JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

ELSEVIER Journal of Computational and Applied Mathematics 122 (2000) 251–273

www.elsevier.nl/locate/cam

The generalized Richardson extrapolation process GREP$^{(1)}$ and computation of derivatives of limits of sequences with applications to the $d^{(1)}$-transformation

Avram Sidi

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

Received 3 May 1999; received in revised form 15 December 1999

Abstract

Let $\{S_m\}$ be an infinite sequence whose limit or antilimit S can be approximated very efficiently by applying a suitable extrapolation method E_0 to $\{S_m\}$. Assume that the S_m and hence also S are differentiable functions of some parameter ξ, $(\mathrm{d}/\mathrm{d}\xi)S$ being the limit or antilimit of $\{(\mathrm{d}/\mathrm{d}\xi)S_m\}$, and that we need to approximate $(\mathrm{d}/\mathrm{d}\xi)S$. A direct way of achieving this would be by applying again a suitable extrapolation method E_1 to the sequence $\{(\mathrm{d}/\mathrm{d}\xi)S_m\}$, and this approach has often been used efficiently in various problems of practical importance. Unfortunately, as has been observed at least in some important cases, when $(\mathrm{d}/\mathrm{d}\xi)S_m$ and S_m have essentially different asymptotic behaviors as $m \to \infty$, the approximations to $(\mathrm{d}/\mathrm{d}\xi)S$ produced by this approach, despite the fact that they are good, do not converge as quickly as those obtained for S, and this is puzzling. In a recent paper (A. Sidi, Extrapolation methods and derivatives of limits of sequences, *Math. Comp.*, 69 (2000) 305–323) we gave a rigorous mathematical explanation of this phenomenon for the cases in which E_0 is the Richardson extrapolation process and E_1 is a generalization of it, and we showed that the phenomenon has nothing to do with numerics. Following that we proposed a very effective procedure to overcome this problem that amounts to first applying the extrapolation method E_0 to $\{S_m\}$ and then differentiating the resulting approximations to S. As a practical means of implementing this procedure we also proposed the direct differentiation of the recursion relations of the extrapolation method E_0 used in approximating S. We additionally provided a thorough convergence and stability analysis in conjunction with the Richardson extrapolation process from which we deduced that the new procedure for $(\mathrm{d}/\mathrm{d}\xi)S$ has practically the same convergence properties as E_0 for S. Finally, we presented an application to the computation of integrals with algebraic/logarithmic endpoint singularities via the Romberg integration. In this paper we continue this research by treating Sidi's generalized Richardson extrapolation process GREP$^{(1)}$ in detail. We then apply the new procedure to various infinite series of logarithmic type (whether convergent or divergent) in conjunction with the $d^{(1)}$-transformation of Levin and Sidi. Both the theory and the numerical results of this paper too indicate that this approach is the preferred one for computing derivatives of limits of infinite sequences and series. © 2000 Elsevier Science B.V. All rights reserved.

MSC: 40A25; 41A60; 65B05; 65B10; 65D30

E-mail address: asidi@cs.technion.ac.il (A. Sidi)

PII: S0377-0427(00)00362-9

1. Introduction and review of recent developments

Let $\{S_m\}$ be an infinite sequence whose limit or antilimit S can be approximated very efficiently by applying a suitable extrapolation method E_0 to $\{S_m\}$. Assume that the S_m and hence also S are differentiable functions of some parameter ξ, $(\mathrm{d}/\mathrm{d}\xi)S$ being the limit or antilimit of $\{(\mathrm{d}/\mathrm{d}\xi)S_m\}$, and that we need to approximate $(\mathrm{d}/\mathrm{d}\xi)S$. A direct way of achieving this would be by applying again a suitable extrapolation method E_1 to the sequence $\{(\mathrm{d}/\mathrm{d}\xi)S_m\}$, and this approach has often been used efficiently in various problems of practical importance. When S_m and $(\mathrm{d}/\mathrm{d}\xi)S_m$ have essentially different asymptotic behaviors as $m \to \infty$, the approximations to $(\mathrm{d}/\mathrm{d}\xi)S$ produced by applying E_1 to $\{(\mathrm{d}/\mathrm{d}\xi)S_m\}$ do not converge to $(\mathrm{d}/\mathrm{d}\xi)S$ as quickly as the approximations to S obtained by applying E_0 to $\{S_m\}$ even though they may be good. This is a curious and disturbing phenomenon that calls for an explanation and a befitting remedy, and both of these issues were addressed by the author in the recent paper [14] via the Richardson extrapolation. As far as is known to us [14] is the first work that handles this problem.

The procedure to cope with the problem above that was proposed in [14] amounts to *first applying the extrapolation method E_0 to $\{S_m\}$ and then differentiating the resulting approximations to S*. As far as practical implementation of this procedure is concerned, it was proposed in [14] to *actually differentiate the recursion relations satisfied by the method E_0*.

In the present work we continue this new line of research by extending the approach of [14] to GREP$^{(1)}$ that is the simplest case of the generalized Richardson extrapolation process GREP of Sidi [7]. Following this, we consider the application of the $d^{(1)}$-transformation, the simplest of the d-transformations of Levin and Sidi [6], to computing derivatives of sums of infinite series. Now GREP is a most powerful extrapolation procedure that can be applied to a very large class of sequences and the d-transformations are GREPs that can be applied successfully again to a very large class of infinite series. Indeed, it is known theoretically and has been observed numerically that GREP in general and the d-transformations in particular have scopes larger than most known extrapolation methods.

Before we go on to the main theme of this paper, we will give a short review of the motivation and results of [14]. This will also help establish some of the notation that we will use in the remainder of this work and set the stage for further developments. As we did in [14], here too we will keep the treatment general by recalling that infinite sequences are either directly related to or can be formally associated with a function $A(y)$, where y may be a continuous or discrete variable.

Let a function $A(y)$ be known and hence computable for $y \in (0, b]$ with some $b > 0$, the variable y being continuous or discrete. Assume, furthermore, that $A(y)$ has an asymptotic expansion of the form

$$A(y) \sim A + \sum_{k=1}^{\infty} \alpha_k y^{\sigma_k} \quad \text{as } y \to 0+, \tag{1.1}$$

where σ_k are known scalars satisfying

$$\sigma_k \neq 0,\ k = 1, 2, \ldots; \quad \Re\sigma_1 < \Re\sigma_2 < \cdots; \quad \lim_{k\to\infty} \Re\sigma_k = +\infty, \tag{1.2}$$

and A and α_k, $k = 1, 2, \ldots$, are constants independent of y that are not necessarily known.

From (1.1) and (1.2) it is clear that $A=\lim_{y\to 0+}A(y)$ when this limit exists. When $\lim_{y\to 0+}A(y)$ does not exist, A is the antilimit of $A(y)$ for $y\to 0+$, and in this case $\Re\sigma_1\leqslant 0$ necessarily. In any case, A can be approximated very effectively by the Richardson extrapolation process that is defined via the linear systems of equations

$$A(y_l)=A_n^{(j)}+\sum_{k=1}^{n}\bar{\alpha}_k y_l^{\sigma_k},\quad j\leqslant l\leqslant j+n, \tag{1.3}$$

with the y_l picked as

$$y_l=y_0\omega^l,\ \ l=0,1,\ldots,\quad \text{for some } y_0\in(0,b] \text{ and } \omega\in(0,1). \tag{1.4}$$

Here $A_n^{(j)}$ are the approximations to A and the $\bar{\alpha}_k$ are additional (auxiliary) unknowns. As is well known, $A_n^{(j)}$ can be computed very efficiently by the following algorithm due to Bulirsch and Stoer [2]:

$$A_0^{(j)}=A(y_j),\quad j=0,1,\ldots,$$

$$A_n^{(j)}=\frac{A_{n-1}^{(j+1)}-c_nA_{n-1}^{(j)}}{1-c_n},\quad j=0,1,\ldots,\ \ n=1,2,\ldots, \tag{1.5}$$

where we have defined

$$c_n=\omega^{\sigma_n},\quad n=1,2,\ldots\ . \tag{1.6}$$

Let us now consider the situation in which $A(y)$ and hence A depend on some real or complex parameter ξ and are continuously differentiable in ξ for ξ in some set X of the real line or the complex plane, and we are interested in computing $(\mathrm{d}/\mathrm{d}\xi)A\equiv\dot{A}$. Let us assume in addition to the above that $(\mathrm{d}/\mathrm{d}\xi)A(y)\equiv\dot{A}(y)$ has an asymptotic expansion for $y\to 0+$ that is obtained by differentiating that in (1.1) term by term. (This assumption is satisfied at least in some cases of practical interest as can be shown rigorously.) Finally, let us assume that the α_k and σ_k, as well as $A(y)$ and A, depend on ξ and that they are continuously differentiable for $\xi\in X$. As a consequence of these assumptions we have

$$\dot{A}(y)\sim\dot{A}+\sum_{k=1}^{\infty}(\dot{\alpha}_k+\alpha_k\dot{\sigma}_k\log y)y^{\sigma_k}\quad\text{as } y\to 0+, \tag{1.7}$$

where $\dot{\alpha}_k\equiv(\mathrm{d}/\mathrm{d}\xi)\alpha_k$ and $\dot{\sigma}_k\equiv(\mathrm{d}/\mathrm{d}\xi)\sigma_k$. Obviously, $\dot{A}$ and the $\dot{\alpha}_k$ and $\dot{\sigma}_k$ are independent of y. As a result, the infinite sum on the right-hand side of (1.7) is simply of the form $\sum_{k=1}^{\infty}(\alpha_{k0}+\alpha_{k1}\log y)y^{\sigma_k}$ with α_{k0} and α_{k1} constants independent of y.

Note that when the σ_k do not depend on ξ, we have $\dot{\sigma}_k=0$ for all k, and, therefore, the asymptotic expansion in (1.7) becomes of exactly the same form as that given in (1.1). This means that we can apply the Richardson extrapolation process above directly to $\dot{A}(y)$ and obtain very good approximations to $\dot{A}$. This amounts to replacing $A(y_j)$ in (1.5) by $\dot{A}(y_j)$, keeping everything else the same. However, when the σ_k are functions of ξ, the asymptotic expansion in (1.7) is essentially different from that in (1.1). This is so since $y^{\sigma_k}\log y$ and y^{σ_k} behave entirely differently as $y\to 0+$. In this case the application of the Richardson extrapolation process directly to $\dot{A}(y)$ does not produce approximations to $\dot{A}$ that are of practical value.

The existence of an asymptotic expansion for $\dot{A}(y)$ of the form given in (1.7), however, suggests immediately that a generalized Richardson extrapolation process can be applied to produce approximations to $\dot{A}$ in an efficient manner. In keeping with the convention introduced by the author in [12], this extrapolation process is defined via the linear systems

$$B(y_l) = B_n^{(j)} + \sum_{k=1}^{\lfloor (n+1)/2 \rfloor} \bar{\alpha}_{k0} y_l^{\sigma_k} + \sum_{k=1}^{\lfloor n/2 \rfloor} \bar{\alpha}_{k1} y_l^{\sigma_k} \log y_l, \quad j \leqslant l \leqslant j+n, \tag{1.8}$$

where $B(y) \equiv \dot{A}(y)$, $B_n^{(j)}$ are the approximations to $B \equiv \dot{A}$, and $\bar{\alpha}_{k0}$ and $\bar{\alpha}_{k1}$ are additional (auxiliary) unknowns. (This amounts to "eliminating" from (1.7) the functions $y^{\sigma_1}, y^{\sigma_1} \log y, y^{\sigma_2}, y^{\sigma_2} \log y, \ldots,$ in this order.) With the y_l as in (1.4), the approximations $B_n^{(j)}$ can be computed very efficiently by the following algorithm developed in Sidi [12] and denoted the SGRom-algorithm there:

$$B_0^{(j)} = B(y_j), \quad j = 0, 1, \ldots,$$

$$B_n^{(j)} = \frac{B_{n-1}^{(j+1)} - \lambda_n B_{n-1}^{(j)}}{1 - \lambda_n}, \quad j = 0, 1, \ldots, \; n = 1, 2, \ldots, \tag{1.9}$$

where we have now defined

$$\lambda_{2k-1} = \lambda_{2k} = c_k, \quad k = 1, 2, \ldots, \tag{1.10}$$

with the c_n as defined in (1.6).

Before going on, we would like to mention that the problem we have described above arises naturally in the numerical evaluation of integrals of the form $B = \int_0^1 (\log x) x^{\xi} g(x)\, \mathrm{d}x$, where $\Re \xi > -1$ and $g \in C^\infty[0,1]$. It is easy to see that $B = (\mathrm{d}/\mathrm{d}\xi) A$, where $A = \int_0^1 x^{\xi} g(x)\, \mathrm{d}x$. Furthermore, the trapezoidal rule approximation $B(h)$ to B with stepsize h has an Euler–Maclaurin (E–M) expansion that is obtained by differentiating with respect to ξ the E–M expansion of the trapezoidal rule approximation $A(h)$ to A. With this knowledge available, B can be approximated by applying a generalized Richardson extrapolation process to $B(h)$. Traditionally, this approach has been adopted in multidimensional integration of singular functions as well. For a detailed discussion see [3,9].

If we arrange the $A_n^{(j)}$ and $B_n^{(j)}$ in two-dimensional arrays of the form

$$\begin{array}{lllll} Q_0^{(0)} & & & & \\ Q_0^{(1)} & Q_1^{(0)} & & & \\ Q_0^{(2)} & Q_1^{(1)} & Q_2^{(0)} & & \\ Q_0^{(3)} & Q_1^{(2)} & Q_2^{(1)} & Q_3^{(0)} & \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{array} \tag{1.11}$$

then the diagonal sequences $\{Q_n^{(j)}\}_{n=0}^{\infty}$ with fixed j have much better convergence properties than the column sequences $\{Q_n^{(j)}\}_{j=0}^{\infty}$ with fixed n. In particular, the following convergence results are known:

1. The column sequences satisfy

$$A_n^{(j)} - A = \mathrm{O}(|c_{n+1}|^j) \quad \text{as } j \to \infty,$$

$$B_{2m+s}^{(j)} - B = \mathrm{O}(j^{1-s}|c_{m+1}|^j) \quad \text{as } j \to \infty,\ s = 0, 1. \tag{1.12}$$

2. Under the additional condition that

$$\Re\sigma_{k+1} - \Re\sigma_k \geqslant d > 0,\ \ k = 1, 2, \ldots, \quad \text{for some fixed } d \tag{1.13}$$

and assuming that α_k, $\dot{\alpha}_k$, and $\alpha_k\dot{\sigma}_k$ grow with k at most like $\exp(\beta k^\eta)$ for some $\beta \geqslant 0$ and $\eta < 2$, the diagonal sequences satisfy, for all practical purposes,

$$A_n^{(j)} - A \doteq \mathrm{O}\left(\prod_{i=1}^{n} |c_i|\right) \quad \text{as } n \to \infty,$$

$$B_n^{(j)} - B \doteq \mathrm{O}\left(\prod_{i=1}^{n} |\lambda_i|\right) \quad \text{as } n \to \infty. \tag{1.14}$$

The results pertaining to $A_n^{(j)}$ in (1.12) and (1.14), with real σ_k, are due to Bulirsch and Stoer [2]. The case of complex σ_k is contained in [12], and so are the results on $B_n^{(j)}$. Actually, [12] gives a complete treatment of the general case in which

$$A(y) \sim A + \sum_{k=1}^{\infty} \left[\sum_{i=0}^{q_k} \alpha_{ki} (\log y)^i \right] y^{\sigma_k} \quad \text{as } y \to 0+, \tag{1.15}$$

where q_k are known arbitrary nonnegative integers, and α_{ki} are constants independent of y, and the σ_k satisfy the condition

$$\sigma_k \neq 0,\ \ k = 1, 2, \ldots, \quad \Re\sigma_1 \leqslant \Re\sigma_2 \leqslant \cdots, \quad \lim_{k\to\infty} \Re\sigma_k = +\infty \tag{1.16}$$

that is much weaker than that in (1.2). Thus, the asymptotic expansions in (1.1) and (1.7) are special cases of that in (1.15) with $q_k = 0$, $k = 1, 2, \ldots,$ and $q_k = 1$, $k = 1, 2, \ldots,$ respectively.

Comparison of the diagonal sequences $\{A_n^{(j)}\}_{n=0}^{\infty}$ and $\{B_n^{(j)}\}_{n=0}^{\infty}$ (with j fixed) with the help of (1.14) reveals that the latter has inferior convergence properties, even though the computational costs of $A_n^{(j)}$ and $B_n^{(j)}$ are almost identical. (They involve the computation of $A(y_l)$, $j \leqslant l \leqslant j+n$, and $B(y_l)$, $j \leqslant l \leqslant j+n$, respectively). As a matter of fact, from (1.6), (1.10), and (1.13) it follows that the bound on $|A_{2m}^{(j)} - A|$ is smaller than that of $|B_{2m}^{(j)} - B|$ by a factor of $\mathrm{O}(\prod_{i=1}^{m} |c_{m+i}/c_i|) = \mathrm{O}(\omega^{dm^2})$ as $m \to \infty$. This theoretical observation is also supported by numerical experiments. Judging from (1.14) again, we see that, when $\Re\sigma_{k+1} - \Re\sigma_k = d$ for all k in (1.13), $B_{\lfloor\sqrt{2}\,n\rfloor}^{(j)}$ will have an accuracy comparable to that of $A_n^{(j)}$. This, however, increases the cost of the extrapolation substantially, as the cost of computing $A(y_l)$ and $B(y_l)$ increases drastically with increasing l in most cases of interest. This quantitative discussion makes it clear that the inferiority of $B_n^{(j)}$ relative to $A_n^{(j)}$ is actually mathematical and has nothing to do with numerics.

From what we have so far it is easy to identify the Richardson extrapolation of (1.3) as method E_0 and the generalized Richardson extrapolation of (1.8) as method E_1. We now turn to the new procedure "$(\mathrm{d}/\mathrm{d}\xi)E_0$".

Let us now approximate $\dot{A}$ by $(\mathrm{d}/\mathrm{d}\xi)A_n^{(j)} = \dot{A}_n^{(j)}$. This can be achieved computationally by differentiating the recursion relation in (1.5), the result being the following recursive algorithm:

$$A_0^{(j)} = A(y_j) \quad \text{and} \quad \dot{A}_0^{(j)} = \dot{A}(y_j), \ \ j = 0, 1, \ldots,$$

$$A_n^{(j)} = \frac{A_{n-1}^{(j+1)} - c_n A_{n-1}^{(j)}}{1 - c_n} \quad \text{and}$$

$$\dot{A}_n^{(j)} = \frac{\dot{A}_{n-1}^{(j+1)} - c_n \dot{A}_{n-1}^{(j)}}{1 - c_n} + \frac{\dot{c}_n}{1 - c_n}(A_n^{(j)} - A_{n-1}^{(j)}), \quad j = 0, 1, \ldots, \ n = 1, 2, \ldots \ . \tag{1.17}$$

Here $\dot{c}_n \equiv (\mathrm{d}/\mathrm{d}\xi)c_n$, $n = 1, 2, \ldots$. This shows that we need two tables of the form given in (1.11), one for $A_n^{(j)}$ and another for $\dot{A}_n^{(j)}$. We also see that the computation of the $\dot{A}_n^{(j)}$ involves both $\dot{A}(y)$ and $A(y)$.

The column sequences $\{\dot{A}_n^{(j)}\}_{j=0}^{\infty}$ converge to $\dot{A}$ almost in the same way the corresponding sequences $\{A_n^{(j)}\}_{j=0}^{\infty}$ converge to A, cf. (1.12). We have

$$\dot{A}_n^{(j)} - \dot{A} = \mathrm{O}(j|c_{n+1}|^j) \quad \text{as } j \to \infty. \tag{1.18}$$

The diagonal sequences $\{\dot{A}_n^{(j)}\}_{n=0}^{\infty}$ converge to $\dot{A}$ also practically the same way the corresponding $\{A_n^{(j)}\}_{n=0}^{\infty}$ converge to A, subject to the mild conditions that $\sum_{i=1}^{\infty} |\dot{c}_i| < \infty$ and $\sum_{i=1}^{n} |\dot{c}_i / c_i| = \mathrm{O}(n^a)$ as $n \to \infty$ for some $a \geqslant 0$, in addition to (1.13). We have for all practical purposes, cf. (1.14),

$$\dot{A}_n^{(j)} - \dot{A} \doteq \mathrm{O}\left(\prod_{i=1}^{n} |c_i| \right) \text{ as } n \to \infty. \tag{1.19}$$

The stability properties of the column and diagonal sequences of the $\dot{A}_n^{(j)}$ are likewise analyzed in [14] and are shown to be very similar to those of the $A_n^{(j)}$. We refer the reader to [14] for details.

This completes our review of the motivation and results of [14]. In the next section we present the extension of the procedure of [14] to GREP$^{(1)}$. We derive the recursive algorithm for computing the approximations and for assessing their numerical stability. In Section 3 we discuss the stability and convergence properties of the new procedure subject to a set of appropriate sufficient conditions that are met in many cases of interest. The main results of this section are Theorem 3.3 on stability and Theorem 3.4 on convergence and both are optimal asymptotically. In Section 4 we show how the method and theory of Sections 2 and 3 apply to the summation of some infinite series of logarithmic type via the $d^{(1)}$-transformation. Finally, in Section 5 we give two numerical examples that illustrate the theory and show the superiority of the new approach to derivatives of limits over the direct one. In the first example we apply the new approach to the computation of the derivative of the Riemann zeta function. In the second example we compute $(\mathrm{d}/\mathrm{d}\xi)F(\xi, \frac{1}{2}; \frac{3}{2}; 1)$, where $F(a, b; c; z)$ is the Gauss hypergeometric function. This example shows clearly that our approach is very effective for computing derivatives of special functions such as the hypergeometric functions with respect to their parameters.

2. GREP$^{(1)}$ and its derivative

2.1. General preliminaries on GREP$^{(1)}$

As GREP$^{(1)}$ applies to functions $A(y)$ that are in the class $\mathrm{F}^{(1)}$, we start by describing $\mathrm{F}^{(1)}$.

Definition 2.1. We shall say that a function $A(y)$, defined for $0 < y \leqslant b$, for some $b > 0$, where y can be a discrete or continuous variable, belongs to the set $\mathrm{F}^{(1)}$, if there exist functions $\phi(y)$ and $\beta(y)$ and a constant A, such that

$$A(y) = A + \phi(y)\beta(y), \tag{2.1}$$

where $\beta(x)$, as a function of the continuous variable x and for some $\eta \leqslant b$, is continuous for $0 \leqslant x \leqslant \eta$, and, for some constant $r > 0$, has a Poincaré-type asymptotic expansion of the form

$$\beta(x) \sim \sum_{i=0}^{\infty} \beta_i x^{ir} \quad \text{as } x \to 0+. \tag{2.2}$$

If, in addition, the function $B(t) \equiv \beta(t^{1/r})$, as a function of the continuous variable t, is infinitely differentiable for $0 \leqslant t \leqslant \eta^r$, we shall say that $A(y)$ belongs to the set $\mathrm{F}^{(1)}_{\infty}$. Note that $\mathrm{F}^{(1)}_{\infty} \subset \mathrm{F}^{(1)}$.

Remark. $A = \lim_{y\to 0+} A(y)$ whenever this limit exists. If $\lim_{y\to 0+} A(y)$ does not exist, then A is said to be the antilimit of $A(y)$. In this case $\lim_{y\to 0+} \phi(y)$ does not exist as is obvious from (2.1) and (2.2).

It is assumed that the functions $A(y)$ and $\phi(y)$ are computable for $0 < y \leqslant b$ (keeping in mind that y may be discrete or continuous depending on the situation) and that the constant r is known. The constants A and β_i are not assumed to be known. The problem is to find (or approximate) A whether it is the limit or the antilimit of $A(y)$ as $y \to 0+$, and GREP$^{(1)}$, the extrapolation procedure that corresponds to $\mathrm{F}^{(1)}$, is designed to tackle precisely this problem.

Definition 2.2. Let $A(y) \in \mathrm{F}^{(1)}$, with $\phi(y)$, $\beta(y)$, A, and r being exactly as in Definition 2.1. Pick $y_l \in (0, b]$, $l = 0, 1, 2, \ldots$, such that $y_0 > y_1 > y_2 > \cdots$, and $\lim_{l\to\infty} y_l = 0$. Then $A_n^{(j)}$, the approximation to A, and the parameters $\bar{\beta}_i$, $i = 0, 1, \ldots, n-1$, are defined to be the solution of the system of $n+1$ linear equations

$$A_n^{(j)} = A(y_l) + \phi(y_l) \sum_{i=0}^{n-1} \bar{\beta}_i y_l^{ir}, \quad j \leqslant l \leqslant j+n, \tag{2.3}$$

provided the matrix of this system is nonsingular. It is this process that generates the approximations $A_n^{(j)}$ that we call GREP$^{(1)}$.

As is seen, GREP$^{(1)}$ produces a two-dimensional table of approximations of the form given in (1.1).

Before going on we let $t=y^r$ and $t_l=y_l^r$, $l=0,1,\ldots,$ and define $a(t)\equiv A(y)$ and $\varphi(t)\equiv\phi(y)$. Then the equations in (2.3) take on the more convenient form

$$A_n^{(j)}=a(t_l)+\varphi(t_l)\sum_{i=0}^{n-1}\bar{\beta}_i t_l^i,\quad j\leqslant l\leqslant j+n. \tag{2.4}$$

A closed-form expression for $A_n^{(j)}$ can be obtained by using divided differences. In the sequel we denote by $D_k^{(s)}$ the divided difference operator of order k over the set of points $t_s,t_{s+1},\ldots,t_{s+k}$. Thus, for any function $g(t)$ defined at these points we have

$$D_k^{(s)}\{g(t)\}=g[t_s,t_{s+1},\ldots,t_{s+k}]=\sum_{l=s}^{s+k}\left(\prod_{\substack{i=s\\ i\neq l}}^{s+k}\frac{1}{t_l-t_i}\right)g(t_l)\equiv\sum_{i=0}^{k}c_{ki}^{(s)}g(t_{s+i}). \tag{2.5}$$

Then $A_n^{(j)}$ is given by

$$A_n^{(j)}=\frac{D_n^{(j)}\{a(t)/\varphi(t)\}}{D_n^{(j)}\{1/\varphi(t)\}}. \tag{2.6}$$

As is clear from (2.6), $A_n^{(j)}$ can be expressed also in the form

$$A_n^{(j)}=\sum_{i=0}^{n}\gamma_{ni}^{(j)}a(t_{j+i}), \tag{2.7}$$

where $\gamma_{ni}^{(j)}$ are constants that are independent of $a(t)$ and that depend solely on the t_l and $\varphi(t_l)$ and satisfy $\sum_{i=0}^{n}\gamma_{ni}^{(j)}=1$. The quantity $\Gamma_n^{(j)}$ defined by

$$\Gamma_n^{(j)}=\sum_{i=0}^{n}|\gamma_{ni}^{(j)}| \tag{2.8}$$

(note that $\Gamma_n^{(j)}\geqslant 1$) plays an important role in assessing the stability properties of the approximation $A_n^{(j)}$ with respect to errors (roundoff or other) in the $a(t_l)$. As has been noted in various places, if ε_l is the (absolute) error committed in the computation of $a(t_l)$, $l=0,1,\ldots,$ then $|A_n^{(j)}-\bar{A}_n^{(j)}|\leqslant\Gamma_n^{(j)}(\max_{j\leqslant l\leqslant j+n}|\varepsilon_l|)$, where $\bar{A}_n^{(j)}$ is the computed (as opposed to exact) value of $A_n^{(j)}$. Concerning $\Gamma_n^{(j)}$ we have a result analogous to (2.6), namely,

$$\Gamma_n^{(j)}=\sum_{i=0}^{n}|\gamma_{ni}^{(j)}|=\frac{|D_n^{(j)}\{u(t)\}|}{|D_n^{(j)}\{1/\varphi(t)\}|}, \tag{2.9}$$

where $u(t)$ is arbitrarily defined for all t except for $t_0,t_1,\ldots,$ where it is defined by

$$u(t_l)=(-1)^l/|\varphi(t_l)|,\quad l=0,1,\ldots\ . \tag{2.10}$$

This is a result of the following lemma that will be used again later in this paper.

Lemma 2.1. *With* $D_k^{(s)}\{g(t)\}$ *as in* (2.5), *we have*

$$\sum_{i=0}^{k} |c_{ki}^{(s)}| h_{s+i} = (-1)^s D_k^{(s)}\{u(t)\}, \tag{2.11}$$

where h_l *are arbitrary scalars and*

$$u(t_l) = (-1)^l h_l, \quad l = 0, 1, \ldots, \tag{2.12}$$

but $u(t)$ *is arbitrary otherwise.*

Proof. The validity of (2.11) follows from (2.5) and from the fact that $c_{ki}^{(s)} = (-1)^i |c_{ki}^{(s)}|,\ i = 0, 1, \ldots, k$. □

The results in (2.6) and (2.9) form the basis of the W-algorithm that is used in computing both the $A_n^{(j)}$ and the $\Gamma_n^{(j)}$ in a very efficient way. For this we define for all j and n

$$M_n^{(j)} = D_n^{(j)}\{a(t)/\varphi(t)\}, \ N_n^{(j)} = D_n^{(j)}\{1/\varphi(t)\} \quad \text{and} \quad H_n^{(j)} = D_n^{(j)}\{u(t)\} \tag{2.13}$$

with $u(t_l)$ as in (2.10), and recall the well-known recursion relation for divided differences, namely,

$$D_n^{(j)}\{g(t)\} = \frac{D_{n-1}^{(j+1)}\{g(t)\} - D_{n-1}^{(j)}\{g(t)\}}{t_{j+n} - t_j}. \tag{2.14}$$

(See, e.g., [15, p. 45].) Here are the steps of the W-Algorithm:

1. For $j = 0, 1, \ldots,$ set

$$M_0^{(j)} = a(t_j)/\varphi(t_j), \ N_0^{(j)} = 1/\varphi(t_j) \quad \text{and} \quad H_0^{(j)} = (-1)^j/|\varphi(t_j)|. \tag{2.15}$$

2. For $j = 0, 1, \ldots,$ and $n = 1, 2, \ldots,$ compute $M_n^{(j)}$, $N_n^{(j)}$, and $H_n^{(j)}$ recursively from

$$Q_n^{(j)} = \frac{Q_{n-1}^{(j+1)} - Q_{n-1}^{(j)}}{t_{j+n} - t_j} \tag{2.16}$$

   with $Q_n^{(j)}$ equal to $M_n^{(j)}$, $N_n^{(j)}$, and $H_n^{(j)}$.
3. For all j and n set

$$A_n^{(j)} = \frac{M_n^{(j)}}{N_n^{(j)}} \quad \text{and} \quad \Gamma_n^{(j)} = \frac{|H_n^{(j)}|}{|N_n^{(j)}|}. \tag{2.17}$$

Note that the W-Algorithm for $A_n^{(j)}$ was originally developed in [8]. The recursion for $\Gamma_n^{(j)}$ was given recently in [10]. Stability and convergence studies for GREP$^{(1)}$ can be found in [10], and more recently in [13].

Let us now assume that $A(y)$ and A depend on a real or complex parameter ξ and that we would like to compute $(\mathrm{d}/\mathrm{d}\xi)A \equiv \dot{A}$ assuming that $\dot{A}$ is the limit or antilimit of $(\mathrm{d}/\mathrm{d}\xi)A(y)$ as $y \to 0+$. We also assume that $\phi(y)$ and β_i in (2.1) are differentiable functions of ξ and that $\dot{A}(y)$ has an asymptotic expansion as $y \to 0+$ obtained by differentiating that of $A(y)$ given in (2.1) and (2.2) term by term. Thus

$$\dot{A}(y) \sim \dot{A} + \dot{\phi}(y)\sum_{i=0}^{\infty}\beta_i y^{ir} + \phi(y)\sum_{i=0}^{\infty}\dot{\beta}_i y^{ir} \quad \text{as } y \to 0+. \tag{2.18}$$

Here $\dot{\phi}(y) \equiv (\mathrm{d}/\mathrm{d}\xi)\phi(y)$ and $\dot{\beta}_i \equiv (\mathrm{d}/\mathrm{d}\xi)\beta_i$ in keeping with the convention of the previous section.

We can now approximate $\dot{A}$ by applying the extrapolation process GREP$^{(2)}$ to (2.18). The approximations $B_n^{(j)}$ to $B \equiv \dot{A}$ that result from this are defined via the linear systems

$$B(y_l) = B_n^{(j)} + \phi(y_l)\sum_{i=0}^{\lfloor (n-1)/2\rfloor}\bar{\beta}_{1i}y_l^{ir} + \dot{\phi}(y_l)\sum_{i=0}^{\lfloor n/2\rfloor -1}\bar{\beta}_{2i}y_l^{ir}, \quad j \leqslant l \leqslant j+n, \tag{2.19}$$

where $B(y) \equiv \dot{A}(y)$ as before. (Compare (2.18) and (2.19) with (1.7) and (1.8), respectively.) Now the $B_n^{(j)}$ converge to $\dot{A}$, but their rate of convergence to $\dot{A}$ is inferior to that of the corresponding $A_n^{(j)}$ to A. We, therefore, would like to employ the approach of [14] hoping that it will produce better results also with GREP$^{(1)}$.

2.2. (d/dξ)GREP$^{(1)}$ *and its implementation*

Let us differentiate (2.7) with respect to ξ. We obtain

$$\dot{A}_n^{(j)} = \sum_{i=0}^{n}\gamma_{ni}^{(j)}\dot{a}(t_{j+i}) + \sum_{i=0}^{n}\dot{\gamma}_{ni}^{(j)}a(t_{j+i}), \tag{2.20}$$

where $\dot{\gamma}_{ni}^{(j)} \equiv (\mathrm{d}/\mathrm{d}\xi)\gamma_{ni}^{(j)}$ and $\dot{a}(t) \equiv (\mathrm{d}/\mathrm{d}\xi)a(t) \equiv \dot{A}(y)$.

It is clear that, unlike $B_n^{(j)}$ in (2.19) that depends only on $\dot{a}(t)$, $\dot{A}_n^{(j)}$ depends on both $\dot{a}(t)$ and $a(t)$. Also the stability of $\dot{A}_n^{(j)}$ is affected by errors both in $a(t_l)$ and $\dot{a}(t_l)$. In particular, if ε_l and η_l are the (absolute) errors in $a(t_l)$ and $\dot{a}(t_l)$, respectively, then $|\dot{A}_n^{(j)} - \bar{\dot{A}}_n^{(j)}| \leqslant \Omega_n^{(j)}[\max_{j\leqslant l\leqslant j+n}\max(|\varepsilon_l|, |\eta_l|)]$, where $\bar{\dot{A}}_n^{(j)}$ is the computed (as opposed to exact) value of $\dot{A}_n^{(j)}$, and

$$\Omega_n^{(j)} = \sum_{i=0}^{n}|\gamma_{ni}^{(j)}| + \sum_{i=0}^{n}|\dot{\gamma}_{ni}^{(j)}|. \tag{2.21}$$

We shall call this extension of GREP$^{(1)}$ simply (d/dξ)GREP$^{(1)}$.

2.2.1. *Computation of* $\dot{A}_n^{(j)}$

Let us start by differentiating $A_n^{(j)} = M_n^{(j)}/N_n^{(j)}$. Upon denoting $(\mathrm{d}/\mathrm{d}\xi)M_n^{(j)} = \dot{M}_n^{(j)}$ and $(\mathrm{d}/\mathrm{d}\xi)N_n^{(j)} = \dot{N}_n^{(j)}$, we have

$$\dot{A}_n^{(j)} = \frac{\dot{M}_n^{(j)}}{N_n^{(j)}} - \frac{M_n^{(j)}\dot{N}_n^{(j)}}{[N_n^{(j)}]^2}. \tag{2.22}$$

Now $M_n^{(j)}$ and $N_n^{(j)}$ are already available from the W-algorithm. We need only compute $\dot{M}_n^{(j)}$ and $\dot{N}_n^{(j)}$, and these can be computed by direct differentiation of (2.16) along with the appropriate initial conditions in (2.15).

2.2.2. Computation of an upper bound on $\Omega_n^{(j)}$

The assessment of stability of $\dot{A}_n^{(j)}$ turns out to be much more involved than that of $A_n^{(j)}$, and it requires a good understanding of the nature of $\dot{M}_n^{(j)}$.

First, we note that, as the t_l are independent of ξ, $D_n^{(j)}$ and $(\mathrm{d}/\mathrm{d}\xi)$ commute, i.e., $(\mathrm{d}/\mathrm{d}\xi)D_n^{(j)}\{g(t)\}=D_n^{(j)}\{(\mathrm{d}/\mathrm{d}\xi)g(t)\}$. Consequently, from (2.16) we have

$$\dot{M}_n^{(j)} = D_n^{(j)}\left\{\frac{\mathrm{d}}{\mathrm{d}\xi}\frac{a(t)}{\varphi(t)}\right\} = D_n^{(j)}\left\{\frac{\dot{a}(t)}{\varphi(t)} - \frac{a(t)\dot{\varphi}(t)}{[\varphi(t)]^2}\right\}. \tag{2.23}$$

Next, substituting (2.23) in (2.22), and using the fact that $D_n^{(j)}$ is a linear operator, we obtain

$$\dot{A}_n^{(j)} = Y_1 + Y_2 + Y_3, \tag{2.24}$$

where

$$Y_1 = \frac{D_n^{(j)}\{\dot{a}(t)/\varphi(t)\}}{N_n^{(j)}} = \sum_{i=0}^{n} \gamma_{ni}^{(j)}\dot{a}(t_{j+i}),$$

$$Y_2 = -\frac{\dot{N}_n^{(j)}D_n^{(j)}\{a(t)/\varphi(t)\}}{[N_n^{(j)}]^2} = -\frac{\dot{N}_n^{(j)}}{N_n^{(j)}}\sum_{i=0}^{n}\gamma_{ni}^{(j)}a(t_{j+i}),$$

$$Y_3 = -\frac{D_n^{(j)}\{a(t)\dot{\varphi}(t)/[\varphi(t)]^2\}}{N_n^{(j)}} = -\sum_{i=0}^{n}\delta_{ni}^{(j)}a(t_{j+i}) \tag{2.25}$$

with $\delta_{ni}^{(j)} = \gamma_{ni}^{(j)}\dot{\varphi}(t_{j+i})/\varphi(t_{j+i})$. Here we have used the fact that

$$\frac{D_n^{(j)}\{h(t)/\varphi(t)\}}{D_n^{(j)}\{1/\varphi(t)\}} = \sum_{i=0}^{n}\gamma_{ni}^{(j)}h(t_{j+i}) \quad \text{for any } h(t). \tag{2.26}$$

Recalling (2.20), we identify

$$\dot{\gamma}_{ni}^{(j)} = -\frac{\dot{N}_n^{(j)}}{N_n^{(j)}}\gamma_{ni}^{(j)} - \delta_{ni}^{(j)}, \quad i = 0, 1, \ldots, n. \tag{2.27}$$

Therefore,

$$\Omega_n^{(j)} = \sum_{i=0}^{n}|\gamma_{ni}^{(j)}| + \sum_{i=0}^{n}\left|\frac{\dot{N}_n^{(j)}}{N_n^{(j)}}\gamma_{ni}^{(j)} + \delta_{ni}^{(j)}\right| = \sum_{i=0}^{n}|\gamma_{ni}^{(j)}| + \sum_{i=0}^{n}|\gamma_{ni}^{(j)}|\left|\frac{\dot{N}_n^{(j)}}{N_n^{(j)}} + \frac{\dot{\varphi}(t_{j+i})}{\varphi(t_{j+i})}\right|. \tag{2.28}$$

Now even though the first summation is simply $\Gamma_n^{(j)}$, and hence can be computed very inexpensively, the second sum cannot, as its general term depends also on $\dot{N}_n^{(j)}/N_n^{(j)}$, hence on j and n. We can, however, compute, again very inexpensively, an upper bound $\tilde{\Omega}_n^{(j)}$ on $\Omega_n^{(j)}$, defined by

$$\tilde{\Omega}_n^{(j)} = \Gamma_n^{(j)} + \frac{|\dot{N}_n^{(j)}|}{|N_n^{(j)}|}\Gamma_n^{(j)} + \Theta_n^{(j)} \quad \text{where } \Theta_n^{(j)} \equiv \sum_{i=0}^{n} |\delta_{ni}^{(j)}| \tag{2.29}$$

which is obtained by manipulating the second summation in (2.28) appropriately. This can be achieved by first realizing that

$$\Theta_n^{(j)} = \frac{|D_n^{(j)}\{v(t)\}|}{|N_n^{(j)}|}, \tag{2.30}$$

where $v(t)$ is arbitrarily defined for all t except for $t_0, t_1, \ldots,$ for which it is defined by

$$v(t_l) = (-1)^l |\dot{\varphi}(t_l)|/|\varphi(t_l)|^2, \quad l = 0, 1, \ldots \tag{2.31}$$

and then by applying Lemma 2.1.

2.2.3. The $(\mathrm{d}/\mathrm{d}\xi)W$-algorithm for $\dot{A}_n^{(j)}$

Combining all of the developments above, we can now extend the W-algorithm to compute $\dot{A}_n^{(j)}$ and $\tilde{\Omega}_n^{(j)}$. We shall denote the resulting algorithm the $(\mathrm{d}/\mathrm{d}\xi)W$-algorithm. Here are the steps of this algorithm.

1. For $j = 0, 1, \ldots,$ set

$$M_0^{(j)} = \frac{a(t_j)}{\varphi(t_j)}, \qquad N_0^{(j)} = \frac{1}{\varphi(t_j)}, \qquad H_0^{(j)} = (-1)^j |N_0^{(j)}|, \qquad \text{and}$$

$$\dot{M}_0^{(j)} = \frac{\dot{a}(t_j)}{\varphi(t_j)} - \frac{a(t_j)\dot{\varphi}(t_j)}{[\varphi(t_j)]^2}, \qquad \dot{N}_0^{(j)} = -\frac{\dot{\varphi}(t_j)}{[\varphi(t_j)]^2}, \qquad \tilde{H}_0^{(j)} = (-1)^j |\dot{N}_0^{(j)}|. \tag{2.32}$$

2. For $j=0, 1, \ldots,$ and $n=1, 2, \ldots,$ compute $M_n^{(j)}$, $N_n^{(j)}$, $H_n^{(j)}, \dot{M}_n^{(j)}$, $\dot{N}_n^{(j)}$, and $\tilde{H}_n^{(j)}$ recursively from

$$Q_n^{(j)} = \frac{Q_{n-1}^{(j+1)} - Q_{n-1}^{(j)}}{t_{j+n} - t_j}. \tag{2.33}$$

3. For all j and n set

$$A_n^{(j)} = \frac{M_n^{(j)}}{N_n^{(j)}}, \qquad \Gamma_n^{(j)} = \frac{|H_n^{(j)}|}{|N_n^{(j)}|}, \qquad \text{and}$$

$$\dot{A}_n^{(j)} = \frac{\dot{M}_n^{(j)}}{N_n^{(j)}} - A_n^{(j)}\frac{\dot{N}_n^{(j)}}{N_n^{(j)}}, \qquad \tilde{\Omega}_n^{(j)} = \frac{|\tilde{H}_n^{(j)}|}{|N_n^{(j)}|} + \left(1 + \frac{|\dot{N}_n^{(j)}|}{|N_n^{(j)}|}\right)\Gamma_n^{(j)}. \tag{2.34}$$

It is interesting to note that we need six tables of the form (1.11) in order to carry out the $(\mathrm{d}/\mathrm{d}\xi)W$-algorithm. This is *twice* the number of tables needed to carry out the W-algorithm. Note also that no tables need to be saved for $A_n^{(j)}$, $\Gamma_n^{(j)}$, $\dot{A}_n^{(j)}$, and $\tilde{\Omega}_n^{(j)}$. This seems to be the situation for all extrapolation methods.

3. Column convergence for $(\mathrm{d}/\mathrm{d}\xi)$GREP$^{(1)}$

In this section we shall give a detailed analysis of the column sequences $\{\dot{A}_n^{(j)}\}_{j=0}^{\infty}$ with n fixed for the case in which the t_l are picked such that

$$t_0 > t_1 > \cdots > 0 \quad \text{and} \quad \lim_{m\to\infty} \frac{t_{m+1}}{t_m} = \omega \quad \text{for some } \omega \in (0,1). \tag{3.1}$$

We also assume that

$$\lim_{m\to\infty} \frac{\varphi(t_{m+1})}{\varphi(t_m)} = \omega^{\delta} \quad \text{for some (complex) } \delta \neq 0, -1, -2, \ldots . \tag{3.2}$$

Recalling from Definition 2.1 that $\beta(y) \equiv B(t) \sim \sum_{i=0}^{\infty} \beta_i t^i$ as $t \to 0+$, we already have the following optimal convergence and stability results for $A_n^{(j)}$ and $\Gamma_n^{(j)}$, see Theorems 2.1 and 2.2 in [10].

Theorem 3.1. *Under the conditions given in* (3.1) *and* (3.2), *we have*

$$A_n^{(j)} - A \sim \left(\prod_{i=1}^{n} \frac{c_{n+\mu+1} - c_i}{1 - c_i} \right) \beta_{n+\mu} \varphi(t_j) t_j^{n+\mu} \quad \text{as } j \to \infty, \tag{3.3}$$

where $\beta_{n+\mu}$ *is the first nonzero* β_i *with* $i \geqslant n$, *and*

$$\lim_{j\to\infty} \sum_{i=0}^{n} \gamma_{ni}^{(j)} z^i = \prod_{i=1}^{n} \frac{z - c_i}{1 - c_i} \equiv U_n(z) \equiv \sum_{i=0}^{n} \tilde{\gamma}_{ni} z^i, \tag{3.4}$$

so that for each fixed n

$$\lim_{j\to\infty} \Gamma_n^{(j)} = \prod_{i=1}^{n} \frac{1 + |c_i|}{|1 - c_i|} \quad \text{hence} \ \sup_j \Gamma_n^{(j)} < \infty. \tag{3.5}$$

Here

$$c_k = \omega^{\delta+k-1}, \quad k = 1, 2, \ldots . \tag{3.6}$$

We shall see below that what we need for the analysis of $(\mathrm{d}/\mathrm{d}\xi)$GREP$^{(1)}$ are the asymptotic behaviors of $\gamma_{ni}^{(j)}$ and $\dot{\gamma}_{ni}^{(j)}$. Now that we know the behavior of $\gamma_{ni}^{(j)}$ as $j \to \infty$ from (3.4), we turn to the study of $\dot{\gamma}_{ni}^{(j)}$. We start with

$$\sum_{i=0}^{n} \gamma_{ni}^{(j)} z^i = \frac{T_n^{(j)}(z)}{T_n^{(j)}(1)} \quad \text{with } T_n^{(j)}(z) = \sum_{i=0}^{n} \frac{c_{ni}^{(j)}}{\varphi(t_{j+i})} z^i, \tag{3.7}$$

which follows from the fact that $\gamma_{ni}^{(j)} = [c_{ni}^{(j)}/\varphi(t_{j+i})]/D_n^{(j)}\{1/\varphi(t)\}$. Of course, $T_n^{(j)}(1) = D_n^{(j)}\{1/\varphi(t)\}$.

Differentiating (3.7) with respect to ξ, and denoting $\dot{T}_n^{(j)}(z) = (\mathrm{d}/\mathrm{d}\xi) T_n^{(j)}(z)$, we obtain

$$\sum_{i=0}^{n} \dot{\gamma}_{ni}^{(j)} z^i = \frac{\dot{T}_n^{(j)}(z) T_n^{(j)}(1) - T_n^{(j)}(z) \dot{T}_n^{(j)}(1)}{[T_n^{(j)}(1)]^2}. \tag{3.8}$$

Obviously,

$$\dot{T}_n^{(j)}(z) = -\sum_{i=0}^{n} c_{ni}^{(j)} \frac{\dot{\varphi}(t_{j+i})}{[\varphi(t_{j+i})]^2} z^i, \tag{3.9}$$

as a result of which we have

$$\frac{\dot{T}_n^{(j)}(z)}{T_n^{(j)}(1)} = -\sum_{i=0}^{n} \gamma_{ni}^{(j)} \frac{\dot{\varphi}(t_{j+i})}{\varphi(t_{j+i})} z^i. \tag{3.10}$$

Substituting (3.10) in (3.8) and using the fact that $\sum_{i=0}^{n} \gamma_{ni}^{(j)} = 1$, we finally get

$$\sum_{i=0}^{n} \dot{\gamma}_{ni}^{(j)} z^i = -\sum_{i=0}^{n} \gamma_{ni}^{(j)} \frac{\dot{\varphi}(t_{j+i})}{\varphi(t_{j+i})} z^i + \left(\sum_{i=0}^{n} \gamma_{ni}^{(j)} z^i\right)\left(\sum_{i=0}^{n} \gamma_{ni}^{(j)} \frac{\dot{\varphi}(t_{j+i})}{\varphi(t_{j+i})}\right). \tag{3.11}$$

We have now come to the point where we have to make a suitable assumption on $\dot{\varphi}(t)$. The following assumption seems to be quite realistic for many examples that involve logarithmically convergent sequences and some others as well:

$$\dot{\varphi}(t) = \varphi(t)[K \log t + L + o(1)] \quad \text{as } t \to 0+ \quad \text{for some constants } K \neq 0 \text{ and } L. \tag{3.12}$$

Now the condition $\lim_{m\to\infty}(t_{m+1}/t_m) = \omega$ in (3.1) implies that $t_{m+1}/t_m = \omega(1+\varepsilon_m)$, where $\lim_{m\to\infty} \varepsilon_m = 0$. Therefore, $t_{j+i} = t_j \omega^i \prod_{s=0}^{i-1}(1+\varepsilon_{j+s})$, and hence, for each fixed $i \geqslant 0$

$$\log t_{j+i} = \log t_j + i \log \omega + \varepsilon_i^{(j)}, \quad \lim_{j\to\infty} \varepsilon_i^{(j)} = 0, \tag{3.13}$$

since $\varepsilon_i^{(j)} = O(\max\{|\varepsilon_j|, |\varepsilon_{j+1}|, \ldots, |\varepsilon_{j+1-i}|\})$. Next, (3.12) and (3.13) imply that, for each fixed $i \geqslant 0$,

$$\frac{\dot{\varphi}(t_{j+i})}{\varphi(t_{j+i})} = (K \log t_j + L) + Ki \log \omega + \sigma_i^{(j)}, \quad \lim_{j\to\infty} \sigma_i^{(j)} = 0, \tag{3.14}$$

since $\lim_{m\to\infty} t_m = 0$.

Substituting (3.14) in (3.11), we see that the problematic term $(K \log t_j + L)$ that is unbounded as $j \to \infty$ disappears altogether, and we obtain

$$\sum_{i=0}^{n} \dot{\gamma}_{ni}^{(j)} z^i = -\sum_{i=0}^{n} \gamma_{ni}^{(j)} (Ki \log \omega + \sigma_i^{(j)}) z^i + \left(\sum_{i=0}^{n} \gamma_{ni}^{(j)} z^i\right)\left(\sum_{i=0}^{n} \gamma_{ni}^{(j)} (Ki \log \omega + \sigma_i^{(j)})\right). \tag{3.15}$$

Letting $j \to \infty$ in (3.15) and invoking $\lim_{j\to\infty} \sigma_i^{(j)} = 0$ and recalling from Theorem 3.1 that $\lim_{j\to\infty} \gamma_{ni}^{(j)} = \tilde{\gamma}_{ni}$, we obtain the finite limit

$$\lim_{j\to\infty} \sum_{i=0}^{n} \dot{\gamma}_{ni}^{(j)} z^i = K \log\omega \left\{ \left(\sum_{i=0}^{n} \tilde{\gamma}_{ni} z^i \right) \left(\sum_{i=0}^{n} i\tilde{\gamma}_{ni} \right) - \left(\sum_{i=0}^{n} i\tilde{\gamma}_{ni} z^i \right) \right\}. \tag{3.16}$$

The following theorem summarizes the developments of this section up to this point.

Theorem 3.2. *Subject to the conditions concerning the t_l and $\varphi(t)$ that are given in* (3.1), (3.2), and (3.12), $\sum_{i=0}^{n} \dot{\gamma}_{ni}^{(j)} z^i$ *has a finite limit as $j \to \infty$ that is given by*

$$\lim_{j\to\infty} \sum_{i=0}^{n} \dot{\gamma}_{ni}^{(j)} z^i = K \log\omega [U_n(z)U_n'(1) - zU_n'(z)] \equiv W_n(z) \equiv \sum_{i=0}^{n} \tilde{\dot{\gamma}}_{ni} z^i, \tag{3.17}$$

where $U_n(z) = \prod_{i=1}^{n} \frac{(z-c_i)}{(1-c_i)}$ *and* $c_i = \omega^{\delta+i-1}$, $i = 1, 2, \ldots,$ *and* $U_n'(z) = (\mathrm{d}/\mathrm{d}z)U_n(z)$.

Theorem 3.2 is the key to the study of stability and convergence of column sequences $\{\dot{A}_n^{(j)}\}_{j=0}^{\infty}$ that follows.

3.1. Stability of column sequences $\{\dot{A}_n^{(j)}\}_{j=0}^{\infty}$

Theorem 3.3. *Under the conditions of Theorem* 3.2, *the sequences* $\{\dot{A}_n^{(j)}\}_{j=0}^{\infty}$ *are stable in the sense that* $\sup_j \Omega_n^{(j)} < \infty$.

Proof. The result follows from the facts that $\lim_{j\to\infty} \gamma_{ni}^{(j)} = \tilde{\gamma}_{ni}$ and $\lim_{j\to\infty} \dot{\gamma}_{ni}^{(j)} = \tilde{\dot{\gamma}}_{ni}$ for all n and i, which in turn follow from Theorems 3.1 and 3.2, respectively. □

3.2. Convergence of column sequences $\{\dot{A}_n^{(j)}\}_{j=0}^{\infty}$

Theorem 3.4. *Under the conditions of Theorem* 3.2 *and with the notation therein we have*

$$\dot{A}_n^{(j)} - \dot{A} = \mathrm{O}(\varphi(t_j) t_j^n \log t_j) \quad \text{as } j \to \infty. \tag{3.18}$$

A more refined result can be stated as follows: If $\beta_{n+\mu}$ is the first nonzero β_i with $i \geqslant n$ in (2.2) *and if $\dot{\beta}_{n+\nu}$ is the first nonzero $\dot{\beta}_i$ with $i \geqslant n$, then*

$$\begin{aligned} \dot{A}_n^{(j)} - \dot{A} &= \dot{\beta}_{n+\nu} U_n(c_{n+\nu+1}) \varphi(t_j) t_j^{n+\nu} [1 + \mathrm{o}(1)] \\ &\quad + K\beta_{n+\mu} U_n(c_{n+\mu+1}) \varphi(t_j) t_j^{n+\mu} \log t_j [1 + \mathrm{o}(1)] \quad \text{as } j \to \infty. \end{aligned} \tag{3.19}$$

Thus, when $\mu \leqslant \nu$ the second term dominates in $\dot{A}_n^{(j)} - \dot{A}$, while the first one does when $\mu > \nu$. In particular, if $\beta_n \neq 0$, we have

$$\dot{A}_n^{(j)} - \dot{A} \sim K\beta_n U_n(c_{n+1}) \varphi(t_j) t_j^n \log t_j \quad \text{as } j \to \infty. \tag{3.20}$$

Proof. We start with the fact that

$$A_n^{(j)} - A = \sum_{i=0}^{n} \gamma_{ni}^{(j)} [a(t_{j+i}) - A] = \sum_{i=0}^{n} \gamma_{ni}^{(j)} \varphi(t_{j+i}) B_n(t_{j+i}), \tag{3.21}$$

where

$$B_n(t) = B(t) - \sum_{i=0}^{n-1} \beta_i t^i \sim \sum_{i=n}^{\infty} \beta_i t^i \quad \text{as } t \to 0+. \tag{3.22}$$

Differentiating (3.21) with respect to ξ, we obtain

$$\dot{A}_n^{(j)} - \dot{A} = E_{n,1}^{(j)} + E_{n,2}^{(j)} + E_{n,3}^{(j)} \tag{3.23}$$

with

$$E_{n,1}^{(j)} = \sum_{i=0}^{n} \dot{\gamma}_{ni}^{(j)} \varphi(t_{j+i}) B_n(t_{j+i}),$$

$$E_{n,2}^{(j)} = \sum_{i=0}^{n} \gamma_{ni}^{(j)} \varphi(t_{j+i}) \dot{B}_n(t_{j+i}),$$

$$E_{n,3}^{(j)} = \sum_{i=0}^{n} \gamma_{ni}^{(j)} \dot{\varphi}(t_{j+i}) B_n(t_{j+i}). \tag{3.24}$$

By the conditions in (3.1) and (3.2), and by (3.14) that follows from the condition in (3.12), it can be shown that

$$t_{j+i} \sim t_j \omega^i, \quad \varphi(t_{j+i}) \sim \omega^{i\delta} \varphi(t_j), \quad \text{and} \quad \dot{\varphi}(t_{j+i}) \sim K\omega^{i\delta} \varphi(t_j) \log t_j \quad \text{as } j \to \infty. \tag{3.25}$$

Substituting these in (3.24), noting that $B_n(t) \sim \beta_{n+\mu} t^{n+\mu}$ and $\dot{B}_n(t) \sim \dot{\beta}_{n+v} t^{n+v}$ as $t \to 0+$, and recalling (3.4) and (3.17), we obtain

$$E_{n,1}^{(j)} = \beta_{n+\mu} W_n(c_{n+\mu+1}) \varphi(t_j) t_j^{n+\mu} + \mathrm{o}(\varphi(t_j) t_j^{n+\mu}) \quad \text{as } j \to \infty,$$

$$E_{n,2}^{(j)} \sim \dot{\beta}_{n+v} U_n(c_{n+v+1}) \varphi(t_j) t_j^{n+v} \quad \text{as } j \to \infty,$$

$$E_{n,3}^{(j)} \sim K\beta_{n+\mu} U_n(c_{n+\mu+1}) \varphi(t_j) t_j^{n+\mu} \log t_j \quad \text{as } j \to \infty, \tag{3.26}$$

with $W_n(z)$ as defined in (3.17). Note that we have written the result for $E_{n,1}^{(j)}$ differently than for $E_{n,2}^{(j)}$ and $E_{n,3}^{(j)}$ since we cannot be sure that $W_n(c_{n+\mu+1}) \neq 0$. The asymptotic equalities for $E_{n,2}^{(j)}$ and $E_{n,3}^{(j)}$, however, are valid as $U_n(c_i) \neq 0$ for all $i \geqslant n+1$. The result now follows by substituting (3.26) in (3.23) and observing also that $E_{n,1}^{(j)} = \mathrm{o}(E_{n,3}^{(j)})$ as $j \to \infty$, so that either $E_{n,2}^{(j)}$ or $E_{n,3}^{(j)}$ determines the asymptotic nature of $\dot{A}_n^{(j)} - \dot{A}$. We leave the details to the reader. □

Remark. Comparing (3.19) pertaining to $\dot{A}_n^{(j)} - \dot{A}$ with (3.3) pertaining to $A_n^{(j)} - A$, we realize that, subject to the additional assumption in (3.12), the two behave practically the same way asymptotically. In addition, their computational costs are generally similar. (In many problems of interest $A(y)$

and $\dot{A}(y)$ can be computed simultaneously, the total cost of this being almost the same as that of computing $A(y)$ only or $\dot{A}(y)$ only. An immediate example is that of numerical integration discussed in Section 1.) In contrast, the convergence of $\{B_n^{(j)}\}_{j=0}^{\infty}$ obtained by applying GREP$^{(2)}$ directly to $\dot{A}(y) \equiv \dot{a}(t)$ (recall (2.18) and (2.19)), is inferior to that of $\{A_n^{(j)}\}_{j=0}^{\infty}$. This can be shown rigorously for the case in which $\dot{\phi}(y) \equiv \dot{\varphi}(t) = K\varphi(t)(\log t + \text{constant})$ exactly. In this case the asymptotic expansion in (2.18) assumes the form $\dot{a}(t) \sim \dot{A} + \sum_{k=1}^{\infty} \varphi(t)(\alpha_{k0} + \alpha_{k1}\log t)t^k$ as $t \to 0+$. Therefore, under the additional condition that $\lim_{m\to\infty} \varepsilon_m \log t_m = 0$, where ε_m is as defined following (3.12), Theorem 2.2 of [11] applies and we have

$$B_{2m}^{(j)} - B = O(\varphi(t_j)t_j^m \log t_j) \quad \text{as } j \to \infty. \tag{3.27}$$

Now the computational costs of $\dot{A}_{2m}^{(j)}$ and $B_{2m}^{(j)}$ are similar, but $\{\dot{A}_{2m}^{(j)}\}_{j=0}^{\infty}$ converges to $\dot{A}$ much faster than $\{B_{2m}^{(j)}\}_{j=0}^{\infty}$. Again, we have verified the superiority of our new approach to the direct approach, at least with respect to column sequences.

We would like to add that the theory of [11] applies to the more general class of functions $A(y)$ that have asymptotic expansions of the form $A(y) \sim A + \sum_{k=1}^{\infty} \psi_k(y)(\sum_{i=0}^{q_k} \alpha_{ki}(\log y)^i)$ as $y \to 0+$, where q_k are arbitrary nonnegative integers.

4. Application to infinite series via the $d^{(1)}$-transformation: the $(d/d\xi)d^{(1)}$-transformation

4.1. General usage

Let $\{S_m\}$ be the sequence of partial sums of the infinite series $\sum_{k=1}^{\infty} v_k$, namely,

$$S_m = \sum_{k=1}^{m} v_k, \quad m = 1, 2, \ldots . \tag{4.1}$$

Assume that

$$v_m \sim \sum_{i=0}^{\infty} \theta_i m^{\rho - i} \quad \text{as } m \to \infty,\ \theta_0 \neq 0,\ \rho + 1 \neq 0, 1, 2, \ldots . \tag{4.2}$$

As is known, $\lim_{m\to\infty} S_m$ exists and is finite if and only if $\Re\rho + 1 < 0$. When $\Re\rho + 1 \geqslant 0$ but $\rho + 1 \neq 0, 1, 2, \ldots$, $\{S_m\}$ diverges but has a well defined and useful antilimit as has been shown in Theorem 4.1 of [10]. For all ρ in (4.2) this theorem reads as follows:

Theorem 4.1. *With S_m as in* (4.1) *and* (4.2), *we have*

$$S_m \sim S + m v_m \sum_{i=0}^{\infty} \beta_i m^{-i} \quad \text{as } m \to \infty,\ \beta_0 \neq 0. \tag{4.3}$$

Here $S = \lim_{m\to\infty} S_m$ when $\Re\rho + 1 < 0$, and S is the antilimit of $\{S_m\}$ otherwise.

The part of Theorem 4.1 concerning convergent sequences $\{S_m\}$ is already contained in Theorem 2 of [6].

From Theorem 4.1 it is clear that GREP$^{(1)}$ can be applied to the sequence $\{S_m\}$ by drawing the analogy $a(t) \leftrightarrow S_m$, $t \leftrightarrow m^{-1}$, $\varphi(t) \leftrightarrow mv_m$, and $A \leftrightarrow S$, and by picking $t_l = 1/R_l$ for some positive integers R_l, $1 \leqslant R_0 < R_1 < R_2 < \cdots$, and the W-algorithm can be used to implement it. This GREP$^{(1)}$ is simply the Levin–Sidi $d^{(1)}$-transformation, and we denote its $A_n^{(j)}$ by $S_n^{(j)}$.

As already explained in [10,4], for the type of sequences considered here we should pick the R_l such that $\{R_l\}$ increases exponentially to ensure the best stability and convergence properties in the $S_n^{(j)}$. Exponential increase in the R_l can be achieved by picking them, for example, as in

$$R_0 = 1 \quad \text{and} \quad R_{l+1} = \lfloor \sigma R_l \rfloor + 1, \;\; l = 0, 1, \ldots, \quad \text{for some } \sigma > 1. \tag{4.4}$$

(With $\sigma = 1$ we have $R_l = l + 1$, $l = 0, 1, \ldots,$ for which the $d^{(1)}$-transformation becomes the Levin [5] u-transformation.) This gives $R_l = \mathrm{O}(\sigma^l)$ as $l \to \infty$. Needless to say, σ should not be picked too far from 1 to avoid too quick a growth in the R_l. We have found that σ between 1.1 and 1.5 is sufficient for most purposes. Since $t_l = 1/R_l$, (4.4) implies that

$$\frac{t_l}{\sigma + t_l} \leqslant t_{l+1} < \frac{t_l}{\sigma}, \quad l = 0, 1, \ldots \tag{4.5}$$

as a result of which $\{t_l\}$ satisfies (3.1) with $\omega = 1/\sigma \in (0, 1)$. Therefore, Theorem 3.1 applies to the approximations $S_m^{(j)}$ to S obtained via the $d^{(1)}$-transformation, as has been shown in [10]. Clearly, $\delta = -\rho - 1$ in (3.2) and (3.6) for this case.

If, in addition, v_m and S are differentiable functions of a parameter ξ, $\dot{S}$ is the limit or antilimit of $\{\dot{S}_m\}$, and

$$\dot{v}_m = v_m[K' \log m + L' + \mathrm{o}(1)] \quad \text{as } m \to \infty, \quad \text{for some constants } K' \neq 0 \text{ and } L' \tag{4.6}$$

and the asymptotic expansion in (4.3) can be differentiated with respect to ξ term by term, then Theorems 3.2–3.4 apply to $\{\dot{S}_n^{(j)}\}_{j=0}^{\infty}$ without any modifications. We shall denote this method that produces the $\dot{S}_n^{(j)}$ the $(\mathrm{d}/\mathrm{d}\xi)d^{(1)}$-transformation for short. The rate of convergence of the $\dot{S}_n^{(j)}$ to $\dot{S}$ is almost identical to the rate of convergence of the $S_n^{(j)}$ to S as we have observed in many numerical examples, and as we have proved in Theorem 3.4 for the column sequences.

To summarize the relevant convergence results for the $d^{(1)}$- and $(\mathrm{d}/\mathrm{d}\xi)d^{(1)}$-transformations as these are applied to $\{S_m\}$ and $\{\dot{S}_m\}$ above, we have from Theorems 3.1 and 3.4

$$S_n^{(j)} - S = \mathrm{O}(v_{R_j} R_j^{-n+1}) = \mathrm{O}(\sigma^{(\rho+1-n)j}) \quad \text{as } j \to \infty,$$

$$\dot{S}_n^{(j)} - \dot{S} = \mathrm{O}(v_{R_j} R_j^{-n+1} \log R_j) = \mathrm{O}(j\sigma^{(\rho+1-n)j}) \quad \text{as } j \to \infty. \tag{4.7}$$

Of course, these results are not optimal. Optimal results follow from (3.3) and (3.19), and we leave them to the reader. The results for $\Gamma_n^{(j)}$ and $\Omega_n^{(j)}$ that pertain to stability can be obtained from Theorems 3.1–3.3.

For the sake of completeness we note that the $(\mathrm{d}/\mathrm{d}\xi)W$-algorithm takes $t_j = 1/R_j$, $a(t_j) = \sum_{k=1}^{R_j} v_k$, $\dot{a}(t_j) = \sum_{k=1}^{R_j} \dot{v}_k$, $\varphi(t_j) = R_j v_{R_j}$, and $\dot{\varphi}(t_j) = R_j \dot{v}_{R_j}$ as input for this problem.

It is worth mentioning that we can also compute $\dot{S}$ by applying the $d^{(2)}$-transformation directly to $\{\dot{S}_m\}$. The $d^{(2)}$-transformation is a GREP$^{(2)}$. As we mentioned earlier, this is less effective than the application of the $(\mathrm{d}/\mathrm{d}\xi)d^{(1)}$-transformation to $\{S_m\}$. We shall see this also through numerical examples in the next section.

4.2. A special application

We next turn to an interesting application of the $(\mathrm{d}/\mathrm{d}\xi)d^{(1)}$-transformation to the summation of a class of infinite series $\sum_{k=1}^{\infty} \tilde{v}_k$, where $\tilde{v}_m$ has the form

$$\tilde{v}_m = [\log \mu(m)]v_m, \quad \mu(m) \sim \sum_{i=0}^{\infty} \mu_i m^{\alpha-i} \quad \text{as } m \to \infty,\ \mu_0 \neq 0 \text{ and } \alpha \neq 0, \tag{4.8}$$

with v_m as in (4.2). (When $\alpha = 0$ the $d^{(1)}$-transformation is very effective on the series $\sum_{k=1}^{\infty} \tilde{v}_k$.) To this end first let us consider the infinite series $\sum_{k=1}^{\infty} u_k(\xi)$, where

$$u_m(\xi) = v_m[\mu(m)]^{\xi}, \quad m = 1, 2, \ldots . \tag{4.9}$$

(Here v_m and $\mu(m)$ do not depend on ξ). Now it can be shown that $[\mu(m)]^{\xi} \sim \sum_{i=0}^{\infty} \mu_i' m^{\varepsilon-i}$ as $m \to \infty$, where $\mu_0' = \mu_0^{\xi} \neq 0$ and $\varepsilon = \alpha\xi$. Consequently, $u_m(\xi) \sim \sum_{i=0}^{\infty} \theta_i' m^{\rho'-i}$ as $m \to \infty$, where $\theta_0' = \theta_0 \mu^{\xi} \neq 0$ and $\rho' = \rho + \alpha\xi$, so that $u_m(\xi)$ is of the form described in (4.1) for all ξ. That is to say, the $d^{(1)}$-transformation can be applied to sum $\sum_{k=1}^{\infty} u_k(\xi)$ for any ξ. Next, $\dot{u}_m(\xi) = u_m(\xi) \log \mu(m) \sim u_m(\xi)[\alpha \log m + \log \mu_0 + o(1)]$ as $m \to \infty$, cf. (4.6). Therefore, the $(\mathrm{d}/\mathrm{d}\xi)d^{(1)}$-transformation can be used for summing $\sum_{k=1}^{\infty} \dot{u}_k(\xi)$ for any ξ. Finally, $u_m(0) = v_m$ and $\dot{u}_m(0) = \tilde{v}_m$, and hence the $(\mathrm{d}/\mathrm{d}\xi)d^{(1)}$-transformation can be used for summing $\sum_{k=1}^{\infty} \tilde{v}_k$ in particular. This can be done by setting $t_j = 1/R_j$, $a(t_j) = \sum_{k=1}^{R_j} v_k$, $\dot{a}(t_j) = \sum_{k=1}^{R_j} \tilde{v}_k$, $\varphi(t_j) = R_j v_{R_j}$, and $\dot{\varphi}(t_j) = R_j \tilde{v}_{R_j}$ in the $(\mathrm{d}/\mathrm{d}\xi)W$-algorithm.

5. Numerical examples

In this section we wish to demonstrate numerically the effectiveness of $(\mathrm{d}/\mathrm{d}\xi)\mathrm{GREP}^{(1)}$ via the $(\mathrm{d}/\mathrm{d}\xi)d^{(1)}$-transformation on some infinite series, convergent or divergent. We will do this with two examples. The first one of these examples has already been treated in [14] within the framework of the Richardson extrapolation process.

Example 5.1. Consider the series $\sum_{k=1}^{\infty} k^{-\xi-1}$ that converges for $\Re\xi > 0$ and defines the Riemann zeta function $\zeta(\xi+1)$. As is known, $\zeta(z)$ can be continued analytically to the entire complex plane except $z = 1$, where it has a simple pole. As the term $v_m = m^{-\xi-1}$ is of the form described in the previous section, Theorem 4.1 applies to $S_m = \sum_{k=1}^{m} k^{-\xi-1}$ with $S = \zeta(\xi+1)$ and $\delta = \xi$, whether $\lim_{m\to\infty} S_m$ exists or not. Furthermore, the asymptotic expansion of $\dot{S}_m = \sum_{k=1}^{m} (-\log k) k^{-\xi-1}$ can be obtained by term-by-term differentiation of the expansion in (4.3), as has already been mentioned in [14]. This implies that the $(\mathrm{d}/\mathrm{d}\xi)d^{(1)}$-transformation can be applied to the computation of $\dot{S} = \zeta'(\xi+1)$, and Theorems 3.2–3.4 are valid with $\delta = \xi$. In particular, (4.7) is valid with $\rho = -\xi - 1$ there.

We applied the $(\mathrm{d}/\mathrm{d}\xi)d^{(1)}$-transformation to this problem to compute $\dot{S} = \zeta'(\xi+1)$. We picked the integers R_l as in (4.4) with $\sigma = 1.2$ there. We considered the two cases (i) $\xi = 1$ and (ii) $\xi = -0.5$. Note that in case (i) both $\lim_{m\to\infty} S_m$ and $\lim_{m\to\infty} \dot{S}_m$ exist and are $S = \zeta(2)$ and $\dot{S} = \zeta'(2)$, respectively, while in case (ii) these limits do not exist and $S = \zeta(0.5)$ and $\dot{S} = \zeta'(0.5)$ are the corresponding antilimits. We also applied the $d^{(2)}$-transformation directly to $\{\dot{S}_m\}$ with the same R_l's, the resulting approximations being denoted $B_n^{(j)}$, as in (2.19). The numerical results are shown in Tables 1–3.

Table 1
Numerical results on Process I for $\zeta(z)$ in Example 5.1, where $\zeta(z)$ is the Riemann zeta function, with $z=2$. The $d^{(1)}$- and $(\mathrm{d}/\mathrm{d}\xi)d^{(1)}$-transformations on $\{S_m\}$ and $\{\dot{S}_m\}$ and the $d^{(2)}$-transformation on $\{\dot{S}_m\}$ are implemented with $\sigma=1.2$ in (4.4). Here $P_n^{(j)}=|S_n^{(j+1)}-S|/|S_n^{(j)}-S|$, $Q_n^{(j)}=|\dot{S}_n^{(j+1)}-\dot{S}|/|\dot{S}_n^{(j)}-\dot{S}|$, and $Z_n^{(j)}=|B_n^{(j+1)}-\dot{S}|/|B_n^{(j)}-\dot{S}|$, where $S_n^{(j)}$, $\dot{S}_n^{(j)}$, and $B_n^{(j)}$ are the approximations obtained from the $d^{(1)}$-, $(\mathrm{d}/\mathrm{d}\xi)d^{(1)}$-, and $d^{(2)}$-transformations, respectively. All six columns are tending to $\sigma^{-7}=0.279\ldots$

j	$P_5^{(j)}$	$Q_5^{(j)}$	$Z_{10}^{(j)}$	$P_6^{(j)}$	$Q_6^{(j)}$	$Z_{12}^{(j)}$
0	$1.53D-01$	$1.62D-01$	$3.18D-01$	$1.09D-03$	$2.25D-02$	$3.08D-02$
2	$1.94D-01$	$2.09D-01$	$2.01D-01$	$3.23D-01$	$3.19D-01$	$1.06D-01$
4	$1.97D-01$	$2.10D-01$	$1.58D-01$	$2.30D-01$	$2.41D-01$	$1.58D-02$
6	$2.33D-01$	$2.46D-01$	$2.02D-01$	$2.44D-01$	$2.56D-01$	$4.27D-01$
8	$2.45D-01$	$2.57D-01$	$4.95D-01$	$2.51D-01$	$2.63D-01$	$2.93D-01$
10	$2.50D-01$	$2.61D-01$	$3.56D-01$	$2.56D-01$	$2.66D-01$	$2.65D-01$
12	$2.65D-01$	$2.75D-01$	$3.22D-01$	$2.66D-01$	$2.75D-01$	$2.58D-01$
14	$2.67D-01$	$2.76D-01$	$3.07D-01$	$2.68D-01$	$2.77D-01$	$2.60D-01$
16	$2.70D-01$	$2.79D-01$	$3.03D-01$	$2.71D-01$	$2.79D-01$	$2.69D-01$
18	$2.70D-01$	$2.79D-01$	$2.99D-01$	$2.71D-01$	$2.79D-01$	$2.78D-01$
20	$2.74D-01$	$2.82D-01$	$2.97D-01$	$2.74D-01$	$2.82D-01$	$2.85D-01$

Table 2
Numerical results on Process II for $\zeta(z)$ in Example 5.1, where $\zeta(z)$ is the Riemann zeta function, with $z=2$. The $d^{(1)}$- and $(\mathrm{d}/\mathrm{d}\xi)d^{(1)}$-transformations on $\{S_m\}$ and $\{\dot{S}_m\}$ and the $d^{(2)}$-transformation on $\{\dot{S}_m\}$ are implemented with $\sigma=1.2$ in (4.4). Here $S_n^{(j)}$, $\dot{S}_n^{(j)}$, and $B_n^{(j)}$ are the approximations obtained from the $d^{(1)}$-, $(\mathrm{d}/\mathrm{d}\xi)d^{(1)}$-, and $d^{(2)}$-transformations, respectively. (The infinite series converge.)

n	R_n	$\lvert S_{R_n}-S\rvert$	$\lvert S_n^{(0)}-S\rvert$	$\lvert \dot{S}_{R_n}-\dot{S}\rvert$	$\lvert \dot{S}_n^{(0)}-\dot{S}\rvert$	$\lvert B_n^{(0)}-\dot{S}\rvert$
0	1	$6.45D-01$	$1.64D+00$	$9.38D-01$	$9.38D-01$	$9.38D-01$
2	3	$2.84D-01$	$1.99D-02$	$6.42D-01$	$3.67D-02$	$4.56D-01$
4	5	$1.81D-01$	$3.12D-05$	$4.91D-01$	$1.07D-04$	$3.28D-03$
6	9	$1.05D-01$	$7.08D-07$	$3.42D-01$	$1.56D-06$	$6.19D-04$
8	14	$6.89D-02$	$8.18D-09$	$2.53D-01$	$2.35D-08$	$3.26D-05$
10	21	$4.65D-02$	$3.71D-11$	$1.89D-01$	$1.25D-10$	$8.26D-07$
12	32	$3.08D-02$	$6.95D-14$	$1.38D-01$	$2.70D-13$	$5.11D-07$
14	47	$2.11D-02$	$2.55D-17$	$1.02D-01$	$1.44D-16$	$4.17D-09$
16	69	$1.44D-02$	$8.28D-20$	$7.54D-02$	$3.03D-19$	$1.60D-11$
18	100	$9.95D-03$	$1.14D-22$	$5.58D-02$	$4.90D-22$	$4.32D-13$
20	146	$6.83D-03$	$5.75D-26$	$4.09D-02$	$2.72D-25$	$2.14D-16$
22	212	$4.71D-03$	$1.52D-29$	$2.99D-02$	$4.53D-29$	$4.53D-17$
24	307	$3.25D-03$	$2.44D-30$	$2.19D-02$	$3.52D-29$	$1.97D-19$

Table 1 shows the validity of the theory for Process I given in Sections 2–4 very clearly. The results of this table that have been computed with $\xi=1$ can be understood as follows:

Table 3
Numerical results on Process II for $\zeta(z)$ in Example 5.1, where $\zeta(z)$ is the Riemann zeta function, with $z = 0.5$. The $d^{(1)}$- and $(\mathrm{d}/\mathrm{d}\xi)d^{(1)}$-transformations on $\{S_m\}$ and $\{\dot{S}_m\}$ and the $d^{(2)}$-transformation on $\{\dot{S}_m\}$ are implemented with $\sigma = 1.2$ in (4.4). Here $S_n^{(j)}$, $\dot{S}_n^{(j)}$, and $B_n^{(j)}$ are the approximations obtained from the $d^{(1)}$-, $(\mathrm{d}/\mathrm{d}\xi)d^{(1)}$-, and $d^{(2)}$-transformations, respectively. (The infinite series diverge.)

n	R_n	$\lvert S_{R_n} - S\rvert$	$\lvert S_n^{(0)} - S\rvert$	$\lvert \dot{S}_{R_n} - \dot{S}\rvert$	$\lvert \dot{S}_n^{(0)} - \dot{S}\rvert$	$\lvert B_n^{(0)} - \dot{S}\rvert$
0	1	2.46D + 00	1.46D + 00	3.92D + 00	3.92D + 00	3.92D + 00
2	3	3.74D + 00	1.28D − 01	2.80D + 00	1.65D − 01	5.33D − 01
4	5	4.69D + 00	1.01D − 03	1.39D + 00	4.64D − 04	9.62D + 00
6	9	6.17D + 00	4.71D − 06	1.55D + 00	9.73D − 06	2.50D + 00
8	14	7.62D + 00	2.32D − 07	5.13D + 00	8.13D − 08	1.05D + 00
10	21	9.27D + 00	2.24D − 09	9.90D + 00	4.19D − 10	2.01D − 01
12	32	1.14D + 01	8.85D − 12	1.69D + 01	5.88D − 12	4.02D − 02
14	47	1.38D + 01	1.33D − 14	2.56D + 01	1.71D − 14	1.79D − 03
16	69	1.67D + 01	2.51D − 18	3.74D + 01	8.66D − 18	1.41D − 05
18	100	2.00D + 01	2.74D − 20	5.23D + 01	2.88D − 20	1.50D − 07
20	146	2.42D + 01	2.76D − 23	7.23D + 01	4.34D − 23	1.91D − 09
22	212	2.92D + 01	6.72D − 27	9.79D + 01	3.13D − 26	2.21D − 11
24	307	3.51D + 01	6.38D − 27	1.31D + 02	1.54D − 26	1.41D − 13

Since

$$S_{n-1} \sim \zeta(\xi+1) - \frac{n^{\xi}}{\xi}\sum_{i=0}^{\infty}\binom{-\xi}{i}B_i n^{-i} \quad \text{as } n\to\infty,$$

and since $B_0 = 1,\ B_1 = -\frac{1}{2}$, while $B_{2i} \neq 0,\ B_{2i+1} = 0,\ i = 1,2,\ldots,$ we have that with the exception of $\beta_{2i+1},\ i = 1,2,\ldots,$ all the other β_i are nonzero, and that exactly the same applies to the $\dot{\beta}_i$. (Here B_i are the Bernoulli numbers and should not be confused with $B_n^{(j)}$.) Consequently, (3.19) of Theorem 3.4 holds with $\mu = v$ there. Thus, whether $\lim_{m\to\infty} S_m$ exists or not, as $j\to\infty$, $|S_n^{(j+1)} - S|/|S_n^{(j)} - S|$ is $O(\sigma^{-1})$ for $n = 0$, $O(\sigma^{-2})$ for $n = 1$, $O(\sigma^{-3})$ for $n = 2$, and $O(\sigma^{-(2i+1)})$ for both $n = 2i-1$ and $n = 2i$, with $i = 2,3,\ldots$. Similarly, whether $\lim_{m\to\infty} \dot{S}_m$ exists or not, as $j\to\infty$, $|\dot{S}_n^{(j+1)} - \dot{S}|/|\dot{S}_n^{(j)} - \dot{S}|$ is $O(\sigma^{-1})$ for $n = 0$, $O(\sigma^{-2})$ for $n = 1$, $O(\sigma^{-3})$ for $n = 2$, and $O(\sigma^{-(2i+1)})$ for both $n = 2i-1$ and $2i$, with $i = 2,3,\ldots$.

As for the approximations $B_n^{(j)}$ to $\dot{S}$ obtained from the $d^{(2)}$-transformation on $\{\dot{S}_m\}$, Theorem 2.2 in [11] implies that, as $j\to\infty$, $|B_n^{(j+1)} - \dot{S}|/|B_n^{(j)} - \dot{S}|$ is $O(\sigma^{-1})$ for $n = 0$, $O(\sigma^{-2})$ for $n = 2$, $O(\sigma^{-3})$ for $n = 4$, and $O(\sigma^{-(2i+1)})$ for both $n = 2(2i-1)$ and $n = 4i$, with $i = 2,3,\ldots$.

The numerical results of Tables 2 and 3 pertain to Process II and show clearly that our approach to the computation of derivatives of limits is a very effective one.

Example 5.2. Consider the summation of the infinite series $\sum_{k=0}^{\infty}\dot{v}_k$, where $v_m = b_m(\xi)_m/m!$ and $(\xi)_m = \prod_{i=0}^{m-1}(\xi+i)$, and $b_m \sim \sum_{i=0}^{\infty}\kappa_i m^{\eta-i}$ as $m\to\infty$. By the fact that $(\xi)_m = \Gamma(\xi+m)/\Gamma(\xi)$ and by formula 6.1.47 in [1] we have that $(\xi)_m/m! \sim \sum_{i=0}^{\infty}\lambda_i m^{\xi-i-1}$ as $m\to\infty$. Consequently, $v_m \sim \sum_{i=0}^{\infty}\theta_i m^{\eta+\xi-1-i}$ as $m\to\infty$, so that the $d^{(1)}$-transformation can be applied successfully to sum

Table 4
Numerical results on Process II for $F(\xi, \frac{1}{2}; \frac{3}{2}; 1)$ in Example 5.2, where $F(a, b; c; z)$is the Gauss hypergeometric function, with $\xi = 0.5$. The $d^{(1)}$- and $(d/d\xi)d^{(1)}$-transformations on $\{S_m\}$ and $\{\dot{S}_m\}$ and the $d^{(2)}$-transformation on $\{\dot{S}_m\}$ are implemented with $\sigma = 1.2$ in (4.4). Here $S_n^{(j)}$, $\dot{S}_n^{(j)}$, and $B_n^{(j)}$are the approximations obtained from the $d^{(1)}$-, $(d/d\xi)d^{(1)}$-, and $d^{(2)}$-transformations, respectively. (The infinite series converge.)

n	R_n	$\lvert S_{R_n} - S\rvert$	$\lvert S_n^{(0)} - S\rvert$	$\lvert \dot{S}_{R_n} - \dot{S}\rvert$	$\lvert \dot{S}_n^{(0)} - \dot{S}\rvert$	$\lvert B_n^{(0)} - \dot{S}\rvert$
0	1	5.71D − 01	1.57D + 00	2.18D + 00	2.18D + 00	2.18D + 00
2	3	3.29D − 01	4.70D − 02	1.64D + 00	2.18D − 01	6.79D − 01
4	5	2.54D − 01	4.06D − 05	1.41D + 00	4.06D − 04	1.51D − 01
6	9	1.89D − 01	1.69D − 06	1.16D + 00	1.22D − 05	4.59D − 02
8	14	1.51D − 01	1.95D − 08	9.96D − 01	1.39D − 07	3.76D − 03
10	21	1.23D − 01	1.11D − 10	8.63D − 01	7.94D − 10	1.47D − 04
12	32	9.99D − 02	3.11D − 13	7.41D − 01	2.20D − 12	4.42D − 06
14	47	8.24D − 02	3.99D − 16	6.43D − 01	2.61D − 15	9.14D − 08
16	69	6.80D − 02	1.20D − 19	5.57D − 01	1.41D − 19	1.26D − 09
18	100	5.64D − 02	2.04D − 22	4.84D − 01	2.38D − 21	1.57D − 11
20	146	4.67D − 02	2.03D − 25	4.18D − 01	2.03D − 24	2.21D − 13
22	212	3.88D − 02	6.77D − 29	3.61D − 01	7.81D − 28	1.86D − 15
24	307	3.22D − 02	2.41D − 29	3.12D − 01	1.51D − 28	1.13D − 18

$\sum_{k=0}^{\infty} v_k$, as described in the previous section. Now $\dot{v}_m = v_m[\sum_{i=0}^{m-1} 1/(\xi + i)]$, and $\sum_{i=0}^{m-1} 1/(\xi + i) \sim \log m + \sum_{k=0}^{\infty} e_i m^{-i}$ as $m \to \infty$. Therefore, we can apply the $(d/d\xi)d^{(1)}$-transformation to sum $\sum_{k=0}^{\infty} \dot{v}_k$ provided that the asymptotic expansion of $\dot{S}_m$ can be obtained by term-by-term differentiation by Theorem 4.1. (We have not shown that this last condition is satisfied).

We have applied the $(d/d\xi)d^{(1)}$-transformation with $b_m = 1/(2m + 1)$. With this b_m the series $\sum_{k=0}^{\infty} v_k$ and $\sum_{k=0}^{\infty} \dot{v}_k$ both converge. Actually we have $\sum_{k=0}^{\infty} v_k = F(\xi, \frac{1}{2}; \frac{3}{2}; 1)$. By formula 15.1.20 in [1], $\sum_{k=0}^{\infty} v_k = (\sqrt{\pi}/2)(\Gamma(1-\xi)/\Gamma(3/2-\xi)) = S$. Differentiating both sides with respect to ξ, we obtain $\sum_{k=0}^{\infty} \dot{v}_k = (\sqrt{\pi}/2)(\Gamma(1 - \xi)/\Gamma(3/2 - \xi))\{\psi(\frac{3}{2} - \xi) - \psi(1 - \xi)\} = \dot{S}$, where $\psi(z) = (d/dz)\Gamma(z)/\Gamma(z)$. Letting now $\xi = \frac{1}{2}$ throughout, we get $S = \pi/2$ and $\dot{S} = \pi \log 2$, the latter following from formulas 6.3.2 and 6.3.3 in [1].

In our computations we picked the R_l as in the first example. We also applied the $d^{(2)}$-transformation directly to $\{\dot{S}_m\}$ with the same R_l's.

Table 4 contains numerical results pertaining to Process II.

Acknowledgements

This research was supported in part by the Fund for the Promotion of Research at the Technion.

References

[1] M. Abramowitz, I.A. Stegun, Handbook of Mathematical Functions, Vol. 55 of National Bureau of Standards, Applied Mathematics Series, Government Printing Office, Washington, DC, 1964.

[2] R. Bulirsch, J. Stoer, Fehlerabschätzungen und extrapolation mit rationalen Funktionen bei Verfahren vom Richardson-Typus, Numer. Math. 6 (1964) 413–427.
[3] P.J. Davis, P. Rabinowitz, Methods of Numerical Integration, Academic Press, New York, 1984. 2nd Edition.
[4] W.F. Ford, A. Sidi, An algorithm for a generalization of the Richardson extrapolation process, SIAM J. Numer. Anal. 24 (1987) 1212–1232.
[5] D. Levin, Development of nonlinear transformations for improving convergence of sequences, Int. J. Comput. Math. Series B 3 (1973) 371–388.
[6] D. Levin, A. Sidi, Two new classes of nonlinear transformations for accelerating the convergence of infinite integrals and series, Appl. Math. Comput. 9 (1981) 175–215.
[7] A. Sidi, Some properties of a generalization of the Richardson extrapolation process, J. Inst. Math. Appl. 24 (1979) 327–346.
[8] A. Sidi, An algorithm for a special case of a generalization of the Richardson extrapolation process, Numer. Math. 38 (1982) 299–307.
[9] A. Sidi, Generalizations of Richardson extrapolation with applications to numerical integration, in: H. Brass, G. Hämmerlin (Eds.), Numerical Integration III, ISNM, Vol. 85, Birkhäuser, Basel, Switzerland, 1988, pp. 237–250.
[10] A. Sidi, Convergence analysis for a generalized Richardson extrapolation process with an application to the $d^{(1)}$-transformation on convergent and divergent logarithmic sequences, Math. Comput. 64 (1995) 1627–1657.
[11] A. Sidi, Further results on convergence and stability of a generalization of the Richardson extrapolation process, BIT 36 (1996) 143–157.
[12] A. Sidi, A complete convergence and stability theory for a generalized Richardson extrapolation process, SIAM J. Numer. Anal. 34 (1997) 1761–1778.
[13] A. Sidi, Further convergence and stability results for the generalized Richardson extrapolation process $\mathrm{GREP}^{(1)}$ with an application to the $D^{(1)}$-transformation for infinite integrals, J. Comput. Appl. Math. 112 (1999) 269–290.
[14] A. Sidi, Extrapolation methods and derivatives of limits of sequences, Math. Comput. 69 (2000) 305–323.
[15] J. Stoer, R. Bulirsch, Introduction to Numerical Analysis, Springer, New York, 1980.