to my own experience, two-point Padé approximants can be extremely useful [7]. There are, however, two serious practical problems: Firstly, it is often not so easy to obtain expansions around two or more expansion points. Secondly, computational algorithms for two- and multi-point Padé are not nearly as sophisticated and convenient as those for ordinary Padé approximants.

It should be clear that the ability of Padé approximants that are constructed from an expansion about zero of providing reasonably accurate approximations at infinity depends crucially on the asymptotics of the function under consideration. Let us for instance assume that a function possesses the following asymptotic behavior:

$$f(z) \sim z^{\alpha}, \qquad z \to \infty, \quad \alpha \in \mathbb{R}. \tag{45}$$

If the decay parameter α happens to be not a positive or negative integer, but a real number, then Padé approximants to $f(z)$ constructed from the expansion about zero cannot provide good approximations at infinity.

A possible remedy would be to construct a power series expansion of the function

$$g(z) = z^{-\alpha} f(z). \tag{46}$$

Since $g(z)$ should approach in general a nonzero constant as $z \to \infty$, diagonal Padé approximants – or also other rational approximants obtained by applying sequence transformations – should be well suited to describe the behavior of $g(z)$ at infinity.

The idea of constructing rational approximants not to the function $f(z)$, but to a related function $g(z)$ with a more appropriate asymptotic behavior, was the essential step that made the otherwise very difficult summation calculations described in [86, 88, 98, 99] successful.

[13] A striking example of the power, which Padé approximation can have under favorable circumstances, is the function given in [3, Eq. (1.1) on p. 3]. As shown in [3, pp. 4 - 5], already very small Padé approximants to this function reproduce its behavior at infinity with remarkable accuracy, although the radius of convergence of its power series is just $1/2$.

In [68, p. 247], you give general references on Padé approximants. In my opinion, the most complete source on this topic is the monograph by Baker and Graves-Morris [7]. This is an excellent book by two authors who work both *on* and *with* Padé approximants. Because of the wealth of information it contains, anybody intending to work seriously either *on* or *with* Padé approximants will have to study it. However, I have to concede that this book has a serious drawback: Its sheer size (746 pages) may look intimidating, in particular for novices. As a more condensed presentation I can recommend another book by Baker on critical phenomena [4] whose Part III treats the theory of Padé approximation in a sufficiently detailed way.

If you have any question about nonlinear sequence transformations or other related topics, please do not hesitate to contact me.

Yours sincerely,

`Ernst Joachim Weniger`

References

[1] M. Abramowitz and I. A. Stegun (Editors), *Handbook of Mathematical Functions*, National Bureau of Standards, Washington, D. C. (1972).

[2] A. C. Aitken, On Bernoulli's numerical solution of algebraic equations, Proc. Roy. Soc. Edinburgh **46**, 289 – 305 (1926).

[3] G. A. Baker, Jr., *Essentials of Padé Approximants*, Academic Press, New York (1975).

[4] G. A. Baker, Jr., *Quantitative Theory of Critical Phenomena*, Academic Press, San Diego (1990).

[5] G. A. Baker, Jr., The accuracy-through-order and the equivalence properties in the algebraic approximant, in A. Cuyt (Editor), *Nonlinear Numerical Methods and Rational Approximation II*, 283 – 289, Kluwer, Dordrecht (1994).

[6] G. A. Baker, Jr., Algebraic approximants and the numerical solution of parabolic equations, J. Comput. Appl. Math. **105**, 129 – 139 (1999).

[7] G. A. Baker, Jr. and P. Graves-Morris, *Padé Approximants*, 2nd edition, Cambridge U. P., Cambridge (1996).

[8] C. M. Bender and E. J. Weniger, Numerical evidence that the perturbation expansion for a non-Hermitian $\mathscr{PT}$-symmetric Hamiltonian is Stieltjes, J. Math. Phys. **42**, 2167 – 2183 (2001).

[9] L. Berlu and H. Safouhi, Multicentre two-electron Coulomb and exchange integrals over Slater functions evaluated using a generalized algorithm based on nonlinear transformations, J. Phys. A **37**, 3393 – 3410 (2004).

[10] S. Bhowmick, R. Bhattacharya, and D. Roy, Iterations of convergence accelerating nonlinear transforms, Comput. Phys. Commun. **54**, 31 – 36 (1989).

[11] P. Bjørstad, G. Dahlquist, and E. Grosse, Extrapolations of asymptotic expansions by a modified Aitken δ^2-formula, BIT **21**, 56 – 65 (1981).

[12] F. Bornemann, D. Laurie, S. Wagon, and J. Waldvogel, *The SIAM 100-Digit Challenge: A Study in High-Accuracy Numerical Computing*, Society of Industrial Applied Mathematics, Philadelphia (2004).

[13] F. Bornemann, D. Laurie, S. Wagon, and J. Waldvogel, *Vom Lösen numerischer Probleme: Ein Streifzug entlang der "SIAM 10x10-Digit Challenge"*, Springer-Verlag, Berlin (2006).

[14] C. Brezinski, *Accélération de la Convergence en Analyse Numérique*, Springer-Verlag, Berlin (1977).

[15] C. Brezinski, *Algorithmes d'Accélération de la Convergence – Étude Numérique*, Éditions Technip, Paris (1978).

[16] C. Brezinski, *Padé-Type Approximation and General Orthogonal Polynomials*, Birkhäuser, Basel (1980).

[17] C. Brezinski, *History of Continued Fractions and Padé Approximants*, Springer-Verlag, Berlin (1991).

[18] C. Brezinski, *A Bibliography on Continued Fractions, Padé Approximation, Sequence Transformation and Related Subjects*, Prensas Universitarias de Zaragoza, Zaragoza (1991).

[19] C. Brezinski, Extrapolation algorithms and Padé approximations: A historical survey, Appl. Numer. Math. **20**, 299 – 318 (1996).

[20] C. Brezinski, *Projection Methods for Systems of Equations*, Elsevier, Amsterdam (1997).

[21] C. Brezinski, Error estimates and convergence acceleration, in H. Bulgak and C. Zenger (Editors), *Error Control and Adaptivity in Scientific Computing*, 87 – 94, Kluwer, Dordrecht (1999).

[22] C. Brezinski, Difference and differential equations, and convergence acceleration algorithms, CRM Proc. Lect. Notes **25**, 53 – 63 (2000).

[23] C. Brezinski, Convergence acceleration during the 20th century, J. Comput. Appl. Math. **122**, 1 – 21 (2000). Reprinted as [24].

[24] C. Brezinski, Convergence acceleration during the 20th century, in C. Brezinski (Editor), *Numerical Analysis 2000, Vol. 2: Interpolation and Extrapolation*, 1 – 21, Elsevier, Amsterdam (2000).

[25] C. Brezinski and A. C. Matos, A derivation of extrapolation algorithms based on error estimates, J. Comput. Appl. Math. **66**, 5 – 26 (1996).

[26] C. Brezinski and M. Redivo Zaglia, *Extrapolation Methods*, North-Holland, Amsterdam (1991).

[27] C. Brezinski and M. Redivo Zaglia, A general extrapolation procedure revisited, Adv. Comput. Math. **2**, 461 – 477 (1994).

[28] C. Brezinski and M. Redivo Zaglia, On the kernel of sequence transformations, Appl. Numer. Math. **16**, 239 – 244 (1994).

[29] C. Brezinski and M. Redivo Zaglia, Vector and matrix sequence transformations based on biorthogonality, Appl. Numer. Math. **21**, 353 – 373 (1996).

[30] C. Brezinski and A. Salam, Matrix and vector sequence transformations revisited, Proc. Edinb. Math. Soc. **38**, 495 – 510 (1995).

[31] C. Brezinski and E. J. Weniger, Book Review: A. Sidi: "Practical Extrapolation Methods. Theory and Applications", Cambridge University Press, Cambridge, 2003, Numer. Algor. **35**, 134 – 136 (2004).

[32] C. Brezinski and E. J. Weniger, Book Review: "Practical Extrapolation Methods, Theory and Applications" by A. Sidi, Cambridge University Press, Cambridge, UK, 2003, SIAM Rev. **46**, 360 – 362 (2004).

[33] J. Cioslowski and E. J. Weniger, Bulk properties from finite cluster calculations. VIII. Benchmark calculations on the efficiency of extrapolation methods for the HF and MP2 energies of polyacenes, J. Comput. Chem. **14**, 1468 – 1481 (1993).

[34] J. Čížek, E. J. Weniger, P. Bracken, and V. Špirko, Effective characteristic polynomials and two-point Padé approximants as summation techniques for the strongly divergent perturbation expansions of the ground state energies of anharmonic oscillators, Phys. Rev. E **53**, 2925 – 2939 (1996).

[35] J. Čížek, J. Zamastil, and L. Skála, New summation technique for rapidly divergent perturbation series. Hydrogen atom in magnetic field, J. Math. Phys. **44**, 962 – 968 (2003).

[36] A. Cuyt and L. Wuytack, *Nonlinear Methods in Numerical Analysis*, North-Holland, Amsterdam (1987).

[37] J.-P. Delahaye, *Sequence Transformations*, Springer-Verlag, Berlin (1988).

[38] J. E. Drummond, Summing a common type of slowly convergent series of positive terms, J. Austral. Math. Soc. **B 19**, 416 – 421 (1976).

[39] S. Duret and H. Safouhi, The W algorithm and the $\overline{D}$ transformation for the numerical evaluation of three-center nuclear attraction integrals, Int. J. Quantum Chem. **107**, 1060 – 1066 (2007).

[40] G. Evans, *Practical Numerical Integration*, Wiley, Chichester (1993).

[41] T. M. Feil and H. H. H. Homeier, Programs for the approximation of real and imaginary single- and multi-valued functions by means of Hermite-Padé-approximants, Comput. Phys. Commun. **158**, 124 – 135 (2004).

[42] F. M. Fernández, Fast and accurate method for summation of divergent series, Int. J. Quantum Chem. **81**, 268 – 279 (2001).

[43] J. Fischer, On the role of power expansions in quantum field theory, Int. J. Mod. Phys. A **12**, 3625 – 3663 (1997).

[44] A. Gil, J. Segura, and N. M. Temme, *Numerical Methods for Special Functions*, SIAM, Philadelphia (2007).

[45] J. Gilewicz, Numerical detection of the best Padé approximant and determination of the Fourier coefficients of the insufficiently sampled functions, in P. R. Graves-Morris (Editor), *Padé Approximants and Their Applications*, 99 – 103, Academic Press, London (1973).

[46] D. Z. Goodson, Convergent summation of Møller-Plesset perturbation theory, J. Chem. Phys. **112**, 4901 – 4909 (2000).

[47] D. Z. Goodson, A summation procedure that improves the accuracy of the fourth-order Møller-Plesset perturbation theory, J. Chem. Phys. **113**, 6461 – 6464 (2000).

[48] D. Z. Goodson, Improving the accuracy of ab initio methods with summation approximants and singularity analysis, Int. J. Quantum Chem. **92**, 35 – 46 (2003).

[49] D. Z. Goodson and A. V. Sergeev, On the use of algebraic approximants to sum divergent series for Fermi resonances in vibrational spectroscopy, J. Chem. Phys. **110**, 8205 – 8206 (1999).

[50] P. R. Graves-Morris, D. E. Roberts, and A. Salam, The epsilon algorithm and related topics, J. Comput. Appl. Math. **122**, 51 – 80 (2000). Reprinted as [51].

[51] P. R. Graves-Morris, D. E. Roberts, and A. Salam, The epsilon algorithm and related topics, in C. Brezinski (Editor), *Numerical Analysis 2000, Vol. 2: Interpolation and Extrapolation*, 51 – 80, Elsevier, Amsterdam (2000).

[52] H. H. H. Homeier, Scalar Levin-type sequence transformations, J. Comput. Appl. Math. **122**, 81 – 147 (2000). Reprinted as [53].

[53] H. H. H. Homeier, Scalar Levin-type sequence transformations, in C. Brezinski (Editor), *Numerical Analysis 2000, Vol. 2: Interpolation and Extrapolation*, 81 – 147, Elsevier, Amsterdam (2000).

[54] U. D. Jentschura, J. Becher, E. J. Weniger, and G. Soff, Resummation of QED perturbation series by sequence transformations and the prediction of perturbative coefficients, Phys. Rev. Lett. **85**, 2446 – 2449 (2000).

[55] U. D. Jentschura, E. J. Weniger, and G. Soff, Asymptotic improvement of resummations and perturbative predictions in quantum field theory, J. Phys. G **26**, 1545 – 1568 (2000).

[56] K. Knopp, *Theorie und Anwendung der unendlichen Reihen*, Springer-Verlag, Berlin (1964).

[57] P. K. Kythe and M. R. Schäferkotter, *Handbook of Computational Methods for Integration*, Chapman & Hall/CRC Press, Boca Raton (2005).

[58] C. N. Leung and J. A. Murakowski, Padé interpolation: Methodology and application to the quarkonium, J. Math. Phys. **41**, 2700 – 2705 (2000).

[59] C. N. Leung and Y. Y. Y. Wong, Application of Padé interpolation to stationary state problems, Amer. J. Phys. **70**, 1020 – 1024 (2002).

[60] D. Levin, Development of non-linear transformations for improving convergence of sequences, Int. J. Comput. Math. B **3**, 371 – 388 (1973).

[61] C. B. Liem, T. Lü, and T. M. Shih, *The Splitting Extrapolation Method*, World Scientific, Singapore (1995).

[62] S. L. Loi and A. W. McInnes, An algorithm for the quadratic approximation, J. Comput. Appl. Math. **11**, 161 – 174 (1984).

[63] G. I. Marchuk and V. V. Shaidurov, *Difference Methods and Their Extrapolations*, Springer-Verlag, New York (1983).

[64] A. C. Matos, Linear difference operators and acceleration methods, IMA J. Numer. Anal. **20**, 359 – 388 (2000).

[65] T. Meurer, *Feedforward and feedback tracking control of diffusion convection reaction systems using summability methods*, Ph.D. thesis, Fakultät Maschinenbau, Universität Stuttgart (2005).

[66] N. Osada, A convergence acceleration method for some logarithmically convergent sequences, SIAM J. Numer. Anal. **27**, 178 – 189 (1990).

[67] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes*, Cambridge U. P., Cambridge (1986).

[68] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, 3rd edition, Cambridge U. P., Cambridge (2007).

[69] H. Safouhi, The properties of sine, spherical Bessel and reduced Bessel functions for improving convergence of semi-infinite very oscillatory integrals: The evaluation of three-centre nuclear attraction integrals over B functions, J. Phys. A **34**, 2801 – 2818 (2001).

[70] H. Safouhi, Numerical evaluation of three-center two-electron Coulomb and hybrid integrals over B functions using the HD and $H\overline{D}$ methods and convergence properties, J. Math. Chem. **29**, 213 – 232 (2001).

[71] H. Safouhi and A. Bouferguene, Extrapolation methods for improving convergence of spherical Bessel integrals for the two-center Coulomb integrals, Int. J. Quantum Chem. **106**, 2318 – 2323 (2006).

[72] H. Safouhi and A. Bouferguene, Nonlinear transformation methods for accelerating the convergence of Coulomb integrals over exponential type functions, Theor. Chem. Acc. **117**, 213 – 222 (2007).

[73] A. V. Sergeev and D. Z. Goodson, Summation of asymptotic expansions of multi-valued functions using algebraic approximants: Application to the anharmonic oscillators, J. Phys. A **31**, 4301 – 4317 (1998).

[74] R. E. Shafer, On quadratic approximation, SIAM J. Numer. Anal. **11**, 447 – 460 (1974).

[75] A. Sidi, *Practical Extrapolation Methods*, Cambridge U. P., Cambridge (2003).

[76] D. A. Smith and W. F. Ford, Acceleration of linear and logarithmic convergence, SIAM J. Numer. Anal. **16**, 223 – 240 (1979).

[77] D. A. Smith and W. F. Ford, Numerical comparisons of nonlinear convergence accelerators, Math. Comput. **38**, 481 – 499 (1982).

[78] Y. Tourigny and P. Drazin, The asymptotic behaviour of algebraic approximants, Proc. Roy. Soc. London A **456**, 1117 – 1137 (2000).

[79] G. Walz, *Asymptotics and Extrapolation*, Akademie Verlag, Berlin (1996).

[80] E. J. Weniger, *Reduzierte Bessel-Funktionen als LCAO-Basissatz: Analytische und numerische Untersuchungen*, Ph.D. thesis, Fachbereich Chemie und Pharmazie, Universität Regensburg (1982). A short abstract of this thesis was published in Zentralblatt für Mathematik **523**, 444 (1984), abstract no. 65015.

[81] E. J. Weniger, Nonlinear sequence transformations for the acceleration of convergence and the summation of divergent series, Comput. Phys. Rep. **10**, 189 – 371 (1989). Los Alamos Preprint math-ph/0306302 (`http://arXiv.org`).

[82] E. J. Weniger, On the summation of some divergent hypergeometric series and related perturbation expansions, J. Comput. Appl. Math. **32**, 291 – 300 (1990).

[83] E. J. Weniger, On the derivation of iterated sequence transformations for the acceleration of convergence and the summation of divergent series, Comput. Phys. Commun. **64**, 19 – 45 (1991).

[84] E. J. Weniger, Interpolation between sequence transformations, Numer. Algor. **3**, 477 – 486 (1992).

[85] E. J. Weniger, *Verallgemeinerte Summationsprozesse als numerische Hilfsmittel für quantenmechanische und quantenchemische Rechnungen*, Habilitation thesis, Fachbereich Chemie und Pharmazie, Universität Regensburg (1994). Los Alamos Preprint math-ph/0306048 (`http://arXiv.org`).

[86] E. J. Weniger, A convergent renormalized strong coupling perturbation expansion for the ground state energy of the quartic, sextic, and octic anharmonic oscillator, Ann. Phys. (NY) **246**, 133 – 165 (1996).

[87] E. J. Weniger, Computation of the Whittaker function of the second kind by summing its divergent asymptotic series with the help of nonlinear sequence transformations, Comput. Phys. **10**, 496 – 503 (1996).

[88] E. J. Weniger, Construction of the strong coupling expansion for the ground state energy of the quartic, sextic and octic anharmonic oscillator via a renormalized strong coupling expansion, Phys. Rev. Lett. **77**, 2859 – 2862 (1996).

[89] E. J. Weniger, Performance of superconvergent perturbation theory, Phys. Rev. A **56**, 5165 – 5168 (1997).

[90] E. J. Weniger, Prediction properties of Aitken's iterated Δ^2 process, of Wynn's epsilon algorithm, and of Brezinski's iterated theta algorithm, J. Comput. Appl. Math. **122**, 329 – 356 (2000). Reprinted as [91].

[91] E. J. Weniger, Prediction properties of Aitken's iterated Δ^2 process, of Wynn's epsilon algorithm, and of Brezinski's iterated theta algorithm, in C. Brezinski (Editor), *Numerical Analysis 2000, Vol. 2: Interpolation and Extrapolation*, 329 – 356, Elsevier, Amsterdam (2000).

[92] E. J. Weniger, Irregular input data in convergence acceleration and summation processes: General considerations and some special Gaussian hypergeometric series as model problems, Comput. Phys. Commun. **133**, 202 – 228 (2001).

[93] E. J. Weniger, A rational approximant for the digamma function, Numer. Algor. **33**, 499 – 507 (2003).

[94] E. J. Weniger, Mathematical properties of a new Levin-type sequence transformation introduced by Čížek, Zamastil, and Skála. II. Convergence theory (2004). In preparation (to be submitted to Journal of Mathematical Physics).

[95] E. J. Weniger, Mathematical properties of a new Levin-type sequence transformation introduced by Čížek, Zamastil, and Skála. I. Algebraic theory, J. Math. Phys. **45**, 1209 – 1246 (2004).

[96] E. J. Weniger, Asymptotic approximations to truncation errors of series representations for special functions, in A. Iske and J. Levesley (Editors), *Algorithms for Approximation*, 331 – 348, Springer-Verlag, Berlin (2007).

[97] E. J. Weniger and J. Čížek, Rational approximations for the modified Bessel function of the second kind, Comput. Phys. Commun. **59**, 471 – 493 (1990).

[98] E. J. Weniger, J. Čížek, and F. Vinette, Very accurate summation for the infinite coupling limit of the perturbation series expansions of anharmonic oscillators, Phys. Lett. A **156**, 169 – 174 (1991).

[99] E. J. Weniger, J. Čížek, and F. Vinette, The summation of the ordinary and renormalized perturbation series for the ground state energy of the quartic, sextic, and octic anharmonic oscillators using nonlinear sequence transformations, J. Math. Phys. **34**, 571 – 609 (1993).

[100] E. J. Weniger and B. Kirtman, Extrapolation methods for improving the convergence of oligomer calculations to the infinite chain limit of *quasi*-onedimensional stereoregular polymers, Comput. Math. Applic. **45**, 189 – 215 (2003).

[101] E. J. Weniger and C. Liegener, Extrapolation of finite cluster and crystal-orbital calculations on trans-polyacetylene, Int. J. Quantum Chem. **38**, 55 – 74 (1990).

[102] J. Wimp, *Sequence Transformations and Their Applications*, Academic Press, New York (1981).

[103] L. Wuytack, Commented bibliography on techniques for computing Padé approximants, in L. Wuytack (Editor), *Padé Approximations and Its Applications*, 375 – 392, Springer-Verlag, Berlin (1979).

[104] P. Wynn, On a device for computing the $e_m(S_n)$ transformation, Math. Tables Aids Comput. **10**, 91 – 96 (1956).

[105] P. Wynn, On a Procrustean technique for the numerical transformation of slowly convergent sequences and series, Proc. Camb. Phil. Soc. **52**, 663 – 671 (1956).

[106] A. I. Zayed, *Handbook of Function and Generalized Function Transformations*, CRC Press, Boca Raton (1996).