Applied Numerical Mathematics 60 (2010) 1395–1410

Contents lists available at ScienceDirect

Applied Numerical Mathematics

www.elsevier.com/locate/apnum

Survey of numerical stability issues in convergence acceleration

Avram Sidi [1]

Computer Science Department, Technion – Israel Institute of Technology, Haifa 32000, Israel

ARTICLE INFO

Article history:
Received 19 November 2009
Received in revised form 8 March 2010
Accepted 9 March 2010
Available online 18 March 2010

Keywords:
Acceleration of convergence
Extrapolation
Generalized Richardson extrapolation process $\text{GREP}^{(m)}$
Levin transformation
Levin–Sidi $d^{(m)}$ transformation
Logarithmic sequences
Linear sequences
General linear sequences
Numerical stability
Floating-point arithmetic

ABSTRACT

An important issue that arises in application of convergence acceleration (extrapolation) methods is that of stability in the presence of floating-point arithmetic. This issue turns out to be critical because numerical instability is inherent, even built in, when convergence acceleration methods are applied to certain types of sequences that occur commonly in practice. If methods are applied without taking this issue into account, the attainable accuracy is limited, and eventually destroyed completely, as more terms are added in the process. Therefore, it is important to understand the origin of the problem and to propose practical ways to solve it effectively. In this work, we present a general discussion of the issue of stability within the context of a generalization of the Richardson extrapolation process, and review some of the recent developments that have taken place in the theoretical study of many of the known acceleration methods. We discuss approaches that have been proposed to cope with built-in instabilities when applying various methods, and illustrate the effectiveness of these strategies with some numerical examples.

1. Introduction

An important issue that arises when applying convergence acceleration methods (equivalently, extrapolation methods or sequence transformations) to slowly converging or diverging sequences is that of numerical stability in the presence of floating-point arithmetic. As always, we measure the extent of stability, or its lack thereof, in the classical sense, namely, by assessing the rate of propagation of the errors in the input to the error in the output. In our case, the input is the sequence whose convergence we would like to accelerate, while the output is the approximations produced by the acceleration process. In general, errors in the input are caused by the finite precision of the floating-point arithmetic that is being used, although they can be caused in other ways too.

For some classes of sequences, the rate at which input errors propagate into the output are seen to be negligible. For some other classes, however, it may be very significant; in such cases, the accuracy of the approximations obtained by the acceleration process may eventually be lost completely, and this makes the output quite unreliable.

In this work, we would like to survey the experience and knowledge that has been accumulated throughout the recent years concerning this phenomenon. Specifically, we first show how the issue of stability can be understood at least for some extrapolation methods, and give some examples to illustrate the validity of this explanation. We next discuss the theoretical assessment of stability for several methods. Based on these assessments, we propose effective strategies for improving the stability properties of extrapolation methods, and illustrate these with numerical examples.

E-mail address: asidi@cs.technion.ac.il.

[1] Research by this author is partially supported by the United States–Israel Binational Science Foundation under grant number 2008399.

0168-9274/$30.00 © 2010 IMACS. Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.apnum.2010.03.009

2. Understanding stability in extrapolation

2.1. Algebraic structure of extrapolation methods

We start by observing that every sequence $\{A_m\}$ can be related to a function $A(y)$ defined (that is, computable) on an interval $(0, b]$, where the variable y may be continuous or discrete, and $A_m = A(y'_m)$ for some monotonically decreasing sequence $\{y'_m\}$, such that $\lim_{m\to\infty} y'_m = 0$. Of course, in case $\lim_{m\to\infty} A_m = A$, we have $\lim_{y\to 0} A(y) = A$ as well.

Example 2.1. As an example, consider the sequence of trapezoidal rule approximations

$$Q_m = \frac{1}{m}\left[\frac{1}{2}f(0) + \sum_{i=1}^{m-1} f(i/m) + \frac{1}{2}f(1)\right], \quad m = 1, 2, \dots,$$

to the integral $I = \int_0^1 f(x)\,dx$. Clearly, the sequence $\{Q_m\}$ is related to a function $A(y)$, for which, $y'_m = 1/m$ and $A(y'_m) = Q_m$, and y assumes only the discrete values $1, 1/2, 1/3, \dots$. In addition, $A(y)$ has some additional interesting properties when $f(x)$ has a number of continuous derivatives on $[0, 1]$. The following result concerning $A(y)$ is a special case of a more general result that is stated and proved in Sidi [33, Theorem D.4.1, pp. 471–472] and has proved to be very useful in the analysis of Romberg integration, which is a Richardson extrapolation procedure: When $f \in C^\infty[0, 1]$, $A(y)$ can be continued to *all real values* of $y \in [0, 1]$ such that $a(t) = A(t^{1/2})$ is infinitely differentiable on $[0, 1]$ as a function of the continuous variable t, in addition to satisfying the interpolation conditions $a(m^{-2}) = A(m^{-1}) = Q_m$, $m = 1, 2, \dots$.[2]

To put things in proper context, we will treat stability within the formalism of the generalized Richardson extrapolation process. Most of the developments below can be found in [33, Introduction and Chapter 3]. For earlier treatments covering some of the aspects of that given here, see Albrecht [2], Håvie [12], Brezinski [4], Sidi [21,24], and also Brezinski and Redivo Zaglia [5, Chapter 2], for example.

For most known extrapolation methods, some of which we discuss later in this work, the approximations $A_n^{(j)}$ to $\lim_{y\to 0} A(y) = A$ are defined via systems of linear equations, or can be shown to satisfy systems of linear equations, of the form

$$A(y_l) = A_n^{(j)} + \sum_{k=1}^{n} \alpha_k \phi_k(y_l), \quad j \leqslant l \leqslant j+n;\ j = 0, 1, \dots,\ n = 1, 2, \dots, \tag{1}$$

with the y_m chosen such that

$$y_0 > y_1 > \cdots > 0, \quad \lim_{m\to\infty} y_m = 0. \tag{2}$$

Here $\phi_k(y)$ are some functions known for $y > 0$, just as $A(y)$ is, and they are determined uniquely by the relevant extrapolation method. Of course, this linear system can be expressed in matrix form as $\boldsymbol{Mx} = \boldsymbol{b}$, where

$$\boldsymbol{M} = \begin{bmatrix} 1 & \phi_1(y_j) & \dots & \phi_n(y_j) \\ 1 & \phi_1(y_{j+1}) & \dots & \phi_n(y_{j+1}) \\ \vdots & \vdots & & \vdots \\ 1 & \phi_1(y_{j+n}) & \dots & \phi_n(y_{j+n}) \end{bmatrix}, \quad \boldsymbol{x} = \begin{bmatrix} A_n^{(j)} \\ \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} A(y_j) \\ A(y_{j+1}) \\ \vdots \\ A(y_{j+n}) \end{bmatrix}. \tag{3}$$

Solving this system for $A_n^{(j)}$ by Cramer's rule, we obtain the determinant representation

$$A_n^{(j)} = \frac{D\{A(y_j), A(y_{j+1}), \dots, A(y_{j+n})\}}{D\{1, 1, \dots, 1\}}, \tag{4}$$

where

$$D\{v_0, v_1, \dots, v_n\} = \det\begin{bmatrix} v_0 & \phi_1(y_j) & \dots & \phi_n(y_j) \\ v_1 & \phi_1(y_{j+1}) & \dots & \phi_n(y_{j+1}) \\ \vdots & \vdots & & \vdots \\ v_n & \phi_1(y_{j+n}) & \dots & \phi_n(y_{j+n}) \end{bmatrix}.$$

Note that this representation of $A_n^{(j)}$ is very useful for the theoretical analysis of the extrapolation methods. It should not be used for actual computation of $A_n^{(j)}$, however; other less expensive and numerically better ways should be employed

[2] Using this result, it is proved in [33, p. 53, Theorem 2.4.1] that all columns of the Romberg table in this case converge to $I = \int_0^1 f(x)\,dx$, and accelerate the convergence of the sequence $\{Q_{m_k}\}_{k=0}^\infty$ for *arbitrarily chosen* stepsizes $h_k = 1/m_k$, and this is quite surprising.

Table 1
Arrangement of extrapolation table.

$A_0^{(0)}$				
$A_0^{(1)}$	$A_1^{(0)}$			
$A_0^{(2)}$	$A_1^{(1)}$	$A_2^{(0)}$		
$A_0^{(3)}$	$A_1^{(2)}$	$A_2^{(1)}$	$A_3^{(0)}$	
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

for this purpose. Indeed, very efficient recursive algorithms exist for computing the extrapolation tables resulting from most methods. (In case the $\phi_k(y_m)$ do not have any particular structure, one can use the E algorithm of Schneider [16], Brezinski [4] and Håvie [12] or the FS algorithm of Ford and Sidi [10].)

The approximations $A_n^{(j)}$ can be arranged in a two-dimensional array as in Table 1.

In agreement with the arrangement of the $A_n^{(j)}$ in this array, we will call the sequences $\{A_n^{(j)}\}_{j=0}^{\infty}$, with fixed n, "column sequences." Similarly, we will call the sequences $\{A_n^{(j)}\}_{n=0}^{\infty}$, with fixed j, "diagonal sequences." Practical experience and some theoretical results suggest strongly that the diagonal sequences have superior convergence properties.

Denoting the cofactor of $A(y_{j+i})$ in the numerator determinant of (4) by C_i, and expanding both the numerator and denominator determinants with respect to their first columns, we obtain

$$A_n^{(j)} = \frac{\sum_{i=0}^{n} C_i A(y_{j+i})}{\sum_{i=0}^{n} C_i},$$

which can also be written as in

$$A_n^{(j)} = \sum_{i=0}^{n} \gamma_{ni}^{(j)} A(y_{j+i}) \quad \text{with} \quad \sum_{i=0}^{n} \gamma_{ni}^{(j)} = 1; \qquad \gamma_{ni}^{(j)} = \frac{C_i}{\sum_{s=0}^{n} C_s}, \quad i = 0, 1, \ldots, n. \tag{5}$$

Actually, we can say more about the $\gamma_{ni}^{(j)}$: Let us denote the first row of $\boldsymbol{M}^{-1}$ by $[m_0, m_1, \ldots, m_n]$. Then

$$\gamma_{ni}^{(j)} = m_i, \quad i = 0, 1, \ldots, n. \tag{6}$$

Thus, the $\gamma_{ni}^{(j)}$ satisfy the linear system

$$\boldsymbol{M}^{\mathrm{T}} \boldsymbol{\gamma} = \boldsymbol{e}_1; \quad \boldsymbol{\gamma} = \left[\gamma_{n0}^{(j)}, \gamma_{n1}^{(j)}, \ldots, \gamma_{nn}^{(j)}\right]^{\mathrm{T}}, \; \boldsymbol{e}_1 = [1, 0, \ldots, 0]^{\mathrm{T}}. \tag{7}$$

Finally, as shown in [24], the $\gamma_{ni}^{(j)}$ also satisfy

$$\sum_{i=0}^{n} \gamma_{ni}^{(j)} z^i = \frac{H_n^{(j)}(z)}{H_n^{(j)}(1)}, \tag{8}$$

where

$$H_n^{(j)}(z) = \det \begin{bmatrix} z^0 & \phi_1(y_j) & \ldots & \phi_n(y_j) \\ z^1 & \phi_1(y_{j+1}) & \ldots & \phi_n(y_{j+1}) \\ \vdots & \vdots & & \vdots \\ z^n & \phi_1(y_{j+n}) & \ldots & \phi_n(y_{j+n}) \end{bmatrix}. \tag{9}$$

Theoretical studies of extrapolation methods are ultimately carried out by analyzing $H_n^{(j)}(z)$ asymptotically. This analysis seems to be possible for column sequences $\{A_n^{(j)}\}_{j=0}^{\infty}$ ($j \to \infty$ and n fixed) in quite a few cases. It is extremely difficult for diagonal sequences $\{A_n^{(j)}\}_{n=0}^{\infty}$ ($n \to \infty$ and j fixed), and there are very few such cases. Nevertheless, conclusions drawn from $j \to \infty$ asymptotics and that are relevant for column sequences seem to be generally just as relevant and valid for diagonal sequences.

2.2. Assessing stability in extrapolation

As mentioned earlier, we assess the stability of the extrapolation method by analyzing the rate at which errors in the $A(y_m)$ propagate into $A_n^{(j)}$. We turn to this analysis next. Writing $A_m = A(y_m)$ for simplicity of notation, we let $\bar{A}_m = \bar{A}(y_m)$, the computed A_m, be $\bar{A}_m = A_m + \epsilon_m$, where ϵ_m is the error incurred in the computation of A_m. Let also $\bar{\gamma}_{ni}^{(j)} = \gamma_{ni}^{(j)} + \delta_{ni}^{(j)}$ be the computed $\gamma_{ni}^{(j)}$. Then the computed $A_n^{(j)}$ is

$$\bar{A}_n^{(j)} = \sum_{i=0}^{n} \bar{\gamma}_{ni}^{(j)} \bar{A}_{j+i}. \tag{10}$$

Consequently,

$$\left|\bar{A}_n^{(j)} - A\right| \leqslant \left|A_n^{(j)} - A\right| + \sum_{i=0}^{n} \left|\gamma_{ni}^{(j)}\right| |\epsilon_{j+i}| + \sum_{i=0}^{n} \left|\delta_{ni}^{(j)}\right| |\bar{A}_{j+i}|.$$

Now, in case the $\bar{A}_m$ and the $\bar{\gamma}_{ni}^{(j)}$ have been computed with machine precision, we have $|\epsilon_m| = |A_m||\rho_m|$ and $|\delta_{ni}^{(j)}| = |\gamma_{ni}^{(j)}||\eta_{ni}^{(j)}|$, where $|\rho_m| \leqslant \mathbf{u}$ and $|\eta_{ni}^{(j)}| \leqslant \mathbf{u}$, $\mathbf{u}$ being the roundoff unit of the arithmetic being used. Thus,

$$\left|\bar{A}_n^{(j)} - A\right| \leqslant \left|A_n^{(j)} - A\right| + \mathbf{u}\left[\sum_{i=0}^{n} \left|\gamma_{ni}^{(j)}\right| |A_{j+i}| + \sum_{i=0}^{n} \left|\gamma_{ni}^{(j)}\right| |\bar{A}_{j+i}|\right].$$

By the fact that $|\bar{A}_m| \leqslant |A_m| + |\epsilon_m| \leqslant |A_m|(1+\mathbf{u})$, we have $\mathbf{u}|\bar{A}_m| \leqslant \mathbf{u}|A_m| + O(\mathbf{u}^2)$. Therefore,

$$\left|\bar{A}_n^{(j)} - A\right| \lessapprox \left|A_n^{(j)} - A\right| + 2\mathbf{u} \sum_{i=0}^{n} \left|\gamma_{ni}^{(j)}\right| |A_{j+i}|.$$

In case $\{A_m\}$ converges, $A_m \approx A$ for all large m. Therefore,

$$\left|\bar{A}_n^{(j)} - A\right| \lessapprox \left|A_n^{(j)} - A\right| + 2\mathbf{u}|A| \sum_{i=0}^{n} \left|\gamma_{ni}^{(j)}\right|,$$

and if $A \neq 0$,

$$\frac{|\bar{A}_n^{(j)} - A|}{|A|} \lessapprox \frac{|A_n^{(j)} - A|}{|A|} + 2\mathbf{u} \sum_{i=0}^{n} \left|\gamma_{ni}^{(j)}\right|.$$

Now, assuming that $\lim_{j\to\infty} A_n^{(j)} = A$ or that $\lim_{n\to\infty} A_n^{(j)} = A$, we have that the theoretical relative error in $A_n^{(j)}$ satisfies $|A_n^{(j)} - A|/|A| \to 0$ as $j \to \infty$ or as $n \to \infty$. Consequently, for all large j or n,

$$\frac{|\bar{A}_n^{(j)} - A|}{|A|} \lessapprox 2\mathbf{u} \sum_{i=0}^{n} \left|\gamma_{ni}^{(j)}\right| = 2\mathbf{u}\Gamma_n^{(j)}, \tag{11}$$

where

$$\Gamma_n^{(j)} = \sum_{i=0}^{n} \left|\gamma_{ni}^{(j)}\right|. \tag{12}$$

We thus see that the quantity $\Gamma_n^{(j)}$ plays a very important role when applying extrapolation methods in floating-point arithmetic. First, by the fact that $\sum_{i=0}^{n} \gamma_{ni}^{(j)} = 1$, we have $\Gamma_n^{(j)} \geqslant 1$. Next, in case $\Gamma_n^{(j)}$ is of order 10^p and $\mathbf{u}$ is of order 10^{-s}, at most $s - p$ decimal digits of $\bar{A}_n^{(j)}$ can be trusted.

In view of our conclusions above, we want $\sup_j \Gamma_n^{(j)}$ or $\sup_n \Gamma_n^{(j)}$ to be finite and small. Of course, the ideal situation is one in which $\Gamma_n^{(j)} = 1$ or $\lim_{j\to\infty} \Gamma_n^{(j)} = 1$ or $\lim_{n\to\infty} \Gamma_n^{(j)} = 1$. In case $\Gamma_n^{(j)}$ is bounded in j or n but is much larger than 1, we want to be able to reduce it as much as possible. In case $\Gamma_n^{(j)}$ is unbounded in j or n, we want to be able to force it to be bounded or to grow as slowly as possible.

Remarks.

1. The roundoff unit $\mathbf{u}$ is of order 10^{-16} for double precision, and of order 10^{-35} for quadruple precision. Therefore, better accuracy can be obtained from $\bar{A}_n^{(j)}$ by increasing the precision of the floating-point arithmetic, provided the A_m are also computed to machine precision in this arithmetic.
2. In case the $\gamma_{ni}^{(j)}$ have been computed exactly, we have $\delta_{ni}^{(j)} = 0$ for all i, and hence (11) assumes the slightly different form

$$\frac{|\bar{A}_n^{(j)} - A|}{|A|} \lessapprox \mathbf{u} \sum_{i=0}^{n} \left|\gamma_{ni}^{(j)}\right| = \mathbf{u}\Gamma_n^{(j)}.$$

Nevertheless, it is easy to see that the preceding conclusions remain the same.

A slightly different approach to stability has been proposed by Fessler, Ford, and Smith [9], and more recently by Senhadji [17,18]. Fessler et al. consider specifically the stability of the u transformation of Levin [13] as applied to the sequence $\{A_m\}$. They estimate the uncertainty in $\bar{A}_n^{(j)}$ via

$$\left|\delta A_n^{(j)}\right| = \left[\sum_{i=0}^{K}\left(\frac{\partial A_n^{(j)}}{\partial a_i}\delta a_i\right)^2\right]^{1/2},$$

where $a_i = A_i - A_{i-1}$, δa_i is the uncertainty (error) in the computed a_i, and K is the number of the a_i used to construct $A_n^{(j)}$. In [17], Senhadji considers quasi-linear sequence transformations [the general extrapolation procedure described by (1) and all other transformations discussed in Section 5 are quasi-linear] and introduces the quantity

$$\overline{\Gamma}_n^{(j)} = \sum_{i=0}^{n}\left|\frac{\partial A_n^{(j)}}{\partial A_{j+i}}\right|$$

for assessing the stability of $A_n^{(j)}$ instead of $\Gamma_n^{(j)}$ defined in (12). In [18], he applies treatment of [17] to the Shanks transformation. Senhadji calls $\overline{\Gamma}_n^{(j)}$ the condition number of $A_n^{(j)}$.

3. Illustration of numerical instability and its treatment

We now illustrate the developments above with two concrete extrapolation methods. The methods we have chosen for this purpose are the u transformation of Levin [13] (see also [33, Chapter 19]) and the $d^{(1)}$ transformation of Levin and Sidi [14] (see also [33, Chapter 6]).

3.1. The u transformation

The approximations $A_n^{(j)}$ generated by applying the Levin u transformation to a sequence $\{A_m\}_{m=1}^{\infty}$ are defined via the linear systems of equations

$$A_l = A_n^{(j)} + (l+1)a_{l+1}\sum_{i=0}^{n-1}\frac{\beta_i}{(l+1)^i}, \quad j \leqslant l \leqslant j+n;\ j=0,1,\ldots,\ n=1,2,\ldots, \tag{13}$$

with the a_m defined as in

$$a_1 = A_1; \qquad a_m = A_m - A_{m-1}, \quad m=2,3,\ldots. \tag{14}$$

Note, conversely, that

$$A_0 = 0; \qquad A_m = \sum_{i=1}^{m} a_i, \quad m=1,2,\ldots. \tag{15}$$

Practical experience, and some theoretical results by the author [20,22,25], suggest strongly that the diagonal sequences obtained by applying the u transformation to certain logarithmically convergent sequences and linearly convergent sequences have superior convergence properties. In their numerical survey of convergence acceleration methods, Smith and Ford [35,36] and Weniger [38] conclude that the u transformation and the θ algorithm of Brezinski [3] are two of the best sequence transformations for accelerating logarithmic convergence, and these two transformations along with that of Shanks [19] are the best transformations for accelerating linear convergence. Both types of convergence will be described in the next paragraph.

Observe that the equations in (13) that define the u transformation are exactly of the form shown in (1). Exploiting the special nature of the matrix $\boldsymbol{M}$ of these equations, Levin has derived the following simple closed-form expression for the approximation $A_n^{(j)}$:

$$A_n^{(j)} = \frac{\Delta^n[(j+1)^{n-2}A_j/a_{j+1}]}{\Delta^n[(j+1)^{n-2}/a_{j+1}]} = \frac{\sum_{i=0}^{n}(-1)^i\binom{n}{i}(j+i+1)^{n-2}A_{j+i}/a_{j+i+1}}{\sum_{i=0}^{n}(-1)^i\binom{n}{i}(j+i+1)^{n-2}/a_{j+i+1}}. \tag{16}$$

From this expression, we can also extract the $\gamma_{ni}^{(j)}$:

$$\gamma_{ni}^{(j)} = \frac{(-1)^i\binom{n}{i}(j+i+1)^{n-2}/a_{j+i+1}}{\sum_{s=0}^{n}(-1)^s\binom{n}{s}(j+s+1)^{n-2}/a_{j+s+1}}, \quad i=0,1,\ldots,n. \tag{17}$$

Consequently, we also have the following simple expression for $\Gamma_n^{(j)}$:

$$\Gamma_n^{(j)} = \frac{\sum_{i=0}^{n} \binom{n}{i}(j+i+1)^{n-2}/|a_{j+i+1}|}{|\sum_{i=0}^{n}(-1)^i\binom{n}{i}(j+i+1)^{n-2}/a_{j+i+1}|}. \tag{18}$$

It is thus clear that the ideal situation $\Gamma_n^{(j)} = 1$ for all j and n will occur when $a_m a_{m+1} < 0$, $m = 1, 2, \ldots,$ that is, when the series $\sum_{i=1}^{\infty} a_i$ is alternating. The following results follow from [20,22] and [33, Chapters 8, 9]:

1. When $a_m \sim \sum_{i=0}^{\infty} \alpha_i m^{\gamma-i}$ as $m \to \infty$, $\gamma \neq -1, 0, 1, 2, \ldots,$ we have that $\lim_{m\to\infty} a_{m+1}/a_m = 1$ (in this case, we say that $\{A_m\} \in \mathbf{b}^{(1)}/\mathrm{LOG}$, and that it *converges logarithmically*), and the partial sums and the column sequences behave as in
$$A_m - A \sim \sum_{i=0}^{\infty} \tilde{\beta}_i m^{\gamma+1-i} \quad \text{as } m \to \infty;$$
$$A_n^{(j)} - A \sim \sum_{i=0}^{\infty} \delta_i j^{\gamma+1-n-i} \quad \text{and} \quad \Gamma_n^{(j)} \sim \frac{(2j)^n}{|(-\gamma-1)_n|} \quad \text{as } j \to \infty. \tag{19}$$
2. When $a_m \sim \zeta^m \sum_{i=0}^{\infty} \alpha_i m^{\gamma-i}$ as $m \to \infty$, $|\zeta| \leqslant 1$, $\zeta \neq 1$, we have that $\lim_{m\to\infty} a_{m+1}/a_m = \zeta \neq 1$ (in this case, we say that $\{A_m\} \in \mathbf{b}^{(1)}/\mathrm{LIN}$, and that it *converges linearly*), and the partial sums and the column sequences behave as in
$$A_m - A \sim \zeta^m \sum_{i=0}^{\infty} \tilde{\beta}_i m^{\gamma-i} \quad \text{as } m \to \infty;$$
$$A_n^{(j)} - A \sim \zeta^j \sum_{i=0}^{\infty} \delta_i j^{\gamma-2n-i} \quad \text{as } j \to \infty \quad \text{and} \quad \lim_{j\to\infty} \Gamma_n^{(j)} = \left(\frac{1+|\zeta|}{|1-\zeta|}\right)^n. \tag{20}$$

Thus, we see that column sequences converge faster than the sequence of partial sums and their convergence improves with increasing n. The treatment of the diagonal sequences turns out to be extremely difficult. Some convergence results of interest with specific examples can be found in [22]. See also [33, Chapter 19]. We now consider the numerical stability of the u transformation with two examples, that have also been treated for their convergence in [22].

Example 3.1. Let us apply the u transformation to the series $\sum_{i=1}^{\infty} 1/i^2 = \pi^2/6$. The partial sums of this series form a logarithmically convergent sequence. In addition, by invoking $a_m = m^{-2}$ and the fact that

$$\sum_{i=0}^{n}(-1)^{n-i}\binom{n}{i}(j+i+1)^n = \Delta^n\big[(j+1)^n\big] = n!,$$

we can first show that

$$\Gamma_n^{(j)} = \frac{\sum_{i=0}^{n}\binom{n}{i}(j+i+1)^n}{|\sum_{i=0}^{n}(-1)^i\binom{n}{i}(j+i+1)^n|} = \frac{1}{n!}\sum_{i=0}^{n}\binom{n}{i}(j+i+1)^n.$$

Next, using the binomial theorem and Stirling's formula, we can make the following definitive statements about $\Gamma_n^{(j)}$:

$$\Gamma_n^{(j)} \sim \frac{(2j)^n}{n!} \quad \text{as } j \to \infty; \qquad \Gamma_n^{(j)} > \frac{(j+n+1)^n}{n!} \sim e^{j+1}\frac{e^n}{\sqrt{2\pi n}} \quad \text{as } n \to \infty.$$

Thus, we see that both the column and the diagonal sequences are unstable since $\sup_j \Gamma_n^{(j)} = \infty$ and $\sup_n \Gamma_n^{(j)} = \infty$.

Table 2 contains some of the numerical results obtained for the diagonal sequence $\{A_n^{(0)}\}_{n=0}^{\infty}$ in double-precision and quadruple-precision arithmetic. These show clearly that, in either precision, there is no way of achieving machine accuracy, and there is no way of retaining the best accuracy that is achieved. Also, the accuracy achieved for each $A_n^{(0)}$ (in both double and quadruple precision) is consistent with our conclusion following (12).

Example 3.2. Let us now apply the u transformation to the series $\sum_{i=1}^{\infty} \zeta^i/i = \log[(1-\zeta)^{-1}]$, $|\zeta| \leqslant 1$, $\zeta \neq 1$. The partial sums of this series form a linearly convergent sequence. In addition, by invoking $a_m = \zeta^m/m$, we have first

$$\Gamma_n^{(j)} = \frac{\sum_{i=0}^{n}\binom{n}{i}(j+i+1)^{n-1}|\zeta|^{-j-i-1}}{|\sum_{i=0}^{n}(-1)^i\binom{n}{i}(j+i+1)^{n-1}\zeta^{-j-i-1}|}.$$

Next, we can make the following definitive statements about $\Gamma_n^{(j)}$:

$$\lim_{j\to\infty} \Gamma_n^{(j)} = \left(\frac{1+|\zeta|}{|1-\zeta|}\right)^n \quad \text{for all } \zeta \neq 1; \qquad \Gamma_n^{(j)} = 1 \quad \text{for all } j, n \text{ when } \zeta < 0.$$

Table 2
Numerical results for Example 3.1 from the u transformation. $\bar{E}_n^{(0)}(\mathrm{d})$: relative error in $\bar{A}_n^{(0)}$ in double precision ($\approx$ 16 decimal digits). $\bar{E}_n^{(0)}(\mathrm{q})$: relative error in $\bar{A}_n^{(0)}$ in quadruple precision ($\approx$ 35 decimal digits).

n	$\bar{E}_n^{(0)}(\mathrm{d})$	$\bar{E}_n^{(0)}(\mathrm{q})$	$\Gamma_n^{(0)}$
0	3.92D−01	3.92D−01	1.00D+00
2	1.21D−02	1.21D−02	9.00D+00
4	1.90D−05	1.90D−05	9.17D+01
6	6.80D−07	6.80D−07	1.01D+03
8	1.56D−08	1.56D−08	1.15D+04
10	1.85D−10	1.83D−10	1.35D+05
12	1.09D−11	6.38D−13	1.60D+06
14	2.11D−10	2.38D−14	1.92D+07
16	7.99D−09	6.18D−16	2.33D+08
18	6.10D−08	7.78D−18	2.85D+09
20	1.06D−07	3.05D−20	3.50D+10
22	1.24D−05	1.03D−21	4.31D+11
24	3.10D−04	1.62D−22	5.33D+12
26	3.54D−03	4.33D−21	6.62D+13
28	1.80D−02	5.44D−20	8.24D+14
30	1.15D−01	4.74D−19	1.03D+16

Table 3
Numerical results for Example 3.2 from the u transformation. $\bar{E}_n^{(0)}(\zeta=t)$: relative error in $\bar{A}_n^{(0)}$ (with $\zeta=t$) in quadruple precision ($\approx$ 35 decimal digits).

n	$\bar{E}_n^{(0)}$ $(\zeta=-1)$	$\Gamma_n^{(0)}$	n	$\bar{E}_n^{(0)}$ $(\zeta=0.5)$	$\Gamma_n^{(0)}$	n	$\bar{E}_n^{(0)}$ $(\zeta=0.95)$	$\Gamma_n^{(0)}$
0	4.43D−01	1.00D+00	0	2.79D−01	1.00D+00	0	6.83D−01	1.00D+00
2	8.15D−03	1.00D+00	4	8.05D−05	8.87D+00	8	9.65D−03	1.64D+04
4	6.95D−06	1.00D+00	8	4.09D−09	4.14D+01	16	4.88D−05	3.87D+07
6	7.06D−08	1.00D+00	12	2.22D−13	1.95D+02	24	2.73D−07	9.66D+10
8	1.33D−10	1.00D+00	16	1.23D−17	9.24D+02	32	1.56D−09	2.44D+14
10	3.31D−13	1.00D+00	20	6.84D−22	4.37D+03	40	9.02D−12	6.17D+17
12	2.31D−15	1.00D+00	24	3.82D−26	2.07D+04	48	4.47D−14	1.57D+21
14	3.20D−18	1.00D+00	28	8.10D−30	9.82D+04	56	1.45D−11	3.98D+24
16	1.48D−20	1.00D+00	32	9.50D−30	4.65D+05	64	2.80D−07	1.01D+28
18	7.54D−23	1.00D+00	36	8.07D−29	2.21D+06	72	3.36D−04	2.57D+31
20	6.17D−26	1.00D+00	40	2.94D−28	1.05D+07	80	3.19D+00	4.37D+35
22	6.09D−28	1.00D+00	44	1.62D−27	4.96D+07	88	1.22D+00	6.74D+34
24	2.38D−30	1.00D+00	48	4.60D−27	2.35D+08	96	4.76D−01	1.29D+34
26	2.89D−34	1.00D+00	52	4.49D−26	1.12D+09	104	1.56D−01	3.46D+34
28	0.00D+00	1.00D+00	56	1.56D−25	5.29D+09	112	9.33D−02	1.07D+34
30	0.00D+00	1.00D+00	60	4.20D−25	2.51D+10	120	4.29D−02	9.58D+33

Thus, we see that the column sequences are always stable. When ζ is real negative, both column and diagonal sequences are absolutely stable since we now have the ideal situation $\Gamma_n^{(j)}=1$ for all j, n.

Table 3 contains some of the numerical results obtained for the diagonal sequence $\{A_n^{(0)}\}_{n=0}^{\infty}$ in quadruple-precision arithmetic. The following conclusions can be drawn from this table: (i) In case $\zeta<0$, absolute stability prevails and machine accuracy is achieved quickly and retained. (ii) When ζ is not negative but is sufficiently far from 1, the point of singularity of $\log[(1-\zeta)^{-1}]$ (that is, of the sum of $\sum_{i=1}^{\infty}\zeta^i/i$), numerical instability seems to set in quite late, by which time close to machine accuracy has been achieved. (iii) When ζ is very close to 1, numerical instability sets in very early, and this causes the transformation to achieve much less than machine accuracy; in addition, there is no way of retaining the best accuracy that is achieved. The situation becomes worse as ζ approaches 1. These conclusions seem to be valid for complex ζ, as well as real ζ. Clearly, the accuracy achieved for each $A_n^{(0)}$ is consistent with our conclusion following (12).

3.2. The $d^{(1)}$ transformation

The stability problems that we have observed in Examples 3.1 and 3.2 can be overcome by applying the Levin–Sidi $d^{(1)}$ transformation in suitable ways to be described soon. The approximations $A_n^{(j)}$ generated by applying the $d^{(1)}$ transformation to a sequence $\{A_m\}_{m=1}^{\infty}$ are defined via the linear systems of equations

$$A_{R_l}=A_n^{(j)}+R_l a_{R_l}\sum_{i=0}^{n-1}\frac{\beta_i}{R_l^i},\quad j\leqslant l\leqslant j+n, \tag{21}$$

where the integers R_l, $1\leqslant R_0<R_1<R_2<\cdots$, are chosen by the user. As we will see later, the fact that the R_l are at the user's disposal makes this transformation very flexible for achieving improved stability and increased accuracy simultaneously. Comparing (21) with (1), we realize that $A_n^{(j)} = \sum_{i=0}^{n} \gamma_{ni}^{(j)} A_{R_{j+i}}$ with appropriate $\gamma_{ni}^{(j)}$ that satisfy $\sum_{i=0}^{n} \gamma_{ni}^{(j)} = 1$. Again, best results are obtained from diagonal sequences $\{A_n^{(j)}\}_{n=0}^{\infty}$.

With general R_l, the $A_n^{(j)}$ and $\Gamma_n^{(j)}$ can be computed in an efficient way via the recursive W algorithm of Sidi [23] (see also [33, Chapter 7]), which reads as follows:

$$M_0^{(j)} = \frac{A_{R_j}}{R_j a_{R_j}}, \qquad N_0^{(j)} = \frac{1}{R_j a_{R_j}}, \qquad H_0^{(j)} = (-1)^j \left|N_0^{(j)}\right|, \quad j = 0, 1, \ldots, \tag{22}$$

$$M_n^{(j)} = \frac{M_{n-1}^{(j+1)} - M_{n-1}^{(j)}}{R_{j+n}^{-1} - R_j^{-1}}, \qquad N_n^{(j)} = \frac{N_{n-1}^{(j+1)} - N_{n-1}^{(j)}}{R_{j+n}^{-1} - R_j^{-1}}, \qquad H_n^{(j)} = \frac{H_{n-1}^{(j+1)} - H_{n-1}^{(j)}}{R_{j+n}^{-1} - R_j^{-1}}, \quad j = 0, 1, \ldots, \ n = 1, 2, \ldots, \tag{23}$$

$$A_n^{(j)} = \frac{M_n^{(j)}}{N_n^{(j)}}, \qquad \Gamma_n^{(j)} = \left|\frac{H_n^{(j)}}{N_n^{(j)}}\right|, \quad j = 0, 1, \ldots, \ n = 0, 1, \ldots. \tag{24}$$

Remarks.

1. By comparing (21) with (13), we realize that the $d^{(1)}$ transformation reduces to the u transformation when $R_l = l + 1$, $l = 0, 1, \ldots$.
2. The $d^{(1)}$ transformation is the simplest special case of the general $d^{(m)}$ transformation by Levin and Sidi [14]. We will discuss the $d^{(m)}$ transformation briefly in the sequel.

To improve stability it was proposed in various papers by the author and collaborators (see Levin and Sidi [14], Ford and Sidi [10, Appendix B], and Sidi [26]; see also [33, Chapters 10, 12, 13]) to apply the $d^{(m)}$ transformation with specially chosen R_l. Thus,

$$R_0 = 1; \qquad R_l = \max\left\{\lfloor \sigma R_{l-1} \rfloor, R_{l-1} + 1\right\}, \quad l = 1, 2, \ldots; \quad \text{with fixed } \sigma > 1 \text{ (for logarithmic convergence)}, \tag{25}$$

$$R_l = \left\lfloor \kappa(l+1) \right\rfloor, \quad l = 0, 1, \ldots; \quad \text{with } \kappa > 1 \text{ (for linear convergence)}, \tag{26}$$

turn out to be very effective choices. The choice in (25) is called *geometric progression sampling* (*GPS*), while that in (26) is called *arithmetic progression sampling* (*APS*). The use of these R_l has been justified theoretically by the author in various places. Note that, (i) for GPS, $\lim_{l\to\infty}(R_{l+1}/R_l) = \sigma$, which means that R_l increases like σ^l with l, whereas (ii) for APS, $R_l \sim \kappa l$ as $l \to \infty$. (Originally, κ was chosen to be an integer; noninteger values were suggested later in [33].) In case of GPS, we choose $\sigma < 2$ so as to avoid a quick growth in the integers R_l with l.

The next results follow from Sidi ([26,28,31] and [33, Chapters 8, 9, 12]):

1. When $a_m \sim \sum_{i=0}^{\infty} \alpha_i m^{\gamma-i}$ as $m \to \infty$, $\gamma \neq -1, 0, 1, 2, \ldots$, hence A_m is as in (19) (therefore belongs to $\mathbf{b}^{(1)}$/LOG), if the R_l are chosen as in (25), then

$$A_n^{(j)} - A = O\left(|c_{n+1}|^j\right) \quad \text{as } j \to \infty \quad \text{and} \quad \lim_{j\to\infty} \Gamma_n^{(j)} = \prod_{k=1}^{n} \frac{1 + |c_k|}{|1 - c_k|}; \qquad c_k = \sigma^{\gamma+2-k}, \quad k = 1, 2, \ldots. \tag{27}$$

   When $\sigma = 2, 3, \ldots,$ in (25), we have $R_l = \sigma^l$, $l = 0, 1, \ldots$. In this case, with the c_k as defined in (27), we have the following strong results concerning diagonal sequences:

$$A_n^{(j)} - A = O\left(e^{-\lambda n}\right) \quad \text{as } n \to \infty, \ \forall \lambda > 0, \quad \text{and} \quad \lim_{n\to\infty} \Gamma_n^{(j)} = \prod_{k=1}^{\infty} \frac{1 + |c_k|}{|1 - c_k|} < \infty;$$
$$c_k = \sigma^{\gamma+2-k}, \quad k = 1, 2, \ldots. \tag{28}$$

   In words, diagonal sequences can be computed with very high accuracy and absolute stability. In addition, the diagonal sequences accelerate convergence, in the sense that they converge superlinearly, as opposed to $\{A_{R_n}\}_{n=0}^{\infty}$, which converges linearly.
2. When $a_m \sim \zeta^m \sum_{i=0}^{\infty} \alpha_i m^{\gamma-i}$ as $m \to \infty$, $|\zeta| \leqslant 1$, $\zeta \neq 1$, hence A_m is as in (20) (therefore belongs to $\mathbf{b}^{(1)}$/LIN), if the R_l are chosen as in (26) with integer κ, then (20) becomes

$$A_{\kappa m} \sim A + \zeta^{\kappa m} \sum_{i=0}^{\infty} \hat{\beta}_i m^{\gamma-i} \quad \text{as } m \to \infty; \qquad A_n^{(j)} - A \sim \zeta^{\kappa j} \sum_{i=0}^{\infty} \delta_i j^{\gamma-2n-i} \quad \text{as } j \to \infty \quad \text{and}$$
$$\lim_{j\to\infty} \Gamma_n^{(j)} = \left(\frac{1 + |\zeta^{\kappa}|}{|1 - \zeta^{\kappa}|}\right)^n. \tag{29}$$

Table 4
Numerical results for Example 3.1 from the $d^{(1)}$ transformation with GPS using $\sigma = 1.3$ in (25). $\bar{E}_n^{(0)}$(d): relative error in $\bar{A}_n^{(0)}$ in double precision (≈ 16 decimal digits). $\bar{E}_n^{(0)}$(q): relative error in $\bar{A}_n^{(0)}$ in quadruple precision (≈ 35 decimal digits). Note that R_n is the approximate number of terms of the series used by $A_n^{(j)}$.

n	R_n	$\bar{E}_n^{(0)}$(d)	$\bar{E}_n^{(0)}$(q)	$\Gamma_n^{(0)}$
0	1	3.92D−01	3.92D−01	1.00D+00
2	3	1.21D−02	1.21D−02	9.00D+00
4	5	1.90D−05	1.90D−05	9.17D+01
6	7	6.80D−07	6.80D−07	1.01D+03
8	11	1.14D−08	1.14D−08	3.04D+03
10	18	6.58D−11	6.59D−11	3.75D+03
12	29	1.58D−13	1.20D−13	3.36D+03
14	48	1.55D−15	4.05D−17	3.24D+03
16	80	7.11D−15	2.35D−19	2.76D+03
18	135	5.46D−14	1.43D−22	2.32D+03
20	227	8.22D−14	2.80D−26	2.09D+03
22	383	1.91D−13	2.02D−30	1.97D+03
24	646	1.00D−13	4.43D−32	1.90D+03
26	1090	4.21D−14	7.24D−32	1.86D+03
28	1842	6.07D−14	3.27D−31	1.82D+03
30	3112	1.24D−13	2.52D−31	1.79D+03

Table 5
Numerical results for Example 3.2 from the $d^{(1)}$ transformation with APS, comparing $\kappa = 10$ with $\kappa = 1$. $\bar{E}_n^{(0)}$: relative error in $\bar{A}_n^{(0)}$ for $\zeta = 0.95$ in quadruple precision (≈ 35 decimal digits). Note that $R_n = \kappa(n+1)$, and this is the approximate number of terms of the series used by $A_n^{(j)}$. (Note also that the $d^{(1)}$ transformation with $\kappa = 1$ is simply the u transformation.)

n	$\bar{E}_n^{(0)}$ ($\kappa = 1$)	$\Gamma_n^{(0)}$ ($\kappa = 1$)	$\bar{E}_n^{(0)}$ ($\kappa = 10$)	$\Gamma_n^{(0)}$ ($\kappa = 10$)
0	6.83D−01	1.00D+00	1.72D−01	1.00D+00
4	1.70D−01	3.92D+02	1.09D−04	1.49D+01
8	9.65D−03	1.64D+04	2.36D−08	9.92D+01
12	6.69D−04	7.86D+05	5.45D−12	6.71D+02
16	4.88D−05	3.87D+07	1.28D−15	4.55D+03
20	3.63D−06	1.93D+09	3.04D−19	3.09D+04
24	2.73D−07	9.66D+10	7.24D−23	2.10D+05
28	2.06D−08	4.85D+12	1.73D−26	1.43D+06
32	1.56D−09	2.44D+14	1.92D−28	9.73D+06
36	1.19D−10	1.23D+16	3.97D−27	6.62D+07
40	9.02D−12	6.17D+17	1.34D−26	4.51D+08
44	6.87D−13	3.11D+19	3.30D−26	3.07D+09
48	4.47D−14	1.57D+21	3.05D−25	2.09D+10
52	1.45D−12	7.89D+22	5.19D−25	1.42D+11
56	1.45D−11	3.98D+24	2.43D−23	9.67D+11
60	1.75D−09	2.01D+26	1.39D−22	6.58D+12

It is clear that, in case ζ is very close to 1, $\Gamma_n^{(j)}$ has a smaller value with increasing κ since ζ^κ is farther from 1 than ζ is.

In Table 4, we present the results obtained from the $d^{(1)}$ transformation as applied to Example 3.1 (logarithmically converging $\{A_m\}$) using GPS. These results suggest clearly that diagonal sequences are absolutely stable and achieve and retain almost machine accuracy. We note that the accuracy achieved for each $A_n^{(0)}$ is consistent with our conclusion following (12).

In Table 5, we present the results obtained from the $d^{(1)}$ transformation as applied to Example 3.2 (linearly converging $\{A_m\}$) using APS. These results suggest that APS improves stability when $\zeta \approx 1$. Increased stability allows the accuracy of $A_n^{(j)}$ to improve dramatically, much before $\Gamma_n^{(0)}$ has increased appreciably. We note again that the accuracy achieved for each $A_n^{(0)}$ is consistent with our conclusion following (12).

In Table 6, the third column shows the best quadruple-precision results that can be achieved by the diagonal sequence $\{A_n^{(0)}\}_{n=0}^{\infty}$ from the u transformation on the series of Example 3.2, as ζ tends to 1, the point of singularity of the sum $\log[(1-\zeta)^{-1}]$. The fourth column shows the deterioration of the accuracy of $\bar{A}_n^{(0)}$, for a fixed n (here, $n = 28$), as $\zeta \to 1$. The last column of the table is obtained by applying the $d^{(1)}$ transformation with APS. Only, in this case, we vary κ simultaneously with ζ so as to fix the value of ζ^κ at 0.5. The results of this column suggest that, as a function of ζ, an almost uniform error in $A_n^{(0)}$ with fixed n (here, $n = 28$) could be maintained by this strategy with APS. This seems to be an effective way of maintaining uniform error treating the Gibbs phenomenon as well. More generally, this strategy also seems to be an effective way of maintaining a uniform error near points of singularity of the sum of the series in question.

Table 6
Absolute errors $\widetilde{E}_n^{(0)} = |\bar{A}_n^{(0)}(\zeta) - \log[(1-\zeta)^{-1}]|$, with ζ approaching 1, obtained from the $d^{(1)}$ transformation with APS on Example 3.2 in quadruple-precision arithmetic (≈ 35 decimal digits). Results in last column obtained by choosing $\kappa = s$ when $\zeta = 0.5^{1/s}$, $s = 1, \ldots, 7$. Note that $R_n = \kappa(n+1)$, and this is the approximate number of terms of the series used by $A_n^{(j)}$.

s	ζ	$\widetilde{E}_n^{(0)}$ $(\kappa=1)$ "smallest" error	$\widetilde{E}_{28}^{(0)}$ $(\kappa=1)$	$\widetilde{E}_{28}^{(0)}$ $(\kappa=s)$
1	0.5	4.97D−30 $(n=28)$	4.97D−30	4.97D−30
2	$0.5^{1/2} \approx 0.707$	1.39D−25 $(n=35)$	5.11D−21	3.26D−31
3	$0.5^{1/3} \approx 0.794$	5.42D−23 $(n=38)$	4.70D−17	4.79D−31
4	$0.5^{1/4} \approx 0.841$	1.41D−21 $(n=41)$	1.02D−14	5.74D−30
5	$0.5^{1/5} \approx 0.871$	1.31D−19 $(n=43)$	3.91D−13	1.59D−30
6	$0.5^{1/6} \approx 0.891$	1.69D−18 $(n=43)$	5.72D−12	2.60D−30
7	$0.5^{1/7} \approx 0.906$	1.94D−17 $(n=44)$	4.58D−11	4.72D−30

4. The $d^{(m)}$ transformation and its application to general linear sequences using APS

In this section, we treat very briefly the Levin–Sidi [14] $d^{(m)}$ transformation (see also [33, Chapter 6]) and discuss the types of sequences whose convergence it accelerates effectively. In particular, we concentrate on *general linear sequences*. We also discuss its use in conjunction with APS, and show, with an application to a Fourier series, how its stability can be enhanced.

4.1. The $d^{(m)}$ transformation

To facilitate this discussion, we will introduce two definitions (see [33, Chapter 6]):

Definition 4.1. We say a function $\alpha(n)$ is in $\mathbf{A}_0^{(\gamma)}$ if it has an asymptotic expansion of the form

$$\alpha(n) \sim \sum_{i=0}^{\infty} \alpha_i n^{\gamma - i} \quad \text{as } n \to \infty.$$

Definition 4.2. We say a sequence $\{a_n\}$ is in $\mathbf{b}^{(m)}$ if its terms a_n satisfy a linear homogeneous difference equation of the form

$$a_n = \sum_{k=1}^{m} p_k(n)\big(\Delta^k a_n\big); \quad p_k(n) \in \mathbf{A}_0^{(i_k)},\ i_k \text{ integer} \leqslant k,\ k = 1, \ldots, m.$$

When $\{a_n\} \in \mathbf{b}^{(m)}$ and the series $\sum_{i=1}^{\infty} a_i$ converges, and the a_n satisfy some additional mild conditions, Levin and Sidi [14] (see also [33, Chapter 6, Theorem 6.1.12]) prove that

$$A_{n-1} \sim A + \sum_{k=0}^{m-1} n^{\rho_k}\big(\Delta^k a_n\big) \sum_{i=0}^{\infty} \beta_{ki} n^{-i} \quad \text{as } n \to \infty, \tag{30}$$

where $A_n = \sum_{i=1}^{n} a_i$, $n = 1, 2, \ldots$, and ρ_k are integers satisfying

$$\rho_k \leqslant \bar{\rho}_k = \max\{i_{k+1}, i_{k+2} - 1, \ldots, i_m - m + k + 1\} \leqslant k + 1, \quad k = 0, 1, \ldots, m-1. \tag{31}$$

Those sequences $\{a_m\}$ mentioned just before Example 3.1 in relation to the logarithmically and linearly converging sequences $A_m = \sum_{i=1}^{m} a_i$, are all in $\mathbf{b}^{(1)}$ and have served as test cases for convergence acceleration methods in almost all surveys. Actually, they make up part of the class $\mathbf{b}^{(1)}$ (see [33, Chapter 6, Theorems 6.7.1 and 6.7.2]) and also serve as building blocks for many sequences in $\mathbf{b}^{(m)}$ with $m \geqslant 2$. Fourier series and generalized Fourier series, such as orthogonal polynomial expansions (and expansions in other special functions) of functions with algebraic and/or logarithmic singularities and/or with finite jump discontinuities have terms that are in the classes $\mathbf{b}^{(m)}$ with $m \geqslant 2$. This subject is discussed in detail, and with many examples, in [33, Chapter 6], where simple rules of thumb for determining what to take for m are also provided. It is thus clear that, when taken together, the classes $\mathbf{b}^{(m)}$, $m = 1, 2, \ldots$, form a very comprehensive set of sequences.

The $d^{(m)}$ transformation for approximating the sum of the infinite series $\sum_{i=1}^{\infty} a_i$, where $\{a_n\} \in \mathbf{b}^{(m)}$, is based on the asymptotic expansion in (30). The approximations $A_n^{(m,j)}$ generated by it are defined via the linear systems

$$A_{R_l} = A_n^{(m,j)} + \sum_{k=1}^{m} R_l^k\big(\Delta^{k-1} a_{R_l}\big) \sum_{i=0}^{n_k - 1} \frac{\beta_{ki}}{R_l^i}, \quad j \leqslant l \leqslant j + N; \qquad N = \sum_{k=1}^{m} n_k, \quad n \equiv (n_1, \ldots, n_m), \tag{32}$$

where the integers $n_1,\dots,n_m$ and $1\leqslant R_0<R_1<R_2<\cdots$, are chosen by the user. (The default is $R_l=l+1$, $l=0,1,\dots$.) Note that, in this definition, the ρ_k of (30) have been replaced by their respective largest upper bounds, namely, $k+1$. Of course, when the $\bar{\rho}_k$ are known, they can be used in the definition of (32); namely, R_l^k in (32) can be replaced by $R_l^{\bar{\rho}_{k-1}}$. The way we have defined $A_n^{(m,j)}$ in (32) covers all possible sets of ρ_k, hence is robust. Comparing (32) with (1), we realize that $A_n^{(j)}=\sum_{i=0}^N\gamma_{ni}^{(j)}A_{R_{j+i}}$ with appropriate $\gamma_{ni}^{(j)}$ that satisfy $\sum_{i=0}^N\gamma_{ni}^{(j)}=1$.

Even though the definition of $A_n^{(m,j)}$ is rather complicated, a whole sequence of approximations that includes sequences of "diagonal" approximations $\{A_{(\nu,\dots,\nu)}^{(m,j)}\}_{\nu=0}^\infty$ as subsequences can be computed in a most efficient way by using the $\mathrm{W}^{(m)}$ algorithm of Ford and Sidi [10] that is recursive in nature. FORTRAN 77 codes for this algorithm, with examples of applications of the $d^{(m)}$ transformation, can be found in [10, Appendix B] and [33, Appendix I]. "Diagonal" sequences $\{A_{(\nu,\dots,\nu)}^{(m,j)}\}_{\nu=0}^\infty$ have been observed to have the best convergence properties. The FORTRAN 77 code just mentioned provides the sequence $\{A_{(\nu,\dots,\nu)}^{(m,0)}\}_{\nu=0}^\infty$.

4.2. Application to general linear sequences and Fourier series

The class $\mathbf{b}^{(m)}$ turns out to be rather rich, as mentioned earlier. It contains, among others, sequences $\{a_n\}$ whose associated partial sums $A_n=\sum_{i=1}^n a_i$ form what we will call *general linear sequences*. These sequences are characterized by asymptotic expansions of the form (see [33, Chapter 6, Theorem 6.8.8])

$$a_n\sim\sum_{k=1}^m\zeta_k^n\sum_{i=0}^\infty\alpha_{ki}n^{\gamma_k-i}\quad\text{as } n\to\infty\quad\Rightarrow\quad A_n\sim A+\sum_{k=1}^m\zeta_k^n\sum_{i=0}^\infty\beta_{ki}n^{\gamma_k-i}\quad\text{as } n\to\infty, \tag{33}$$

where

$$\zeta_k\neq 1,\quad k=1,\dots,m;\qquad \zeta_1,\dots,\zeta_m\quad\text{distinct but arbitrary;}\qquad \gamma_1,\dots,\gamma_m\quad\text{arbitrary.} \tag{34}$$

Such sequences arise from power series of functions that are analytic at the origin but have algebraic branch singularities in the complex plane. Stability problems arise in convergence acceleration when one or more of the ζ_k are near 1 in the complex plane. When used with APS, the $d^{(m)}$ transformation turns out to be very effective in achieving improved stability and accuracy. (See [33, Chapter 12, Section 12.9], for example.) Sequences as in (33) arise also from Fourier series of functions that have algebraic singularities and/or finite jump discontinuities. In this case, we have $|\zeta_1|=\cdots=|\zeta_m|=1$, in addition to (34). When used with APS, the $d^{(m)}$ transformation is very effective on these sequences as well. (See [26] and [33, Chapter 13], for example.) This can be guessed at by realizing that $A_{\kappa n}$ has an asymptotic expansion of the form

$$A_{\kappa n}\sim A+\sum_{k=1}^m\zeta_k^{\kappa n}\sum_{i=0}^\infty\hat{\beta}_{ki}n^{\gamma_k-i}\quad\text{as } n\to\infty, \tag{35}$$

which is of the same form as (33), with each ζ_k replaced by ζ_k^κ.

So far, there is no convergence or stability theory for this transformation, as it is applied to Fourier series, for example. Nevertheless, what we have learned from the theory of the $d^{(1)}$ transformation concerning linearly convergent sequences (such as Example 3.2) seems to be just as relevant for the $d^{(m)}$ transformation on Fourier and generalized Fourier series. The next example illustrates this.

Example 4.3. Consider the application of the $d^{(m)}$ transformation to the slowly converging Fourier series

$$\frac{2}{\pi}\sum_{k=1}^\infty\frac{\sin kh}{k}\cos kx=f(x)=H\big(h-|x|\big)-\frac{h}{\pi},\quad -\pi<x<\pi;\qquad 0<h<\pi,\quad H(y)=\begin{cases}1 & \text{for } y>0,\\ 0 & \text{for } y<0.\end{cases}$$

Now, $H(h-|x|)$ has two points of jump discontinuity in $[-\pi,\pi]$, and these are $x=\pm h$. In fact, at these points, the Gibbs phenomenon takes place. By applying the $d^{(m)}$ transformation with APS, we are able to overcome this problem effectively. When using APS, we are careful to make sure that *all* ζ_k^κ remain away from 1.

Using the results of [33, Chapter 6, Section 6.4], we are able to conclude that the sequence of the series coefficients $\{\frac{2}{\pi}\frac{\sin kh}{k}\cos kx\}$ is in the class $\mathbf{b}^{(4)}$ and hence that the right transformation for this series is $d^{(4)}$. (As a rule, for such series, m is twice the number of points of singularity.) In addition, $\bar{\rho}_k\leqslant 0$ for all k, which can be used in the definition of (32). Actually, A_n, the nth partial sum of the Fourier series, has an asymptotic expansion of the form given in (33) and (34), with $\zeta_1=e^{\mathrm{i}(x-h)}=\zeta_2^{-1}$ and $\zeta_3=e^{\mathrm{i}(x+h)}=\zeta_4^{-1}$ and $\gamma_1=\cdots=\gamma_4=-1$.

In Table 7, we present the results obtained for the "diagonal" sequence $\{A_{(\nu,\dots,\nu)}^{(m,0)}\}_{\nu=0}^\infty$. We take $h=1$ and probe the sum $f(x)$ at the point $x=0.9$ that is close to the location of the jump at $x=1$. Then ζ_1 and ζ_2 are close to 1. This explains the poorness of the results obtained with $\kappa=1$, that is for $R_l=l+1$, $l=0,1,\dots$. Of course, by choosing $R_l=\lfloor\kappa(l+1)\rfloor$, $l=0,1,\dots$, with $\kappa>1$, we can cause ζ_1^κ and ζ_2^κ, to be away from 1. At the same time, we want ζ_3^κ and ζ_4^κ to remain

Table 7
Numerical results for Example 4.3, with $h = 1$ and $x = 0.9$, from the $d^{(4)}$ transformation with APS. $\widetilde{E}_s^{(0)}$: relative error in $\bar{A}_n^{(4,0)}$ with $n = (s/4, s/4, s/4, s/4)$ in quadruple precision (≈ 35 decimal digits). The number of terms of the series used for each entry in the table is approximately $R_s = \lfloor\kappa(s+1)\rfloor$.

s	R_s	$\widetilde{E}_s^{(0)}$ ($\kappa = 1$)	R_s	$\widetilde{E}_s^{(0)}$ ($\kappa = 4.95$)	R_s	$\widetilde{E}_s^{(0)}$ ($\kappa = 8.27$)	R_s	$\widetilde{E}_s^{(0)}$ ($\kappa = 11.55$)
0	1	0.51D+00	4	0.49D+00	8	0.32D+00	11	0.21D+00
8	9	0.68D+00	44	0.46D−01	74	0.38D−01	103	0.59D−02
16	17	0.32D+00	84	0.11D−01	140	0.12D−02	196	0.16D−01
24	25	0.32D+00	123	0.71D−04	206	0.53D−06	288	0.53D−05
32	33	0.23D+00	163	0.34D−05	272	0.51D−09	381	0.65D−07
40	41	0.32D+00	202	0.16D−07	339	0.24D−11	473	0.15D−11
48	49	0.24D−01	242	0.74D−09	405	0.96D−14	565	0.23D−14
56	57	0.31D−02	282	0.16D−09	471	0.10D−16	658	0.91D−17
64	65	0.43D−03	321	0.10D−12	537	0.60D−19	750	0.67D−20
72	73	0.34D−02	361	0.16D−15	603	0.48D−21	843	0.31D−22
80	81	0.82D−04	400	0.25D−16	669	0.23D−24	935	0.11D−24
88	89	0.19D−04	440	0.80D−18	736	0.14D−24	1027	0.14D−28
96	97	0.21D−02	480	0.17D−20	802	0.50D−30	1120	0.14D−30
104	105	0.17D−03	519	0.43D−21	868	0.34D−30	1212	0.49D−32
112	113	0.12D−03	559	0.49D−20	934	0.21D−32	1305	0.22D−31
120	121	0.15D−01	598	0.34D−21	1000	0.35D−32	1397	0.17D−31

away from 1. For this, we have taken κ such that $\zeta_k^\kappa \approx -1$, $k = 3, 4$, which means $\kappa(x + h) \approx (2r + 1)\pi$ for integer r. We have taken $r = 1, 2, 3$, in our computations, for which we have set $\kappa = 4.95$, 8.27, 11.55, respectively. With these increasing κ values, ζ_1 and ζ_2 get further away from 1, while ζ_3 and ζ_4 are approximately −1, and this is an appropriate situation that helps to increase stability and to achieve almost machine accuracy. Table 7 shows that this aim is achieved very effectively with the APS strategy by choosing κ as described.

5. Partial review of literature on stability in extrapolation

In Section 3, we summarized the convergence and stability properties of the u and $d^{(1)}$ transformations that are available so far. In this section, we will present a brief and partial review of the convergence and stability properties of other sequence transformations for summation of infinite series that have been published so far. For additional results, we refer the reader to [33] and the references therein.

5.1. Classical Richardson extrapolation process

Our first example concerns sequences $\{A_n\}$ that are related to a function $A(y)$ that has an asymptotic expansion of the form

$$A(y) \sim A + \sum_{k=1}^{\infty} \alpha_k y^{\sigma_k} \quad \text{as } y \to 0; \qquad \Re\sigma_1 < \Re\sigma_2 < \cdots, \qquad \lim_{m\to k\infty} \Re\sigma_k = \infty.$$

The trapezoidal rule of Example 2.1 is an example of such sequences. Choosing $y_0 > y_1 > \cdots > 0$, the classical Richardson extrapolation then is defined via the linear systems

$$A(y_l) = A_n^{(j)} + \sum_{k=1}^{n} \bar{\alpha}_k y_l^{\sigma_k}, \quad j \leqslant l \leqslant j + n. \tag{36}$$

Comparing (36) with (1), we see that $\phi_k(y_l) = y_l^{\sigma_k}$ in (1).

1. In case $y_l = y_0\omega^l$ (that is, $y_{l+1}/y_l = \omega$) for some fixed $\omega \in (0, 1)$,

$$\sum_{i=0}^{n} \gamma_{ni}^{(j)} z^i = \prod_{k=1}^{n} \frac{z - c_k}{1 - c_k}, \qquad \Gamma_n^{(j)} \leqslant \prod_{k=1}^{n} \frac{1 + |c_k|}{|1 - c_k|}, \quad \forall j, n; \qquad c_k = \omega^{\sigma_k}.$$

Thus, $\Gamma_n^{(j)}$ is independent of, and hence bounded, in j. Provided $\Re\sigma_{k+1} - \Re\sigma_k \geqslant d > 0$ for all k, we have that $\Gamma_n^{(j)}$ is bounded in n; indeed, $\Gamma_n^{(j)}$ satisfies

$$\lim_{n\to\infty} \Gamma_n^{(j)} \leqslant \prod_{k=1}^{\infty} \frac{1 + |c_k|}{|1 - c_k|} < \infty.$$

The relevant convergence results are as follows:

$$A_n^{(j)} - A = O\big(|c_{n+1}|^j\big) \quad \text{as } j \to \infty; \qquad A_n^{(j)} - A = O\big(e^{-\lambda n}\big) \quad \text{as } n \to \infty, \quad \forall \lambda > 0.$$

2. In case $\lim_{l\to\infty}(y_{l+1}/y_l) = \omega \in (0,1)$, but $y_{l+1}/y_l \neq \omega$, $l = 0, 1, \ldots,$ we have that $\Gamma_n^{(j)}$ is bounded in j. Actually,

$$A_n^{(j)} - A = O\big(|c_{n+1}|^j\big) \quad \text{as } j \to \infty; \qquad \lim_{j\to\infty} \sum_{i=0}^{n} \gamma_{ni}^{(j)} z^i = \prod_{k=1}^{n} \frac{z - c_k}{1 - c_k}, \qquad \lim_{j\to\infty} \Gamma_n^{(j)} \leqslant \prod_{k=1}^{n} \frac{1 + |c_k|}{|1 - c_k|} < \infty;$$

$c_k = \omega^{\sigma_k}$.

3. In case $y_l = c/(l+\eta)^q$, $c, \eta, q > 0$, we have

$$A_n^{(j)} - A = O\big(j^{-q\sigma_{n+1}}\big) \quad \text{as } j \to \infty; \qquad \Gamma_n^{(j)} \sim \left(\prod_{i=1}^{n} |\sigma_i|\right)^{-1} \left(\frac{2j}{q}\right)^n \quad \text{as } j \to \infty.$$

That is, $\lim_{j\to\infty} \Gamma_n^{(j)} = \infty$. In addition, when $\sigma_k = k$, we have $\lim_{n\to\infty} \Gamma_n^{(j)} = \infty$. Actually, in this case, there holds

$$\Gamma_n^{(j)} = O\big(n^{\mu}\big) \quad \text{as } n \to \infty, \ \forall \mu > 0.$$

For the above, see Bulirsch and Stoer [6–8], and Sidi [29,30], [33, Chapters 1, 2]. See also Albrecht [2] for a matrix approach to the issue of stability for the Richardson extrapolation that is essentially the same as that described in Section 2.

5.2. *Iterated Aitken Δ^2 process and the Shanks transformation*

The iterated Aitken Δ^2 process (see Stoer and Bulirsch [37, Chapter 5] or [33, Chapter 15], for example) is defined via the recursion

$$B_0^{(j)} = A_j, \quad j \geqslant 0; \qquad B_{n+1}^{(j)} = \phi_j\big(\big\{B_n^{(s)}\big\}\big), \quad j, n \geqslant 0; \qquad \phi_j\big(\{C_s\}\big) \equiv \frac{\Delta(C_j/\Delta C_j)}{\Delta(1/\Delta C_j)}. \tag{37}$$

The Δ^2 process of Aitken [1] transforms the sequence $\{A_m\}$ to $\{\hat{A}_m\}$, where

$$\hat{A}_m = \phi_m\big(\{A_s\}\big) = \frac{\Delta(A_m/\Delta A_m)}{\Delta(1/\Delta A_m)}.$$

The Shanks [19] transformation is defined via the linear equations

$$A_l = e_k(A_n) + \sum_{k=1}^{n} \bar{\beta}_k(\Delta A_{l+k-1}), \quad j \leqslant l \leqslant j+n. \tag{38}$$

Here $B_n^{(j)}$ and $e_k(A_n)$ are the approximations to A; we will denote both of them by $A_n^{(j)}$ for simplicity of notation. The Shanks transformation can be implemented in a most efficient way via the epsilon algorithm of Wynn [39]. It can also be implemented, practically at the same cost, via the FS/qd algorithm of the author [33, Chapter 21]. Comparing (38) with (1), we realize that $A_n^{(j)} = \sum_{i=0}^{n} \gamma_{ni}^{(j)} A_{j+i}$ with appropriate $\gamma_{ni}^{(j)}$ that satisfy $\sum_{i=0}^{n} \gamma_{ni}^{(j)} = 1$. Using (37), it can be shown that such a relation holds for the iterated Δ^2 process as well.

As shown in [33, Chapter 15, Theorem 15.3.4 and Chapter 16, Theorem 16.5.4], neither the iterated Δ^2 process nor the Shanks transformation accelerate logarithmic convergence, in the sense that, if $\{A_m\} \in \mathbf{b}^{(1)}/\mathrm{LOG}$, then $\{A_n^{(j)}\}_{j=0}^{\infty}$ converges at the same rate as $\{A_m\}$.

1. Assuming A_m is as in (20) with the notation therein, provided $\gamma \neq 0, 1, \ldots,$ for both transformations, we have (see [33, Chapter 15, Theorem 15.3.4] and Garibotti and Grinstein [11])

$$A_n^{(j)} - A = O\big(\zeta^j j^{\gamma-2n}\big) \quad \text{as } j \to \infty \quad \text{and} \quad \lim_{j\to\infty} \sum_{i=0}^{n} \gamma_{ni}^{(j)} z^i = \left(\frac{z-\zeta}{1-\zeta}\right)^n, \qquad \lim_{j\to\infty} \Gamma_n^{(j)} = \left(\frac{1+|\zeta|}{|1-\zeta|}\right)^n.$$

2. In case $A_m \sim A + \sum_{k=1}^{\infty} \alpha_k \lambda_k^m$ as $m \to \infty$, where $\lambda_k \neq 1$ are distinct, $|\lambda_1| \geqslant |\lambda_2| \geqslant \cdots$, and $\lim_{k\to\infty} \lambda_k = 0$, provided $|\lambda_n| > |\lambda_{n+1}|$, for the Shanks transformation, we have (see Wynn [40] and Sidi [33, Chapter 15, Theorems 16.4.3, 16.4.5])

$$A_n^{(j)} - A = O\big(|\lambda_{n+1}|^j\big) \quad \text{as } j \to \infty \quad \text{and} \quad \lim_{j\to\infty} \sum_{i=0}^{n} \gamma_{ni}^{(j)} z^i = \prod_{i=1}^{n} \frac{z-\lambda_i}{1-\lambda_i}, \qquad \lim_{j\to\infty} \Gamma_n^{(j)} \leqslant \prod_{i=1}^{n} \frac{1+|\lambda_i|}{|1-\lambda_i|}.$$

3. More generally, let $A_m \sim A + \sum_{k=1}^{\infty} P_k(m)\lambda_k^m$ as $m \to \infty$, where $\lambda_k \neq 1$ are distinct, $|\lambda_1| \geqslant |\lambda_2| \geqslant \cdots$, and $P_k \in \pi_{\omega_k - 1}$ for all k. Let $\hat{\omega} = \max\{\omega_k\colon\ |\lambda_k| = |\lambda_{t+1}|\}$. Provided $|\lambda_t| > |\lambda_{t+1}|$ and $n = \sum_{k=1}^{t} \omega_k$, for the Shanks transformation, we have (see Sidi [27], [33, Chapter 16, Theorems 16.4.6, 16.4.8])

$$A_n^{(j)} - A = O\big(j^{\hat{\omega}-1}|\lambda_{t+1}|^j\big) \quad \text{as } j \to \infty \quad \text{and}$$

$$\lim_{j\to\infty} \sum_{i=0}^{n} \gamma_{ni}^{(j)} z^i = \prod_{i=1}^{t} \left(\frac{z-\lambda_i}{1-\lambda_i}\right)^{\omega_i}, \qquad \lim_{j\to\infty} \Gamma_n^{(j)} \leqslant \prod_{i=1}^{t} \left(\frac{1+|\lambda_i|}{|1-\lambda_i|}\right)^{\omega_i}.$$

For additional results that concern the interesting case of $n \neq \sum_{i=1}^{t} \omega_i$, see [27] and [33, Chapter 16, Theorem, 16.4.9]. Convergence takes place only for certain values of n in this case.

For additional results concerning the application of the Shanks transformation to totally monotonic and totally oscillating sequences $\{A_m\}$, see [18].

5.3. Iterated Lubkin transformation and the theta algorithm

The iterated Lubkin transformation is defined via the recursion

$$B_0^{(j)} = A_j, \quad j \geqslant 0; \qquad B_{n+1}^{(j)} = W_j\big(\big\{B_n^{(s)}\big\}\big), \quad j, n \geqslant 0; \qquad W_j\big(\{C_s\}\big) \equiv \frac{\Delta^2(C_j/\Delta C_j)}{\Delta^2(1/\Delta C_j)}.$$

The method of Lubkin [15] transforms the sequence $\{A_m\}$ to $\{\hat{A}_m\}$, where

$$\hat{A}_m = W_m\big(\{A_s\}\big) = \frac{\Delta^2(A_m/\Delta A_m)}{\Delta^2(1/\Delta A_m)}.$$

The theta algorithm is defined via the recursions

$$\theta_{-1}^{(j)} = 0, \qquad \theta_0^{(j)} = A_j, \quad j \geqslant 0; \qquad \theta_{2n+1}^{(j)} = \theta_{2n-1}^{(j+1)} + D_{2n}^{(j)}, \qquad \theta_{2n+2}^{(j)} = \theta_{2n}^{(j+1)} - \frac{\Delta\theta_{2n}^{(j+1)}}{\Delta D_{2n+1}^{(j)}} D_{2n+1}^{(j)}, \quad j, n \geqslant 0;$$

$$D_k^{(j)} = 1/\Delta\theta_k^{(j)}.$$

Note that $\theta_2^{(j)} = W_j(\{A_s\})$. For both transformations, the forward difference operator Δ operates on the index j only. Also, $B_n^{(j)}$ and $\theta_{2n}^{(j)}$ are the approximations to A; we will denote both of them by $A_n^{(j)}$ for simplicity of notation. Using the recursions for both methods, we can show that $A_n^{(j)} = \sum_{i=0}^{n} \gamma_{ni}^{(j)} A_{j+n+i}$ with appropriate $\gamma_{ni}^{(j)}$ that satisfy $\sum_{i=0}^{n} \gamma_{ni}^{(j)} = 1$. The following results are known (see Sidi [32] and [33, Chapters 15, 20]) for the column sequences obtained from $\{A_m\}$ in $\mathbf{b}^{(m)}$/LOG and $\mathbf{b}^{(m)}$/LIN:

1. Assuming A_m is as in (19) with the notation therein, we have

$$A_n^{(j)} - A = O\big(j^{\gamma_n}\big) \quad \text{and} \quad \sum_{i=0}^{n} \gamma_{ni}^{(j)} z^i \sim \left(\prod_{k=1}^{n} \gamma_k\right)^{-1} (1-z)^n j^n, \qquad \Gamma_n^{(j)} \sim \left|\prod_{k=1}^{n} \gamma_k\right|^{-1} (2j)^n \quad \text{as } j \to \infty,$$

where $\gamma_0 = \gamma$ and γ_k are such that $\gamma_{k+1} - \gamma_k \leqslant -2$, $k = 0, 1, \ldots$. The γ_k for the two methods considered are not necessarily the same. In any case, we have $\gamma_n \leqslant \gamma - 2n$ for all n.
2. Assuming A_m is as in (20) with the notation therein, we have

$$A_n^{(j)} - A = O\big(\zeta^j j^{\gamma_n}\big) \quad \text{as } j \to \infty \quad \text{and} \quad \lim_{j\to\infty} \sum_{i=0}^{n} \gamma_{ni}^{(j)} z^i = \left(\frac{z-\zeta}{1-\zeta}\right)^n, \qquad \lim_{j\to\infty} \Gamma_n^{(j)} = \left(\frac{1+|\zeta|}{|1-\zeta|}\right)^n,$$

where $\gamma_0 = \gamma$ and γ_k are such that $\gamma_{k+1} - \gamma_k \leqslant -3$, $k = 0, 1, \ldots$. The γ_k for the two methods considered are not necessarily the same. In any case, we have $\gamma_n \leqslant \gamma - 3n$ for all n. By the form of the asymptotic expansion of $A_{\kappa m}$ in (29), it is clear that APS can be applied in conjunction with these methods, in the sense that the methods are applied to the subsequence $\{A_{\kappa m}\}$. This is likely to improve the stability of the $A_n^{(j)}$ substantially when ζ is very close to 1.

5.4. General linear sequences and GREP$^{(m)}$

When $\{A_n\}$ is a general linear sequence precisely as described in (33) and (34), with the notation therein, if we know the ζ_k and γ_k, we can use the generalized Richardson extrapolation process GREP$^{(m)}$ by the author (see [21] and [33, Chapter 4]) on $\{A_n\}$ by defining the approximation $A_n^{(m,j)}$ to A as in

$$A_l = A_n^{(j)} + \sum_{k=1}^{m} \zeta_k^l \sum_{i=0}^{n_k-1} \bar{\beta}_{ki}(l+1)^{\gamma_k - i}, \quad j \leqslant l \leqslant j+N; \qquad N = \sum_{k=1}^{m} n_k, \quad n \equiv (n_1, \ldots, n_m). \tag{39}$$

Comparing (39) with (1), we realize that $A_n^{(j)} = \sum_{i=0}^{N} \gamma_{ni}^{(j)} A_{j+i}$ with appropriate $\gamma_{ni}^{(j)}$ that satisfy $\sum_{i=0}^{N} \gamma_{ni}^{(j)} = 1$. As shown recently in Sidi [34],

$$A_n^{(j)} - A = \sum_{k=1}^{m} O\left(\zeta_k^j j^{\gamma_k - 2n_k}\right) \quad \text{as } j \to \infty \quad \text{and}$$

$$\lim_{j\to\infty} \sum_{i=0}^{n} \gamma_{ni}^{(j)} z^i = \prod_{k=1}^{m} \left(\frac{z-\zeta_k}{1-\zeta_k}\right)^{n_k}, \qquad \lim_{j\to\infty} \Gamma_n^{(j)} \leqslant \prod_{i=1}^{m} \left(\frac{1+|\zeta_k|}{|1-\zeta_k|}\right)^{n_k}.$$

This means that if $\zeta_k \approx 1$ for some k, then we can apply GREP$^{(m)}$ to $\{A_{\kappa n}\}$ with some $\kappa > 1$ integer, making sure that all ζ_k^κ stay away from 1. Because the $d^{(m)}$ transformation is a GREP$^{(m)}$, we have used this APS strategy successfully in conjunction with the $d^{(m)}$ transformation in Section 4.

In case the ζ_k and γ_k are not known, it has been observed that the only known nonlinear sequence transformations that accelerate the convergence of general linear sequences are the Shanks transformation and the $d^{(m)}$ transformation. So far, there are no theoretical results to explain the success of these methods, however.

References

[1] A.C. Aitken, On Bernoulli's numerical solution of algebraic equations, Proc. Roy. Soc. Edinburgh 46 (1926) 289–305.
[2] P. Albrecht, Simplified error considerations for Richardson's method, Z. Angew. Math. Mech. 57 (1977) 49–51.
[3] C. Brezinski, Accélération de suites à convergence logarithmique, C. R. Acad. Sci. Paris 273 A (1971) 727–730.
[4] C. Brezinski, A general extrapolation algorithm, Numer. Math. 35 (1980) 175–187.
[5] C. Brezinski, M. Redivo Zaglia, Extrapolation Methods: Theory and Practice, North-Holland, Amsterdam, 1991.
[6] R. Bulirsch, J. Stoer, Fehlerabschätzungen und Extrapolation mit rationalen Funktionen bei Verfahren vom Richardson-Typus, Numer. Math. 6 (1964) 413–427.
[7] R. Bulirsch, J. Stoer, Numerical treatment of ordinary differential equations by extrapolation methods, Numer. Math. 8 (1966) 1–13.
[8] R. Bulirsch, J. Stoer, Numerical quadrature by extrapolation, Numer. Math. 9 (1967) 271–278.
[9] T. Fessler, W.F. Ford, D.A. Smith, HURRY: An acceleration algorithm for scalar sequences and series, ACM Trans. Math. Software 9 (1983) 346–354.
[10] W.F. Ford, A. Sidi, An algorithm for a generalization of the Richardson extrapolation process, SIAM J. Numer. Anal. 24 (1987) 1212–1232.
[11] C.R. Garibotti, F.F. Grinstein, Recent results relevant to the evaluation of infinite series, J. Comp. Appl. Math. 9 (1983) 193–200.
[12] T. Håvie, Generalized Neville type extrapolation schemes, BIT 19 (1979) 204–213.
[13] D. Levin, Development of non-linear transformations for improving convergence of sequences, Int. J. Comput. Math. B 3 (1973) 371–388.
[14] D. Levin, A. Sidi, Two new classes of nonlinear transformations for accelerating the convergence of infinite integrals and series, Appl. Math. Comp. 9 (1981) 175–215, originally appeared as a Tel Aviv University preprint in 1975.
[15] S. Lubkin, A method of summing infinite series, J. Res. Nat. Bur. Standards 48 (1952) 228–254.
[16] C. Schneider, Vereinfachte Rekursionen zur Richardson-Extrapolation in Spezialfällen, Numer. Math. 24 (1975) 177–184.
[17] M.N. Senhadji, On condition numbers of some quasi-linear sequence transformations, J. Comp. Appl. Math. 104 (1999) 1–17.
[18] M.N. Senhadji, On condition numbers of the Shanks transformation, J. Comp. Appl. Math. 135 (2001) 41–61.
[19] D. Shanks, Nonlinear transformations of divergent and slowly convergent sequences, J. Math. Phys. 34 (1955) 1–42.
[20] A. Sidi, Convergence properties of some nonlinear sequence transformations, Math. Comp. 33 (1979) 315–326.
[21] A. Sidi, Some properties of a generalization of the Richardson extrapolation process, J. Inst. Math. Applics. 24 (1979) 327–346.
[22] A. Sidi, Analysis of convergence of the T-transformation for power series, Math. Comp. 35 (1980) 833–850.
[23] A. Sidi, An algorithm for a special case of a generalization of the Richardson extrapolation process, Numer. Math. 38 (1982) 299–307.
[24] A. Sidi, On a generalization of the Richardson extrapolation process, Numer. Math. 57 (1990) 365–377.
[25] A. Sidi, On rates of acceleration of extrapolation methods for oscillatory infinite integrals, BIT 30 (1990) 347–357.
[26] A. Sidi, Acceleration of convergence of (generalized) Fourier series by the d-transformation, Ann. Numer. Math. 2 (1995) 381–406.
[27] A. Sidi, Extension and completion of Wynn's theory on convergence of columns of the epsilon table, J. Approx. Theory 86 (1996) 21–40.
[28] A. Sidi, Further convergence and stability results for the generalized Richardson extrapolation process GREP$^{(1)}$ with an application to the $D^{(1)}$-transformation for infinite integrals, J. Comp. Appl. Math. 112 (1999) 269–290.
[29] A. Sidi, Extrapolation methods and derivatives of limits of sequences, Math. Comp. 69 (2000) 305–323.
[30] A. Sidi, The Richardson extrapolation process with a harmonic sequence of collocation points, SIAM J. Numer. Anal. 37 (2000) 1729–1746.
[31] A. Sidi, New convergence results on the generalized Richardson extrapolation process GREP$^{(1)}$ for logarithmic sequences, Math. Comp. 71 (2002) 1569–1596.
[32] A. Sidi, A convergence and stability study of the iterated Lubkin transformation and the θ-algorithm, Math. Comp. 72 (2003) 419–433.
[33] A. Sidi, Practical Extrapolation Methods: Theory and Applications, Cambridge Monographs on Applied and Computational Mathematics, vol. 10, Cambridge University Press, Cambridge, 2003.
[34] A. Sidi, Asymptotic analysis of a generalized Richardson extrapolation process on linear sequences, Math. Comp., in press; published electronically on November 30, 2009.

[35] D.A. Smith, W.F. Ford, Acceleration of linear and logarithmic convergence, SIAM J. Numer. Anal. 16 (1979) 223–240.
[36] D.A. Smith, W.F. Ford, Numerical comparisons of nonlinear convergence accelerators, Math. Comp. 38 (1982) 481–499.
[37] J. Stoer, R. Bulirsch, Introduction to Numerical Analysis, 3rd ed., Springer-Verlag, New York, 2002.
[38] E.J. Weniger, Nonlinear sequence transformations for the acceleration of convergence and the summation of series, Comput. Phys. Rep. 10 (1989) 189–371.
[39] P. Wynn, On a device for computing the $e_m(S_n)$ transformation, Mathematical Tables and Other Aids to Computation 10 (1956) 91–96.
[40] P. Wynn, On the convergence and stability of the epsilon algorithm, SIAM J. Numer. Anal. 3 (1966) 91–122.