NONLINEAR SEQUENCE TRANSFORMATIONS FOR THE ACCELERATION OF CONVERGENCE AND THE SUMMATION OF DIVERGENT SERIES

Ernst Joachim WENIGER

Institut für Physikalische und Theoretische Chemie, Universität Regensburg, D-8400 Regensburg, Federal Republic of Germany

1989

NORTH-HOLLAND – AMSTERDAM

The manuscript for this Computer Physics Reports issue was submitted as a TeX compuscript and converted to an STI file from which the typesetting was done.

Contents

Computer Physics Reports 10 (1989) 189–371
North-Holland, Amsterdam

NONLINEAR SEQUENCE TRANSFORMATIONS FOR THE ACCELERATION OF CONVERGENCE AND THE SUMMATION OF DIVERGENT SERIES

Ernst Joachim WENIGER

Institut für Physikalische und Theoretische Chemie, Universität Regensburg, D-8400 Regensburg, Federal Republic of Germany

Received 17 August 1989

Slowly convergent series and sequences as well as divergent series occur quite frequently in the mathematical treatment of scientific problems. In this report, a large number of mainly nonlinear sequence transformations for the acceleration of convergence and the summation of divergent series are discussed. Some of the sequence transformations of this report as for instance Wynn's ϵ algorithm or Levin's sequence transformation are well established in the literature on convergence acceleration, but the majority of them is new. Efficient algorithms for the evaluation of these transformations are derived. The theoretical properties of the sequence transformations in convergence acceleration and summation processes are analyzed. Finally, the performance of the sequence transformations of this report are tested by applying them to certain slowly convergent and divergent series, which are hopefully realistic models for a large part of the slowly convergent or divergent series that can occur in scientific problems and in applied mathematics.

1. Introduction

1.1. Infinite series and their evaluation

Infinite series are ubiquitous in the mathematical analysis of scientific problems. They naturally appear in the evaluation of integrals, in the solutions of differential and integral equations, or as Fourier series. They are also used for both the definition and the evaluation of many of the special functions of mathematical physics. The conventional approach for the evaluation of an infinite series consists in computing a finite sequence of partial sums

$$s_n = \sum_{k=0}^{n} a_k \tag{1.1-1}$$

by adding up one term after the other. Then, the magnitude of the truncation error is estimated. If the sequence of partial sums $s_0, \ldots, s_n$ has not converged yet to the desired accuracy, additional terms must be added until convergence has finally been achieved. With this approach it is at least in principle possible to determine the value of an infinite series as accurately as one likes provided that one is able to compute a sufficiently large number of terms accurately enough to overcome eventual numerical instabilities.

However, in many scientific problems one will only be able to compute a relatively small number of terms. In addition, particularly the series terms with higher summation indices are often affected by serious inaccuracies which may lead to a catastrophic accumulation of round-off errors.

Consequently, if an infinite series is to be evaluated by adding one term after the other, an infinite series will be of practical use only if it converges after a sufficiently small number of terms. Unfortunately, many counterexamples are known in which alternative methods for the evaluation of infinite series must be used since in these cases the conventional approach of evaluating an infinite series does not suffice.

For instance, when Haywood and Morgan [1] performed a discrete basis-set calculation of the Bethe logarithm of the 1s state of the hydrogen atom, they found that even 120 basis functions gave no more than 2–3 decimal digits and they estimated that approximately 10^{10} basis functions would be needed to obtain an accuracy of more than 10 decimal digits. Haywood and Morgan also showed that with the help of a suitable convergence acceleration method an accuracy of more than 13 decimal digits can be extracted from their data.

A good mathematical model for the convergence problems which Haywood and Morgan [1] encountered in their calculation of the Bethe logarithm is the following series expansion for the Riemann zeta function:

$$\zeta(z) = \sum_{n=0}^{\infty} (n+1)^{-z}. \tag{1.1-2}$$

It is well known that this infinite series converges if $\mathrm{Re}(z) > 1$ holds. However, if $\mathrm{Re}(z)$ is only slightly larger than one, the rate of convergence becomes extremely slow. For instance, Bender and Orszag remark in their book (see p. 379 of ref. [2]) that about 10^{20} terms of the above series expansion would be needed to compute $\zeta(1.1)$ accurate to one percent. Bender and Orszag also show that only 10 terms of the series in connection with a specially designed acceleration method are needed to compute $\zeta(1.1)$ to 26 decimal digits (see table 8.7 on p. 380 of ref. [2]).

Even more striking examples for the inadequacy of the conventional approach towards the evaluation of infinite series are some Rayleigh-Schrödinger perturbation expansions of elementary quantum mechanical systems. For instance, if the following normalization for the Hamiltonian of the quartic anharmonic oscillator is used,

$$\hat{H} = \hat{p}^2 + \hat{x}^2 + \beta \hat{x}^4, \tag{1.1-3}$$

then it follows from the results obtained by Bender and Wu (see eq. (1.8) of ref. [3]) that the coefficients c_n of the power series in the coupling constant β for the ground state energy eigenvalue $E_0(\beta)$ of the quartic anharmonic oscillator,

$$E_0(\beta) = \sum_{n=0}^{\infty} c_n \beta^n, \tag{1.1-4}$$

possesses the following asymptotic behaviour:

$$c_n \sim (-1)^{n+1} (3/2)^n \Gamma(n + 1/2), \qquad n \to \infty. \tag{1.1-5}$$

The radius of convergence of the above Rayleigh-Schrödinger perturbation series is obviously zero, i.e., it diverges for all nonzero values of β and summation techniques have to be applied to give this series any meaning beyond a mere formal expansion.

A good mathematical model for the kind of divergence, which occurs in the perturbation series of the quartic anharmonic oscillator, is the so-called Euler integral

$$E(z) = \int_0^{\infty} \frac{e^{-t}\,dt}{1 + zt}, \tag{1.1-6}$$

and its associated asymptotic series, the so-called Euler series

$$E(z) \sim \sum_{n=0}^{\infty} (-1)^n n! z^n = {}_2F_0(1, 1; -z), \qquad z \to \infty. \tag{1.1-7}$$

The radius of convergence of the Euler series is obviously zero. Consequently, this series diverges quite wildly for all $z \neq 0$ and appropriate summation techniques have to be applied if numerical values for the Euler integral are to be computed with the help of this asymptotic series. In fact, the Euler series (1.1-7) will be used quite frequently in this report to test the ability of a sequence transformation to sum wildly divergent series.

1.2. A short history of sequence transformations

In this section, a short sketch of the historical development of sequence transformations will be given. A more complete treatment of the history would be beyond the scope of this report. Consequently, the emphasis will be on those developments which laid the foundations for the sequence transformations which are discussed in this report.

The idea of applying suitable transformations for the acceleration of the convergence of a series or for the summation of a divergent series is almost as old as analysis itself. According to Knopp (see p. 249 of ref. [4]) the first series transformation was published by Stirling [5] already in 1730, and in 1755 Euler [6] published the series transformation which now bears his name.

These early ideas were later extended and refined as well as supplemented by convergence proofs, and they finally led to the theory of regular matrix transformations. Let $\{s_n\}$ be a sequence of partial sums of a series according to eq. (1.1-1). Then, a new sequence $\{s_n'\}$ with hopefully better convergence is obtained by forming weighted means of the elements of the original sequence,

$$s_n' = \sum_{k=0}^{n} \mu_{nk} s_k. \tag{1.2-1}$$

The main appeal of these matrix transformations lies in the fact that for the weights μ_{nk} with k, $n \in \mathbb{N}_0$, which define such a transformation, some necessary and sufficient conditions could be formulated which ensure the regularity of the transformation. This implies that such a regular matrix transformation can safely be applied to any convergent sequence $\{s_n\}$ since the transformed sequence $\{s_n'\}$ will also converge to the same limit. A fairly complete survey of the properties of such matrix transformations can be found in books by Knopp [4], Hardy [7], Petersen [8], Peyerimhoff [9], Zeller and Beekmann [10], and Powell and Shah [11].

This general applicability of regular matrix transformations to all convergent sequences is undoubtedly quite advantageous from a theoretical point of view. However, for the practical usefulness of a transformation in actual computations this general applicability is more likely a hindrance rather than an advantage. This may sound paradoxical. But one cannot realistically expect that a given method will be particularly efficient in a special case if it is simultaneously required that this method should also be able to work in all cases.

Consequently, in recent years emphasis has shifted towards the more powerful but also more specialized nonlinear transformations. Theoretically, nonlinear transformations are much more complicated than matrix transformations and their properties are by no means completely understood. In addition, nonlinear transformations are usually nonregular, i.e., it is not guaranteed that the transformed sequence $\{s_n'\}$ will converge at all, let alone to the same limit as the original sequence $\{s_n\}$. Hence, unless additional information about the sequence to be transformed is available, the use of a nonlinear sequence transformation may be risky. However, this undeniable disadvantage is compensated by the empirical fact that if a nonlinear transformation works, it frequently does so in a spectacular fashion.

The probably oldest nonlinear sequence transformation is the famous Δ^2 process,

$$s_n' = s_n - \frac{[\Delta s_n]^2}{\Delta^2 s_n} = s_n - \frac{[s_{n+1} - s_n]^2}{s_{n+2} - 2s_{n+1} + s_n}, \qquad n \in \mathbb{N}_0. \tag{1.2-2}$$

This sequence transformation is named after Aitken [12] who published this transformation in 1926 but there are indications that it is in fact much older. For instance, Todd (see p. 5 of ref. [13]) claims that this transformation was in principle already known to Kummer [14].

It is generally accepted that the current interest in nonlinear transformations is due to two articles by Shanks [15] and Wynn [16], respectively. Shanks rediscovered in 1955 a sequence transformation which was originally derived in 1941 by Schmidt [17]. Wynn showed only one year later how this sequence transformation, which was originally defined as the ratio of two determinants, can be computed quite efficiently by a nonlinear recursive scheme which is now commonly called the ϵ algorithm. These two articles by Shanks [15] and Wynn [16] had an enormous impact since they stimulated a large amount of research not only in various branches of mathematics but also in theoretical physics and in other sciences. This is amply demonstrated by the long lists of references in books by Baker [18], Brezinski [19–21], Baker and Graves-Morris [22], and Wimp [23], and also by a recent review article by Brezinski [24].

This active research on nonlinear transformations contributed significantly not only to the understanding of Padé approximants or the ϵ algorithm, but also led to the discovery of several other sequence transformations. For instance, in 1956 Wynn [25] introduced the so-called ρ algorithm which is essentially an intelligent and efficient way of computing and extrapolating even-order convergents of an interpolating continued fraction.

In 1971 Brezinski [26] introduced his so-called ϑ algorithm which may be interpreted to be some kind of improved and accelerated ϵ algorithm. Brezinski's derivation of this powerful algorithm was purely heuristic. It was emphasised by Brezinski [27] that this heuristic approach is not restricted to Wynn's ϵ algorithm and can also be used in the case of other sequence transformations. Some examples of new sequence transformations, which were derived in that way, will be given later.

Another class of sequence transformations was introduced by Levin [28] in 1973. According to Smith and Ford [29,30] who compared the performances of several linear as well as nonlinear sequence transformations, some variants of the Levin transformation are probably the most powerful and most versatile convergence accelerators currently known and they are also able to sum even wildly divergent series. The sequence transformations introduced by Levin are also the basis of a large part of this article since they are the starting point for the derivation of several new sequence transformations which offer in some cases computational advantages, in particular if wildly divergent series are to be summed. In addition, a theoretical analysis of the properties of the new transformations can also often be done more easily than in the case of the Levin transformation.

A general extrapolation algorithm, which encompasses the majority of the currently known extrapolation methods and also many of the new sequence transformations of this report as special cases, was developed independently by Brezinski [31] and Håvie [32].

Finally, Germain-Bonne [33] developed in 1973 a formal theory of convergence acceleration which is of considerable importance not only for this report. By means of Germain-Bonne's theory it can in some cases be decided whether a given transformation is regular, i.e., whether the convergence of a sequence $\{s_n\}$ to some limit s implies the convergence of the transformed sequence $\{s'_n\}$ to the same limit s. Also, in some cases it can be decided by a priori considerations whether the transformed sequence $\{s'_n\}$ will converge faster than the original sequence $\{s_n\}$.

1.3. Organization of this report

No attempt is made to treat all aspects of the acceleration of convergence and the summation of divergent series. The emphasis of this report is on convergence acceleration and summation by means of nonlinear sequence transformations. Linear sequence transformations are only treated if they are special cases of nonlinear sequence transformations. The nonlinear sequence transformations, which occur in this report, are designed to handle convergent or divergent sequences of partial sums of infinite series as they occur in scientific applications or in the theory of special functions. However, the specific problems, which arise in connection with the acceleration of the convergence of Fourier series or of orthogonal expansions, are not treated. Also, the acceleration or summation of multidimensional sequences – or vector sequences, as they are called in the literature on convergence acceleration – is not considered in this report.

Several nonlinear sequence transformations as for instance Aitken's Δ^2 process [12], Wynn's ϵ algorithm [16], Wynn's ρ algorithm [25], Brezinski's ϑ algorithm [26], and Levin's sequence transformation [28] are now relatively well known and many applications of these transformations have been reported in the more recent literature. The properties of these nonlinear sequence transformations are reviewed shortly in this report and efficient algorithms for their computation are discussed. However, the emphasis of this report is on the derivation of new nonlinear sequence transformations, on the construction of efficient algorithms for their computation, and on the analysis of their properties in convergence acceleration and summation processes.

In this report, the sequence transformations are always computed with the help of linear or nonlinear 2-dimensional recurrence formulas. Also, it is always tried to find computational schemes for these recursions which are optimal with respect to storage requirements. Such an optimization is actually not necessary if the sequence transformations are programmed in FORTRAN because then storage space would not be a problem even if less efficient computational algorithms would be used. If, however, sequence transformations are programmed in a formal manipulation language such as REDUCE, MACSYMA or MAPLE, it is probably a good idea to use such an optimized computational scheme since storage restrictions may then be much more severe.

Some listings of FORTRAN 77 programs are included in the text. In order to save space, all comments and also all IF statements, which check the validity of the input data, were removed from the programs. Consequently, these FORTRAN 77 programs are not "good" programs which comply with the recommendations of books on programming style. The sole purpose of these program listings is to facilitate the understanding of the sometimes relatively intricate computational algorithms which are discribed in this report.

In order to make this report more selfcontained, in section 2 the mathematical terminology, which is specific for this report, as well as the most important mathematical concepts and techniques, which are needed for the derivation and understanding of sequence transformations, are introduced.

In section 3 general properties of nonlinear sequence transformations are discussed. In addition, it has been attempted to give a motivation for some of the most important concepts and assumptions, which are the basis for the construction of a large class of nonlinear sequence transformations.

Section 4 deals with Wynn's ϵ algorithm [16], which is an efficient algorithm for the computation of the Shanks transformation [15] or – if the elements of the sequence to be accelerated or summed are the partial sums of a power series – of Padé approximants. Section 5 deals with Aitken's Δ^2 process [12] and its iteration, which are both close relatives of Wynn's ϵ algorithm. Section 6 deals with Wynn's ρ algorithm [25], which is structurally almost identical with Wynn's ϵ algorithm. Also, a new sequence transformation is constructed by iterating the explicit expression for $\rho_2^{(n)}$ along the lines of Aitken's iterated Δ^2 process.

Section 7 deals with with Levin's sequence transformation [28], and several other sequence transformations which are either special cases or generalizations of Levin's sequence transformation. Levin's sequence transformation is also the starting point for two new classes of sequence transformations which are treated in sections 8 and 9. The difference between Levin's sequence transformation and the new sequence transformations is that Levin's sequence transformation is based upon the assumption that the remainders of the partial sums can be approximated by truncated Poincaré-type asymptotic expansions whereas the new sequence transformations assume that the remainders can be approximated by truncations of factorial series and related expansions which are also based upon Pochhammer symbols.

Section 10 deals with Brezinski's ϑ algorithm [26], which was derived by modifying the recursive scheme for Wynn's ϵ algorithm, and a closely related sequence transformation which is obtained by iterating the expression for $\vartheta_2^{(n)}$. In section 11, the recursive schemes of several other linear and nonlinear sequence transformations are modified along the lines of Brezinski's ϑ algorithm and several new nonlinear sequence transformations are derived.

The practical usefulness of the original version of Germain-Bonne's formal theory of convergence acceleration [33] is quite limited since it can only analyze the properties of a sequence transformation if its recursive scheme satisfies some very restrictive conditions. Many sequence transformations of this report do not satisfy these conditions. Consequently, in section 12 Germain-Bonne's formal theory of convergence acceleration is modified in such a way that the properties of the sequence transformations of this report can also be analyzed.

Unfortunately, Germain-Bonne's theory cannot be applied in all cases of interest. In particular, it cannot be used for the analysis of the summation of wildly divergent Stieltjes series as they for instance occur in the Rayleigh-Schrödinger perturbation expansion for the energy eigenvalues of the quartic anharmonic oscillator. In section 13 the transformation of sequences of partial sums of convergent and divergent Stieltjes series is analyzed. The estimates, which are obtained in this way, indicate that some variants of the new sequence transformations, which are discussed in sections 8 and 9, should sum a divergent Stieltjes series as good or even somewhat better than the analogous variants of Levin's sequence transformation, and that they should all be far more efficient than Padé approximants. These theoretical estimates are supported quite convincingly by some numerical examples.

One of the most complicated computational problems, which can occur in this context, is the acceleration of the convergence of infinite series with terms a_n that all have the same sign and that decay like a fixed power $n^{-\alpha}$ with $\alpha > 1$ as $n \to \infty$. A good example of such an extremely slowly convergent infinite series with positive terms is the series (1.1-2) for the Riemann ζ function. In the case of such slowly convergent series with positive terms, Germain-Bonne's formal theory of convergence acceleration [33] also does not help. Section 14 deals with the acceleration of the convergence of series of that type. Some exactness results and also some error estimates are derived. However, it is relatively difficult to obtain theoretical results. Consequently, the emphasis in section 14 is on numerical testing.

Finally, section 15 contains a condensed review of the properties of the sequence transformations which are treated in this report.

2. Terminology

2.1. Special mathematical symbols and special functions

In this report, essentially standard mathematical terminology will be used. In particular, $\mathbb{N}$ stands for the set of positive integers $n = 1, 2, 3, \ldots,$ and $\mathbb{N}_0$ stands for the set of nonnegative integers $n = 0, 1, 2, \ldots$. Also, $\mathbb{R}$ and $\mathbb{C}$ denote the sets of real and complex numbers, respectively. The following notations are, however, nonstandard:

$\{\!\{s_n\}\!\}$: Sequence of elements s_n with $n \in \mathbb{N}_0$. It is always tacitly assumed that the sequence elements with negative indices, $s_{-1}, s_{-2}, s_{-3}, \ldots,$ are zero.

$[\![x]\!]$: Integral part of $x \in \mathbb{R}$, i.e., the largest integer m satisfying the inequality $m \leq x$.

$\mathbb{F}^n$: Set of all vectors $(x_1, \ldots, x_n) \in \mathbb{R}^n$ with all components being different from zero, i.e., $x_j \neq 0$ for all $j = 1, 2, \ldots, n$.

$\mathbb{D}^n$: Set of all vectors $(x_1, \ldots, x_n) \in \mathbb{R}^n$ with all components being distinct, i.e., $i \neq j$ implies $x_i \neq x_j$ for all $i, j = 1, 2, \ldots, n$.

$\mathbb{H}^n$: Intersection of $\mathbb{F}^n$ and $\mathbb{D}^n$, i.e., the set of vectors $(x_1, \ldots, x_n) \in \mathbb{R}^n$ with all components being nonzero and distinct.

Sometimes sums or products will occur in which the lower limit is greater than the upper limit. In this report, we shall always use the convention that such an *empty sum* will be interpreted as zero, i.e.,

$$\sum_{k=m}^{n} a_k = 0, \qquad \text{if } m > n, \tag{2.1-1}$$

and that such an *empty product* will be interpreted as one, i.e.,

$$\prod_{k=m}^{n} a_k = 1, \qquad \text{if } m > n. \tag{2.1-2}$$

For the commonly occurring special functions of mathematical physics the notation and the conventions of Magnus, Oberhettinger, and Soni [34] will be used in this report unless explicitly stated.

2.2. Order symbols

Let $f(z)$ and $g(z)$ be two functions defined on some domain D in the complex plane and let z_0 be a limit point of D, possibly the point at infinity. Then,

$$f(z) = \mathrm{O}(g(z)), \qquad z \to z_0 \tag{2.2-1}$$

means that there is a positive constant A and a neighbourhood U of z_0 such that

$$|f(z)| \leq A|g(z)| \tag{2.2-2}$$

for all $z \in U \cap D$. If $g(z)$ does not vanish on $U \cap D$ this simply means that $f(z)/g(z)$ is bounded on $U \cap D$. Also,

$$f(z) = o(g(z)), \qquad z \to z_0 \tag{2.2-3}$$

means that for any positive number $\epsilon \in \mathbb{R}$ there exists a neighbourhood U of z_0 such that

$$|f(z)| \leq \epsilon |g(z)| \tag{2.2-4}$$

for all $z \in U \cap D$. If $g(z)$ does not vanish on $U \cap D$ this simply means that $f(z)/g(z)$ approaches zero as $z \to z_0$.

2.3. Asymptotic sequences and asymptotic expansions

A finite or infinite sequence of functions $\{\Phi_n(z)\}$ with $n \in \mathbb{N}_0$, which are defined on some domain D of complex numbers on which all $\Phi_n(z)$ are nonzero except possibly at z_0, is called an *asymptotic sequence* as $z \to z_0$ if, for all $n \in \mathbb{N}_0$,

$$\Phi_{n+1}(z) = o(\Phi_n(z)), \qquad z \to z_0. \tag{2.3-1}$$

Examples for asymptotic sequences with $n \in \mathbb{N}_0$ are $\{(z - z_0)^n\}$ as $z \to z_0$ or $\{(\log z)^{-n}\}$ as $z \to \infty$.

The formal series

$$f(z) \sim \sum_{n=0}^{\infty} c_n \Phi_n(z), \tag{2.3-2}$$

which need not be convergent, is called an *asymptotic expansion* of $f(z)$ with respect to the asymptotic sequence $\{\Phi_n(z)\}$ in the *sense of Poincaré* if, for every $m \in \mathbb{N}_0$,

$$f(z) - \sum_{n=0}^{m} c_n \Phi_n(z) = o(\Phi_m(z)), \qquad z \to z_0. \tag{2.3-3}$$

If such a Poincaré-type asymptotic expansion exists, it is unique, and its coefficients c_n can be computed recursively,

$$c_m = \lim_{z \to z_0} \left\{ \left[f(z) - \sum_{n=0}^{m-1} c_n \Phi_n(z) \right] \Big/ \Phi_m(z) \right\}, \qquad m \in \mathbb{N}_0. \tag{2.3-4}$$

The first term of the asymptotic expansion (2.3-2) is usually called the *dominant* or *leading term* and one frequently writes

$$f(z) \sim \Phi_0(z), \tag{2.3-5}$$

indicating that $f(z)$ tends to c_0 as $z \to z_0$.

A particularly simple asymptotic sequence as $z \to \infty$ is the set $\{\Psi(z)/z^n\}$, $n \in \mathbb{N}_0$, with $\Psi(z)$ being a suitable function. If a given function $f(z)$ possesses an asymptotic expansion with respect to this sequence,

$$f(z) \sim \Psi(z) \sum_{n=0}^{\infty} c_n/z^n, \qquad z \to \infty, \tag{2.3-6}$$

then the ratio $f(z)/\Psi(z)$ can be expressed as an *asymptotic power series* in $1/z$,

$$f(z)/\Psi(z) \sim \sum_{n=0}^{\infty} c_n/z^n, \qquad z \to \infty. \tag{2.3-7}$$

2.4. Finite differences

Let f be a function defined on the set of integers $\mathbb{N}_0$. Then, the *forward difference* $\Delta f(n)$ is defined by the relationship

$$\Delta f(n) = f(n+1) - f(n), \qquad n \in \mathbb{N}_0. \tag{2.4-1}$$

Higher powers of the difference operator Δ can be defined recursively, i.e.,

$$\Delta^k f(n) = \Delta\left[\Delta^{k-1} f(n)\right], \qquad k \in \mathbb{N}, \tag{2.4-2}$$

$$\Delta^0 f(n) = f(n). \tag{2.4-3}$$

The *shift operator* E is defined by the relationship

$$Ef(n) = f(n+1). \tag{2.4-4}$$

Higher powers of E can again be defined recursively. Obviously, we have

$$E^k f(n) = f(n+k), \qquad k \in \mathbb{N}, \tag{2.4-5}$$

$$E^0 f(n) = f(n). \tag{2.4-6}$$

It follows at once from their definitions that the operators Δ and E are connected by the relationship

$$\Delta = E - 1. \tag{2.4-7}$$

This relationship can be combined with the binomial theorem to give

$$\Delta^k f(n) = (-1)^k \sum_{j=0}^{k} (-1)^j \binom{k}{j} f(n+j), \qquad k \in \mathbb{N}_0. \tag{2.4-8}$$

In the following text it will always be tacitly assumed that in the case of several indices the difference operator Δ and the shift operator E will only act upon n and not on other indices.

2.5. Special sequences

Let us assume that the sequence $\{s_n\}$ either converges to some limit s, or, if it diverges, can be summed by an appropriate summation method to give s. In the case of divergence s is frequently called *antilimit*. Then, the partition of a sequence element s_n into the *limit* or *antilimit* s and the *remainder* r_n according to

$$s_n = s + r_n \tag{2.5-1}$$

makes sense for all $n \in \mathbb{N}_0$. If s_n is the partial sum of a series,

$$s_n = \sum_{k=0}^{n} a_k, \tag{2.5-2}$$

the remainder r_n obviously satisfies

$$r_n = - \sum_{k=n+1}^{\infty} a_k. \tag{2.5-3}$$

For the ratio of two consecutive terms of an infinite series we write

$$\rho_n = a_{n+1}/a_n, \qquad n \in \mathbb{N}_0. \tag{2.5-4}$$

The magnitude of the remainder r_n is a natural measure for the convergence of a sequence or series. Often, it is also of considerable interest to analyze the asymptotics of the sequence of remainders $\{r_n\}$ as $n \to \infty$. Let $\{\varphi_k(n)\}$, $k \in \mathbb{N}_0$ be a suitable asymptotic sequence as $n \to \infty$ with $\varphi_0(n)$. Then, ω_n denotes the dominant part of r_n with respect to the asymptotic sequence $\varphi_k(n)$, i.e.,

$$r_n/\omega_n \sim \sum_{k=0}^{\infty} c_k \varphi_k(n), \qquad n \to \infty. \tag{2.5-5}$$

Sequences of *remainder estimates* $\{\omega_n\}$ will be of considerable importance in this report. The reason is that it is often possible to obtain at least some structural information about the behaviour of the dominant term of a remainder r_n as $n \to \infty$. It will become clear later that those convergence acceleration or summation methods, which explicitly utilize the information contained in the remainder estimates $\{\omega_n\}$, are frequently particularly efficient.

Many sequence transformations do not only require the input of the sequences $\{s_n\}$ and $\{\omega_n\}$ but also the input of an additional sequence of auxiliary quantities as for instance interpolation points. In this report, such an *auxiliary sequence* will usually be denoted by $\{x_n\}$.

2.6. Types of convergence

It is neither possible nor desirable to set up a complete classification scheme which is able to cover all types of convergence. However, in the majority of all practical applications only a few types of convergence occur. Consequently, special names were given to them in the literature.

Let us assume that the sequence $\{\!\{s_n\}\!\}$, which converges to some limt s, satisfies

$$\lim_{n\to\infty} \frac{s_{n+1}-s}{s_n-s} = \lim_{n\to\infty} \frac{r_{n+1}}{r_n} = \rho. \tag{2.6-1}$$

If $0<|\rho|<1$ holds, we say that the sequence $\{\!\{s_n\}\!\}$ converges *linearly*, if $\rho=1$ holds, we say that $\{\!\{s_n\}\!\}$ converges *logarithmically*, and if $\rho=0$ holds, we say that $\{\!\{s_n\}\!\}$ converges *hyperlinearly*. Of course, $|\rho|>1$ implies that the sequence $\{\!\{s_n\}\!\}$ diverges.

The standard example for linear convergence is the sequence of partial sums of the geometric series,

$$s_n(z) = \sum_{k=0}^{n} z^k = \frac{1-z^{n+1}}{1-z}, \qquad 0<|z|<1, \qquad n\in\mathbb{N}_0. \tag{2.6-2}$$

The sequence of partial sums of the series (1.1-2) for the Riemann zeta function is a good example for logarithmic convergence. Also, it can be shown quite easily that the partial sums of the power series for the exponential function form a sequence which converges hyperlinearly.

The above definitions for hyperlinear, linear and logarithmic convergence do not seem to be particularly well suited for the classification of infinite series because normally only the terms a_k of a series but not the remainders r_n are known. However, Wimp showed on p. 6 of his book [23] that if $0<|\rho|<1$ holds, the two statements

$$\lim_{n\to\infty} (r_{n+1}/r_n) = \rho \tag{2.6-3}$$

and

$$\lim_{n\to\infty} (a_{n+1}/a_n) = \rho \tag{2.6-4}$$

are equivalent. In addition, Clark, Gray, and Adams [35] showed that if the terms a_k of a convergent series are all real and have the same sign, then

$$\lim_{n\to\infty} (r_{n+1}/r_n) = \lim_{n\to\infty} (a_{n+1}/a_n) = \rho. \tag{2.6-5}$$

In eq. (2.6-5), the case $\rho=1$, which corresponds to logarithmic convergence, is not excluded. Hence, it is at least possible to classify linearly and logarithmically convergent series – which are of particular interest in connection with convergence acceleration methods – according to the behaviour of their terms a_n as $n\to\infty$.

Sequences and series, which converge hyperlinearly, often converge so rapidly that not much can be gained by convergence acceleration methods. Consequently, hyperlinear convergence is more or less neglected in the literature on convergence acceleration. This is not entirely justified because in some situations the use of convergence acceleration methods can indeed be quite helpful. A simple and nevertheless striking example, which shows that convergence acceleration methods may be quite useful even in the case of hyperlinear convergence, is the power series for the exponential function. It cannot be used for the computation of e^{-x} if x is positive and large because then large terms with alternating signs would lead to cancellation and to severe numerical instabilities. However, if suitable convergence acceleration methods are used, remarkably accurate results can be obtained after a relatively small number of terms.

Let us assume that two sequences $\{s_n\}$ and $\{s'_n\}$ both converge to the same limit s. We shall say that the sequence $\{s'_n\}$ *converges more rapidly* than $\{s_n\}$ if

$$\lim_{n \to \infty} \frac{s'_n - s}{s_n - s} = 0. \tag{2.6-6}$$

In convergence acceleration processes, this definition is somewhat inconvenient since it requires the knowledge of the limit s which is usually not known. Consequently, it would be desirable to replace eq. (2.6-6) by the alternative condition

$$\lim_{n \to \infty} \frac{s'_{n+1} - s'_n}{s_{n+1} - s_n} = \lim_{n \to \infty} \frac{\Delta s'_n}{\Delta s_n} = 0. \tag{2.6-7}$$

However, it seems that it is not possible to prove the equivalence of eqs. (2.6-6) and (2.6-7) without making explicit assumptions about how fast the sequences $\{s_n\}$ and $\{s'_n\}$ approach their common limit s.

If the sequence $\{s_n\}$ converges linearly, the transformed sequence $\{s'_n\}$ can only converge more rapidly than $\{s_n\}$ if it converges at least linearly or even faster. In this case, the equivalence of the two conditions (2.6-6) and (2.6-7) follows at once from the relationship

$$\frac{s'_{n+1} - s'_n}{s_{n+1} - s_n} = \frac{s'_n - s}{s_n - s} \, \frac{[(s'_{n+1} - s)/(s'_n - s)] - 1}{[(s_{n+1} - s)/(s_n - s)] - 1}. \tag{2.6-8}$$

However, if $\{s_n\}$ converges logarithmically, the denominator of the second term on the right-hand side of eq. (2.6-8) approaches zero as $n \to \infty$. In this case, some additional assumptions about the rate of convergence of $\{s'_n\}$ to s have to be made in order to be able to show that the two conditions (2.6-6) and (2.6-7) are indeed equivalent.

2.7. Sequence transformations

In this report a sequence transformation $\mathscr{T}$ will always be a rule which transforms a given sequence $\{s_n\}$ into a new sequence $\{s'_n\}$,

$$\mathscr{T}: \{s_n\} \mapsto \{s'_n\}, \qquad n \in \mathbb{N}_0. \tag{2.7-1}$$

Since a computational algorithm can only involve a finite number of operations, only finite subsets of a sequence $\{s_n\}$ can be associated to a new sequence element s'_m by a sequence transformation $\mathscr{T}$. In addition, it will always be assumed that the finite subset, which is to be transformed, will entirely consist of consecutive elements. This means that only subsets of the type $\{s_n, s_{n+1}, \ldots, s_{n+l}\}$ with $n, l \in \mathbb{N}_0$ will be considered. Since the subsets, which are to be transformed, contain $l+1$ elements, l will frequently be called the *order* of the transformation $\mathscr{T}$. Hence, if all elements of the sequence $\{\!\{s_n\}\!\}$ are real and if no sequences of interpolation points $\{\!\{x_n\}\!\}$ or remainder estimates $\{\!\{\omega_n\}\!\}$ are needed, a sequence transformation $\mathscr{T}$ of order l is a map of the following type:

$$\mathscr{T}: \mathbb{R}^{l+1} \to \mathbb{R}. \tag{2.7-2}$$

In this report, a sequence transformation $\mathscr{T}$ can always be represented by an infinite set of doubly indexed quantities $T_k^{(n)}$ with $k, n \in \mathbb{N}_0$. The superscript n always indicates the minimal index occurring in the finite subset of sequence elements which are used for the computation of the transform $T_k^{(n)}$, and the subscript k is a measure for the complexity of such a $T_k^{(n)}$.

The quantities $T_k^{(n)}$ are gauged in such a way that $T_0^{(n)}$ will always correspond to an untransformed sequence element, i.e.,

$$T_0^{(n)} = s_n, \qquad n \in \mathbb{N}_0. \tag{2.7-3}$$

Increasing values of k imply that the order l of the transform $T_k^{(n)}$ also increases. This means that for every $k, n \in \mathbb{N}_0$ the sequence transformation $\mathscr{T}$ will produce a new transform for which we shall write

$$T_k^{(n)} = \mathscr{T}(s_n, s_{n+1}, \ldots, s_{n+l}). \tag{2.7-4}$$

Here, the order l is of course a function of k. The exact relationship, which connects the subscript k and the order l, is specific for every sequence transformation $\mathscr{T}$. In this report we shall encounter a variety of different relationships such as $l=k$, $l=k+1$, $l=2k$ or even $l=3k$.

The transforms $T_k^{(n)}$ with $k, n \in \mathbb{N}_0$ can be displayed in a 2-dimensional array which is called the *table* of the sequence transformation $\mathscr{T}$. In this report, the transforms $T_k^{(n)}$ will always be ordered in a rectangular scheme in such a way that the superscript n indicates the row and the subscript k the column of the array. Hence, in this report the table of a transformation $\mathscr{T}$ will always be displayed in the following way:

$$\begin{array}{cccccc}
T_0^{(0)} & T_1^{(0)} & T_2^{(0)} & \ldots & T_n^{(0)} & \ldots \\
T_0^{(1)} & T_1^{(1)} & T_2^{(1)} & \ldots & T_n^{(1)} & \ldots \\
T_0^{(2)} & T_1^{(2)} & T_2^{(2)} & \ldots & T_n^{(2)} & \ldots \\
T_0^{(3)} & T_1^{(3)} & T_2^{(3)} & \ldots & T_n^{(3)} & \ldots \\
\vdots & \vdots & \vdots & \ddots & \vdots & \ddots \\
T_0^{(n)} & T_1^{(n)} & T_2^{(n)} & \ldots & T_n^{(n)} & \ldots \\
\vdots & \vdots & \vdots & \ddots & \vdots & \ddots
\end{array} \tag{2.7-5}$$

In the process of convergence acceleration or summation only those elements of the table of a sequence transformation $\mathscr{T}$ should be computed which will be needed to obtain convergence up to a certain accuracy. Of course, this implies that one has to decide in advance which transforms $T_k^{(n)}$ should be used for that purpose. In this context it will be advantageous to introduce the following terminology.

A sequence $\{(n_j, k_j)\}$ of ordered pairs of integers $n_j, k_j \in \mathbb{N}_0$ is called a *path* if $n_0 = k_0 = 0$ and if for all integers $j \in \mathbb{N}_0$ we have $n_{j+1} \geq n_j$ and $k_{j+1} \geq k_j$ and if for each $j \in \mathbb{N}_0$ either one or both of the two relations $n_{j+1} = n_j + 1$ and $k_{j+1} = k_j + 1$ are true. Obviously, $n_j + k_j \to \infty$ as $j \to \infty$.

Paths where k_j is ultimately constant are called *vertical paths*, and paths where n_j is ultimately constant are called *horizontal paths*.

We shall say that a sequence transformation $\mathscr{T}$ is *regular* on a given path $\mathscr{P} = \{(n_j, k_j)\}$ if for every convergent sequence $\{s_n\}$ we have

$$\lim_{j \to \infty} T_{k_j}^{(n_j)} = s. \tag{2.7-6}$$

Next, we want to define what we mean by saying that a transformation $\mathscr{T}$ is called *accelerative on a path* $\mathscr{P}$ for a sequence $\{s_n\}$. In the literature on convergence acceleration the following definition is the most common one:

$$\lim_{j \to \infty} \frac{T_{k_j}^{(n_j)} - s}{s_{n_j} - s} = 0. \tag{2.7-7}$$

However, since for a given subscript k a sequence transformation $\mathscr{T}$ always acts on $l + 1$ consecutive sequence elements $s_n, s_{n+1}, \ldots, s_{n+l}$ with l being a function of k, it would actually be better to say that $\mathscr{T}$ is accelerative on a path $\mathscr{P}$ for a sequence $\{s_n\}$ if

$$\lim_{j \to \infty} \frac{T_{k_j}^{(n_j)} - s}{s_{n_j + l_j} - s} = 0. \tag{2.7-8}$$

Hence, if the second definition is used a transformation $\mathscr{T}$ will be called accelerative on a path $\mathscr{P} = \{(n_j, k_j)\}$ if the transforms $T_{k_j}^{(n_j)}$ converge faster than the last elements $s_{n_j + l_j}$ of the strings $s_{n_j}, s_{n_j+1}, \ldots, s_{n_j + l_j}$ which are used for the computation of the transforms $T_{k_j}^{(n_j)}$.

In this report both definitions (2.7-7) and (2.7-8) will be used. However, it will always be stated explicitly which of the two different definitions is actually meant.

Let us again assume that a sequence $\{s_n\}$ converges to some limit s. A sequence transformation $\mathscr{T}$ will be called *exact* for the sequence $\{s_n\}$ if some integer $l_0 \in \mathbb{N}_0$ exists such that the application of $\mathscr{T}$ to every *finite* string $s_n, \ldots, s_{n+l}$ of sequence elements with $l \geq l_0$ yields the exact limit s of this sequence.

3. On the derivation of sequence transformations

3.1. General properties of nonlinear sequence transformations

It was remarked earlier that in many cases of physical interest the conventional approach of evaluating an infinite series by adding one term after the other does not suffice. Examples are logarithmically convergent series which may converge so slowly that an evaluation by adding one term after the other would overstep even the potential of modern supercomputers, or divergent series as they for instance occur in Rayleigh-Schrödinger perturbation theory.

In such cases it is necessary to replace the conventional process of evaluating a series by a *generalized summation process* which is able to associate a numerical value even to a prohibitively slowly convergent or divergent series.

The generalized summation processes of this report are transformations which are defined on finite subsets of the sequence $\{s_n\}$ of partial sums. Let $\mathscr{T}_l$ be such a generalized summation process which acts upon $l+1$ partial sums $s_n, \ldots, s_{n+l}$. In view of the fact that in the case of convergent series with real terms we have

$$\sum_{n=0}^{\infty} (\alpha a_n + \beta b_n) = \alpha \sum_{n=0}^{\infty} a_n + \beta \sum_{n=0}^{\infty} b_n, \qquad \alpha, \beta \in \mathbb{R}, \tag{3.1-1}$$

it seems natural to require that such a generalized summation process $\mathscr{T}_l$ should also be *linear*. Therefore, let us assume that $\{s_n\}$ and $\{t_n\}$ are two sequences of partial sums of real terms which converge to s and t, respectively. Thus, a generalized summation process $\mathscr{T}_l$ should satisfy

$$\mathscr{T}_l(\alpha s_n + \beta t_n, \ldots, \alpha s_{n+l} + \beta t_{n+l}) = \alpha \mathscr{T}_l(s_n, \ldots, s_{n+l}) + \beta \mathscr{T}_l(t_n, \ldots, t_{n+l}),$$
$$\alpha, \beta \in \mathbb{R}, \qquad l, n \in \mathbb{N}_0. \tag{3.1-2}$$

Also, such a generalized summation process should preserve the limit of a convergent sequence, i.e., it should be *regular*. Hence, a *linear and regular generalized summation process* $\mathscr{T}_l$ should satisfy for all sequences $\{s_n\}$ and $\{t_n\}$ which converge to s and t, respectively,

$$\lim_{n\to\infty} \mathscr{T}_l(\alpha s_n + \beta t_n, \ldots, \alpha s_{n+l} + \beta t_{n+l}) = \alpha s + \beta t. \tag{3.1-3}$$

Unfortunately, the generalized summation processes considered in this report will in general be neither linear nor regular and we have to content ourselves with a weaker requirement. Let $\{s_n\}$ be a sequence and let α and τ be two constants. We may only assume that a generalized summation process $\mathscr{T}_l$ is *invariant under translation*, i.e., that for all admissible l, $n \in \mathbb{N}_0$

$$\mathscr{T}_l(\alpha s_n + \tau, \ldots, \alpha s_{n+l} + \tau) = \alpha \mathscr{T}_l(s_n, \ldots, s_{n+l}) + \tau. \tag{3.1-4}$$

It must be emphasized that because of the nonregularity of the generalized summation processes $\mathscr{T}_l$ of this report we also cannot assume that either

$$\lim_{n\to\infty} \mathscr{T}_l(s_n, \ldots, s_{n+l}) = s \tag{3.1-5}$$

or – if this limit is defined – that

$$\lim_{l \to \infty} \mathscr{T}_l(s_n, \ldots, s_{n+l}) = s \tag{3.1-6}$$

holds for arbitrary convergent sequences $\{s_n\}$.

Nonlinearity and nonregularity are undeniably unpleasant complications which one would like to avoid instinctively. However, they are essential and indispensable since the power of the sequence transformations, which are discussed in this report, stems from their nonlinearity and nonregularity.

3.2. *An example: convergence acceleration of alternating series*

If we want to construct a generalized summation process, which is able to accelerate the convergence of an infinite series, we are confronted with the practical problem that the information contained in a finite string of partial sums $s_0, s_1, \ldots, s_m$ has to be extracted and utilized in a way which is more efficient than the conventional process of adding up one term after the other. If we again assume that for all $n \in \mathbb{N}_0$ a sequence element s_n can be partitioned into the limit s and the remainder r_n according to

$$s_n = s + r_n, \tag{3.2-1}$$

then this essentially means that we have to find a way of eliminating the remainder r_n and determining the limit s at least approximately by exploiting the information stored in the finite string $s_0, s_1, \ldots, s_m$ of partial sums.

Essentially the same problem of eliminating the remainder r_n and determining the antilimit s at least approximately arises if we try to sum a divergent series. The only difference is that in the case of a divergent series the remainder r_n does not vanish as $n \to \infty$, and that the antilimit s cannot be obtained by simply adding up the terms of the series. Instead, the antilimit s of a sequence can only be determined with the help of a suitable summation method.

Since we cannot assume that it will be possible to obtain the numerical values of the remainders r_n directly, the best we can hope for is something which may be called *structural information*. In order to clarify this concept we will consider a simple example. Let us assume that the sequence elements s_n are partial sums of a series with real and strictly alternating terms,

$$s_n = \sum_{k=0}^{n} (-1)^k b_k, \tag{3.2-2}$$

which means that all b_n with $n \in \mathbb{N}_0$ have the same sign. The remainder r_n of s_n is then given by

$$r_n = -\sum_{k=n+1}^{\infty} (-1)^k b_k. \tag{3.2-3}$$

Let us now also assume that all b_n with $n \in \mathbb{N}_0$ are positive and strictly decreasing with increasing n and that they vanish as $n \to \infty$. This implies that the series converges to some limit

s. In addition, it can be shown that the sequence of remainders $\{r_n\}$ is also strictly alternating and that the magnitude of a remainder r_n is bounded by the first term which was not included in the partial sum s_n (see p. 259 of ref. [4]),

$$|r_n| < b_{n+1}, \qquad n \in \mathbb{N}_0. \tag{3.2-4}$$

Now, we have to find a way of utilizing this structural information about the behaviour of the sequence of remainders $\{r_n\}$. Simply adding the next term $(-1)^{n+1}b_{n+1}$ to s_n would only produce s_{n+1} and we would not gain anything substantial. Consequently, we need an additional assumption which will help us to construct a sequence transformation for alternating series.

It is a relatively natural idea to assume that the ratio $r_n/[(-1)^{n+1}b_{n+1}]$ can be expanded in a Poincaré-type asymptotic power series in the variable $1/(n+1)$, i.e.,

$$r_n \sim (-1)^{n+1} b_{n+1} \sum_{j=0}^{\infty} c_j (n+1)^{-j}, \qquad n \to \infty. \tag{3.2-5}$$

The assumption, that such a Poincaré-type asymptotic expansion exists, will enable us to derive a sequence transformation which is capable of accelerating the convergence of alternating series.

However, the complete elimination of such a remainder r_n on the basis of the asymptotic expansion (3.2-5) will not be possible since any computational algorithm can only determine a finite number of the unknown linear coefficients c_j in eq. (3.2-5). Consequently, we can only construct a sequence transformation which is able to eliminate model remainders of the following type:

$$\tilde{r}_n = (-1)^{n+1} b_{n+1} \sum_{j=0}^{k-1} c_j (n+1)^{-j}, \qquad n \in \mathbb{N}_0. \tag{3.2-6}$$

Model remainders of this type are obtained by truncating the infinite series in eq. (3.2-5) after the first k terms. Hence, at least for sufficiently large values of k and n the model remainders $\tilde{r}_n$ should approximate the actual remainders r_n very well. This implies that the partial sums s_n can also be approximated very well by the elements of the following model sequence, which contain only finitely many terms:

$$\tilde{s}_n = s + (-1)^{n+1} b_{n+1} \sum_{j=l}^{k-1} c_j (n+1)^{-j}, \qquad n \in \mathbb{N}_0. \tag{3.2-7}$$

In eq. (3.2-7) there occur $k+1$ unknowns, the limit s and the k coefficients $c_0, \ldots, c_{k-1}$. Since all unknowns occur *linearly* their determination poses in principle no problems. All that is needed are the numerical values of $k+1$ sequence elements, e.g., the string $\tilde{s}_n$, $\tilde{s}_{n+1}, \ldots, \tilde{s}_{n+k}$, and it is possible to determine the limit s of the model sequence (3.2-7).

Hence, we only have to use Cramer's rule in order to see that the limit s is given by the following ratio of determinants:

$$s = \frac{\begin{vmatrix} \tilde{s}_n & \cdots & \tilde{s}_{n+k} \\ (-1)^{n+1} b_{n+1} & \cdots & (-1)^{n+k+1} b_{n+k+1} \\ \vdots & \ddots & \vdots \\ \dfrac{(-1)^{n+1} b_{n+1}}{(n+1)^{k-1}} & \cdots & \dfrac{(-1)^{n+k+1} b_{n+k+1}}{(n+k+1)^{k-1}} \end{vmatrix}}{\begin{vmatrix} 1 & \cdots & 1 \\ (-1)^{n+1} b_{n+1} & \cdots & (-1)^{n+k+1} b_{n+k+1} \\ \vdots & \ddots & \vdots \\ \dfrac{(-1)^{n+1} b_{n+1}}{(n+1)^{k-1}} & \cdots & \dfrac{(-1)^{n+k+1} b_{n+k+1}}{(n+k+1)^{k-1}} \end{vmatrix}}. \tag{3.2-8}$$

Now we could try to replace the elements $\tilde{s}_n$ of the model sequence (3.2-7) in the first determinant in eq. (3.2-8) by the partial sums s_n. This would certainly not produce the exact limit s of the alternating series since the partial sums s_n satisfy eq. (3.2-7) only approximately. However, if the sequence elements $\tilde{s}_n$ are able to approximate the partial sums s_n with sufficient accuracy then we can hope that the ratio of determinants, in which now the partial sums s_n, $s_{n+1}, \ldots, s_{n+k}$ occur, will be a better approximation to the limit s than the last partial sum s_{n+k} which occurs in the ratio of determinants.

We shall see later that this is indeed the case. Actually, with the help of our simple arguments we found the determinantal representation of the sequence transformation $d_k^{(n)}(\beta, s_n)$ with $\beta = 1$. This transformation, which will be defined later in eq. (7.3-9), is a special case of a very powerful sequence transformation which was introduced by Levin [28]. Levin's general sequence transformation and its numerous variants are discussed quite extensively in section 7 of this report. In section 13 it is also shown that $d_k^{(n)}(\beta, s_n)$, eq. (7.3-9), is able to sum even wildly divergent series and to accelerate the convergence of linearly convergent series.

Our approach did not lead to a representation of this sequence transformation $d_k^{(n)}(\beta, s_n)$, eq. (7.3-9), which is completely satisfactory from a computational point of view. Determinantal representations of sequence transformations are computationally quite unattractive since the reliable and economical evaluation of determinants is a more or less unsolved problem of numerical mathematics. Consequently, it is important to find other methods for the computation of a sequence transformation. In section 7, it will be shown how the determinantal representation (3.2-8) can be replaced by other representations which are better suited for numerical work. However, the concepts and principles, which will be used for the derivation of a large part of the sequence transformations of this report, should now be clear. They can be summarized as follows:

(1): Consider a model sequence with elements $\tilde{s}_n = s + \tilde{r}_n$ and assume that their remainders $\tilde{r}_n$ can be partitioned into a remainder estimate ω_n multiplied by some other quantity z_n. This implies that the elements $\tilde{s}_n$ of the model sequence satisfy:

$$\tilde{s}_n = s + \omega_n z_n, \qquad n \in \mathbb{N}_0. \tag{3.2-9}$$

(2): Assume that an operator $\hat{T}$, which is defined on finite subsets of sequences and which is linear, annihilates the quantities z_n defined in eq. (3.2-9), i.e., $\hat{T}(z_n) = 0$. If we rewrite eq. (3.2-9) in the following way

$$(\tilde{s}_n - s)/\omega_n = z_n, \tag{3.2-10}$$

we see that a sequence transformation $\mathscr{T}(\tilde{s}_n, \omega_n)$, which is exact for the model sequence eq. (3.2-9), i.e., which satisfies $\mathscr{T}(\tilde{s}_n, \omega_n) = s$, is given by the following ratio:

$$\mathscr{T}(\tilde{s}_n, \omega_n) = \frac{\hat{T}(\tilde{s}_n/\omega_n)}{\hat{T}(1/\omega_n)}. \tag{3.2-11}$$

(3): Replace the elements $\tilde{s}_n$ of the model sequence (3.2-9) in the expression defining the sequence transformation $\mathscr{T}(\tilde{s}_n, \omega_n)$ – in this report either the ratio of two determinants, an explicit expression, or a recursive scheme – by the elements of the sequence $\{s_n\}$ which is to be transformed.

The crucial step in this approach is the choice of an appropriate sequence of model remainders $\{\tilde{r}_n\}$ since the $\tilde{r}_n$ should have a mathematical structure which permits the construction of a manageable annihilation operator $\hat{T}$. In addition, the $\tilde{r}_n$ should also be capable of producing good approximations for remainders r_n which occur in actual problems, because only then we may hope that the sequence of transforms will converge more rapidly than the original sequence $\{s_n\}$. These aims are usually accomplished by partitioning $\tilde{r}_n$ into a remainder estimate ω_n multiplied by a finite sum,

$$\tilde{r}_n = \omega_n \sum_{j=0}^{m} c_j \varphi_j(n), \qquad m, n \in \mathbb{N}_0. \tag{3.2-12}$$

Here, the $\{\varphi_j(n)\}$ with $j, n \in \mathbb{N}_0$ are a suitable set of functions – usually an asymptotic sequence as for instance $(n+1)^{-j}$ with $j, n \in \mathbb{N}_0$ – for which a sufficiently simple annihilation operator can be found and the c_j are so far completely unspecified coefficients which are responsible for the flexibility of this ansatz.

Once the asymptotic sequence $\{\varphi_j(n)\}$ is chosen, the crucial problem is the determination of a suitable sequence of remainder estimates $\{\omega_n\}$.

Although it may not be obvious at first sight, the determinantal expression (3.2-8) is exactly of the form of eq. (3.2-11). To see this, one only has to divide both determinants in eq. (3.2-8) by the product $(-1)^{n+1} b_{n+1} \cdots (-1)^{n+k+1} b_{n+k+1}$.

For the sake of simplicity no distinction between the elements of a model sequence $\{\tilde{s}_n\}$ and the elements of a sequence $\{s_n\}$, which is to be transformed, will be made from here on. This means that if a sequence transformation is constructed on the basis of a model sequence, then in the explicit expression or the recursive scheme, which defines this transformation, the sequence elements s_n and not the elements $\tilde{s}_n$ of the model sequence will occur.

3.3. *The general extrapolation algorithm by Brezinski and Håvie*

It is a typical feature of a large part of the modern nonlinear sequence transformations that they are by construction exact for special model sequences. The most general ansatz described in the literature was introduced independently by Brezinski [31] and Håvie [32]. They assume a model sequence of the following type:

$$s_n = s + \sum_{j=0}^{k-1} c_j f_j(n), \qquad k, n \in \mathbb{N}_0. \tag{3.3-1}$$

Concerning the set $\{f_j(n)\}$ with $j, n \in \mathbb{N}_0$ it is assumed that the $f_j(n)$ are known functions of n, but otherwise, they are essentially completely arbitrary. Hence, the ansatz (3.3-1) incorporates convergent as well as divergent sequences, depending upon the behaviour of the functions $f_j(n)$ as $n \to \infty$.

In eq. (3.3-1), there occur $k+1$ unknowns, the limit or antilimit s and the k coefficients $c_0, \ldots, c_{k-1}$. Since all the unknowns in eq. (3.3-1) occur linearly, the numerical values of $k+1$ sequence elements $s_n, s_{n+1}, \ldots, s_{n+k}$ have to be known in order to be able to determine the limit or antilimit s with the help of Cramer's rule. Consequently, the general extrapolation algorithm $E_k(s_n)$ by Brezinski and Håvie, which is by design exact for sequences of the type of eq. (3.3-1), can be formulated as the ratio of two determinants,

$$E_k(s_n) = \frac{\begin{vmatrix} s_n & \ldots & s_{n+k} \\ f_0(n) & \ldots & f_0(n+k) \\ \vdots & \ddots & \vdots \\ f_{k-1}(n) & \ldots & f_{k-1}(n+k) \end{vmatrix}}{\begin{vmatrix} 1 & \ldots & 1 \\ f_0(n) & \ldots & f_0(n+k) \\ \vdots & \ddots & \vdots \\ f_{k-1}(n) & \ldots & f_{k-1}(n+k) \end{vmatrix}}. \tag{3.3-2}$$

Brezinski and Håvie were also able to derive a recursive scheme for the computation of the transforms $E_k(s_n)$, which is, however, relatively complicated. A description of a FORTRAN IV program, which computes the transforms $E_k(s_n)$ via this recursive scheme, can be found in ref. [36].

Brezinski [31] showed that the general extrapolation algorithm $E_k(s_n)$ contains the majority of the currently known sequence transformations as special cases, among them Levin's sequence transformation [28]. It will be seen later that many of the new sequence transformations, which will be discussed in this report, are actually special cases of the general extrapolation algorithm $E_k(s_n)$.

Consequently, it may seem that it is sufficient to consider only the general extrapolation algorithm $E_k(s_n)$ and not its numerous special cases. However, the complicated structure of its recursive scheme [31,32,36] makes the general extrapolation algorithm $E_k(s_n)$ computationally much less efficient than its special cases. Of considerable importance is also the following aspect: In practical applications it is certainly helpful to know that for arbitrary functions $f_j(n)$ the sequence transformation $E_k(s_n)$ can be computed recursively. But it is clearly of greater practical relevance to find out which set $\{f_j(n)\}$ produces the best results for a given sequence $\{s_n\}$. Questions of that kind can only be answered by studying special transformations and by exploiting their specific properties.

3.4. Iterated sequence transformations

In section 3.3, it was mentioned that a large part of the modern nonlinear sequence transformations are constructed in such a way that they are exact for certain model sequences. However, it is also possible to find new sequence transformations which are not constructed on the basis of model sequences.

Let us assume that a sequence $\{s_n\}$ is to be transformed by a sequence transformation $T_k^{(n)}$ with $k,\ n \in \mathbb{N}_0$ and that for some $\kappa \in \mathbb{N}$, which is usually a relatively small number, a transform $T_\kappa^{(n)}$ can be expressed explicitly in terms of the sequence elements $s_n,\ s_{n+1}, \ldots, s_{n+\lambda}$, i.e.,

$$T_\kappa^{(n)} = F(s_n,\ s_{n+1}, \ldots, s_{n+\lambda}). \tag{3.4-1}$$

Then, a new sequence transformation $\Theta_k^{(n)}$ can be obtained by iterating the expression for $T_\kappa^{(n)}$. This means that we define

$$\Theta_0^{(n)} = s_n, \qquad n \in \mathbb{N}_0, \tag{3.4-2}$$

and that eq. (3.4-1) is rewritten in the following way:

$$\Theta_1^{(n)} = F_0\big(\Theta_0^{(n)},\ \Theta_0^{(n+1)}, \ldots, \Theta_0^{(n+\lambda)}\big), \qquad n \in \mathbb{N}_0. \tag{3.4-3}$$

This relationship can now be used to construct a recursive scheme by means of which the transforms $\Theta_k^{(n)}$ with $k \geq 2$ can be computed. One simple possibility of obtaining a recursive scheme would be to assume that eq. (3.4-3) corresponds to the special case $k = 0$ of the following, more general recursive scheme:

$$\Theta_{k+1}^{(n)} = F_k\big(\Theta_k^{(n)},\ \Theta_k^{(n+1)}, \ldots, \Theta_k^{(n+\lambda)}\big), \qquad k,\ n \in \mathbb{N}_0. \tag{3.4-4}$$

Later, we shall encounter several very powerful sequence transformations which are derived by iterating explicit expressions for other sequence transformations. A well known example is Aitken's iterated Δ^2 process which is obtained by iterating the explicit expression for Aitken's Δ^2 process, eq. (1.2-2). Interestingly, it often happens that the properties of the new sequence transformation differ significantly from the properties of the transformation from which it was derived.

4. The epsilon algorithm and related topics

4.1. The Shanks transformation

In his article on nonlinear sequence transformations Shanks [15] considered the following model sequence:

$$s_n = s + \sum_{j=0}^{k-1} c_j \Delta s_{n+j}, \qquad n \in \mathbb{N}_0. \tag{4.1-1}$$

If the sequence elements s_n are partial sums of an infinite series,

$$s_n = \sum_{\nu=0}^{n} a_\nu, \tag{4.1-2}$$

the above model sequence can also be rewritten in the following way:

$$s_n = s + \sum_{j=0}^{k-1} c_j a_{n+j+1}, \qquad n \in \mathbb{N}_0. \tag{4.1-3}$$

Essentially this means that the limit s of the infinite series is approximated by the partial sum s_n plus a weighted sum of the next k terms $a_{n+1}, a_{n+2}, \dots, a_{n+k}$. As in the previous examples the model sequence s_n contains $k+1$ unknowns – the limit or antilimit s and the k linear coefficients $c_0, \dots, c_{k-1}$ – which all occur linearly. Consequently, according to Cramer's rule the sequence transformation $e_k(s_n)$, which is by construction exact for the model sequence (4.1-1), can be defined by the following ratio of determinants:

$$e_k(s_n) = \frac{\begin{vmatrix} s_n & \cdots & s_{n+k} \\ \Delta s_n & \cdots & \Delta s_{n+k} \\ \vdots & \ddots & \vdots \\ \Delta s_{n+k-1} & \cdots & \Delta s_{n+2k-1} \end{vmatrix}}{\begin{vmatrix} 1 & \cdots & 1 \\ \Delta s_n & \cdots & \Delta s_{n+k} \\ \vdots & \ddots & \vdots \\ \Delta s_{n+k-1} & \cdots & \Delta s_{n+2k-1} \end{vmatrix}}. \tag{4.1-4}$$

For the computation of the transform $e_k(s_n)$ the sequence elements $s_n, \dots, s_{n+2k}$ are needed. This implies that $e_k(s_n)$ is a transformation of order $2k$.

The sequence transformation $e_k(s_n)$ is a special case of the general extrapolation algorithm $E_k(s_n)$ introduced by Brezinski [31] and Håvie [32]. To see this we only have to replace $f_j(n)$ in eq. (3.3-2) by Δs_{n+j}. The transformation (4.1-4) was originally introduced in 1941 by Schmidt

[17] who used it for the iterative solution of linear systems. In 1955 it was rediscovered by Shanks [15] who also analyzed some of the mathematical properties of the sequence transformation $e_k(s_n)$ and derived several interesting results. For instance, he was able to show that this transformation is also exact for model sequences with remainders that are sums of exponentials:

$$s_n = s + \sum_{j=0}^{k-1} c_j \lambda_j^n, \qquad n \in \mathbb{N}_0. \tag{4.1-5}$$

Concerning the λ_j it is assumed that they are ordered in such a way that their magnitudes are strictly decreasing, i.e.,

$$|\lambda_0| > |\lambda_1| > \cdots > |\lambda_{k-1}|. \tag{4.1-6}$$

If the condition $|\lambda_0| < 1$ is satisfied the model sequence (4.1-5) converges. In analogy with so-called *physical transients*, which disappear after a sufficiently long time, Shanks called the terms on the right-hand side of eq. (4.1-5) *mathematical transients* since all λ_j^n with $|\lambda_j| < 1$ vanish as $n \to \infty$. However, this concept of a mathematical transient has to be used here in a broader sense since the Shanks transformation $e_k(s_n)$ can also be used for the summation of divergent sequences and series. The model sequence (4.1-5) diverges if at least one of the λ_j satisfies $|\lambda_k| \geq 1$, because then such a term λ_k^n will not vanish as $n \to \infty$.

Shanks [15] also showed that the transformation $e_k(s_n)$ and Padé approximants are closely related. Let us assume that $f(z)$ is analytic in a neighbourhood of $z = 0$,

$$f(z) = \sum_{\nu=0}^{\infty} \alpha_\nu z^\nu. \tag{4.1-7}$$

Following the notation of Baker and Graves-Morris [22] we say that the Padé approximant of $f(z)$ is the ratio of two polynomials $p_l(z)$ and $q_m(z)$ of degrees l and m, respectively, and write

$$[l/m]_f(z) = p_l(z)/q_m(z). \tag{4.1-8}$$

The Padé approximants $[l/m]$ with l, $m \in \mathbb{N}_0$ are displayed in a 2-dimensional rectangular scheme called the *Padé table* in such a way that the first index l indicates the column and the second index m the row of the array. The coefficients of the two polynomials $p_l(z)$ and $q_m(z)$ are defined by the relationship

$$f(z) - p_l(z)/q_m(z) = \mathrm{O}(z^{l+m+1}). \tag{4.1-9}$$

This implies that the coefficients of the Taylor expansion of the Padé approximant $p_l(z)/q_m(z)$ have to agree with the series coefficients α_k up to the coefficient α_{l+m}. Let $f_n(z)$ stand for a partial sum of the power series (4.1-7),

$$f_n(z) = \sum_{\nu=0}^{n} \alpha_\nu z^\nu. \tag{4.1-10}$$

Shanks [15] could show that the application of his transformation to the sequence $\{f_n(z)\}$ produces the following elements of the Padé table:

$$e_k(f_n(z)) = [n+k/k]_f(z), \qquad k, n \in \mathbb{N}_0. \tag{4.1-11}$$

4.2. *Wynn's epsilon algorithm*

As it stands, the Shanks transformation $e_k(s_n)$, eq. (4.1-4), is not particularly useful because of its definition as the ratio of two determinants. Fortunately, only one year after the publication of Shanks' article [15] Wynn [16] found a nonlinear recursive scheme which is now commonly called the ϵ algorithm:

$$\epsilon_{-1}^{(n)} = 0, \qquad \epsilon_0^{(n)} = s_n, \tag{4.2-1a}$$

$$\epsilon_{k+1}^{(n)} = \epsilon_{k-1}^{(n+1)} + 1/\left[\epsilon_k^{(n+1)} - \epsilon_k^{(n)}\right], \quad k, n \in \mathbb{N}_0. \tag{4.2-1b}$$

Wynn [16] was able to show that the elements of the ϵ table with even subscripts give the Shanks transformation,

$$\epsilon_{2k}^{(n)} = e_k(s_n), \qquad k, n \in \mathbb{N}_0, \tag{4.2-2}$$

whereas the elements of the ϵ table with odd subscripts are only auxiliary quantities satisfying

$$\epsilon_{2k+1}^{(n)} = 1/e_k(\Delta s_n), \qquad k, n \in \mathbb{N}_0. \tag{4.2-3}$$

The publication of the ϵ algorithm, which allows a simple and efficient evaluation of the Shanks transformation, stimulated an enormous amount of research. According to Wimp (see p. 120 of ref. [23]) over 50 articles on the ϵ algorithm were published by Wynn alone, and at least 30 articles by Brezinski. As a fairly complete source of references Wimp recommends Brezinski's first book [19]. Since the main concern of this report are other sequence transformations and not Wynn's ϵ algorithm, only those properties of the ϵ algorithm will be discussed which are relevant for an understanding of its power as well as its limitations as a convergence acceleration and summation method.

In a later article, Wynn [37] analyzed the convergence properties of the ϵ algorithm by applying it to several model sequences. For instance, in the case of sequences, which have strictly alternating remainders r_n of the following type,

$$s_n \sim s + (-1)^n \sum_{j=0}^{\infty} c_j/(n+\beta)^{j+1}, \qquad \beta \in \mathbb{R}_+, \qquad n \to \infty, \tag{4.2-4}$$

Wynn [37] obtained, assuming $c_0 \neq 0$, for fixed k an estimate which shows that the ϵ algorithm accelerates convergence:

$$\epsilon_{2k}^{(n)} \sim s + \frac{(-1)^n (k!)^2}{2^{2k}(n+\beta)^{2k+1}} c_0, \quad n \to \infty. \tag{4.2-5}$$

Wynn [37] also considered sequences which generalize the model sequence (4.1-5),

$$s_n \sim s + \sum_{j=0}^{\infty} c_j \lambda_j^n, \qquad 1 > \lambda_0 > \lambda_1 > \cdots > 0, \quad n \to \infty. \tag{4.2-6}$$

For fixed k Wynn [37] obtained the following estimate which shows that the ϵ algorithm accelerates convergence:

$$\epsilon_{2k}^{(n)} \sim s + c_k \frac{\{(\lambda_k - \lambda_0)(\lambda_k - \lambda_1) \ldots (\lambda_k - \lambda_{k-1})\}^2}{\{(1-\lambda_0)(1-\lambda_1) \ldots (1-\lambda_{k-1})\}^2} \lambda_k^n, \qquad n \to \infty. \tag{4.2-7}$$

Wynn [37] also applied the ϵ algorithm to logarithmically convergent sequences of the following type:

$$s_n \sim s + \sum_{j=0}^{\infty} c_j/(n+\beta)^{j+1}, \qquad \beta \in \mathbb{R}_+, \quad n \to \infty. \tag{4.2-8}$$

Assuming $c_0 \neq 0$, Wynn [37] obtained for fixed k an estimate which shows that the ϵ algorithm – or equivalently the Shanks transformation – is not able to accelerate the convergence of sequences of that type:

$$\epsilon_{2k}^{(n)} \sim s + \frac{c_0}{(k+1)(n+\beta)}, \qquad n \to \infty. \tag{4.2-9}$$

This inability of accelerating logarithmic convergence is one of the major defects of the otherwise very powerful ϵ algorithm.

If the elements s_n of the sequence to be transformed are partial sums of a power series as in eq. (4.1-10), the ϵ algorithm produces Padé approximants according to eqs. (4.1-11) and (4.2-2),

$$\epsilon_{2k}^{(n)} = [n+k/k]_f(z), \qquad k, n \in \mathbb{N}_0, \tag{4.2-10}$$

and the convergence theory of Padé approximants can be applied in this case. This convergence theory can be found in the standard references on Padé approximants, as for instance the books by Baker [18] or by Baker and Graves-Morris [22].

4.3. Programming the epsilon algorithm

In this section it will be discussed how the ϵ algorithm can be programmed efficiently. But first, the objectives of such a program, which performs the transformation of a given sequence $\{s_n\}$ with the help of Wynn's ϵ algorithm, should be stated.

Since it is in general not possible to predict by a priori considerations how many sequence elements will be needed until convergence has finally been achieved, such a program should be input-directed. This means it should read in the sequence elements $s_0, s_1, \ldots, s_m, \ldots$ successively starting with s_0. After the input of each new sequence element s_m as many new elements $\epsilon_k^{(n)}$

should be computed as is permitted by the recurrence formula (4.2-1b). That new element $\epsilon_k^{(n)}$, which has the largest even subscript k, should be used as the new approximation to the limit of the sequence.

Let us arrange the elements $\epsilon_k^{(n)}$ of the ϵ table in a rectangular scheme in such a way that the superscript n indicates the row and the subscript k the column of the 2-dimensional array:

$$\begin{array}{cccccc}
\epsilon_0^{(0)} & \epsilon_1^{(0)} & \epsilon_2^{(0)} & \cdots & \epsilon_n^{(0)} & \cdots \\
\epsilon_0^{(1)} & \epsilon_1^{(1)} & \epsilon_2^{(1)} & \cdots & \epsilon_n^{(1)} & \cdots \\
\epsilon_0^{(2)} & \epsilon_1^{(2)} & \epsilon_2^{(2)} & \cdots & \epsilon_n^{(2)} & \cdots \\
\epsilon_0^{(3)} & \epsilon_1^{(3)} & \epsilon_2^{(3)} & \cdots & \epsilon_n^{(3)} & \cdots \\
\vdots & \vdots & \vdots & \ddots & \vdots & \ddots \\
\epsilon_0^{(n)} & \epsilon_1^{(n)} & \epsilon_2^{(n)} & \cdots & \epsilon_n^{(n)} & \cdots \\
\vdots & \vdots & \vdots & \ddots & \vdots & \ddots
\end{array} \tag{4.3-1}$$

The entries in the first column of the array are the starting values $\epsilon_0^{(n)} = s_n$ of the recursion according to eq. (4.2-1a). The remaining elements of the ϵ table can be computed with the help of the recurrence formula (4.2-1b). This nonlinear 4-term recursion connects four elements of the ϵ table which are located at the vertices of a rhombus:

$$\begin{array}{ccc}
 & \epsilon_k^{(n)} & \epsilon_{k+1}^{(n)} \\
\epsilon_{k-1}^{(n+1)} & \epsilon_k^{(n+1)} &
\end{array} \tag{4.3-2}$$

If the sequence elements $s_0, s_1, \ldots, s_m$ are used as initial conditions according to eq. (4.2-1a), the recurrence formula (4.2-1b) is able to compute all elements $\epsilon_j^{(\mu-j)}$ with $0 \leq \mu \leq m$ and $0 \leq j \leq \mu$. Obviously, these elements form an equilateral triangle located in the upper left corner of the ϵ table. If the next sequence element s_{m+1} is also used as a starting value for the recursion (4.2-1b), the triangle will be enlarged by the neighbouring counterdiagonal $\epsilon_j^{(m-j+1)}$ with $0 \leq j \leq m+1$. In this context it will be advantageous to rewrite Wynn's ϵ algorithm, eq. (4.2-1), in the following way:

$$\epsilon_0^{(n)} = s_n, \qquad n \geq 0, \tag{4.3-3a}$$

$$\epsilon_1^{(n-1)} = 1/[s_n - s_{n-1}], \qquad n \geq 1, \tag{4.3-3b}$$

$$\epsilon_j^{(n-j)} = \epsilon_{j-2}^{(n-j+1)} + 1/\left[\epsilon_{j-1}^{(n-j+1)} - \epsilon_{j-1}^{(n-j)}\right], \quad n \geq 2, \quad 2 \leq j \leq n. \tag{4.3-3c}$$

Concerning the approximations to the limit it follows from eqs. (4.2-2) and (4.2-3) that one has to distinguish between even and odd values of the index m of the last sequence element s_m which was used in the recursion. If m is even, $m = 2\mu$, our approximation to the limit of the sequence is the transformation

$$\{s_0, s_1, \ldots, s_{2\mu}\} \to \epsilon_{2\mu}^{(0)} = e_\mu(s_0), \tag{4.3-4}$$

and if m is odd, $m = 2\mu + 1$, we use the transformation

$$\{s_1, s_2, \ldots, s_{2\mu+1}\} \to \epsilon_{2\mu}^{(1)} = e_\mu(s_1). \tag{4.3-5}$$

With the help of the notation $[\![x]\!]$ for the integral part of x, i.e., the largest integer ν satisfying $\nu \leq x$, these two relationships can be combined into a single equation yielding

$$\{s_{m-2[\![m/2]\!]}, s_{m-2[\![m/2]\!]+1}, \ldots, s_m\} \to \epsilon_{2[\![m/2]\!]}^{(m-2[\![m/2]\!])} = e_{[\![m/2]\!]}(s_{m-2[\![m/2]\!]}). \tag{4.3-6}$$

If the sequence elements s_n are partial sums of a power series as in eq. (4.1-10), our approximations to the limit correspond according to eq. (4.2-10) to the following staircase sequence in the Padé table:

$$[0/0], [1/0], [1/1], \ldots, [\nu/\nu], [\nu+1/\nu], [\nu+1/\nu+1], \ldots. \tag{4.3-7}$$

Because of the rhombus structure (4.3-2) of the 4-term recursion in Wynn's ϵ algorithm it appears that a program would either need a single 2-dimensional or at least two 1-dimensional arrays. However, Wynn [38] could show that a single 1-dimensional array is sufficient. Wynn's algorithm, which is called *moving lozenge technique*, is based upon the observation that for the computation of a new element $\epsilon_j^{(m-j+1)}$ only the two neighbouring elements $\epsilon_{j-1}^{(m-j+1)}$ and $\epsilon_{j-2}^{(m-j+2)}$ have to be known but not the whole upper counterdiagonal $\epsilon_\mu^{(m-\mu)}$ with $0 \leq j \leq \mu \leq m$. Hence, in Wynn's moving lozenge technique these quantities are stored in auxiliary variables while the recursion (4.3-3) moves along the current counterdiagonal $\epsilon_j^{(m-j+1)}$ with $0 \leq j \leq m+1$ and overwrites the previous entries $\epsilon_\mu^{(m-\mu)}$ with $0 \leq \mu \leq m$. A good description of Wynn's moving lozenge technique [38] can also be found in Brezinski's second book (see pp. 326–327 of ref. [20]).

Wynn [38] performed the recursion in a 1-dimensional array E in which he stored the elements of the current counterdiagonal of the ϵ table in such a way that the index of the array element coincides with the subscript of the corresponding element of the ϵ table,

$$\epsilon_j^{(m-j)} \to E(j), \qquad m \geq 0, \quad 0 \leq j \leq m. \tag{4.3-8}$$

If the above convention is used three auxiliary variables will be needed. But Wynn's moving lozenge technique can be improved if the elements of the current counterdiagonal of the ϵ table are stored in a 1-dimensional array E in such a way that the superscript of the corresponding element of the ϵ table coincides with the index of the array element,

$$\epsilon_j^{(m-j)} \to E(m-j), \qquad m \geq 0, \quad 0 \leq j \leq m. \tag{4.3-9}$$

If this convention is used only two auxiliary variables will be needed and the structure of the resulting computer program will also be simpler and more elegant. The recursive scheme (4.3-3) can then be reformulated in terms of the elements of the array E in the following way:

$$E(n) \qquad \leftarrow s_n, \qquad n \geq 0, \tag{4.3-10a}$$

$$E(n-1) \qquad \leftarrow 1/[E(n) - E'(n-1)], \qquad n \geq 1, \tag{4.3-10b}$$

$$E(n-j) \qquad \leftarrow E'(n-j+1) + 1/[E(n-j+1) - E'(n-j)], \quad n \geq 2, \quad 2 \leq j \leq n. \tag{4.3-10c}$$

The primed array elements $E'(n-j)$ and $E'(n-j+1)$ have to be stored in auxiliary variables since they will be overwritten during the computation of the current counterdiagonal $\epsilon_j^{(n-j)}$ with $0 \leq j \leq n$. The primes also indicate that the indices of the array elements $E'(n-j)$ and $E'(n-j+1)$ refer to the occupation of E after the previous run, i.e., after the computation of the counterdiagonal $\epsilon_j^{(n-j-1)}$ with $0 \leq j \leq n-1$.

If a counterdiagonal $\epsilon_\mu^{(m-\mu)}$ with $0 \leq \mu \leq m$ is computed with the help of the recursive scheme (4.3-3) and if the elements $\epsilon_{2k}^{(n)}$ converge – which means that the whole process is successful – the computation of the elements $\epsilon_{2k+1}^{(n)}$ will necessarily involve divisions by the small quantities $\epsilon_{2k}^{(n+1)} - \epsilon_{2k}^{(n)}$. This may easily lead to an intolerable magnification of the inevitable rounding errors. Hence, it looks as if Wynn's ϵ algorithm should be extremely susceptible to rounding errors.

Fortunately, this is normally not the case although the elements $\epsilon_{2k+1}^{(n)}$ may become quite large in magnitude and may have a very low relative accuracy due to the numerical problems described above. But in the next step of the recursion the elements $\epsilon_{2k+1}^{(n)}$ serve as divisors which will dampen the rounding errors again. Consequently, it is not clear what the overall effect will be. However, numerical experience indicates that in most cases of practical interest Wynn's ϵ algorithm is remarkably stable. This experimental evidence is supported by a theoretical analysis of the numerical stability of the ϵ algorithm which was performed by Wynn [37] in the case of several model sequences.

In some cases – for instance if the elements of the sequence to be transformed are the partial sums of the Taylor series of a rational function – it may happen that the difference $\epsilon_{2k}^{(n+1)} - \epsilon_{2k}^{(n)}$ vanishes. If pathologies of that kind occur, the so-called *singular rules* of the ϵ algorithm can be used which were also derived by Wynn [39]. A good discussion of these singular rules and of related problems can also be found in section 4.1.2 of Brezinski's second book [20]. There, one can also find listings of FORTRAN IV programs for Wynn's ϵ algorithm which partly use the singular rules mentioned above (see pp. 338–352 of ref. [20]).

According to the limited experience of the author pathologies of that kind occur only rarely in scientific applications. Consequently, Wynn's singular rules are not used in the following FORTRAN 77 program EPSAL which computes the Shanks transformation of a given sequence by means of Wynn's ϵ algorithm. However, a good program should take some precautions against an approximate equality of the elements $\epsilon_k^{(n+1)}$ and $\epsilon_k^{(n)}$ since in this case the reciprocal of the difference of these two elements could exceed the largest floating point number representable on the computer. This would lead to overflow and to an error termination of the program.

This safeguard against overflow can be accomplished by defining two variables HUGE and TINY. Their values should be close to but not identical with the largest and smallest floating point numbers representable on the computer. If the difference $\epsilon_k^{(n+1)} - \epsilon_k^{(n)}$ is smaller in magnitude than TINY, then $\epsilon_{k+1}^{(n)}$ will be set equal to HUGE. If this approximate equality of the elements $\epsilon_k^{(n+1)}$ and $\epsilon_k^{(n)}$ was accidental the program can continue with the computation of the other elements of the ϵ table producing numbers which are normally not noticably affected. It is also possible that this approximate equality of the elements $\epsilon_k^{(n+1)}$ and $\epsilon_k^{(n)}$ was not accidental but due to convergence. However, in this case the program should have been stopped before.

In order to monitor the exact or approximate vanishing of the denominators, it may also be a good idea to define in a FORTRAN program for Wynn's ϵ algorithm an error variable, for instance IFAIL, whose value is changed if one of the differences $\epsilon_k^{(n+1)} - \epsilon_k^{(n)}$ is smaller in magnitude than TINY.

The following FORTRAN 77 program EPSAL uses the modification (4.3-10) of Wynn's moving lozenge technique. It is safeguarded against approximate equality of the elements $\epsilon_k^{(n+1)}$ and $\epsilon_k^{(n)}$ by using two variables HUGE and TINY as described above. The elements s_n with $n = 0, 1, 2, \ldots$ of the sequence to be transformed have to be computed in a DO loop in the calling program. Whenever a new sequence element s_n is computed in the outer DO loop this subroutine EPSAL has to be called again and a new counterdiagonal of the ϵ table will be calculated. The new sequence element s_n is read in via the variable SOFN and the approximation to the limit is returned via the variable ESTLIM.

Finally, it is important to note that this subroutine EPSAL only calculates the approximations to the limit according to eqs. (4.3-3) and (4.3-4) and does not analyze the convergence of the whole process. This has to be done in the calling program.

```
      SUBROUTINE EPSAL(SOFN,N,E,LARRAY,ESTLIM)
      DIMENSION E(0:LARRAY)
      PARAMETER (HUGE = 1.E+60, TINY = 1.E-60, ZERO = 0.E0, ONE = 1.E0)
      E(N)  = SOFN
      IF (N.EQ.0) THEN
        ESTLIM  = SOFN
      ELSE
        AUX2    = ZERO
        DO 10 J = N,1,-1
          AUX1  = AUX2
          AUX2  = E(J-1)
          DIFF  = E(J) - AUX2
          IF (ABS(DIFF) .LE. TINY) THEN
            E(J-1) = HUGE
          ELSE
            E(J-1) = AUX1 + ONE/DIFF
          END IF
10      CONTINUE
        IF ( MOD(N,2) .EQ. 0 ) THEN
          ESTLIM = E(0)
        ELSE
          ESTLIM = E(1)
        END IF
      END IF
      RETURN
      END
```

5. The iteration of Aitken's Δ^2 process

5.1. Aitken's Δ^2 transformation and its iteration

Let us consider the following model sequence which is obtained by setting $k = 1$ in eq. (4.1-5):

$$s_n = s + c\lambda^n, \qquad c \neq 0,\ |\lambda| < 1,\ n \in \mathbb{N}_0. \tag{5.1-1}$$

Each sequence element s_n contains the three unknowns c, λ, and the limit s. Consequently, a sequence transformation will at least require three elements of the above model sequence for the determination of the limit s. In order to derive such a transformation we form the first and second differences of s_n:

$$\Delta s_n = c\lambda^n(\lambda - 1), \tag{5.1-2}$$

$$\Delta^2 s_n = c\lambda^n(\lambda - 1)^2. \tag{5.1-3}$$

A short computation shows that the following sequence transformation is exact for the model sequence (5.1-1):

$$\mathscr{A}_1^{(n)} = s_n - \frac{[\Delta s_n]^2}{\Delta^2 s_n}, \qquad n \in \mathbb{N}_0. \tag{5.1-4}$$

This sequence transformation is Aitken's well-known Δ^2 process [12]. The structure of this transformation explains quite clearly why it bears this name.

It follows at once from the derivation of the sequence transformation $\mathscr{A}_1^{(n)}$ via the model sequence (5.1-1) that it is a special case of the Shanks transformation, eq. (4.1-4), or Wynn's ϵ algorithm, eq. (4.2-1),

$$\mathscr{A}_1^{(n)} = e_1(s_n) = \epsilon_2^{(n)}. \tag{5.1-5}$$

Many other representations for Aitken's Δ^2 process can be derived by suitable manipulations of eq. (5.1-4). Examples are:

$$\mathscr{A}_1^{(n)} = s_{n+1} - \frac{[\Delta s_n][\Delta s_{n+1}]}{\Delta^2 s_n} \tag{5.1-6}$$

$$= s_{n+2} - \frac{[\Delta s_{n+1}]^2}{\Delta^2 s_n} \tag{5.1-7}$$

$$= \frac{s_{n+2}s_n - [s_{n+1}]^2}{\Delta^2 s_n} \tag{5.1-8}$$

$$= \frac{[\Delta s_{n+1}] s_{n+1} - [\Delta s_n] s_{n+2}}{\Delta^2 s_n} \tag{5.1-9}$$

$$= \frac{[\Delta s_{n+1}] s_n - [\Delta s_n] s_{n+1}}{\Delta^2 s_n} \tag{5.1-10}$$

$$= s_{n+1} + \frac{1}{\Delta[1/\Delta s_n]}, \tag{5.1-11}$$

$$= \frac{\Delta[s_{n+1}/\Delta s_n]}{\Delta[1/\Delta s_n]}. \tag{5.1-12}$$

Aitken's Δ^2 process was studied in articles by Shanks [15], Clark, Gray, and Adams [35], Lubkin [40], Tucker [41,42], Cordellier [43], and Bell and Phillips [44]. A multidimensional generalization of Aitken's transformation to vector sequences was discussed by MacLeod [45]. Modifications of Aitken's Δ^2 process were proposed by Drummond [46] and by Bjørstad, Dahlquist, and Grosse [47].

The properties of Aitken's Δ^2 process are discussed in books by Brezinski (see pp. 37–40 of ref. [19] and pp. 43–45 of ref. [20]) and Wimp (see pp. 149–152 of ref. [23]). Those properties which are particularly important for our purposes can be summarized as follows:

(i): The Δ^2 process accelerates linear convergence.

(ii): The Δ^2 process is regular but not accelerative for logarithmically convergent sequences of the type of eq. (4.2-8).

This shows that Aitken's Δ^2 process has similar properties as Wynn's ϵ algorithm. In view of eq. (5.1-5) this is not surprising. However, one cannot expect that Aitken's Δ^2 process will be as powerful as Wynn's ϵ algorithm. The reason is that the transform $\mathscr{A}_1^{(n)}$ is produced by only three sequence elements s_n, s_{n+1}, and s_{n+2} which implies that $\mathscr{A}_1^{(n)}$ is a transformation of order $l = 2$. This will certainly limit the power as well as practical usefulness of this transformation.

If the accelerative power of Aitken's transformation turns out to be insufficient and if it is necessary to use a more powerful sequence transformation one could of course use Wynn's ϵ algorithm which because of eq. (5.1-5) can be considered to be a more complex and also more powerful generalization of Aitken's Δ^2 process. Another alternative, which also produces sequence transformations with higher transformation orders, would be to iterate the Δ^2 process. This means that Aitken's Δ^2 process will be applied to the transformed sequence $\{\mathscr{A}_1^{(n)}\}$ yielding a new sequence $\{\mathscr{A}_2^{(n)}\}$. This process can in principle be repeated indefinitely.

In order to obtain some heuristic motivation for this iteration, let us apply Aitken's Δ^2 process to the following model sequence which generalizes the sequence (5.1-1):

$$s_n = s + ax^n + by^n, \qquad 0 < |y| < |x| < 1, \quad a,\ b \neq 0. \tag{5.1-13}$$

A short calculation shows that Aitken's Δ^2 process eliminates the dominating term ax^n from the model sequence (5.1-13):

$$\mathscr{A}_1^{(n)} = s + \frac{b[(x-y)/(x-1)]^2 y^n}{1 + (b/a)[(y-1)/(x-1)]^2 (y/x)^n}. \tag{5.1-14}$$

Since we have by assumption $0 < |y| < |x| < 1$, the transformed sequence (5.1-14) converges faster than the original sequence (5.1-13). Also, since $(y/x)^n$ vanishes as $n \to \infty$, at least for large values of n the elements of the resulting sequence (5.1-14) have essentially the same structure as the elements of the sequence (5.1-1).

For the iteration of Aitken's Δ^2 process each of the numerous representations for $\mathscr{A}_1^{(n)}$ given above can be used since they are all mathematically equivalent. However, the various representations for $\mathscr{A}_1^{(n)}$ differ considerably in their numerical stability. In the book by Press, Flannery, Teukolsky, and Vetterling (see p. 133 of ref. [48]) it is remarked that Aitken's Δ^2 process should be computed with the help of eq. (5.1-4) since the other equivalent representations are numerically less reliable. Numerical studies performed by the author confirmed this statement. Consequently, in this report an iteration of Aitken's Δ^2 process will always be based upon eq. (5.1-4). If we identify the sequence elements s_n with the initial values $\mathscr{A}_0^{(n)}$ of the recursion we obtain the following nonlinear recursive scheme:

$$\mathscr{A}_0^{(n)} = s_n, \tag{5.1-15a}$$

$$\mathscr{A}_{k+1}^{(n)} = \mathscr{A}_k^{(n)} - \frac{\left[\Delta \mathscr{A}_k^{(n)}\right]^2}{\Delta^2 \mathscr{A}_k^{(n)}} \qquad k, n \in \mathbb{N}_0. \tag{5.1-15b}$$

As usual, the difference operator Δ acts upon the superscript n and not upon the subscript k. It follows from this recurrence formula that the computation of $\mathscr{A}_k^{(n)}$ requires the sequence elements $s_n, s_{n+1}, \ldots, s_{n+2k}$. Consequently, $\mathscr{A}_k^{(n)}$ is a transformation of order $2k$. In this respect $\mathscr{A}_k^{(n)}$ is equivalent to $\epsilon_{2k}^{(n)}$ which needs the same set $s_n, s_{n+1}, \ldots, s_{n+2k}$ of sequence elements for its computation. However, we shall see later that the numerical properties of Wynn's ϵ algorithm and Aitken's iterated Δ^2 process often differ considerably although they are both generalizations of the same sequence transformation $\mathscr{A}_1^{(n)}$, eq. (5.1-4).

The numerical properties of Aitken's iterated Δ^2 process were studied by Smith and Ford [30]. Concerning the theoretical properties of Aitken's iterated Δ^2 process, very little seems to be known. Apparently, there is only one article by Hillion [49] in which the theoretical properties of Aitken's iterated Δ^2 process were studied. Hillion was able to find a model sequence for which the iterated Δ^2 process is exact. He also derived a determinantal representation for the transforms $\mathscr{A}_k^{(n)}$. However, Hillion's expressions contain in both cases explicitly the lower order transforms $\mathscr{A}_0^{(n)}, \ldots, \mathscr{A}_{k-1}^{(n)}, \ldots, \mathscr{A}_0^{(n+k)}, \ldots, \mathscr{A}_{k-1}^{(n+k)}$. Consequently, it seems that not much insight about the properties of Aitken's iterated Δ^2 process can be gained by these results.

5.2. Programming Aitken's iterated Δ^2 process

A program for Aitken's iterated Δ^2 process should have the same features as the subroutine EPSAL which transforms a given sequence $\{s_n\}$ with the help of Wynn's ϵ algorithm. This means it should read in the sequence elements $s_0, s_1, \ldots, s_m, \ldots$ successively starting with s_0. After the input of each new sequence element s_m as many new elements $\mathscr{A}_k^{(n)}$ of the Aitken table should be computed as it is permitted by the recursive scheme (5.1-15). That element $\mathscr{A}_k^{(n)}$, which has the largest subskript k, should be used as the new approximation to the limit of the sequence $\{s_n\}$.

Let us arrange the elements $\mathscr{A}_k^{(n)}$ of the Aitken table in rectangular scheme in such a way that the superscript n indicates the row and the subscript k the column of the 2-dimensional array:

$$\begin{array}{cccccc} \mathscr{A}_0^{(0)} & \mathscr{A}_1^{(0)} & \mathscr{A}_2^{(0)} & \cdots & \mathscr{A}_n^{(0)} & \cdots \\ \mathscr{A}_0^{(1)} & \mathscr{A}_1^{(1)} & \mathscr{A}_2^{(1)} & \cdots & \mathscr{A}_n^{(1)} & \cdots \\ \mathscr{A}_0^{(2)} & \mathscr{A}_1^{(2)} & \mathscr{A}_2^{(2)} & \cdots & \mathscr{A}_n^{(2)} & \cdots \\ \mathscr{A}_0^{(3)} & \mathscr{A}_1^{(3)} & \mathscr{A}_2^{(3)} & \cdots & \mathscr{A}_n^{(3)} & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots \\ \mathscr{A}_0^{(n)} & \mathscr{A}_1^{(n)} & \mathscr{A}_2^{(n)} & \cdots & \mathscr{A}_n^{(n)} & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots \end{array} \tag{5.2-1}$$

The entries in the first column of the array are the starting values $\mathscr{A}_0^{(n)} = s_n$ of the recursion according to eq. (5.1-15a). The remaining elements of the Aitken table can be computed with the help of the recurrence formula (5.1-15b). The 4 elements, which are connected by this nonlinear recursion, form a pattern in the Aitken table which looks like the move of a knight on the chessboard:

$$\begin{array}{cc} \mathscr{A}_k^{(n)} & \mathscr{A}_{k+1}^{(n)} \\ \mathscr{A}_k^{(n+1)} & \\ \mathscr{A}_k^{(n+2)} & \end{array} \tag{5.2-2}$$

This pattern implies that the recursion (5.1-15b) has to proceed along a relatively complicated path in the Aitken table if the elements $s_0, s_1, \ldots, s_m, \ldots$ are read in successively and if one tries to increase the subscript k as much as possible. In this context, it is advantageous to rewrite the recursive scheme (5.1-15) in the following way:

$$\mathscr{A}_0^{(n)} = s_n, \qquad n \geq 0, \tag{5.2-3a}$$

$$\mathscr{A}_j^{(n-2j)} = \mathscr{A}_{j-1}^{(n-2j)} - \frac{\left[\Delta \mathscr{A}_{j-1}^{(n-2j)}\right]^2}{\Delta^2 \mathscr{A}_{j-1}^{(n-2j)}}, \qquad n \geq 2, \quad 1 \leq j \leq [\![n/2]\!]. \tag{5.2-3b}$$

Here, $[\![n/2]\!]$ denotes the integral part of $n/2$, i.e., the largest integer ν satisfying $\nu \leq n/2$. If the sequence elements $s_0, s_1, \ldots, s_m$ are used as starting values, the recursion (5.2-3b) is able to compute all elements $\mathscr{A}_j^{(\mu-2j)}$ with $0 \leq \mu \leq m$ and $0 \leq j \leq [\![\mu/2]\!]$. If the next sequence element s_{m+1} is also used as a starting value for the recursion, this set of elements of the Aitken table will be enlarged by the string $\mathscr{A}_j^{(m-2j+1)}$ with $0 \leq j \leq [\![(m+1)/2]\!]$.

As in Wynn's ϵ algorithm the approximation to the limit depends upon the index m of the last sequence element s_m which was used in the recursion. If m is even, $m = 2\mu$, our approximations to the limit of the sequence are the transformations

$$\{s_0, s_1, \ldots, s_{2\mu}\} \to \mathscr{A}_\mu^{(0)}, \tag{5.2-4}$$

and if m is odd, $m = 2\mu + 1$, the approximation to the limit will be

$$\{s_1, s_2, \ldots, s_{2\mu+1}\} \rightarrow \mathscr{A}_{\mu}^{(1)}. \tag{5.2-5}$$

As in the case of Wynn's ϵ algorithm, these two relationships can be combined into a single equation,

$$\{s_{m-2[\![m/2]\!]}, s_{m-2[\![m/2]\!]+1}, \ldots, s_m\} \rightarrow \mathscr{A}_{[\![m/2]\!]}^{(m-2[\![m/2]\!])}. \tag{5.2-6}$$

Because of the relatively complicated geometrical structure (5.2-2) of the recursion (5.2-3b) it appears that a program, which computes Aitken's iterated Δ^2 process, would need a 2-dimensional array. However, a single 1-dimensional array A is sufficient if the elements of the Aitken table are stored according to the following rule:

$$\mathscr{A}_{[\![\nu/2]\!]}^{(n-\nu)} \rightarrow A(n-\nu), \qquad n \geq 0, \quad 0 \leq \nu \leq n. \tag{5.2-7}$$

With this convention the recurrence formula (5.2-3) can be reformulated in terms of the elements of the 1-dimensional array A:

$$A(n) \leftarrow s_n, \qquad n \geq 0, \tag{5.2-8a}$$

$$A(n-2j) \leftarrow A(n-2\,j) - \frac{[\Delta A(n-2j)]^2}{\Delta^2 A(n-2j)}, \qquad n \geq 2, \quad 1 \leq j \leq [\![n/2]\!]. \tag{5.2-8b}$$

Aitken's iterated Δ^2 transformation makes sense only if the second differences $\Delta^2 \mathscr{A}_{k-1}^{(n)}$ do not vanish for sufficiently large values of k. This will certainly be guaranteed if for fixed k all elements of the sequence $\{\mathscr{A}_{k-1}^{(n)}\}$ are different from zero and strictly alternating in sign.

Unfortunately, the above statement is not particularly helpful since only very little is known about the theoretical properties of Aitken's iterated Δ^2 process. In addition, it does not help at all if the initial sequence $\{s_n\}$ is not alternating. A related problem, which may easily arise in this context, is that some second differences $\Delta^2 \mathscr{A}_{k-1}^{(n)}$ may become so small that division would lead to overflow. Consequently, a good program should be protected against the exact or approximate vanishing of the second differences $\Delta^2 \mathscr{A}_{k-1}^{(n)}$.

As in the case of Wynn's ϵ algorithm this can be accomplished by introducing two variables HUGE and TINY which have values that are close to but not identical with the largest and smallest floating point number representable on the computer. If $\Delta^2 \mathscr{A}_{k-1}^{(n)}$ is smaller in magnitude than TINY, $\mathscr{A}_k^{(n)}$ will be set equal to HUGE and the recursion is continued.

The following FORTRAN 77 subroutine AITKEN performs the recursive computation of the Aitken table in a single 1-dimensional array A according to eq. (5.2-7). It is safeguarded against an exact or approximate vanishing of the second differences $\Delta^2 \mathscr{A}_{k-1}^{(n)}$ by using two variables HUGE and TINY. The elements s_n with $n = 0, 1, 2, \ldots$ of the sequence to be transformed have to be computed in a DO loop in the calling program. Whenever a new sequence element s_n is computed in the outer DO loop this subroutine AITKEN has to be called again and a new string $\mathscr{A}_j^{(n-2j)}$ with $0 \leq j \leq [\![n/2]\!]$ will be calculated. The new sequence element s_n is read in via the variable SOFN and the approximation to the limit is returned via the variable ESTLIM.

It is important to note that this subroutine AITKEN only calculates the approximations to the limit according to eqs. (5.2-4) and (5.2-5). The convergence of the whole process has to be analyzed in the calling program.

```
      SUBROUTINE AITKEN(SOFN,N,A,LARRAY,ESTLIM)
      DIMENSION A(0:LARRAY)
      PARAMETER ( HUGE = 1.E+60 , TINY = 1.E-60 , TWO = 2.E0 )
      A(N) = SOFN
      IF (N.LT.2) THEN
        ESTLIM    = SOFN
      ELSE
        LOWMAX  = N/2
        DO 10 J = 1,LOWMAX
          M     = N - 2*J
          DENOM = A(M+2) - TWO*A(M+1) + A(M)
          IF (ABS(DENOM).LT.TINY) THEN
            A(M) = HUGE
          ELSE
            A(M) = A(M) - (A(M) - A(M+1))**2 / DENOM
          END IF
10      CONTINUE
        IF ( MOD(N,2) .EQ. 0 ) THEN
          ESTLIM = A(0)
        ELSE
          ESTLIM = A(1)
        END IF
      END IF
      RETURN
      END
```

6. Wynn's rho algorithm and related topics

6.1. Polynomial and rational extrapolation

Assume that the values of a function $f(x)$ are only known at some discrete points $x_0 < x_1 < \cdots < x_m$. It is one of the classical problems of numerical analysis to estimate the value of f at some point $\xi \notin \{x_0, x_1 \cdots x_m\}$. If $x_0 < \xi < x_m$, this problem is called *interpolation*, and if either $\xi < x_0$ or $x_m < \xi$, this problem is called *extrapolation*. These problems and their solution are discussed in any book on numerical analysis. More specialized treatments of these topics can be found in a monograph on interpolation by Davis [50] or in a review article on extrapolation processes by Joyce [51].

Extrapolation techniques can be used for the construction of convergence acceleration methods. In this approach, the existence of a function $\mathscr{S}$ of a continuous variable is postulated which coincides on a discrete set of arguments $\{\!\{x_n\}\!\}$ with the elements of the sequence $\{\!\{s_n\}\!\}$ to be transformed,

$$\mathscr{S}(x_n) = s_n, \qquad n \in \mathbb{N}_0. \tag{6.1-1}$$

This ansatz reduces the problem of accelerating the convergence of a sequence $\{s_n\}$ to an extrapolation problem. If a finite string $s_m, s_{m+1}, \ldots, s_{m+k}$ of $k+1$ sequence elements is known one can construct an approximation $\mathscr{S}_k(x)$ to $\mathscr{S}(x)$ which satisfies the $k+1$ interpolation conditions

$$\mathscr{S}_k(x_{m+j}) = s_{m+j}, \qquad 0 \le j \le k. \tag{6.1-2}$$

In the next step one has to determine the value of the approximant $\mathscr{S}_k(x)$ for $x \to x_\infty$. If this can be done and if the assumptions which are implicitly contained in this ansatz – the existence of a function $\mathscr{S}(x)$ which can be approximated at least locally by a suitable set of interpolating functions – are justified, one can expect that the extrapolated value $\mathscr{S}_k(x_\infty)$ will provide a better approximation to the limit s of the sequence $\{s_n\}$ than the last sequence element s_{m+k} which was used for the construction of $\mathscr{S}_k(x)$.

In interpolation and extrapolation problems the function under consideration has to be modelled either in between or beyond a finite set $x_0, x_1, \ldots, x_n$ of interpolation points by a suitable set of interpolating functions. These interpolating functions should be flexible and general enough to produce good approximations for large classes of functions which can occur in practice. In addition, they should also be simple enough to be manageable. The most common interpolating functions are either polynomials or rational functions. These two sets will also lead to different convergence acceleration methods.

If interpolation by polynomials is used as the basis of a convergence acceleration method it is implicitly assumed that the k-th order approximant $\mathscr{S}_k(x)$ is a polynomial of degree k in x,

$$\mathscr{S}_k(x) = c_0 + c_1 x + \cdots + c_k x^k. \tag{6.1-3}$$

For polynomials, the most natural extrapolation point is $x = 0$. Consequently, the interpolation points x_n have to satisfy the conditions

$$x_0 > x_1 > x_2 > \cdots > x_m > x_{m+1} > \cdots > 0, \tag{6.1-4a}$$

$$\lim_{n\to\infty} x_n = 0. \tag{6.1-4b}$$

The choice $x = 0$ as the extrapolation point implies that the approximation to the limit is to be identified with the constant term c_0 of the polynomial (6.1-3).

Several different methods for the computation of interpolating polynomials $\mathscr{S}_k(x)$ are described in the mathematical literature. Since only the constant term of a polynomial $\mathscr{S}_k$ has to be computed and since in most applications it is desirable to compute simultaneously a whole string of approximants $\mathscr{S}_0(0)$, $\mathscr{S}_1(0)$, $\mathscr{S}_2(0), \ldots,$ the most economical choice is Neville's scheme [52] for the recursive computation of interpolating polynomials. In the case $x = 0$ Neville's algorithm reduces to the following 2-dimensional linear recursive scheme (see p. 6 of ref. [20]):

$$\mathscr{N}_0^{(n)}(s_n, x_n) = s_n, \qquad n \in \mathbb{N}_0, \tag{6.1-5a}$$

$$\mathscr{N}_{k+1}^{(n)}(s_n, x_n) = \frac{x_n \mathscr{N}_k^{(n+1)}(s_{n+1}, x_{n+1}) - x_{n+k+1} \mathscr{N}_k^{(n)}(s_n, x_n)}{x_n - x_{n+k+1}}, \quad k, n \in \mathbb{N}_0. \tag{6.1-5b}$$

In the literature on convergence acceleration this variant of Neville's recursive scheme is usually called Richardson extrapolation [53]. Obviously, the linear transformation $\mathscr{N}_k^{(n)}(s_n, x_n)$ is exact if the sequence elements s_n are polynomials of degree k in the interpolation points x_n, i.e., for model sequences of the following type:

$$s_n = s + \sum_{j=0}^{k-1} c_j x_n^{j+1}, \qquad k, n \in \mathbb{N}_0. \tag{6.1-6}$$

The most obvious interpolation points for the Richardson extrapolation scheme, eq. (6.1-5), are $x_n = 1/(n+\beta)$ with $\beta > 0$ or also $x_n = 1/(n+\beta)^2$. These two choices are known to work quite well in a variety of cases. However, if either one of these sets of interpolation points $\{x_n\}$ is used, the Richardson extrapolation scheme (6.1-5) is not regular, i.e., the convergence of a sequence $\{s_n\}$ to some limit s does not imply the convergence of the transformed sequence to the same limit. In Brezinski's second book (see pp. 37–38 of ref. [20]) it is shown that the regularity of the Richardson extrapolation scheme is guaranteed only if some $a > 1$ exists such that the interpolation points $\{x_n\}$ satisfy for all $n \in \mathbb{N}_0$:

$$x_n / x_{n+1} \geq a. \tag{6.1-7}$$

This condition (6.1-7) is obviously fulfilled if the interpolation points $\{x_n\}$ satisfy $x_n = b^n$ with $0 < b < 1$ for all $n \in \mathbb{N}_0$. A good discussion of the properties of the Richardson extrapolation scheme as well as a list of various different sets of interpolation points $\{x_n\}$ can be found in Brezinski's second book (see pp. 36–42 of ref. [20]).

It is well known that some functions can be approximated by polynomials only quite poorly but by rational functions they can be approximated very well. Consequently, it is likely that at least for some sequences $\{s_n\}$ rational extrapolation will give better results than polynomial extrapolation. Let us therefore assume that the approximant $\mathscr{S}_k(x)$ can be written as the ratio of two polynomials of degrees l and m, respectively,

$$\mathscr{S}_k(x) = \frac{a_0 + a_1 x + a_2 x^2 + \cdots + a_l x^l}{b_0 + b_1 x + b_2 x^2 + \cdots + b_m x^m}, \qquad k, l, m \in \mathbb{N}_0. \tag{6.1-8}$$

This rational function contains $l + m + 2$ coefficients $a_0, \ldots, a_l$ and $b_0, \ldots, b_m$. However, only $l + m + 1$ coefficients are independent since they are determined only up to a common nonvanishing factor. Usually, one requires either $b_0 = 1$ or $b_m = 1$. Consequently, the $k + 1$ interpolation conditions (6.1-2) will determine the coefficients $a_0, \ldots, a_l$ and $b_0, \ldots, b_m$ provided that $k = l + m$ holds.

The extrapolation point $x = 0$ is also the most obvious choice in the case of rational extrapolation. Extrapolation to $x = 0$ implies that the interpolation points $\{x_n\}$ have to satisfy eq. (6.1-4) and that the approximation to the limit is to be identified with the ratio a_0/b_0 of the constant terms of the polynomials in eq. (6.1-8).

However, if in eq. (6.1-8) $l = m$ holds, extrapolation to infinity is also possible. In that case the interpolation points $\{x_n\}$ would have to satisfy

$$0 < x_0 < x_1 < x_2 < \cdots < x_m < x_{m+1} < \cdots, \tag{6.1-9a}$$
$$\lim_{n \to \infty} x_n = \infty. \tag{6.1-9b}$$

In the case of extrapolation to infinity only the coefficients a_l and b_l of the polynomials in eq. (6.1-8), which are proportional to highest power x^l, contribute. Consequently, the approximation to the limit has to be identified with the ratio a_l/b_l.

As in the case of polynomial interpolation several different algorithms for the computation of rational interpolants are described in the literature. A discussion of the relative merits of these algorithms as well as a survey of the relevant literature can be found in chapter III of a book by Cuyt and Wuytack [54] which was recently published .

6.2. *Wynn's rho algorithm*

Wynn's ρ algorithm [25] is designed to compute even-order convergents of Thiele's interpolating continued fraction [55] and to extrapolate them to infinity. The even-order convergents are rational functions of the following type:

$$\mathscr{S}_{2k}(x) = \frac{a_k x^k + a_{k-1} x^{k-1} + \cdots + a_1 x + a_0}{b_k x^k + b_{k-1} x^{k-1} + \cdots + b_1 x + b_0}, \qquad k \in \mathbb{N}_0. \tag{6.2-1}$$

This means that the ratio a_k/b_k is to be identified with the approximation to the limit. According to Cuyt and Wuytack (see p. 214 of ref. [54]) Wynn's ρ algorithm performs the computation of the interpolating rational function (6.2-1) and its extrapolation to infinity with a smaller number of arithmetic operations than similar recursive algorithms.

Wynn's ρ algorithm [25] is the following nonlinear recursive scheme which is formally almost identical with Wynn's ϵ algorithm, eq. (4.2-1):

$$\rho_{-1}^{(n)} = 0, \quad \rho_0^{(n)} = s_n, \tag{6.2-2a}$$
$$\rho_{k+1}^{(n)} = \rho_{k-1}^{(n+1)} + \frac{x_{n+k+1} - x_n}{\rho_k^{(n+1)} - \rho_k^{(n)}}, \quad k, n \in \mathbb{N}_0. \tag{6.2-2b}$$

The only difference between Wynn's ϵ and Wynn's ρ algorithm is that the ρ algorithm also involves a sequence of interpolation points $\{x_n\}$ which have to satisfy eq. (6.1-9). As in the case of Wynn's ϵ algorithm only the elements $\rho_{2k}^{(n)}$ with even orders serve as approximations to the limit. The elements $\rho_{2k+1}^{(n)}$ with odd orders are only auxiliary quantities which diverge if the whole process converges.

Despite their formal similarity, the ϵ and ρ algorithm differ significantly in their ability of accelerating convergence. For instance, the ϵ algorithm is exact for the model sequence (4.1-5), and is known to be a very efficient accelerator for linearly convergent sequences. In many cases the ϵ algorithm is also able to sum divergent series. However, the otherwise very powerful ϵ algorithm fails to accelerate logarithmic convergence.

The properties of Wynn's ϵ algorithm and Wynn's ρ algorithm are in some sense complementary. Wynn's ρ algorithm fails to accelerate linear convergence and is not able to sum divergent series. However, it is very powerful for some logarithmically convergent sequences. This can easily be understood on the basis of the following model sequence for which the transform $\rho_{2k}^{(n)}$ is exact:

$$s_n = \frac{s x_n^k + a_1 x_n^{k-1} + \cdots + a_k}{x_n^k + b_1 x_n^{k-1} + \cdots + b_k}, \quad k, n \in \mathbb{N}_0. \tag{6.2-3}$$

Since for fixed $k \in \mathbb{N}_0$ the zeros of the numerator and denominator polynomials in eq. (6.2-3) are contained in a compact set and since the interpolation points $\{x_n\}$ diverge as $n \to \infty$, at least for sufficiently large values of n a rational function of that kind will change only relatively slowly and monotoneously with increasing n. Certainly, such an expression will not oscillate or even diverge. This should explain why the ρ algorithm normally works well in the case of logarithmic convergence but fails in the case of oscillating or divergent sequences.

The properties of Wynn's ρ algorithm are discussed in books by Brezinski (see pp. 102–106 of ref. [19] and pp. 96–102 of ref. [20]) and Wimp (see pp. 168–169 of ref. [23]). In these books the connection of the ρ algorithm with interpolating continued fractions is emphasized and it is also shown that the transforms $\rho_{2k}^{(n)}$ can be represented as the ratio of two determinants. But otherwise, relatively little seems to be known about this sequence transformation.

The most obvious interpolation points $\{x_n\}$ are $x_n = n + \beta$ with $\beta > 0$. With this choice, Wynn's ρ algorithm assumes its standard form:

$$\rho_{-1}^{(n)} = 0, \qquad \rho_0^{(n)} = s_n, \tag{6.2-4a}$$

$$\rho_{k+1}^{(n)} = \rho_{k-1}^{(n+1)} + \frac{k+1}{\rho_k^{(n+1)} - \rho_k^{(n)}}, \quad k, n \in \mathbb{N}_0. \tag{6.2-4b}$$

Other possible sequences of interpolation points as for instance $x_n = (n+\beta)^2$ with $\beta > 0$ are discussed in Brezinski's second book [20].

As in Wynn's ϵ algorithm the approximation to the limit depends upon the index m of the last sequence element s_m which was used in the recursion. If m is even, $m = 2\mu$, our approximation to the limit of the sequence is the transformation

$$\{s_0, x_0; s_1, x_1; \ldots; s_{2\mu}, x_{2\mu}\} \to \rho_{2\mu}^{(0)}, \tag{6.2-5}$$

and if m is odd, $m = 2\mu + 1$, we use the transformation

$$\{s_1, x_1; s_2, x_2; \ldots; s_{2\mu+1}, x_{2\mu+1}\} \to \rho_{2\mu}^{(1)}. \tag{6.2-6}$$

With the help of the notation $[\![x]\!]$ for the integral part of x, i.e., the largest integer ν satisfying $\nu \leq x$, these two relationships can be combined into a single equation yielding

$$\{s_{m-2[\![m/2]\!]}, x_{m-2[\![m/2]\!]}; s_{m-2[\![m/2]\!]+1}, x_{m-2[\![m/2]\!]+1}; \ldots; s_m, x_m\} \to \rho_{2[\![m/2]\!]}^{(m-2[\![m/2]\!])}. \tag{6.2-7}$$

The elements of the ρ table can be arranged in the same rectangular scheme as the elements of the ϵ table in (4.3-1). Since the recurrence relationships for Wynn's ϵ algorithm and Wynn's ρ algorithm are structurally identical, those elements of the ρ table which are connected by the 4-term recurrence formula (6.2-2b), are also located in the ρ table at the vertices of a rhombus:

$$\begin{array}{ccc} & \rho_k^{(n)} & \rho_{k+1}^{(n)} \\ \rho_{k-1}^{(n+1)} & \rho_k^{(n+1)} & \end{array} \tag{6.2-8}$$

Consequently, Wynn's ρ algorithm can be programmed in exactly the same way as Wynn's ϵ algorithm. For that purpose we rewrite the recursive scheme (6.2-2) in the following way:

$$\rho_0^{(n)} = s_n, \qquad n \geq 0, \tag{6.2-9a}$$

$$\rho_1^{(n-1)} = \frac{x_n - x_{n-1}}{s_n - s_{n-1}}, \qquad n \geq 1, \tag{6.2-9b}$$

$$\rho_j^{(n-j)} = \rho_{j-2}^{(n-j+1)} + \frac{x_n - x_{n-j}}{\rho_{j-1}^{(n-j+1)} - \rho_{j-1}^{(n-j)}}, \quad n \geq 2, \quad 2 \leq j \leq n. \tag{6.2-9c}$$

As in the case of the ϵ algorithm the modification (4.3-9) of Wynn's moving lozenge technique (4.3-8) can be used. This means that only a single 1-dimensional array R will be needed if the elements of the current counterdiagonal $\rho_j^{(m-j)}$ with $m \in \mathbb{N}_0$ and $0 \leq j \leq m$ are stored in R in such a way that the superscript of the element of the ρ table coincides with the index of the corresponding array element,

$$\rho_j^{(m-j)} \to R(m-j). \tag{6.2-10}$$

The only difference with Wynn's ϵ algorithm is that also a second 1-dimensional array ξ will be needed in which the interpolation points x_n are stored according to the rule

$$x_n \to \xi(n). \tag{6.2-11}$$

With these two conventions the recursive scheme (6.2-9) can be reformulated in terms of the elements of the 1-dimensional arrays R and ξ,

$$R(n) \leftarrow s_n, \qquad n \geq 0, \tag{6.2-12a}$$

$$R(n-1) \leftarrow \frac{\xi(n) - \xi(n-1)}{R(n) - R'(n-1)}, \qquad n \geq 1, \tag{6.2-12b}$$

$$R(n-j) \leftarrow R'(n-j+1) + \frac{\xi(n) - \xi(n-j)}{R(n-j+1) - R'(n-j)},$$
$$n \geq 2, \qquad 2 \leq j \leq n. \tag{6.2-12c}$$

As in the case of Wynn's ϵ algorithm, the primed array elements $R'(n-j)$ and $R'(n-j+1)$ have to be stored in auxiliary variables. The primes also indicate that the array elements

$R'(n-j)$ and $R'(n-j+1)$ refer to the occupation of R after the previous run, i.e., after the computation of the counterdiagonal ρ_j^{n-j-1} with $0 \le j \le n-1$. The listings of some FORTRAN IV programs, which compute Wynn's ρ algorithm, can be found in Brezinski's second book (see pp. 361–365 of ref. [20]).

6.3. *The iteration of Wynn's rho algorithm*

According to eq. (5.1-5), Aitken's Δ^2 process is identical with the transform $\epsilon_2^{(n)}$,

$$\mathscr{A}_1^{(n)} = s_n - \frac{[\Delta s_n]^2}{\Delta^2 s_n} = \epsilon_2^{(n)}, \qquad n \in \mathbb{N}_0. \tag{6.3-1}$$

If Aitken's Δ^2 process is iterated, a new sequence transformation $\mathscr{A}_k^{(n)}$ results which has a similiar ability of accelerating convergence as Wynn's ϵ algorithm. However, in section 13 we shall encounter some examples in which Aitken's iterated Δ^2 process clearly outperforms Wynn's ϵ algorithm. This alone should justify an analysis of Aitken's iterated Δ^2 transformation $\mathscr{A}_k^{(n)}$.

Since Wynn's ϵ algorithm (4.2-1) and Wynn's ρ algorithm (6.2-2) are formally almost identical, one can construct a new sequence transformation by proceeding as in the case of Aitken's iterated Δ^2 process. This means that first the transform $\rho_2^{(n)}$ is expressed in terms of some sequence elements s_n and interpolation points x_n The resulting expression for $\rho_2^{(n)}$ will then be iterated.

From eqs. (6.2-2a) and (6.2-2b) we obtain the following expression for the ρ analogue of Aitken's Δ^2 process:

$$\rho_2^{(n)} = s_{n+1} + \frac{(x_{n+2} - x_n)[\Delta s_{n+1}][\Delta s_n]}{[\Delta x_{n+1}][\Delta s_n] - [\Delta x_n][\Delta s_{n+1}]}, \qquad n \in \mathbb{N}_0. \tag{6.3-2}$$

An iteration of this expression, which may be considered to be a kind of weighted Δ^2 process, can be done in a variety of ways. The problem is that there is no unique way of choosing the indices of the interpolation points x_n. However, if we take into account that in Wynn's ρ algorithm, eq. (6.2-2), the differences of the indices of the interpolation points x_n increase with increasing k, we see that the following nonlinear recursive scheme should be the most natural iteration of the transform (6.3-2):

$$\mathscr{W}_0^{(n)} = s_n, \qquad n \in \mathbb{N}_0, \tag{6.3-3a}$$

$$\mathscr{W}_{k+1}^{(n)} = \mathscr{W}_k^{(n+1)} + \frac{(x_{n+2k+2} - x_n)\left[\Delta \mathscr{W}_k^{(n+1)}\right]\left[\Delta \mathscr{W}_k^{(n)}\right]}{(x_{n+2k+2} - x_{n+1})\left[\Delta \mathscr{W}_k^{(n)}\right] - (x_{n+2k+1} - x_n)\left[\Delta \mathscr{W}_k^{(n+1)}\right]},$$

$$k,\ n \in \mathbb{N}_0. \tag{6.3-3b}$$

As usual it is assumed that the difference operator Δ acts upon n and not upon k. The most obvious interpolation points are as in Wynn's ρ algorithm $x_n = n + \beta$ with $\beta > 0$. With this choice, the iterated ρ_2 transformation assumes its standard form:

$$\mathscr{W}_0^{(n)} = s_n, \qquad n \in \mathbb{N}_0, \tag{6.3-4a}$$

$$\mathscr{W}_{k+1}^{(n)} = \mathscr{W}_k^{(n+1)} - \frac{(2k+2)\left[\Delta \mathscr{W}_k^{(n+1)}\right]\left[\Delta \mathscr{W}_k^{(n)}\right]}{(2k+1)\Delta^2 \mathscr{W}_k^{(n)}}, \quad k,\, n \in \mathbb{N}_0. \tag{6.3-4b}$$

The table of this transformation can be arranged in the same rectangular scheme as the Aitken table (5.2-1). Also, the recurrence formulas for Aitken's iterated Δ^2 process, eq. (5.1-14), and the recurrence formulas for the transforms $\mathscr{W}_k^{(n)}$ are structurally identical. This implies that the four elements, which are connected by the nonlinear recurrence formulas (4.3-2) or (4.3-3), also form a pattern in the $\mathscr{W}$ table which looks like the move of a knight on the chessboard:

$$\begin{matrix} \mathscr{W}_k^{(n)} & \mathscr{W}_{k+1}^{(n)} \\ \mathscr{W}_k^{(n+1)} & \\ \mathscr{W}_k^{(n+2)} & \end{matrix} \tag{6.3-5}$$

Consequently, this iterated ρ_2 process can be programmed in exactly the same way as Aitken's iterated Δ^2 process and only a few minor alterations have to be done in the subroutine AITKEN. In this context, it is advantageous to reformulate the recurrence scheme (6.3-4) in the following way:

$$\mathscr{W}_0^{(n)} = s_n, \qquad n \geq 0, \tag{6.3-6a}$$

$$\mathscr{W}_j^{(n-2j)} = \mathscr{W}_{j-1}^{(n-2j+1)} + \frac{(x_n - x_{n-2j})\left[\Delta \mathscr{W}_{j-1}^{(n-2j+1)}\right]\left[\Delta \mathscr{W}_{j-1}^{(n-2j)}\right]}{(x_n - x_{n-2j+1})\left[\Delta \mathscr{W}_{j-1}^{(n-2j)}\right] - (x_{n-1} - x_{n-2j})\left[\Delta \mathscr{W}_{j-1}^{(n-2j+1)}\right]},$$

$$n \geq 2, \quad 1 \leq j \leq [\![n/2]\!]. \tag{6.3-6b}$$

As usual, $[\![n/2]\!]$ stands for the integral part of $n/2$, i.e., the largest integer ν satisfying $\nu \leq n/2$. If the sequence elements $s_0, s_1, \ldots, s_m$ are used as starting values, the recursive scheme (6.3-6) is able to compute all elements $\mathscr{W}_j^{(\mu-2j)}$ with $0 \leq \mu \leq m$ and $0 \leq j \leq [\![\mu/2]\!]$.

As in the case of Aitken's iterated Δ^2 process the approximation to the limit depends upon the index m of the last sequence element s_m which was used in the recursion. If m is even, $m = 2\mu$, our approximations to the limit of the sequence are the transformations

$$\{s_0, x_0;\, s_1, x_1;\ldots,\, s_{2\mu}, x_{2\mu}\} \to \mathscr{W}_\mu^{(0)}, \tag{6.3-7}$$

and if m is odd, $m = 2\mu + 1$, the approximation to the limit will be

$$\{s_1, x_1;\, s_2, x_2;\ldots,\, s_{2\mu+1}, x_{2\mu+1}\} \to \mathscr{W}_\mu^{(1)}. \tag{6.3-8}$$

These two relationships can be combined into a single equation,

$$\{s_{m-2[\![m/2]\!]}, x_{m-2[\![m/2]\!]};\, s_{m-2[\![m/2]\!]+1}, x_{m-2[\![m/2]\!]+1};\ldots;\, s_m, x_m\} \to \mathscr{W}_{[\![m/2]\!]}^{(m-2[\![m/2]\!])}. \tag{6.3-9}$$

Only two 1-dimensional arrays w and ξ are needed if the interpolation points x_n are stored in ξ according to eq. (6.2-11) and if the elements $\mathscr{W}_k^{(n)}$ are stored in w according to the following rule:

$$\mathscr{W}_{[\![\nu/2]\!]}^{(n-\nu)} \to w(n-\nu), \qquad n \geq 0, \quad 0 \leq \nu \leq n. \tag{6.3-10}$$

With this convention the recursive scheme (6.3-6) can be reformulated in terms of the elements of the 1-dimensional arrays w and ξ:

$$w(n) \leftarrow s_n, \qquad n \geq 0, \tag{6.3-11a}$$

$$w(l) \leftarrow w(l+1) + \frac{[\xi(n) - \xi(l+1)][\Delta w(l+1)][\Delta w(l)]}{[\xi(n) - \xi(l+1)][\Delta w(l)] - [\xi(n-1) - \xi(l)][\Delta w(l+1)]},$$

$$l = n - 2j, \qquad n \geq 2, \qquad 1 \leq j \leq [\![n/2]\!]. \tag{6.3-11b}$$

7. The Levin transformation

7.1. The derivation of Levin's sequence transformation

Levin's sequence transformation [28] is designed to be exact for model sequences of the following type:

$$s_n = s + \omega_n \sum_{j=0}^{k-1} c_j/(n+\beta)^j, \qquad k,\, n \in \mathbb{N}_0. \tag{7.1-1}$$

Here, the remainder estimates ω_n are essentially arbitrary functions of n. It is only assumed that they are different from zero for all finite values of n. In addition, it would not make much sense to consider in convergence acceleration and summation processes remainder estimates which are constant. Consequently, we shall also assume that for all finite values of n the elements of the sequence $\{\!\{\omega_n\}\!\}$ are all distinct. Depending upon the behaviour of the remainder estimates $\{\!\{\omega_n\}\!\}$ as $n \to \infty$, the sequence $\{\!\{s_n\}\!\}$ may either converge or diverge. In eq. (7.1-1) it has to be required that $\beta + n$ must not be zero. This implies that the parameter β must not be zero or a negative integer. However, the elements of the model sequence (7.1-1) will serve as finite approximations to Poincaré-type asymptotic expansions of the following kind:

$$s_n \sim s + \omega_n \sum_{j=0}^{\infty} c_j/(n+\beta)^j, \qquad n \to \infty. \tag{7.1-2}$$

In expansions of that kind negative values of β will lead to different signs of the terms if either $n + \beta < 0$ or $n + \beta > 0$ holds. Since model sequences of the type of eq. (7.1-1) will be used as approximations for asymptotic expansions of the above type and since these approximations should be uniformly valid over a wide range of admissible values of n – preferably for all $n \geq 0$ – it is necessary to require that the sign pattern of the terms of the sum in eq. (7.1-1) must not depend upon n. This rules out $\beta < 0$ and we have to require $\beta > 0$. But otherwise, the parameter β is in principle completely arbitrary. In the literature on Levin's sequence transformation, only the case $\beta = 1$, which is the most obvious choice, has been considered so far.

In eq. (7.1-1) there occur $k+1$ unknown quantities, the limit or antilimit s and the k linear coefficients $c_0, \ldots, c_{k-1}$. Hence, $k+1$ sequence elements $s_n, \ldots, s_{n+k}$ are needed for the determination of s, and according to Cramer's rule the general Levin transformation $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$ can be defined by the following ratio of determinants:

$$\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n) = \frac{\begin{vmatrix} s_n & \ldots & s_{n+k} \\ \omega_n & \ldots & \omega_{n+k} \\ \vdots & \ddots & \vdots \\ \omega_n/(\beta+n)^{k-1} & \ldots & \omega_{n+k}/(\beta+n+k)^{k-1} \end{vmatrix}}{\begin{vmatrix} 1 & \ldots & 1 \\ \omega_n & \ldots & \omega_{n+k} \\ \vdots & \ddots & \vdots \\ \omega_n/(\beta+n)^{k-1} & \ldots & \omega_{n+k}/(\beta+n+k)^{k-1} \end{vmatrix}}. \tag{7.1-3}$$

If the sequence elements $s_n, s_{n+1}, \ldots, s_{n+k}$ satisfy eq. (7.1-1) then Levin's general sequence transformation is exact by construction, i.e.,

$$\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n) = s. \tag{7.1-4}$$

The representation of the general Levin transformation as the ratio of two determinants is not well suited for practical applications because the fast and reliable evaluation of large order determinants is a not yet satisfactorily solved problem of numerical analysis. Thus, alternative expressions for the general Levin transformation are highly desirable. Fortunately, they can be derived quite easily.

Levin's original derivation [28] of nondeterminantal expressions for his sequence transformation was based upon the observation that the determinants in eq. (7.1-3) may be expressed in terms of Vandermonde determinants. However, for our purposes it is advantageous to follow Sidi's approach [56] which exploits properties of the difference operator Δ, since this approach can easily be extended to other sequence transformations which will be treated later in this report. For that purpose, eq. (7.1-1) is rewritten in the following way:

$$(n+\beta)^{k-1}[s_n - s]/\omega_n = \sum_{j=0}^{k-1} c_j (n+\beta)^{k-j-1}. \tag{7.1-5}$$

The highest power of n, which occurs on the right-hand side of eq. (7.1-5), is n^{k-1}. We now utilize the well-known fact that any polynomial of degree $k-1$ in n will be annihilated by the difference operator Δ^k. Since the difference operator Δ^k is linear, we may conclude from eqs.

(7.1-4) and (7.1-5) that the general Levin transformation $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$ is given by the following ratio:

$$\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n) = \frac{\Delta^k\left\{(n+\beta)^{k-1} s_n/\omega_n\right\}}{\Delta^k\left\{(n+\beta)^{k-1}/\omega_n\right\}}. \tag{7.1-6}$$

With the help of eq. (2.4-8) the action of the difference operators in eq. (7.1-6) can be expressed in closed form. This leads to a representation of the general Levin transformation as the ratio of two finite sums:

$$\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n) = \frac{\sum\limits_{j=0}^{k} (-1)^j \binom{k}{j} \frac{(\beta+n+j)^{k-1}}{(\beta+n+k)^{k-1}} \frac{s_{n+j}}{\omega_{n+j}}}{\sum\limits_{j=0}^{k} (-1)^j \binom{k}{j} \frac{(\beta+n+j)^{k-1}}{(\beta+n+k)^{k-1}} \frac{1}{\omega_{n+j}}}, \qquad k, n \in \mathbb{N}_0. \tag{7.1-7}$$

The common factor $(\beta+n+k)^{k-1}$ in eq. (7.1-7) was introduced in order to decrease the magnitude of the terms of the numerator and denominator sums, because otherwise overflow may happen too easily for larger values of k.

A mild extension of the general Levin transformation, eq. (7.1-7), will also be considered:

$$\mathscr{L}_{k,l}^{(n)}(\beta, s_n, \omega_n) = \frac{\sum\limits_{j=0}^{k} (-1)^j \binom{k}{j} \frac{(\beta+n+j)^{k-l-1}}{(\beta+n+k)^{k-1}} \frac{s_{n+j}}{\omega_{n+j}}}{\sum\limits_{j=0}^{k} (-1)^j \binom{k}{j} \frac{(\beta+n+j)^{k-l-1}}{(\beta+n+k)^{k-1}} \frac{1}{\omega_{n+j}}}, \qquad k, l, n \in \mathbb{N}_0. \tag{7.1-8}$$

For $l = 0$ this transformation reduces to the general Levin transformation, eq. (7.1-7). An alternative representation for this generalization of Levin's sequence transformation can be derived with the help of eq. (2.4-8),

$$\mathscr{L}_{k,l}^{(n)}(\beta, s_n, \omega_n) = \frac{\Delta^k\left\{(n+\beta)^{k-l-1} s_n/\omega_n\right\}}{\Delta^k\left\{(n+\beta)^{k-l-1}/\omega_n\right\}}. \tag{7.1-9}$$

This relationship implies that this generalization of Levin's sequence transformation is exact for sequences of the following type:

$$s_n = s + (\beta+n)^l \omega_n \sum_{j=0}^{k-1} c_j/(\beta+n)^j, \qquad k, l, n \in \mathbb{N}_0. \tag{7.1-10}$$

7.2. Recursive computation of the Levin transformation

Another advantage of Sidi's approach [56] for the derivation of nondeterminantal expressions for Levin's sequence transformation is that starting from eq. (7.1-6) Fessler, Ford, and Smith [57] could derive a recursive scheme which allows a convenient computation of both the numerator and the denominator of the general Levin transformation, eq. (7.1-7). In eq. (7.1-6), both numerator and denominator are of the general form

$$P_k^{(n)}(\beta) = \Delta^k X_k^{(n)}(\beta), \qquad k, n \in \mathbb{N}_0. \tag{7.2-1}$$

As usual, it is assumed here that the difference operator Δ as well as the shift operator E, which is defined in eq. (2.4-4), act only upon n and not upon k. Comparison with eq. (7.1-6) shows that the quantities $X_k^{(n)}(\beta)$ satisfy the following 2-term recursion in k:

$$X_k^{(n)}(\beta) = (\beta + n) X_{k-1}^{(n)}(\beta), \qquad k \geq 1, \quad n \geq 0. \tag{7.2-2}$$

The following commutator relationship can be proved by complete induction with respect to k [57],

$$\Delta^k(\beta + n) - (\beta + n)\Delta^k = kE\Delta^{k-1}. \tag{7.2-3}$$

Combination of eqs. (2.4-4), (7.2-1), (7.2-2), and (7.2-3) yields:

$$P_k^{(n)}(\beta) = \{kE + (\beta + n)\Delta\}\Delta^{k-1} X_{k-1}^{(n)}(\beta) \tag{7.2-4}$$

$$= \{kE + (\beta + n)\Delta\} P_{k-1}^{(n)}(\beta) \tag{7.2-5}$$

$$= (\beta + n + k) P_{k-1}^{(n+1)}(\beta) - (\beta + n) P_{k-1}^{(n)}(\beta). \tag{7.2-6}$$

With the help of the 3-term recurrence formula (7.2-6) the numerator as well as the denominator of the general Levin transformation $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$ can be computed for $k \geq 1$. However, for the sake of numerical stability and in order to make overflow less likely it is preferable to scale the quantities $P_k^{(n)}(\beta)$ by defining

$$L_k^{(n)}(\beta) = P_k^{(n)}(\beta)/(\beta + n + k)^{k-1}. \tag{7.2-7}$$

Inserting this into eq. (7.2-6) yields the following 3-term recurrence formula for the scaled quantities $L_k^{(n)}(\beta)$:

$$L_{k+1}^{(n)}(\beta) = L_k^{(n+1)}(\beta) - \frac{(\beta + n)(\beta + n + k)^{k-1}}{(\beta + n + k + 1)^k} L_k^{(n)}(\beta), \qquad k, n \in \mathbb{N}_0. \tag{7.2-8}$$

If we use the starting values

$$L_0^{(n)}(\beta) = s_n/\omega_n, \qquad n \in \mathbb{N}_0, \tag{7.2-9}$$

the recurrence formula (7.2-8) produces the numerator of the general Levin transformation, eq. (7.1-7) and if we use the starting values

$$L_0^{(n)}(\beta) = 1/\omega_n, \qquad n \in \mathbb{N}_0, \tag{7.2-10}$$

we obtain the denominator of the general Levin transformation.

With the help of the 3-term recurrence formula (7.2-8) it is also possible to compute both the numerator and the denominator of the generalized Levin transformation, eq. (7.1-8). If the starting values

$$L_0^{(n)}(\beta) = s_n/\left[(\beta+n)^l \omega_n\right], \qquad l,\ n \in \mathbb{N}_0, \tag{7.2-11}$$

are used, eq. (7.2-8) produces the numerator of the transformation (7.1-8), and the starting values

$$L_0^{(n)}(\beta) = 1/\left[(\beta+n)^l \omega_n\right], \qquad l,\ n \in \mathbb{N}_0, \tag{7.2-12}$$

produce the denominator of the transformation (7.1-8).

The 3-term recurrence formula (7.2-6) was according to the knowledge of the author first published by Longman [58]. However, Longman's derivation of the recurrence formula (7.2-6) is based upon Sister Celine's technique [59] and not on properties of the difference operator Δ as the derivation by Fessler, Ford, and Smith [57].

7.3. Remainder estimates for the Levin transformation

Until now, we have completely ignored the sequence $\{\omega_n\}$ of remainder estimates and its rôle in the process of convergence acceleration or summation. In order to deal with this question we take into account that the elements of the model sequence (7.1-1), for which the general Levin transformation is exact, can be obtained from the elements of the sequence (7.1-2) by truncating the asymptotic power series in $1/(\beta+n)$ after the first k terms.

This indicates that the Levin transformation (7.1-7) should work very well for a given sequence $\{s_n\}$ if the sequence $\{\omega_n\}$ of remainder estimates is chosen in such a way that ω_n is proportional to the dominant term of an asymptotic expansion of the remainder r_n,

$$r_n = s_n - s = \omega_n\left[c + O(n^{-1})\right], \qquad n \to \infty. \tag{7.3-1}$$

Now, one is confronted with the practical problem of finding such a sequence $\{\omega_n\}$ of remainder estimates for a given sequence $\{s_n\}$. Here, it must be emphasized that a sequence $\{\omega_n\}$ of remainder estimates is not determined uniquely by the asymptotic condition (7.3-1). Consequently, it should at least in principle always be possible to find a variety of different sequences of remainder estimates which all satisfy eq. (7.3-1).

In some exceptional cases it is possible to derive explicit analytical expressions for the remainder estimates ω_n which satisfy eq. (7.3-1) – for instance if the s_n are partial sums of a series and if the series terms a_k have a sufficiently simple analytical structure. If such an explicit expression for ω_n is used in eq. (7.1-7), the general Levin transformation is a *linear* sequence transformation.

However, in most practical applications no information about the analytical structure of the sequence of remainders $\{r_n\}$ will be available and only the numerical values of a relatively small number of sequence elements $s_m, s_{m+1}, \ldots, s_{m+l}$ will be known. Consequently, it is necessary to find a way of obtaining the sequence of remainder estimates $\{\omega_n\}$ directly from the numerical values of the elements of the sequence $\{s_n\}$. If such a sequence of remainder estimates is used in eq. (7.1-7), the Levin transformation is a *nonlinear* sequence transformation because each remainder estimate ω_m depends explicitly upon at least one element of $\{s_n\}$.

On the basis of purely heuristic arguments Levin [28] suggested for sequences of partial sums

$$s_n = \sum_{\nu=0}^{n} a_\nu, \qquad n \in \mathbb{N}_0, \tag{7.3-2}$$

some simple remainder estimates which according to experience nevertheless work remarkably well. In the case of logarithmic convergence, i.e., if the elements of the sequence of partial sums s_n satisfy

$$\lim_{n\to\infty} \frac{s_{n+1} - s}{s_n - s} = 1, \tag{7.3-3}$$

Levin [28] suggested the remainder estimate

$$\omega_n = (\beta + n) a_n, \qquad n \in \mathbb{N}_0. \tag{7.3-4}$$

The use of this remainder estimate in eq. (7.1-7) yields Levin's u transformation:

$$u_k^{(n)}(\beta, s_n) = \frac{\displaystyle\sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{(\beta+n+j)^{k-2}}{(\beta+n+k)^{k-1}} \frac{s_{n+j}}{a_{n+j}}}{\displaystyle\sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{(\beta+n+j)^{k-2}}{(\beta+n+k)^{k-1}} \frac{1}{a_{n+j}}}. \tag{7.3-5}$$

In the case of alternating series Levin [28] suggested the remainder estimate

$$\omega_n = a_n, \qquad n \in \mathbb{N}_0. \tag{7.3-6}$$

This gives Levin's t transformation:

$$t_k^{(n)}(\beta, s_n) = \frac{\displaystyle\sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{(\beta+n+j)^{k-1}}{(\beta+n+k)^{k-1}} \frac{s_{n+j}}{a_{n+j}}}{\displaystyle\sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{(\beta+n+j)^{k-1}}{(\beta+n+k)^{k-1}} \frac{1}{a_{n+j}}}. \tag{7.3-7}$$

However, Smith and Ford [29] rightly remarked that the best simple remainder estimate for a convergent series with strictly alternating terms a_ν would be

$$\omega_n = a_{n+1}, \qquad n \in \mathbb{N}_0. \tag{7.3-8}$$

Using this in eq. (7.1-7) gives Smith and Ford's [29] modification of Levin's t transformation:

$$d_k^{(n)}(\beta, s_n) = \frac{\sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{(\beta+n+j)^{k-1}}{(\beta+n+k)^{k-1}} \frac{s_{n+j}}{a_{n+j+1}}}{\sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{(\beta+n+j)^{k-1}}{(\beta+n+k)^{k-1}} \frac{1}{a_{n+j+1}}}. \tag{7.3-9}$$

As a third simple remainder estimate Levin [28] suggested

$$\omega_n = \frac{a_n a_{n+1}}{a_n - a_{n+1}}, \qquad n \in \mathbb{N}_0. \tag{7.3-10}$$

Comparison with eq. (5.1-6) shows that this remainder estimate is based upon Aitken's Δ^2 process. It gives Levin's v transformation:

$$v_k^{(n)}(\beta, s_n) = \frac{\sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{(\beta+n+j)^{k-1}}{(\beta+n+k)^{k-1}} \frac{a_{n+j} - a_{n+j+1}}{a_{n+j} a_{n+j+1}} s_{n+j}}{\sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{(\beta+n+j)^{k-1}}{(\beta+n+k)^{k-1}} \frac{a_{n+j} - a_{n+j+1}}{a_{n+j} a_{n+j+1}}}. \tag{7.3-11}$$

The remainder estimates (7.3-4), (7.3-6), (7.3-8), and (7.3-10) can also be used if the sequence $\{s_n\}$, which is to be transformed, is not a sequence of partial sums. It is only necessary to replace in eqs. (7.3-5), (7.3-7), (7.3-9), and (7.3-11) a_0 by s_0 and a_n with $n \geq 1$ by Δs_{n-1}.

Levin's remainder estimates (7.3-4), (7.3-6), and (7.3-10) as well as Smith and Ford's modification (7.3-8) were derived using simple heuristic arguments. However, experience indicates that these remainder estimates nevertheless give rise to very powerful sequence transformations [29,30,57,60–64].

In some cases a more rigorous derivation of the remainder estimates mentioned above can be given. For instance, in Wimp's book (see p. 19 of ref. [23]) it is shown that if the terms a_n of a series satisfy

$$a_n \sim \lambda^n n^\Theta \left\{ \alpha_0 + \frac{\alpha_1}{n} + \frac{\alpha_2}{n} + \cdots \right\}, \qquad n \to \infty, \tag{7.3-12}$$

with $\alpha_0 \neq 0$, then there exist constants β_j and γ_j such that the remainders r_n of the partial sums of this series satisfy

$$r_n \sim \frac{\lambda^{n+1} n^\Theta}{\lambda - 1} \left\{ \alpha_0 + \frac{\beta_1}{n} + \frac{\beta_2}{n^2} + \cdots \right\}, \qquad n \to \infty, \tag{7.3-13}$$

if $|\lambda| < 1$, and

$$r_n \sim -\frac{n^{\Theta+1}}{\Theta+1}\left\{\alpha_0 + \frac{\gamma_1}{n} + \frac{\gamma_2}{n^2} + \cdots\right\}, \qquad n \to \infty, \tag{7.3-14}$$

if $\lambda = 1$ and $\mathrm{Re}(\Theta) < 1$.

A comparison of eqs. (7.3-12) and (7.3-13) gives us essentially the remainder estimate (7.3-6) which leads to the t transformation, eq. (7.3-7), and a comparison of eqs. (7.3-12) and (7.3-14) gives us essentially the remainder estimate (7.3-4) which leads to the u transformation, eq. (7.3-5).

If we replace in the model sequence (6.1-6), for which the Richardson extrapolation scheme, eq. (6.1-5), is exact, x_n by $1/(n+\beta)$, we obtain the model sequence

$$s_n = s + \sum_{j=0}^{k-1} c_j/(n+\beta)^{j+1}, \qquad k, n \in \mathbb{N}_0. \tag{7.3-15}$$

This model sequence can be obtained from the model sequence (7.1-2), for which Levin's sequence transformation, eq. (7.1-7) is exact, by choosing $\omega_n = 1/(n+\beta)$. Hence, if we define

$$\Lambda_k^{(n)}(\beta, s_n) = \mathscr{L}_k^{(n)}(\beta, s_n, 1/(n+\beta)), \qquad k, n \in \mathbb{N}_0, \tag{7.3-16}$$

we see that that the sequence transformation $\Lambda_k^{(n)}(\beta, s_n)$ is obviously exact for the model sequence (7.3-15). If we now use eq. (7.1-6), we see that the transformation $\Lambda_k^{(n)}(\beta, s_n)$ satisfies

$$\Lambda_k^{(n)}(\beta, s_n) = \frac{\Delta^k\{(n+\beta)^k s_n\}}{\Delta^k (n+\beta)^k}, \qquad k, n \in \mathbb{N}_0. \tag{7.3-17}$$

The denominator in eq. (7.3-17) can be expressed in closed form. If we use (see p. 4 of ref. [65])

$$\sum_{j=0}^{k} (-1)^j \binom{k}{j} (\beta + n + j)^k = (-1)^k k!, \qquad k, n \in \mathbb{N}_0, \tag{7.3-18}$$

together with eq. (2.4-8), we find:

$$\Delta^k (n+\beta)^k = k!. \tag{7.3-19}$$

Combination of eqs. (7.3-17) and (7.3-19) with eq. (2.4-8) gives us an explicit expression for this sequence transformation:

$$\Lambda_k^{(n)}(\beta, s_n) = (-1)^k \sum_{j=0}^{k} (-1)^j \frac{(\beta+n+j)^k}{j!(k-j)!} s_{n+j}, \qquad k, n \in \mathbb{N}_0. \tag{7.3-20}$$

If we combine eq. (7.2-6) with eqs. (7.3-17) and (7.3-19), we can derive the following recursive scheme for the sequence transformation $\Lambda_k^{(n)}(\beta, s_n)$:

$$\Lambda_0^{(n)}(\beta, s_n) = s_n, \qquad n \in \mathbb{N}_0, \tag{7.3-21a}$$

$$\Lambda_{k+1}^{(n)}(\beta, s_n) = \Lambda_k^{(n+1)}(\beta, s_{n+1}) + \frac{\beta + n}{k+1} \Delta \Lambda_k^{(n)}(\beta, s_n), \qquad k, n \in \mathbb{N}_0. \tag{7.3-21b}$$

This recursive scheme for the sequence transformation $\Lambda_k^{(n)}(\beta, s_n)$ corresponds to the special case $x_n = 1/(\beta + n)$ in the recursive scheme (6.1-5) which computes the Richardson extrapolation scheme with arbitrary interpolation points $\{x_n\}$.

A more complete discussion of the properties of the linear but nonregular sequence transformation $\Lambda_k^{(n)}(\beta, s_n)$ can be found in articles by Salzer [66, 67], Salzer and Kimbro [68], and Wimp [69] as well as in Wimp's book (see pp. 35–38 of ref. [23]).

The sequence transformations $u_k^{(n)}(\beta, s_n)$, eq. (7.3-5), and $t_k^{(n)}(\beta, s_n)$, eq. (7.3-7), require the sequence elements $s_{n-1}, s_n, s_{n+1}, \ldots, s_{n+k}$ for their computation, whereas $d_k^{(n)}(\beta, s_n)$, eq. (7.3-9), requires the sequence elements $s_n, s_{n+1}, \ldots, s_{n+k+1}$. Hence, they are all transformations of order $k+1$. The sequence transformation $v_k^{(n)}(\beta, s_n)$, eq. (7.3-11), requires the sequence elements s_{n-1}, $s_n, s_{n+1}, \ldots, s_{n+k+1}$ which implies that it is a transformation of order $k+2$. The linear sequence transformation $\Lambda_k^{(n)}(\beta, s_n)$, eq. (7.3-20), requires the sequence elements $s_n, s_{n+1}, \ldots, s_{n+k}$, i.e., it is a transformation of order k.

The situation is somewhat different if the transforms with superscript $n = 0$ are computed because then $u_k^{(0)}(\beta, s_0)$ and $t_k^{(0)}(\beta, s_0)$ are transformations of order k, whereas $d_k^{(0)}(\beta, s_0)$ and $v_k^{(0)}(\beta, s_0)$ are transformations of order $k+1$.

7.4. Sidi's generalization of Levin's sequence transformation

As discussed in section 6.1, the Richardson extrapolation scheme, eq. (6.1-5), is exact for model sequences of the following type:

$$s_n = s + \sum_{j=0}^{k-1} c_j x_n^{j+1}, \qquad k, n \in \mathbb{N}_0. \tag{7.4-1}$$

The interpolation points $\{x_n\}$ have to satisfy eq. (6.1-4). Very natural interpolation points for the Richardson extrapolation scheme are $x_n = 1/(n + \beta)$ with $\beta > 0$. If this set of extrapolation points is used in eq. (7.4-1) we obtain the model sequence (7.3-15) for which the sequence transformation $\Lambda_k^{(n)}(\beta, s_n)$, eq. (7.3-17), is exact. This implies that the Richardson extrapolation scheme (6.1-5) with the interpolation points $x_n = 1/(n + \beta)$ is a special case of Levin's sequence transformation with $\omega_n = 1/(n + \beta)$.

However, the Richardson extrapolation scheme is also in some sense more general than Levin's sequence transformation since the interpolation points $\{x_n\}$ only have to satisfy eq. (6.1-4) but otherwise they are completely arbitrary. In Levin's sequence transformation it is tacitly assumed that the interpolation points $\{x_n\}$ always satisfy $x_n = 1/(n + \beta)$ with $\beta > 0$.

Now, one can try to construct a sequence transformation which combines the advantageous features of the Levin transformation and the Richardson extrapolation scheme. This was accomplished by Sidi [70] who constructed a sequence transformation, which he called *generalized Richardson extrapolation process*, on the basis of the following model sequence:

$$s_n = s + \omega_n \sum_{j=0}^{k-1} c_j x_n^j, \qquad k, n \in \mathbb{N}_0. \tag{7.4-2}$$

This model sequence combines the features of the model sequence (7.4-1) for the Richardson extrapolation scheme, eq. (6.1-5), since it contains arbitrary interpolation points $\{x_n\}$, and of the model sequence (7.1-1) for the Levin transformation, eq. (7.1-7), since it contains arbitrary remainder estimates $\{\omega_n\}$.

For the construction of such a sequence transformation, which is exact for the above model sequence, eq. (7.4-2) is rewritten in the following way:

$$[s_n - s]/\omega_n = \sum_{j=0}^{k-1} c_j x_n^j. \tag{7.4-3}$$

Obviously, the right-hand side of eq. (7.4-3) is a polynomial of degree $k-1$ in the variable x_n. Hence, the limit or antilimit s of this sequence can be determined if a linear operator can be found which annihilates the polynomial on the right-hand side of eq. (7.4-3).

This annihilation of a polynomial can be accomplished with the help of *divided differences* which for instance occur in Newton's interpolation formula. A discussion of divided differences and their properties can be found in any book on numerical analysis or also in books on finite differences such as Nörlund [71] or Milne-Thomson [72].

Let $\{x_n\}$ with $n \in \mathbb{N}_0$ be a set of distinct interpolation points. Then the divided differences of orders $0, 1, \ldots, k, k+1, \ldots$ of a given function f are defined recursively by the relations

$$f[x_n] = f(x_n), \tag{7.4-4a}$$

$$f[x_n, \ldots, x_{n+k+1}] = \frac{f[x_{n+1}, \ldots, x_{n+k+1}] - f[x_n, \ldots, x_{n+k}]}{x_{n+k+1} - x_n}, \qquad k, n \in \mathbb{N}_0. \tag{7.4-4b}$$

The divided differences $f[x_n, \ldots, x_{n+k}]$ can also be expressed in closed form:

$$f[x_n, \ldots, x_{n+k}] = \sum_{j=0}^{k} f(x_{n+j}) \prod_{\substack{i=0 \\ i \neq j}}^{k} \frac{1}{x_{n+j} - x_{n+i}}, \qquad k, n \in \mathbb{N}_0. \tag{7.4-5}$$

It follows either from this expression or from the recursive scheme (7.4-4) that the divided differences $f[x_n, \ldots, x_{n+k}]$ are linear functions of the initial values $f(x_n), \ldots, f(x_{n+k})$. In addition, it can be shown that if $p_m(x)$ is a polynomial of degree m in x,

$$p_m(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_m x^m, \tag{7.4-6}$$

then all its divided differences with $k > m$ satisfy

$$p_m[x_n, \ldots, x_{n+k}] = 0, \qquad k > m. \tag{7.4-7}$$

The divided differences $f[x_n, \ldots, x_{n+k}]$ with k, $n \in \mathbb{N}_0$ can be identified with the elements $\Delta_k^{(n)}(f)$ of a 2-dimensional rectangular array. With this convention, the recursive scheme (7.4-4) for the computation of the divided differences of a given function f can be rewritten in the following way:

$$\Delta_0^{(n)}(f) = f(x_n), \qquad n \in \mathbb{N}_0, \tag{7.4-8a}$$

$$\Delta_{k+1}^{(n)}(f) = \frac{\Delta_k^{(n+1)}(f) - \Delta_k^{(n)}(f)}{x_{n+k+1} - x_n}, \qquad k, n \in \mathbb{N}_0. \tag{7.4-8b}$$

If we assume that two functions $\mathscr{S}(x)$ and $\Omega(x)$ of a continuous variable x exist, which coincide at the interpolation points x_n with s_n and ω_n, respectively,

$$\mathscr{S}(x_n) = s_n, \qquad \Omega(x_n) = \omega_n, \tag{7.4-9}$$

and which satisfy

$$[\mathscr{S}(x) - s]/\Omega(x) = \sum_{j=0}^{k-1} c_j x^j, \tag{7.4-10}$$

we see that we only have to compute the divided difference of order k for the left-hand side of eq. (7.4-10) at the interpolation points $x_n, \ldots, x_{n+k}$ in order to obtain the limit or antilimit s. Hence, Sidi's generalized Richardson extrapolation process $\mathscr{R}_k^{(n)}(s_n, \omega_n, x_n)$ can be defined in terms of divided differences in the following way:

$$\mathscr{R}_k^{(n)}(s_n, \omega_n, x_n) = \frac{\{\mathscr{S}(x)/\Omega(x)\}[x_n, \ldots, x_{n+k}]}{\{1/\Omega(x)\}[x_n, \ldots, x_{n+k}]}, \qquad k, n \in \mathbb{N}_0. \tag{7.4-11}$$

It follows from eq. (7.4-8) that both numerator and denominator of this transformation can be computed with the help of the same 3-term recurrence formula:

$$R_{k+1}^{(n)} = \frac{R_k^{(n+1)} - R_k^{(n)}}{x_{n+k+1} - x_n}, \qquad k, n \in \mathbb{N}_0. \tag{7.4-12}$$

If we use in eq. (7.4-12) the starting values

$$R_0^{(n)} = s_n/\omega_n, \qquad n \in \mathbb{N}_0, \tag{7.4-13}$$

we obtain the numerator of Sidi's generalized Richardson extrapolation process, eq. (7.4-11), and if we use the starting values

$$R_0^{(n)} = 1/\omega_n, \qquad n \in \mathbb{N}_0, \tag{7.4-14}$$

we obtain the denominator in eq. (7.4-11).

Obviously, the Richardson extrapolation scheme (6.1-5) is a special case of Sidi's generalized Richardson extrapolation process $\mathscr{R}_k^{(n)}(s_n, \omega_n, x_n)$. To see this one only has to specialize $\Omega(x) = x$ in eq. (7.4-11) which also implies $\omega_n = x_n$.

With this specialization, the Richardson extrapolation scheme (6.1-5) follows from eqs. (7.4-11) and (7.4-12). It is only necessary to compute the divided differences for $1/x$ at the interpolation points which can be done quite easily. The result is (see p. 8 of ref. [72]):

$$\frac{1}{x}[x_n, \ldots, x_{n+k}] = \frac{(-1)^k}{x_n \cdots x_{n+k}}. \tag{7.4-15}$$

This implies that for $\omega_n = x_n$ the denominator of the generalized Richardson extrapolation process $\mathscr{R}_k^{(n)}(s_n, \omega_n, x_n)$ can be expressed in closed form. If we now set

$$R_0^{(n)} = s_n/x_n, \tag{7.4-16}$$

and

$$\mathscr{N}_k^{(n)} = (-1)^k x_n \cdots x_{n+k} R_k^{(n)}, \qquad k, n \in \mathbb{N}_0, \tag{7.4-17}$$

we immediately obtain from eqs. (7.4-11) and (7.4-12) the Richardson extrapolation scheme, eq. (6.1-5).

Further generalizations of Sidi's generalized Richardson extrapolation are possible. For instance, let us consider the following model sequence:

$$s_n = s + \sum_{\mu=0}^{m} \omega_n^{(\mu)} \sum_{j=0}^{k_\mu - 1} c_j^{(\mu)} x_n^j, \qquad m, n, k_\mu \in \mathbb{N}_0. \tag{7.4-18}$$

This model sequence contains $m+1$ different sets of remainder estimates $\{\omega_n^{(0)}\}, \ldots, \{\omega_n^{(m)}\}$. If we set $m = 0$ in eq. (7.4-18), we obtain the model sequence (7.4-2) as a special case.

Sidi [73] constructed a sequence transformation which is exact for the above model sequence. Originally, this sequence transformation was defined as the ratio of determinants, which is computationally quite unattractive. But recently, Ford and Sidi [74] could derive an algorithm which permits a recursive computation of Sidi's sequence transformation.

From Sidi's very general sequence transformation [73], which is exact for the model sequence (7.4-18), other sequence transformations can be obtained by specializing the interpolation points $\{x_n\}$. For instance, in earlier articles Levin and Sidi [75,76] had chosen the interpolation points according to the rule $x_n = 1/(n+1)$ and had obtained a generalization of Levin's sequence transformation with several sets of remainder estimates.

The typical feature of these sequence transformations is that they involve several sets of remainder estimates $\{\omega_n^{(0)}\}, \ldots, \{\omega_n^{(m)}\}$. Consequently, it is to be expected that these sequence transformations should be particularly powerful if sequences are to be accelerated which are superpositions of sequences with different convergence types. The numerical examples presented in the literature [74–76] confirm this opinion and it seems that the sequence transformations, which are based upon variants of the model sequence (7.4-18), are able to accelerate convergence even in cases in which other transformations fail.

The power of these transformations stems from the occurrence of several different sets of remainder estimates. This is at the same time also the major disadvantage of these transformations. This may sound paradoxical. However, one should take into consideration that the popularity and the success of Levin's sequence transformation, eq. (7.1-7), is largely due to the fact that the simple remainder estimates (7.3-4), (7.3-6), (7.3-8), and (7.3-10) work remarkably well even in situations in which only the numerical values of a relatively small number of sequence elements $s_m, \ldots, s_{m+l}$ are known. If we would try to use in such a situation a sequence transformation, which is based upon a variant of the model sequence (7.4-18), we would first have to find a way of determining numerically more than a single set of remainder estimates. Unfortunately, no simple and manageable theory is in sight which would yield more than a single set of remainder estimates if only the numerical values of a few sequence elements are known.

Consequently, if little or no information about the analytical behaviour of the remainders $\{r_n\}$ is available and if only a relatively small number of sequence elements $s_m, \ldots, s_{m+l}$ are known, it is normally not possible to use sequence transformations, which are based upon a variant of the model sequence (7.4-18), although they are potentially very powerful. The reason is that such a sequence transformation has too many degrees of freedom since it requires the input of $m+1$ different sets of remainder estimates $\{\omega_n^{(0)}\}, \ldots, \{\omega_n^{(m)}\}$ as well as a set of interpolation points $\{x_n\}$.

In such a situation, one is more or less forced to use a simpler and probably also less efficient sequence transformation which, however, does not require such a detailed knowledge about the sequence to be transformed.

7.5. Programming the Levin transformation

In this section it will be discussed how the general Levin transformation (7.1-7) and its variants (7.1-8), (7.3-5), (7.3-7), (7.3-9), and (7.3-11) can be programmed efficiently. It is a typical feature of the general Levin transformation $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$ and its variants that they can all be expressed as the ratio of two finite sums and that both the numerator and the denominator sum can be computed by the same 3-term recursion (7.2-8).

Consequently, a program for the general Levin transformation or any of its variants has to compute simultaneously the numerator and denominator sums of the transformation. In the case of the general Levin transformation a program requires not only the input of the sequence elements s_n, but also the remainder estimates ω_n. But otherwise, such a program should have essentially the same features as the subroutines EPSAL and AITKEN, which were described in sections 4.3 and 5.2, respectively.

This means that such a program for the general Levin transformation should read in the sequence elements $s_0, s_1, \ldots, s_m, \ldots$ and remainder estimates $\omega_0, \omega_1, \ldots, \omega_m, \ldots$ successively, starting with s_0 and ω_0. After the input of each new pair s_m and ω_m as many new numerator and denominator sums of the Levin transformation (7.1-7) or its variants should be computed as possible.

The elements $L_k^{(n)}(\beta)$, which either represent numerators or denominators of the general Levin transformation and which are computed with the help of the 3-term recurrence formula (7.2-8), can be arranged in a rectangular scheme in such a way that the superscript n indicates the row and the subscript k the column of the 2-dimensional array:

$$\begin{array}{cccccc}
L_0^{(0)}(\beta) & L_1^{(0)}(\beta) & L_2^{(0)}(\beta) & \dots & L_n^{(0)}(\beta) & \dots \\
L_0^{(1)}(\beta) & L_1^{(1)}(\beta) & L_2^{(1)}(\beta) & \dots & L_n^{(1)}(\beta) & \dots \\
L_0^{(2)}(\beta) & L_1^{(2)}(\beta) & L_2^{(2)}(\beta) & \dots & L_n^{(2)}(\beta) & \dots \\
L_0^{(3)}(\beta) & L_1^{(3)}(\beta) & L_2^{(3)}(\beta) & \dots & L_n^{(3)}(\beta) & \dots \\
\vdots & \vdots & \vdots & \ddots & \vdots & \ddots \\
L_0^{(n)}(\beta) & L_1^{(n)}(\beta) & L_2^{(n)}(\beta) & \dots & L_n^{(n)}(\beta) & \dots \\
\vdots & \vdots & \vdots & \ddots & \vdots & \ddots
\end{array} \tag{7.5-1}$$

The entries in the first column of the array are the starting values of the recursion. If the starting values $L_0^{(n)}(\beta)$ are chosen according to eq. (7.2-9), the 3-term recurrence formula (7.2-8) will produce a table $L_k^{(n)}(\beta)$ of numerators of the general Levin transformation, and if the starting values are chosen according to eq. (7.2-10), a table of denominators will be computed. The 3 elements, which are connected by the linear 3-term recursion (7.2-8), form a triangle in the Levin table:

$$\begin{array}{ll}
L_k^{(n)}(\beta) & L_{k+1}^{(n)}(\beta) \\
L_k^{(n+1)}(\beta) &
\end{array} \tag{7.5-2}$$

This pattern implies that the recursion (7.2-8) has to proceed along counterdiagonals in the Levin table. Consequently, it is advantageous to rewrite the recurrence formula (7.2-8) in the following way:

$$L_j^{(n-j)}(\beta) = L_{j-1}^{(n-j+1)}(\beta) - \frac{(\beta+n-j)(\beta+n-1)^{j-2}}{(\beta+n)^{j-1}} L_{j-1}^{(n-j)}(\beta),$$

$$n \geq 1, \qquad 1 \leq j \leq n. \tag{7.5-3}$$

It also follows from the triangular structure of this recursion that the sequence elements s_0, $s_1, \dots, s_m$ and the remainder estimates $\omega_0, \omega_1, \dots, \omega_m$ have to be known for the computation of all elements $L_\nu^{(n-\nu)}$ with $0 \leq n \leq m$ and $0 \leq \nu \leq n$ of the numerator and denominator tables. Since the transforms with the highest values of the subscript normally give the best results, our approximation to the limit s of the sequence $\{s_n\}$ to be transformed will be:

$$\{s_0, \omega_0; s_1, \omega_1; \dots; s_m, \omega_m\} \to \mathscr{L}_m^{(0)}(\beta, s_0, \omega_0), \qquad m \in \mathbb{N}_0. \tag{7-5.4}$$

Essentially this means that we shall use the following sequence of Levin transforms with minimal superscripts and maximal subscripts as approximations to the limit s:

$$\mathscr{L}_0^{(0)}(\beta, s_0, \omega_0), \mathscr{L}_1^{(0)}(\beta, s_0, \omega_0), \dots, \mathscr{L}_m^{(0)}(\beta, s_0, \omega_0), \dots . \tag{7.5-5}$$

Because of the triangular structure (7.5-2) of the recurrence formula (7.5-3) and since the computation proceeds along a counterdiagonal of the Levin table, a single 1-dimensional array will be sufficient for the computation of the $L_j^{(n-j)}(\beta)$ which are either numerator or denominator sums of the general Levin transformation. For that purpose the $L_j^{(n-j)}(\beta)$ are stored in a 1-dimensional array l in such a way that the superscript coincides with the index of the corresponding array element:

$$L_\nu^{(n-\nu)}(\beta) \to l(n-\nu), \qquad n \geq 0, \quad 0 \leq \nu \leq n. \tag{7.5-6}$$

With this convention the recurrence formula (7.5-3) can be reformulated in terms of the elements of the array l:

$$l(n-j) \leftarrow l(n-j+1) - \frac{(\beta+n-j)(\beta+n-1)^{j-2}}{(\beta+n)^{j-1}} l(n-j),$$

$$n \geq 1, \qquad 1 \leq j \leq n. \tag{7.5-7}$$

This computational scheme is simpler than the corresponding scheme for Wynn's ϵ algorithm, eq. (4.3-10), since no auxiliary variables are needed here.

Essentially the same approach works also in the case of the Richardson extrapolation scheme, eq. (6.1-5), or in the case of Sidi's generalized Richardson extrapolation process, eq. (7.4-11). For instance, the recurrence formula (7.4-12) for Sidi's generalized Richardson extrapolation process can be rewritten in the following way:

$$R_j^{(n-j)} = \frac{R_{j-1}^{(n-j+1)} - R_{j-1}^{(n-j)}}{x_n - x_{n-j}}, \qquad n \geq 1, \quad 1 \leq j \leq n. \tag{7.5-8}$$

If the elements $R_j^{(n-j)}$ and the interpolation points x_n are stored in 1-dimensional arrays r and ξ according to the rules

$$R_j^{(n-j)} \to r(n-j), \qquad n \geq 0, \quad 0 \leq j \leq n, \tag{7.5-9}$$

$$x_n \to \xi(n), \qquad n \geq 0, \tag{7.5-10}$$

the recurrence formula (7.5-8) can be reformulated in terms of the elements of the arrays r and ξ:

$$r(n-j) \leftarrow \frac{r(n-j+1) - r(n-j)}{\xi(n) - \xi(n-j)}, \qquad n \geq 1, \quad 1 \leq j \leq n. \tag{7.5-11}$$

Similarly, the Richardson extrapolation scheme, eq. (6.1-5), can be rewritten in the following way:

$$\mathcal{N}_0^{(n)}(s_n, x_n) = s_n, \qquad n \geq 0, \tag{7.5-12a}$$

$$\mathcal{N}_j^{(n-j)}(s_{n-j}, x_{n-j}) = \frac{x_{n-j}\mathcal{N}_{j-1}^{(n-j+1)}(s_{n-j+1}, x_{n-j+1}) - x_n \mathcal{N}_{j-1}^{(n-j)}(s_{n-j}, x_{n-j})}{x_{n-j} - x_n},$$

$$n \geq 1, \qquad 1 \leq j \leq n. \tag{7.5-12b}$$

If the elements $\mathcal{N}_j^{(n-j)}(s_{n-j}, x_{n-j})$ and the interpolation points x_n are stored in 1-dimensional arrays N and ξ according to the rules

$$\mathcal{N}_j^{(n-j)}(s_{n-j}, x_{n-j}) \to N(n-j), \qquad n \geq 0, \quad 0 \leq j \leq n, \tag{7.5-13a}$$

$$x_n \to \xi(n), \qquad n \geq 0, \tag{7.5-13b}$$

the recurrence formula (7.5-12) can be reformulated in terms of the elements of the arrays N and ξ:

$$N(n) \leftarrow s_n, \qquad n \geq 0, \tag{7.5-14a}$$

$$N(n-j) \leftarrow \frac{\xi(n)N(n-j+1) - \xi(n-j)N(n-j)}{\xi(n) - \xi(n-j)}, \qquad n \geq 1, \quad 1 \leq j \leq n. \tag{7.5-14b}$$

Obviously, the computational scheme (7.5-11) for Sidi's generalized Richardson extrapolation process and the computational scheme (7.5-14) for the Richardson extrapolation process are structurally identical with the computational scheme (7.5-7) for Levin's general sequence transformation. This implies that programs for the Richardson extrapolation process and Sidi's generalization of the Richardson extrapolation process would have the same features as a program for Levin's general sequence transformation.

A program which computes the general Levin transformation (7.1-7) or any of its variants has to take precautions against an exact or approximate vanishing of the denominator sum. Again, this can be accomplished by introducing two variables HUGE and TINY which have values that are close to but not identical with the largest and smallest floating point number representable on the computer. If the denominator sum of the transform $\mathcal{L}_m^{(0)}(\beta, s_0, \omega_0)$ is smaller in magnitude than TINY, then $\mathcal{L}_m^{(0)}(\beta, s_0, \omega_0)$ will be set equal to HUGE. This check is only necessary if the approximation to the limit according to eq. (7.5-5) is computed.

The following FORTRAN 77 subroutine GLEVIN computes the numerator and denominator sum of the general Levin transformation $\mathcal{L}_m^{(0)}(\beta, s_0, \omega_0)$ with the help of the recurrence formula (7.5-3) in two 1-dimensional arrays ARUP and ARLO. It is safeguarded against an exact or approximate vanishing of the denominator sum by using the variables HUGE and TINY described above. The sequence elements s_n and the remainder estimates ω_n with $n = 0, 1, 2, \ldots$ have to be computed in a DO loop in the calling program. Whenever a new pair s_n and ω_n is computed in the outer DO loop this subroutine GLEVIN has to be called again and a new string of transforms $\mathscr{L}_j^{(n-j)}(\beta, s_{n-j}, \omega_{n-j})$ with $0 \le j \le n$ is computed. The new sequence element s_n is read in via the variable SOFN, and the new remainder estimate ω_n is read in via the variable ROFN. The approximation to the limit, which is given by the ratio ARUP(0)/ARLO(0), is returned via the variable ESTLIM.

Again, it should be noted that GLEVIN only calculates the approximation to the limit according to eq. (7.5-4). The convergence of the whole process has to be analyzed in the calling program.

Finally, it should be noted that the description of a FORTRAN IV program, which computes Levin's u transformation, eq. (7.3-5), with $\beta = 1$, can be found in ref. [57].

```
      SUBROUTINE GLEVIN(SOFN,ROFN,BETA,N,ARUP,ARLO,LARRAY,ESTLIM)
      DIMENSION ARUP(0:LARRAY),ARLO(0:LARRAY)
      PARAMETER ( HUGE = 1.E+60 , TINY = 1.E-60 )
      ARLO(N)   = 1.0 / ((FLOAT(N)+BETA)*ROFN)
      ARUP(N)   = SOFN * ARLO(N)
      IF (N.GT.0) THEN
        BN1     = BETA + FLOAT(N-1)
        BN2     = BETA + FLOAT(N)
        DO 10 J = 1,N
          FACT  = (BETA+FLOAT(N-J))*BN1**(J-1)/BN2**J
          ARUP(N-J)  = ARUP(N-J+1) - FACT*ARUP(N-J)
          ARLO(N-J)  = ARLO(N-J+1) - FACT*ARLO(N-J)
10      CONTINUE
      END IF
        IF (ABS(ARLO(0)).LT.TINY) THEN
          ESTLIM = HUGE
        ELSE
          ESTLIM = ARUP(0)/ARLO(0)
        END IF
      RETURN
      END
```

8. Sequence transformations based upon factorial series

8.1. Factorial series

In extensive numerical studies performed by Smith and Ford [29,30] and also in other articles [57, 60–64] it was demonstrated that the general Levin transformation (7.1-7) and its variants (7.3-5), (7.3-7), (7.3-9), and (7.3-11) are remarkably powerful sequence transformations. Consequently, if one tries to derive alternative sequence transformations it should definitely be worthwhile to try to retain as many of the advantageous features of the Levin transformation as possible.

It is the conviction of the author that the power of the Levin transformation is due to the fact that a sequence $\{\omega_n\}$ of remainder estimates is explicitly included in the transformation. This is not necessarily an advantage because if the remainder estimates ω_n, which are used, are poor approximants of the actual remainders r_n, the Levin transformation will lose much of its efficiency. However, if the remainder estimates are good approximants, it is likely that the Levin transformation will produce excellent results.

In section 7.4, it was shown that Sidi's generalized Richardson extrapolation process (7.4-11) – which is also a generalization of Levin's sequence transformation – is by construction exact if the remainders r_n of the sequence to be transformed can be written as an remainder estimate ω_n multiplied by a polynomial of degree $k-1$ in x_n,

$$r_n = \omega_n \sum_{j=0}^{k-1} c_j x_n^j, \qquad k, n \in \mathbb{N}_0. \tag{8.1-1}$$

The interpolation points x_n have to satisfy eq. (6.1-4) which means that they have to approach zero as $n \to \infty$. If we choose in eq. (8.1-1) $x_n = 1/(n+\beta)$ we obtain the remainder of the model sequence (7.1-1) which is the basis for the construction of Levin's sequence transformation. Model sequences with remainders of the above type can be viewed to be finite approximations of sequence elements s_n which can be written as Poincaré-type asymptotic expansions with respect to the asymptotic sequence $\{\!\{\omega_n x_n^j\}\!\}$ with $n,\ j \in \mathbb{N}_0$,

$$s_n \sim s + \omega_n \sum_{j=0}^{\infty} c_j x_n^j, \qquad n \to \infty. \tag{8.1-2}$$

Essentially this means that the sequence of remainder estimates $\{\!\{\omega_n\}\!\}$ should be chosen in such a way that the ratio $(s_n - s)/\omega_n$ can be written as an asymptotic power series in the interpolation points $\{\!\{x_n\}\!\}$,

$$(s_n - s)/\omega_n \sim \sum_{j=0}^{\infty} c_j x_n^j, \qquad n \to \infty. \tag{8.1-3}$$

If one tries to construct alternative sequence transformations, which also incorporate explicit remainder estimates via the auxiliary sequence $\{\!\{\omega_n\}\!\}$, the simplest approach would be to replace the asymptotic power series on the right-hand side of eq. (8.1-3) by some other kind of expansion. This means that in eq. (8.1-3) instead of the powers $\{\!\{x_n^j\}\!\}$ some other asymptotic sequence $\{\!\{\varphi_j(n)\}\!\}$ with $n,\ j \in \mathbb{N}_0$ would have to be used. Consequently, it would be necessary to construct a transformation which is exact for the following class of model sequences:

$$s_n = s + \omega_n \sum_{j=0}^{k-1} c_j \varphi_j(n), \qquad k, n \in \mathbb{N}_0. \tag{8.1-4}$$

Such a transformation would also be a special case of the general extrapolation algorithm $E_k(s_n)$, eq. (3.3-2), which was introduced by Brezinski [31] and Håvie [32]. This follows immediately if $f_j(n)$ in eq. (3.3-1) is replaced by $\omega_n \varphi_j(n)$.

In principle, every set $\{\!\{\varphi_j(n)\}\!\}$ of functions of n could be used in eq. (8.1-4) which satisfies

$$\varphi_0(n) = 1, \qquad n \in \mathbb{N}_0, \tag{8.1-5a}$$

$$\varphi_{j+1}(n) = o(\varphi_j(n)), \qquad j \in \mathbb{N}_0, \quad n \to \infty. \tag{8.1-5b}$$

However, such a minimal requirement on the set $\{\varphi_j(n)\}$ would not suffice to make a new transformation practically useful, let alone to give it any advantage over already existing transformations.

In order to be practically useful, a new sequence transformation should produce excellent numerical results in convergence acceleration and summation processes. Preferably, it should be as good as the Levin transformation or maybe even better. However, this would not be enough. Since the evaluation of large order determinants, as they for instance occur in eq. (3.3-2), is computationally very unattractive, a comparatively simple recursive scheme, which allows a fast and reliable computation of the transformation, would also be of considerable importance. The derivation of an explicit expression of the type of eq. (7.1-7) for the new transformation would also be desirable since this would give us a better chance of understanding the mechanism as well as the shortcomings of the new transformation.

It is not a simple task to find an alternative asymptotic sequence other than powers $\{x_n^j\}$ with $n,\ j \in \mathbb{N}_0$ which leads to a sequence transformation satisfying the requirements mentioned above. However, it will become clear later that a new class of sequence transformations with most of the advantageous features of the Levin transformation and some new ones can be derived quite easily if it is assumed that the ratio $(s_n - s)/\omega_n$ is expressed as a factorial series and not as an asymptotic power series as in eq. (8.1-3).

Let $\Omega(z)$ be a function which vanishes as $|z| \to \infty$. Then, a factorial series for $\Omega(z)$ is an expansion of the following type,

$$\Omega(z) = \frac{c_0}{z} + \frac{c_1}{z(z+1)} + \frac{c_2}{z(z+1)(z+2)} + \cdots = \sum_{\nu=0}^{\infty} \frac{c_\nu}{(z)_{\nu+1}}. \tag{8.1-6}$$

Here, $(z)_{\nu+1}$ is a Pochhammer symbol which is commonly defined as the ratio of two gamma functions (see p. 3 of ref. [34]),

$$(z)_{\nu+1} = \Gamma(z+\nu+1)/\Gamma(z) = z(z+1)\ldots(z+\nu), \qquad \nu \in \mathbb{N}_0. \tag{8.1-7}$$

Factorial series have a long tradition in mathematics. For instance, a large part of Stirling's book [5], which was published in 1730, deals with factorial series. In the nineteenth century the theory of factorial series was developed and refined by a variety of authors. A fairly complete survey of the older literature on this subject can be found in books by Nielsen [77] and Nörlund [71]. In these two books good treatments of the fundamental properties of factorial series can be found.

Factorial series have a remarkable property which will also be utilized quite profitably in this report: it is extremely simple to apply higher powers of the difference operator Δ to a factorial series. Consequently, factorial series play a similar rôle in the theory of difference equations as power series in the theory of differential equations. This explains why factorial series were often treated in the classical literature on finite differences, e.g., in books by Nörlund [71,78] and Milne-Thomson [72].

Quite interesting in the context of convergence acceleration and summation is also Borel's book on divergent series [79] in which the connection between factorial series and summability is emphasized.

However, it seems that in recent years mathematicians have lost interest in factorial series. This can be concluded from the fact that only quite rarely references dealing with factorial series can be found in the more modern mathematical literature. Notable exceptions are a book by Wasow [80], which contains a chapter on factorial series, and an article by Iseki and Iseki [81] on remainder estimates of truncated factorial series. In the opinion of the author this declining interest in factorial series is quite deplorable because the numerical potential of factorial series has not yet been fully exploited.

The fact, that the argument z of a factorial series occurs in Pochhammer symbols and not in the form of inverse powers as in asymptotic power series, has some far-reaching consequences for the convergence properties of factorial series.

A power series converges in the interior of a circle which may coincide with the whole complex plane $\mathbb{C}$ or which may shrink to a single point as in the case of divergent asymptotic series. However, if a factorial series converges then according to Landau [82] it converges in a half-plane. This means that if a factorial series converges for some $z_0 \in \mathbb{C}$ it also converges with the possible exception of the points $z = 0, -1, -2, \ldots$ for all $z \in \mathbb{C}$ with $\mathrm{Re}(z) > \mathrm{Re}(z_0)$.

The different convergence properties of power series and factorial series are demonstrated quite drastically by the following two infinite series which both have the same numerical coefficients $c_m = (-1)^m m!$:

$$\frac{1}{x} - \frac{1}{x^2} + \frac{2}{x^3} - \frac{6}{x^4} + \cdots = \sum_{m=0}^{\infty} \frac{(-1)^m m!}{x^{m+1}}, \tag{8.1-8}$$

$$\frac{1}{x} - \frac{1}{(x)_2} + \frac{2}{(x)_3} - \frac{6}{(x)_4} + \cdots = \sum_{m=0}^{\infty} \frac{(-1)^m m!}{(x)_{m+1}}. \tag{8.1-9}$$

The power series diverges for all finite values of $x \in \mathbb{R}$, whereas the factorial series converges for all $x > 0$.

Because of the different convergence properties of factorial and power series it may happen that a given function $\Omega(z)$, which possesses a representation as a divergent asymptotic power series,

$$\Omega(z) \sim \frac{c_0'}{z} + \frac{c_1'}{z^2} + \frac{c_3'}{z^3} + \cdots, \qquad z \to \infty, \tag{8.1-10}$$

possesses also a representation as a convergent factorial series according to eq. (8.1-6).

The algebraic processes, by means of which the two series expansions (8.1-6) and (8.1-10) can be transformed into each other, were already described by Stirling [5] in 1730. A more modern description of Stirling's method can be found in Nielsen's book (see pp. 272–282 of ref. [77]). A detailed investigation of the problems associated with the transformation of an asymptotic series into a convergent factorial series can be found in a long article by Watson [83].

8.2. *A factorial series analogue of Levin's transformation*

The following model sequence will be the basis for the new class of sequence transformations which will be discussed in this section:

$$s_n = s + \omega_n \sum_{j=0}^{k-1} c_j/(n+\beta)_j, \qquad k,\, n \in \mathbb{N}_0. \tag{8.2-1}$$

This sequence is formally almost identical with the model sequence (7.1-1) which is the basis of the Levin transformation. The only difference is that the powers $(n+\beta)^j$ in eq. (7.1-1) are replaced by Pochhammer symbols $(n+\beta)_j$. Concerning the sequence $\{\omega_n\}$ of remainder estimates it is again assumed that the ω_n are known functions of n which are different from zero and distinct for all finite values of n. But otherwise, the ω_n are in principle completely arbitrary.

The parameter β in eq. (8.2-1) is subject to the restriction that the Pochhammer symbols $(n+\beta)_j$ must not be zero for all n, $j \in \mathbb{N}_0$. This is certainly guaranteed if β is not a negative integer or zero. However, the elements of the model sequence (8.2-1) will serve as finite approximations to factorial series of the following kind:

$$s_n \sim s + \omega_n \sum_{j=0}^{\infty} c_j/(n+\beta)_j, \qquad n \to \infty. \tag{8.2-2}$$

In expansions of that kind negative values of β will lead to different signs of the terms of this factorial series if either $n+\beta < 0$ or $n+\beta > 0$ holds. Since the sign pattern of the terms of such a factorial series should not change as n increases, we see that as in the case of the Levin transformation the additional restriction $\beta > 0$ is necessary. But otherwise, β is in principle completely arbitrary.

In eq. (8.2-1) there occur $k+1$ unknown quantities, the limit or antilimit s and the k linear coefficients $c_0, \ldots, c_{k-1}$. Hence, if $k+1$ sequence elements $s_n, \ldots, s_{n+k}$ are known, the sequence transformation $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$ can be defined according to Cramer's rule by the following ratio of determinants:

$$\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n) = \frac{\begin{vmatrix} s_n & \cdots & s_{n+k} \\ \omega_n & \cdots & \omega_{n+k} \\ \vdots & \ddots & \vdots \\ \omega_n/(\beta+n)_{k-1} & \cdots & \omega_{n+k}/(\beta+n+k)_{k-1} \end{vmatrix}}{\begin{vmatrix} 1 & \cdots & 1 \\ \omega_n & \cdots & \omega_{n+k} \\ \vdots & \ddots & \vdots \\ \omega_n/(\beta+n)_{k-1} & \cdots & \omega_{n+k}/(\beta+n+k)_{k-1} \end{vmatrix}}. \tag{8.2-3}$$

If the sequence elements $s_n, \ldots, s_{n+k}$ satisfy eq. (8.2-1), then obviously the sequence transformation $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$ is exact by construction, i.e.,

$$\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n) = s. \tag{8.2-4}$$

As in the case of the Levin transformation it would be desirable to find some alternative representation for the transformation $\mathcal{S}_k^{(n)}(\beta, s_n, \omega_n)$. Fortunately, this can be accomplished as easily as in the case of the Levin transformation. For that purpose eq. (8.2-1) is rewritten in the following way:

$$(\beta+n)_{k-1}[s_n - s]/\omega_n = \sum_{j=0}^{k-1} c_j (\beta+n+j)_{k-j-1}. \tag{8.2-5}$$

The highest power of n, which occurs on the right-hand side of eq. (8.2-5), is n^{k-1}. Hence, if we apply the difference operator Δ^k to eq. (8.2-5), the sum on the right-hand side, which is a polynomial of degree $k-1$ in n, will be annihilated and we may conclude from eqs. (8.2-4) and (8.2-5) that the sequence transformation $\mathcal{S}_k^{(n)}(\beta, s_n, \omega_n)$ is given by the following ratio:

$$\mathcal{S}_k^{(n)}(\beta, s_n, \omega_n) = \frac{\Delta^k\{(\beta+n)_{k-1} s_n/\omega_n\}}{\Delta^k\{(\beta+n)_{k-1}/\omega_n\}}. \tag{8.2-6}$$

With the help of eq. (2.4-8) we obtain a representation of this transformation as the ratio of two finite sums:

$$\mathcal{S}_k^{(n)}(\beta, s_n, \omega_n) = \frac{\sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{(\beta+n+j)_{k-1}}{(\beta+n+k)_{k-1}} \frac{s_{n+j}}{\omega_{n+j}}}{\sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{(\beta+n+j)_{k-1}}{(\beta+n+k)_{k-1}} \frac{1}{\omega_{n+j}}}, \qquad k, n \in \mathbb{N}_0. \tag{8.2-7}$$

The common factor $(\beta+n+k)_{k-1}$ in eq. (8.2-7) was introduced in order to decrease the magnitude of the terms of the numerator and denominator sums, because otherwise overflow may happen too easily for larger values of k.

The transformation (8.2-7) had already been treated by Sidi (see eq. (1.9) of ref. [84]) who used this as well as some other transformations for the derivation of explicit expressions for Padé approximants of some special hypergeometric series. However, it seems that Sidi did not consider the transformation (8.2-7) to be a sequence transformation in its own right. This is certainly an undeserved neglect. It will become clear later that the transformation (8.2-7) is very powerful. We shall see in section 13 that for divergent Stieltjes series, as they for instance occur in the perturbation expansion of the quartic anharmonic oscillator [3,85–88], it is certainly one of the most efficient summation methods which is currently known.

As in the case of the Levin transformation, we also consider the following mild extension of the sequence transformation $\mathcal{S}_k^{(n)}(\beta, s_n, \omega_n)$:

$$\mathcal{S}_{k,l}^{(n)}(\beta, s_n, \omega_n) = \frac{\sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{(\beta+n+l+j)_{k-l-1}}{(\beta+n+k)_{k-1}} \frac{s_{n+j}}{\omega_{n+j}}}{\sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{(\beta+n+l+j)_{k-l-1}}{(\beta+n+k)_{k-1}} \frac{1}{\omega_{n+j}}}, \qquad k, l, n \in \mathbb{N}_0. \tag{8.2-8}$$

For $l = 0$ this transformation reduces to the transformation $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7). An alternative representation for the generalized transformation (8.2-8) can be derived with the help of eq. (2.4-8),

$$\mathscr{S}_{k,l}^{(n)}(\beta, s_n, \omega_n) = \frac{\Delta^k\{(\beta+n+l)_{k-l-1} s_n/\omega_n\}}{\Delta^k\{(\beta+n+l)_{k-l-1}/\omega_n\}}. \tag{8.2-9}$$

From this relationship we may deduce immediately that the transformation (8.2-8) is exact for model sequences of the following type:

$$s_n = s + (\beta+n)_l \omega_n \sum_{j=0}^{k-1} c_j/(\beta+n)_j, \qquad k, l, n \in \mathbb{N}_0. \tag{8.2-10}$$

8.3. Recurrence formulas

Next, it will be shown how the numerator and denominator of the transformation (8.2-7) can be computed recursively. It will turn out that virtually the same technique can be used as in the case of the Levin transformation. In eq. (8.2-7) both numerator and denominator are of the general form

$$Q_k^{(n)}(\beta) = \Delta^k Y_k^{(n)}(\beta), \qquad k, n \in \mathbb{N}_0. \tag{8.3-1}$$

As usual, it is assumed that the difference operator Δ as well as the shift operator E, which is defined in eq. (2.4-4), act only upon n and not upon k. The quantities $Y_k^{(n)}(\beta)$ satisfy the following 2-term recursion in k:

$$Y_k^{(n)}(\beta) = (\beta+n+k-2) Y_{k-1}^{(n)}(\beta), \qquad k \geq 1, \quad n \geq 0. \tag{8.3-2}$$

Combination of eqs. (2.4-4), (7.2-3), (8.3-1) and (8.3-2) yields:

$$Q_k^{(n)}(\beta) = \{kE + (\beta+n+k-2)\Delta\}\Delta^{k-1} Y_{k-1}^{(n)}(\beta) \tag{8.3-3}$$

$$= \{kE + (\beta+n+k-2)\Delta\} Q_{k-1}^{(n)}(\beta) \tag{8.3-4}$$

$$= (\beta+n+2k-2) Q_{k-1}^{(n+1)}(\beta) - (\beta+n+k-2) Q_{k-1}^{(n)}(\beta). \tag{8.3-5}$$

With the help of the 3-term recurrence formula (8.3-5) both the numerator as well as the denominator of $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$ can be computed for $k \geq 1$. However, as in the case of the Levin transformation it is preferable to compute instead the scaled quantities

$$S_k^{(n)}(\beta) = Q_k^{(n)}(\beta)/(\beta+n+k)_{k-1}. \tag{8.3-6}$$

If we insert this into eq. (8.3-5), we obtain the following recurrence formula for the scaled quantities $S_k^{(n)}(\beta)$:

$$S_{k+1}^{(n)}(\beta) = S_k^{(n+1)}(\beta) - \frac{(\beta+n+k)(\beta+n+k-1)}{(\beta+n+2k)(\beta+n+2k-1)} S_k^{(n)}(\beta), \qquad k, n \geq 0. \tag{8.3-7}$$

If we use the starting values

$$S_0^{(n)}(\beta) = s_n/\omega_n, \qquad n \in \mathbb{N}_0, \tag{8.3-8}$$

the 3-term recursion (8.3-7) produces the numerator of the transformation (8.2-7), and if we use the starting values

$$S_0^{(n)}(\beta) = 1/\omega_n, \qquad n \in \mathbb{N}_0, \tag{8.3-9}$$

we obtain the denominator of the transformation (8.2-7).

With the help of the 3-term recursion (8.3-7) it is also possible to compute both the numerator and the denominator of the generalized transformation (8.2-8). If the starting values

$$S_0^{(n)}(\beta) = s_n/[(\beta+n)_l \omega_n], \qquad l, n \in \mathbb{N}_0, \tag{8.3-10}$$

are used, eq. (8.3-7) produces the numerator of the generalized transformation (8.2-8), and the starting values

$$S_0^{(n)}(\beta) = 1/[(\beta+n)_l \omega_n], \qquad l, n \in \mathbb{N}_0, \tag{8.3-11}$$

give the denominator of the transformation (8.2-8).

Since the transforms with the highest values of the subscript normally give the best results, our approximation to the limit s of the sequence $\{s_n\}$ to be transformed will be the same as in the case of the Levin transformation,

$$\{s_0, \omega_0; s_1, \omega_1; \ldots; s_m, \omega_m\} \to \mathscr{S}_m^{(0)}(\beta, s_0, \omega_0), \qquad m \in \mathbb{N}_0. \tag{8.3-12}$$

Essentially this means that we shall use the following sequence of transforms with minimal superscripts and maximal subscripts as approximations to the limit s:

$$\mathscr{S}_0^{(0)}(\beta, s_0, \omega_0), \mathscr{S}_1^{(0)}(\beta, s_0, \omega_0), \ldots, \mathscr{S}_m^{(0)}(\beta, s_0, \omega_0), \ldots. \tag{8.3-13}$$

The recursive computation of the sequence transformation $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$ can be done in virtually the same way as in the case of the Levin transformation. For that purpose it is recommendable to reformulate the 3-term recursion (8.3-7) in the following way:

$$S_j^{(n-j)}(\beta) = S_{j-1}^{(n-j+1)}(\beta) - \frac{(\beta+n-1)(\beta+n-2)}{(\beta+n+j-2)(\beta+n+j-3)} S_{j-1}^{(n-j)}(\beta),$$
$$n \geq 1, \qquad 1 \leq j \leq n. \tag{8.3-14}$$

As in the case of the Levin transformation, only a single 1-dimensional array will be needed for the computation of the $S_j^{(n-j)}(\beta)$ which are either numerator or denominator sums of the transformation (8.2-7). For that purpose the $S_j^{(n-j)}(\beta)$ are stored in a 1-dimensional array s in such a way that the superscript coincides with the index of the corresponding array element:

$$S_\nu^{(n-\nu)}(\beta) \rightarrow s(n-\nu), \qquad n \geq 0, \quad 0 \leq \nu \leq n. \tag{8.3-15}$$

With this convention the recursive scheme (8.3-14) can be reformulated in terms of the elements of the array s:

$$s(n-j) \leftarrow s(n-j+1) - \frac{(\beta+n-1)(\beta+n-2)}{(\beta+n+j-2)(\beta+n+j-3)} s(n-j),$$
$$n \geq 1, \qquad 1 \leq j \leq n. \tag{8.3-16}$$

8.4. Explicit remainder estimates

It still has to be discussed how the auxiliary sequence $\{\omega_n\}$ in eq. (8.2-7) should be chosen. The simplest approach would be to proceed as in the case of the Levin transformation. There, it was argued that the auxiliary sequence $\{\omega_n\}$ should be chosen in such a way that ω_n is proportional to the dominant term of the asymptotic expansion of the remainder r_n,

$$r_n = s_n - s = \omega_n \left[c + O(n^{-1}) \right], \qquad n \rightarrow \infty. \tag{8.4-1}$$

Since the dominant term will not be affected if an asymptotic expansion is transformed into a factorial series or vice versa, it should be possible to use the same simple remainder estimates for sequences of partial sums as in the case of the Levin transformation.

Hence, the remainder estimate (7.3-4) will be used in eq. (8.2-7). This gives an analogue of Levin's u transformation, eq. (7.3-5):

$$y_k^{(n)}(\beta, s_n) = \frac{\displaystyle\sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{(\beta+n+j+1)_{k-2}}{(\beta+n+k)_{k-1}} \frac{s_{n+j}}{a_{n+j}}}{\displaystyle\sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{(\beta+n+j+1)_{k-2}}{(\beta+n+k)_{k-1}} \frac{1}{a_{n+j}}}. \tag{8.4-2}$$

In the same way, the remainder estimate (7.3-6) can be used in eq. (8.2-7). This gives an analogue of Levin's t transformation, eq. (7.3-7):

$$\tau_k^{(n)}(\beta, s_n) = \frac{\displaystyle\sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{(\beta+n+j)_{k-1}}{(\beta+n+k)_{k-1}} \frac{s_{n+j}}{a_{n+j}}}{\displaystyle\sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{(\beta+n+j)_{k-1}}{(\beta+n+k)_{k-1}} \frac{1}{a_{n+j}}}. \tag{8.4-3}$$

The use of the remaider estimate (7.3-8) in eq. (8.2-7) gives an analogue of Levin's d transformation, eq. (7.3-9):

$$\delta_k^{(n)}(\beta, s_n) = \frac{\sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{(\beta+n+j)_{k-1}}{(\beta+n+k)_{k-1}} \frac{s_{n+j}}{a_{n+j+1}}}{\sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{(\beta+n+j)_{k-1}}{(\beta+n+k)_{k-1}} \frac{1}{a_{n+j+1}}}. \tag{8.4-4}$$

Finally, the remainder estimate (7.3-10) gives an analogue of Levin's v transformation, eq. (7.3-11):

$$\varphi_k^{(n)}(\beta, s_n) = \frac{\sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{(\beta+n+j)_{k-1}}{(\beta+n+k)_{k-1}} \frac{a_{n+j}-a_{n+j+1}}{a_{n+j}a_{n+j+1}} s_{n+j}}{\sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{(\beta+n+j)_{k-1}}{(\beta+n+k)_{k-1}} \frac{a_{n+j}-a_{n+j+1}}{a_{n+j}a_{n+j+1}}}. \tag{8.4-5}$$

If one of the remainder estimates (7.3-4), (7.3-6), (7.3-8), and (7.3-10) is used in eq. (8.2-7), $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$ is a *nonlinear* sequence transformation. If, however, remainder estimates $\{\omega_n\}$ are used that do not depend explicitly upon the elements of the sequence $\{s_n\}$, $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$ is a *linear* sequence transformation.

Next, a factorial series analogue of the linear sequence transformation $\Lambda_k^{(n)}(\beta, s_n)$, eq. (7.3-20), will be constructed. A factorial series analogue of the model sequence (7.3-15), for which the sequence transformation $\Lambda_k^{(n)}(\beta, s_n)$ is exact, would be

$$s_n = s + \sum_{j=0}^{k-1} c_j/(n+\alpha)_{j+1}, \qquad k, n \in \mathbb{N}_0. \tag{8.4-6}$$

This model sequence can be obtained from the model sequence (8.2-1), for which the sequence transformation (8.2-7) is exact, by choosing $\beta = \alpha + 1$ and $\omega_n = 1/(n+\alpha)$. Hence, if we define

$$\mathscr{F}_k^{(n)}(\alpha, s_n) = \mathscr{S}_k^{(n)}(\alpha+1, s_n, 1/(n+\alpha)), \qquad k, n \in \mathbb{N}_0, \tag{8.4-7}$$

we see that that the sequence transformation $\mathscr{F}_k^{(n)}(\alpha, s_n)$ is obviously exact for the model sequence (8.4-6). If we now use eq. (8.2-6), we see that the transformation $\mathscr{F}_k^{(n)}(\alpha, s_n)$ satisfies

$$\mathscr{F}_k^{(n)}(\alpha, s_n) = \frac{\Delta^k\{(n+\alpha)_k s_n\}}{\Delta^k(n+\alpha)_k}, \qquad k, n \in \mathbb{N}_0. \tag{8.4-8}$$

The denominator in eq. (8.4-8) can be expressed in closed form. We only have to use

$$\Delta^k(n+\alpha)_k = k!. \tag{8.4-9}$$

This is a special case of the following general relationship which can be proved by complete induction in k,

$$\Delta^k \frac{\Gamma(a+n)}{\Gamma(b+n)} = (-1)^k (b-a)_k \frac{\Gamma(a+n)}{\Gamma(b+n+k)}. \tag{8.4-10}$$

Combining eqs. (8.4-8) and (8.4-9) with eq. (2.4-8) gives us

$$\mathscr{F}_k^{(n)}(\alpha, s_n) = (-1)^k \sum_{j=0}^{k} (-1)^j \frac{(\alpha+n+j)_k}{j!(k-j)!} s_{n+j}, \qquad k, n \in \mathbb{N}_0. \tag{8.4-11}$$

If we combine eq. (8.3-5) with eqs. (8.4-8) and (8.4-9), we obtain the following recursive scheme for the sequence transformation $\mathscr{F}_k^{(n)}(\alpha, s_n)$:

$$\mathscr{F}_0^{(n)}(\alpha, s_n) = s_n, \qquad n \in \mathbb{N}_0, \tag{8.4-12a}$$

$$\mathscr{F}_{k+1}^{(n)}(\alpha, s_n) = \mathscr{F}_k^{(n+1)}(\alpha, s_{n+1}) + \frac{\alpha+n+k}{k+1} \Delta \mathscr{F}_k^{(n)}(\alpha, s_n), \qquad k, n \in \mathbb{N}_0. \tag{8.4-12b}$$

The transformation $\mathscr{F}_k^{(n)}(\alpha, s_n)$ can be computed in essentially the same way as the Richardson extrapolation process. For that purpose it is recommendable to rewrite the above recursive scheme in the following way:

$$\mathscr{F}_0^{(n)}(\alpha, s_n) = s_n, \qquad n \geq 0, \tag{8.4-13a}$$

$$\mathscr{F}_j^{(n-j)}(\alpha, s_{n-j}) = \mathscr{F}_{j-1}^{(n-j+1)}(\alpha, s_{n-j+1}) + \frac{\alpha+n-1}{j} \Delta \mathscr{F}_{j-1}^{(n-j)}(\alpha, s_{n-j}),$$

$$n \geq 1, \quad 1 \leq j \leq n. \tag{8.4-13b}$$

If the $\mathscr{F}_j^{(n)}(\alpha, s_n)$ are stored in a 1-dimensional array f according to the rule

$$\mathscr{F}_j^{(n-j)}(\alpha, s_{n-j}) \to f(n-j), \qquad n \geq 0, \quad 0 \leq j \leq n, \tag{8.4-14}$$

we see that the recursive scheme (8.4-13) can be reformulated in terms of the elements of the array f:

$$f(n) = s_n, \qquad n \geq 0, \tag{8.4-15a}$$

$$f(n-j) = f(n-j+1) + \frac{\alpha+n-1}{j} \Delta f(n-j), \qquad n \geq 1, \quad 1 \leq j \leq n. \tag{8.4-15b}$$

A discussion of the linear but nonregular sequence transformation $\mathscr{F}_k^{(n)}(\alpha, s_n)$ can be found in Wimp's book (see pp. 38–40 of ref. [23]). However, the recursive scheme (8.4-12) for the computation of this transformation seems to be new.

The sequence transformations $y_k^{(n)}(\beta, s_n)$, eq. (8.4-2), and $\tau_k^{(n)}(\beta, s_n)$, eq. (8.4-3), require the sequence elements $s_{n-1}, s_n, s_{n+1}, \ldots, s_{n+k}$ for their computation, whereas $\delta_k^{(n)}(\beta, s_n)$, eq. (8.4-4), requires the sequence elements $s_n, s_{n+1}, \ldots, s_{n+k+1}$. Hence, they are all transformations of order $k+1$. The sequence transformation $\varphi_k^{(n)}(\beta, s_n)$, eq. (8.4-5), requires the sequence elements s_{n-1}, $s_n, s_{n+1}, \ldots, s_{n+k+1}$ which implies that it is a transformation of order $k+2$. The linear sequence transformation $\mathscr{F}_k^{(n)}(\alpha, s_n)$, eq. (8.4-11), requires the sequence elements $s_n, s_{n+1}, \ldots, s_{n+k}$, i.e., it is a transformation of order k.

The situation is somewhat different if the transforms with superscript $n = 0$ are computed because then $y_k^{(0)}(\beta, s_0)$ and $\tau_k^{(0)}(\beta, s_0)$ are transformations of order k, whereas $\delta_k^{(0)}(\beta, s_0)$ and $\varphi_k^{(0)}(\beta, s_0)$ are transformations of order $k+1$.

9. Other generalizations of Levin's sequence transformation

9.1. Asymptotic approximations based upon Pochhammer symbols

In the last section it was demonstrated how a new class of sequence transformations can be derived in exactly the same way as the Levin transformation which is generally accepted to be a very powerful convergence acceleration and summation method [29,30,57,60–64].

The only difference between the Levin transformation and the new transformation discussed in the last section is that the Levin transformation assumes that the ratio r_n/ω_n can be expressed as an asymptotic power series whereas the new transformation assumes that r_n/ω_n can be expressed as a factorial series. Consequently, the analytical expressions for the various Levin transformations and those for the analogous variants of the new transformation can easily be transformed into each other. For instance, one only has to replace the powers $(\beta + n + j)^{k-1}$ in the expression for the general Levin transformation $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), by Pochhammer symbols $(\beta + n + j)_{k-1}$ in order to obtain the analogous new transformation $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7).

However, these new transformations, which were discussed in the last section, do not yet exhaust all possibilities of constructing other simple generalizations of the Levin transformation, which nevertheless retain most of the advantages of the Levin transformation. For instance, in recent articles on large order perturbation theory asymptotic approximations of the following general type were considered [89–91]:

$$f(z) \sim \frac{c_0}{z} + \frac{c_1}{z(z-1)} + \frac{c_2}{z(z-1)(z-2)} + \cdots + \frac{c_n}{z(z-1)(z-2)\ldots(z-n)}$$

$$= \sum_{\nu=0}^{n} (-1)^{\nu+1} \frac{c_\nu}{(-z)_{\nu+1}}, \qquad |z| \to \infty. \tag{9.1-1}$$

Superficially, such an expression looks very much like a truncated factorial series since the argument z occurs also in Pochhammer symbols. However, the fact that the Pochhammer symbols in eq. (9.1-1) are of the type

$$(-z)_{\nu+1} = (-z)(-z+1)\dots(-z+\nu)$$
$$= (-1)^{\nu+1} z(z-1)\dots(z-\nu), \qquad \nu \in \mathbb{N}_0, \tag{9.1-2}$$

has some far-reaching consequences. For instance, if z is a positive real number, eq. (9.1-1) makes sense only if $z > n$ holds. For $n > z$, either the later terms in the sum will show irregular sign patterns, or, if z happens to be a positive integer, some Pochhammer symbols will then be zero. Consequently, for a fixed value of z the summation limit n in eq. (9.1-1) cannot be extended to infinity, and such an expression cannot be considered to be the truncation of an asymtotic series after a finite number of terms. Instead, an expression such as eq. (9.1-1) has to be interpreted to be some kind of asymptotic approximation involving only a finite number of terms.

The author is not aware of any reference in the mathematical literature, in which expressions like the one in eq. (9.1-1) are treated and their properties are analyzed. Consequently, the material in this section is somewhat experimental and its mathematical basis is not as solid as in the other sections.

However, it must be emphasized that these objections do not exclude the possibility that finite sums of the type of eq. (9.1-1) may yield excellent approximations if suitable restrictions on z and n are made. We shall see later in section 13 that sequence transformations which are based upon asymptotic approximations of the type of eq. (9.1-1) are indeed able to produce excellent results in convergence acceleration and summation processes.

9.2. *New sequence transformations based upon Pochhammer symbols*

The following model sequence will be the basis for the new class of sequence transformations which will be derived in this section:

$$s_n = s + \omega_n \sum_{j=0}^{k-1} c_j/(-\gamma-n)_j, \qquad k,\, n \in \mathbb{N}_0. \tag{9.2-1}$$

This sequence is formally almost identical with the model sequence (8.2-1). The only difference is that the Pochhammer symbols $(n+\beta)_j$ in eq. (8.2-1) are replaced by Pochhammer symbols $(-\gamma-n)_j$. Concerning the sequence $\{\omega_n\}$ of remainder estimates it is again assumed that the ω_n are known functions of n which have to be different from zero and distinct for all finite values of n. But otherwise, the ω_n are in principle completely arbitrary.

The parameter γ in eq. (9.2-1) is not only subject to the restriction that the Pochhammer symbols $(-\gamma-n)_j$ must not be zero for all admissible values of n and j. Also, the regular sign pattern of the Pochhammer symbols in eq. (9.2-1) must not be destroyed. These two restrictions suggest that γ should be a positive number satisfying $\gamma \geq k-1$.

In eq. (9.2-1) there occur $k+1$ unknown quantities, the limit or antilimit s and the k linear coefficients $c_0, \dots, c_{k-1}$. Hence, if $k+1$ sequence elements $s_n, \dots, s_{n+k}$ are known, the sequence transformation $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$ can be defined according to Cramer's rule by the following ratio of determinants:

$$\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n) = \frac{\begin{vmatrix} s_n & \dots & s_{n+k} \\ \omega_n & \dots & \omega_{n+k} \\ \vdots & \ddots & \vdots \\ \omega_n/(-\gamma-n)_{k-1} & \dots & \omega_{n+k}/(-\gamma-n-k)_{k-1} \end{vmatrix}}{\begin{vmatrix} 1 & \dots & 1 \\ \omega_n & \dots & \omega_{n+k} \\ \vdots & \ddots & \vdots \\ \omega_n/(-\gamma-n)_{k-1} & \dots & \omega_{n+k}/(-\gamma-n-k)_{k-1} \end{vmatrix}}. \tag{9.2-2}$$

If the sequence elements $s_n, \ldots, s_{n+k}$ satisfy eq. (9.2-1), then obviously the sequence transformation $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$ is exact by construction, i.e.,

$$\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n) = s. \tag{9.2-3}$$

Again it would be desirable to have some alternative representation for the transformation $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$. Fortunately, this can be accomplished quite easily. For that purpose eq. (9.2-1) is rewritten in the following way:

$$(-\gamma-n)_{k-1}[s_n - s]/\omega_n = \sum_{j=0}^{k-1} c_j(-\gamma-n+j)_{k-j-1}. \tag{9.2-4}$$

The highest power of n, which occurs on the right-hand side of eq. (9.2-4), is n^{k-1}. Hence, if we apply the difference operator Δ^k to eq. (9.2-4), the sum on the right-hand side, which is a polynomial of degree $k-1$ in n, will be annihilated and we may conclude from eqs. (9.2-3) and (9.2-4) that the sequence transformation $\mathscr{M}_k^{(n)}(\beta, s_n, \omega_n)$ is given by the following ratio:

$$\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n) = \frac{\Delta^k\{(-\gamma-n)_{k-1}s_n/\omega_n\}}{\Delta^k\{(-\gamma-n)_{k-1}/\omega_n\}}. \tag{9.2-5}$$

If we use eq. (2.4-8), we see that this transformation can be represented as the ratio of two finite sums:

$$\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n) = \frac{\displaystyle\sum_{j=0}^{k}(-1)^j\binom{k}{j}\frac{(-\gamma-n-j)_{k-1}}{(-\gamma-n-k)_{k-1}}\frac{s_{n+j}}{\omega_{n+j}}}{\displaystyle\sum_{j=0}^{k}(-1)^j\binom{k}{j}\frac{(-\gamma-n-j)_{k-1}}{(-\gamma-n-k)_{k-1}}\frac{1}{\omega_{n+j}}}, \qquad k, n \in \mathbb{N}_0. \tag{9.2-6}$$

The common factor $(-\gamma-n-k)_{k-1}$ in eq. (9.2-6) was introduced in order to decrease the magnitude of the terms of the numerator and denominator sums, because otherwise overflow may happen too easily for larger values of k.

As in the previous sections, we also consider a mild extension of the sequence transformation $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$:

$$\mathscr{M}_{k,l}^{(n)}(\gamma, s_n, \omega_n) = \frac{\sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{(-\gamma-n-j+l)_{k-l-1}}{(-\gamma-n-k)_{k-1}} \frac{s_{n+j}}{\omega_{n+j}}}{\sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{(-\gamma-n-j+l)_{k-l-1}}{(-\gamma-n-k)_{k-1}} \frac{1}{\omega_{n+j}}}, \qquad k, l, n \in \mathbb{N}_0. \tag{9.2-7}$$

For $l=0$ this transformation reduces to the transformation $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-6). An alternative representation for the generalized transformation (9.2-7) can be derived with the help of eq. (2.4-8),

$$\mathscr{M}_{k,l}^{(n)}(\gamma, s_n, \omega_n) = \frac{\Delta^k\{(-\gamma-n+l)_{k-l-1} s_n/\omega_n\}}{\Delta^k\{(-\gamma-n+l)_{k-l-1}/\omega_n\}}. \tag{9.2-8}$$

From this relationship we may deduce immediately that the transformation (9.2-7) is exact for model sequences of the following type:

$$s_n = s + (-\gamma-n)_l \omega_n \sum_{j=0}^{k-1} c_j/(-\gamma-n)_j, \qquad k, l, n \in \mathbb{N}_0. \tag{9.2-9}$$

9.3. Recurrence formulas

Next, it will be shown how the numerator and denominator of the transformation (9.2-6) can be computed recursively. It will turn out that virtually the same technique can be used as in sections 7.2 and 8.3. In eq. (9.2-6) both numerator and denominator are of the general form

$$R_k^{(n)}(\gamma) = \Delta^k Z_k^{(n)}(\gamma), \qquad k, n \in \mathbb{N}_0. \tag{9.3-1}$$

As usual, it is assumed that the difference operator Δ as well as the shift operator E, which is defined in eq. (2.4-4), act only upon n and not upon k. The quantities $Z_k^{(n)}(\gamma)$ satisfy the following 2-term recursion in k:

$$Z_k^{(n)}(\gamma) = (-\gamma-n+k-2) Z_{k-1}^{(n)}(\gamma), \qquad k \geq 1, \quad n \geq 0. \tag{9.3-2}$$

Combination of eqs. (2.4-4), (7.2-3), (9.3-1) and (9.3-2) yields:

$$R_k^{(n)}(\gamma) = \{(-\gamma-n+k-2)\Delta - kE\} \Delta^{k-1} Z_{k-1}^{(n)}(\gamma) \tag{9.3-3}$$

$$= \{(-\gamma-n+k-2)\Delta - kE\} R_{k-1}^{(n)}(\gamma) \tag{9.3-4}$$

$$= (\gamma+n-k+2) R_{k-1}^{(n)}(\gamma) - (\gamma+n+2) R_{k-1}^{(n+1)}(\gamma). \tag{9.3-5}$$

With the help of the 3-term recurrence formula (9.3-5) both the numerator as well as the denominator of the transformation $\mathcal{M}_k^{(n)}(\gamma, s_n, \omega_n)$ can be computed for $k \geq 1$. However, it is again preferable to compute instead the scaled quantities

$$M_k^{(n)}(\gamma) = R_k^{(n)}(\gamma)/(-\gamma - n - k)_{k-1}. \tag{9.3-6}$$

If we insert this into eq. (9.3-5), we obtain the following recurrence formula for the scaled quantities $M_k^{(n)}(\gamma)$:

$$M_{k+1}^{(n)}(\gamma) = M_k^{(n+1)}(\gamma) - \frac{\gamma + n - k + 1}{\gamma + n + k + 1} M_k^{(n)}(\gamma), \qquad k, n \geq 0. \tag{9.3-7}$$

If we use the starting values

$$M_0^{(n)}(\gamma) = s_n/\omega_n, \qquad n \in \mathbb{N}_0, \tag{9.3-8}$$

the 3-term recursion (9.3-7) produces the numerator of the transformation (9.2-6), and if we use the starting values

$$M_0^{(n)}(\gamma) = 1/\omega_n, \qquad n \in \mathbb{N}_0, \tag{9.3-9}$$

we obtain the denominator of the transformation (9.2-7).

With the help of the 3-term recursion (9.3-7) it is also possible to compute both the numerator and the denominator of the generalized transformation (9.2-7). If the starting values

$$M_0^{(n)}(\gamma) = s_n/[(-\gamma - n)_l \omega_n], \qquad l, n \in \mathbb{N}_0, \tag{9.3-10}$$

are used, the 3-term recursion (9.3-7) produces the numerator of the generalized transformation (9.2-7), and the starting values

$$M_0^{(n)}(\gamma) = 1/[(-\gamma - n)_l \omega_n], \qquad l, n \in \mathbb{N}_0, \tag{9.3-11}$$

produce the denominator of the transformation (9.2-7).

Since the transforms with the highest values of the subscript normally give the best results, our approximation to the limit s of the sequence $\{s_n\}$ to be transformed will be the same as in the previous two sections, i.e.,

$$\{s_0, \omega_0; s_1, \omega_1; \ldots; s_m, \omega_m\} \rightarrow \mathcal{M}_m^{(0)}(\gamma, s_0, \omega_0), \qquad m \in \mathbb{N}_0. \tag{9.3-12}$$

Essentially this means that we shall use the following sequence of transforms with minimal superscripts and maximal subscripts as approximations to the limit s:

$$\mathcal{M}_0^{(0)}(\gamma, s_0, \omega_0), \mathcal{M}_1^{(0)}(\gamma, s_0, \omega_0), \ldots, \mathcal{M}_m^{(0)}(\gamma, s_0, \omega_0), \ldots. \tag{9.3-13}$$

The recursive computation of the sequence transformations $\mathcal{M}_k^{(n)}(\gamma, s_n, \omega_n)$ can be done in virtually the same way as in the previous two sections. For that purpose it is recommendable to reformulate the 3-term recursion (9.3-7) in the following way:

$$M_j^{(n-j)}(\gamma) = M_{j-1}^{(n-j+1)}(\gamma) - \frac{\gamma+n-2j+2}{\gamma+n} M_{j-1}^{(n-j)}(\gamma), \qquad n \geq 1, \quad 1 \leq j \leq n. \tag{9.3-14}$$

Again, only a single 1-dimensional array will be needed for the computation of the $M_j^{(n-j)}(\gamma)$ which are either numerator or denominator sums of the transformation (9.2-6). For that purpose the $M_j^{(n-j)}(\gamma)$ are stored in a 1-dimensional array m in such a way that the superscript coincides with the index of the corresponding array element:

$$M_\nu^{(n-\nu)}(\gamma) \to m(n-\nu), \qquad n \geq 0, \quad 0 \leq \nu \leq n. \tag{9.3-15}$$

With this convention the recursive scheme (9.3-14) can be reformulated in terms of the elements of the array m:

$$m(n-j) \leftarrow m(n-j+1) - \frac{\gamma+n-2j+2}{\gamma+n} m(n-j), \qquad n \geq 1, \quad 1 \leq j \leq n. \tag{9.3-16}$$

9.4. *Explicit remainder estimates*

It still must be discussed how the auxiliary sequence $\{\omega_n\}$ in eq. (9.2-6) should be chosen. The simplest approach would again consist of using the same simple remainder estimates for sequences of partial sums as in the case of the Levin transformation.

Hence, if we use in eq. (9.2-6) the remainder estimate

$$\omega_n = (-\gamma-n)a_n, \tag{9.4-1}$$

we obtain an analogue of Levin's u transformation, eq. (7.3-5):

$$Y_k^{(n)}(\gamma, s_n) = \frac{\sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{(-\gamma-n-j+1)_{k-2}}{(-\gamma-n-k)_{k-1}} \frac{s_{n+j}}{a_{n+j}}}{\sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{(-\gamma-n-j+1)_{k-2}}{(-\gamma-n-k)_{k-1}} \frac{1}{a_{n+j}}}. \tag{9.4-2}$$

In the same way, the remainder estimate (7.3-6) can be used in eq. (9.2-6). This gives an analogue of Levin's t transformation, eq. (7.3-7):

$$T_k^{(n)}(\gamma, s_n) = \frac{\sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{(-\gamma-n-j)_{k-1}}{(-\gamma-n-k)_{k-1}} \frac{s_{n+j}}{a_{n+j}}}{\sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{(-\gamma-n-j)_{k-1}}{(-\gamma-n-k)_{k-1}} \frac{1}{a_{n+j}}}. \tag{9.4-3}$$

The use of the remainder estimate (7.3-8) in eq. (9.2-6) gives an analogue of Levin's d transformation, eq. (7.3-9):

$$\Delta_k^{(n)}(\gamma, s_n) = \frac{\sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{(-\gamma-n-j)_{k-1}}{(-\gamma-n-k)_{k-1}} \frac{s_{n+j}}{a_{n+j+1}}}{\sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{(-\gamma-n-j)_{k-1}}{(-\gamma-n-k)_{k-1}} \frac{1}{a_{n+j+1}}}. \tag{9.4-4}$$

Finally, the remainder estimate (7.3-10) gives an analogue of Levin's v transformation, eq. (7.3-11):

$$\Phi_k^{(n)}(\gamma, s_n) = \frac{\sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{(-\gamma-n-j)_{k-1}}{(-\gamma-n-k)_{k-1}} \frac{a_{n+j}-a_{n+j+1}}{a_{n+j}a_{n+j+1}} s_{n+j}}{\sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{(-\gamma-n-j)_{k-1}}{(-\gamma-n-k)_{k-1}} \frac{a_{n+j}-a_{n+j+1}}{a_{n+j}a_{n+j+1}}}. \tag{9.4-5}$$

If one of the remainder estimates (7.3-6), (7.3-8), (7.3-10), and (9.4-1) is used in eq. (9.2-6), $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$ is a *nonlinear* sequence transformation. If, however, remainder estimates $\{\omega_n\}$ are used that do not depend explicitly upon the elements of the sequence $\{s_n\}$, $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$ is a *linear* sequence transformation.

Next, an analogue of the Richardson extrapolation scheme (6.1-5) will be introduced which is based upon asymptotic approximations involving Pochhammer symbols of the type of eq. (9.1-1). For that purpose we consider the following model sequence:

$$s_n = s + \sum_{j=0}^{k-1} c_j/(-\zeta-n)_{j+1}, \qquad k, n \in \mathbb{N}_0. \tag{9.4-6}$$

This model sequence can be obtained from the model sequence (9.2-1), for which the sequence transformation (9.2-6) is exact, by choosing $\zeta = \gamma + 1$ and $\omega_n = -1/(\zeta + n)$. Hence, if we define

$$\mathscr{P}_k^{(n)}(\zeta, s_n) = \mathscr{M}_k^{(n)}(\zeta - 1, s_n, -1/(n+\zeta)), \qquad k, n \in \mathbb{N}_0, \tag{9.4-7}$$

we see that that the sequence transformation $\mathscr{P}_k^{(n)}(\zeta, s_n)$ is obviously exact for the model sequence (9.4-6). If we now use eq. (9.2-5), we see that the transformation $\mathscr{P}_k^{(n)}(\zeta, s_n)$ satisfies

$$\mathscr{P}_k^{(n)}(\zeta, s_n) = \frac{\Delta^k\{(-\zeta-n)_k s_n\}}{\Delta^k(-\zeta-n)_k}, \qquad k, n \in \mathbb{N}_0. \tag{9.4-8}$$

The denominator in eq. (9.4-8) can be expressed in closed form. We only have to use

$$\Delta^k(-\zeta-n)_k = (-1)^k k!. \tag{9.4-9}$$

This is a special case of the following general relationship which can be proved by complete induction in k,

$$\Delta^k \frac{\Gamma(a-n)}{\Gamma(b-n)} = (b-a)_k \frac{\Gamma(a-n-k)}{\Gamma(b-n)}. \tag{9.4-10}$$

Combination of eqs. (9.4-8) and (9.4-9) with eq. (2.4-8) yields

$$\mathscr{P}_k^{(n)}(\zeta, s_n) = \sum_{j=0}^{k} (-1)^j \frac{(-\zeta-n-j)_k}{j!(k-j)!} s_{n+j}, \qquad k, n \in \mathbb{N}_0. \tag{9.4-11}$$

If we combine eq. (9.3-5) with eqs. (9.4-8) and (9.4-9), we obtain the following recursive scheme for the sequence transformation $\mathscr{P}_k^{(n)}(\zeta, s_n)$:

$$\mathscr{P}_0^{(n)}(\zeta, s_n) = s_n, \qquad n \in \mathbb{N}_0, \tag{9.4-12a}$$

$$\mathscr{P}_{k+1}^{(n)}(\zeta, s_n) = \mathscr{P}_k^{(n+1)}(\zeta, s_{n+1}) + \frac{\zeta+n-k}{k+1} \Delta \mathscr{P}_k^{(n)}(\zeta, s_n), \qquad k, n \in \mathbb{N}_0. \tag{9.4-12b}$$

The transformation $\mathscr{P}_k^{(n)}(\zeta, s_n)$ can be computed in essentially the same way as the Richardson extrapolation process. For that purpose it is recommendable to rewrite the above recursive scheme in the following way:

$$\mathscr{P}_0^{(n)}(\zeta, s_n) = s_n, \qquad n \geq 0, \tag{9.4-13a}$$

$$\mathscr{P}_j^{(n-j)}(\zeta, s_{n-j}) = \mathscr{P}_{j-1}^{(n-j+1)}(\zeta, s_{n-j+1}) + \frac{\zeta+n-2j+1}{j} \Delta \mathscr{P}_{j-1}^{(n-j)}(\zeta, s_{n-j}),$$

$$n \geq 1, \qquad 1 \leq j \leq n. \tag{9.4-13b}$$

If the $\mathscr{P}_j^{(n)}(\zeta, s_n)$ are stored in a 1-dimensional array p according to the rule

$$\mathscr{P}_j^{(n-j)}(\zeta, s_{n-j}) \to p(n-j), \qquad n \geq 0, \quad 0 \leq j \leq n, \tag{9.4-14}$$

we see that the recursive scheme (9.4-14) can be reformulated in terms of the elements of the array p:

$$p(n) \leftarrow s_n, \qquad n \geq 0, \tag{9.4-15a}$$

$$p(n-j) \leftarrow p(n-j+1) + \frac{\zeta+n-2j+1}{j} \Delta p(n-j), \qquad n \geq 1, \quad 1 \leq j \leq n. \tag{9.4-15b}$$

The sequence transformations $Y_k^{(n)}(\gamma, s_n)$, eq. (9.4-2), and $T_k^{(n)}(\gamma, s_n)$, eq. (9.4-3), require the sequence elements $s_{n-1}, s_n, s_{n+1}, \ldots, s_{n+k}$ for their computation, whereas $\Delta_k^{(n)}(\gamma, s_n)$, eq. (9.4-4), requires the sequence elements $s_n, s_{n+1}, \ldots, s_{n+k+1}$. Hence, they are all transformations of order $k+1$. The sequence transformation $\Phi_k^{(n)}(\gamma, s_n)$, eq. (9.4-5), requires the sequence elements s_{n-1},

$s_n, s_{n+1}, \ldots, s_{n+k+1}$ which implies that it is a transformation of order $k+2$. The linear sequence transformation $\mathscr{P}_k^{(n)}(\zeta, s_n)$, eq. (9.4-11), requires the sequence elements $s_n, s_{n+1}, \ldots, s_{n+k}$, i.e., it is a transformation of order k.

The situation is somewhat different if the transforms with superscript $n=0$ are computed because then $Y_k^{(0)}(\gamma, s_0)$ and $T_k^{(0)}(\gamma, s_0)$ are transformations of order k, whereas $\Delta_k^{(0)}(\gamma, s_0)$ and $\Phi_k^{(0)}(\gamma, s_0)$ are transformations of order $k+1$.

9.5. Drummond's sequence transformation

Let us assume that $P_{k-1}(n)$ is a polynomial of degree $k-1$ in n. We want to derive a sequence transformation $\mathscr{D}_k^{(n)}(s_n, \omega_n)$, which is by construction exact for the following model sequence:

$$s_n = s + \omega_n P_{k-1}(n), \qquad k, n \in \mathbb{N}_0. \tag{9.5-1}$$

Concerning the sequence $\{\omega_n\}$ of remainder estimates it is again assumed that the ω_n are different from zero for all finite values of n, but otherwise they are in principle completely arbitrary. For the derivation of a sequence transformation, which is exact for the above model sequence, we rewrite eq. (9.5-1) in the following way:

$$[s_n - s]/\omega_n = P_{k-1}(n), \qquad k, n \in \mathbb{N}_0. \tag{9.5-2}$$

Since $P_{k-1}(n)$ is by assumption a polynomial of degree $k-1$ in n, the right-hand side of eq. (9.5-2) will be annihilated by the difference operator Δ^k. Hence, we can define the sequence transformation $\mathscr{D}_k^{(n)}(s_n, \omega_n)$ by the following ratio:

$$\mathscr{D}_k^{(n)}(s_n, \omega_n) = \frac{\Delta^k\{s_n/\omega_n\}}{\Delta^k\{1/\omega_n\}}. \tag{9.5-3}$$

With the help of eq. (2.4-8) we obtain a representation of this transformation as the ratio of two finite sums:

$$\mathscr{D}_k^{(n)}(s_n, \omega_n) = \frac{\sum\limits_{j=0}^{k} (-1)^j \binom{k}{j} \dfrac{s_{n+j}}{\omega_{n+j}}}{\sum\limits_{j=0}^{k} (-1)^j \binom{k}{j} \dfrac{1}{\omega_{n+j}}}, \qquad k, n \in \mathbb{N}_0. \tag{9.5-4}$$

The special case $\omega_n = a_n$ of this sequence transformation was originally derived by Drummond [92]. Later, it was rederived by Sidi [84].

Both numerator and denominator of the sequence transformation (9.5-4) can be computed with the help of the following 3-term recurrence formula, which is an immediate consequence of eq. (9.5-3):

$$D_{k+1}^{(n)} = D_k^{(n+1)} - D_k^{(n)}, \qquad k, n \geq 0. \tag{9.5-5}$$

If we use the starting values

$$D_0^{(n)} = s_n/\omega_n, \qquad n \in \mathbb{N}_0, \tag{9.5-6}$$

the recurrence formula (9.5-5) produces the numerator of Drummond's transformation (9.5-4), and if we use the starting values

$$D_0^{(n)} = 1/\omega_n, \qquad n \in \mathbb{N}_0, \tag{9.5-7}$$

we obtain the denominator of Drummond's transformation (9.5-4).

Essentially the same computational algorithm, which was used for the other sequence transformations of sections 7, 8, and 9 can also be used in the case of Drummond's sequence transformation $\mathscr{D}_k^{(n)}(s_n, \omega_n)$. Consequently, our approximation to the limit s of the sequence $\{s_n\}$ to be transformed will be

$$\{s_0, \omega_0; s_1, \omega_1; \ldots; s_m, \omega_m\} \to \mathscr{D}_m^{(0)}(s_0, \omega_0), \qquad m \in \mathbb{N}_0. \tag{9.5-8}$$

This means that we shall again use a sequence of transforms with minimal superscripts and maximal subscripts as approximations to the limit s:

$$\mathscr{D}_0^{(0)}(s_0, \omega_0), \mathscr{D}_1^{(0)}(s_0, \omega_0), \ldots, \mathscr{D}_m^{(0)}(s_0, \omega_0), \ldots . \tag{9.5-9}$$

Since the recursive computation of the sequence transformation $\mathscr{D}_k^{(n)}(s_n, \omega_n)$ can be done in virtually the same way as in the case of the Levin transformation, it is recommendable to reformulate the recursive scheme (9.5-5) in the following way:

$$D_j^{(n-j)} = D_{j-1}^{(n-j+1)} - D_{j-1}^{(n-j)}, \qquad n \geq 1, \quad 1 \leq j \leq n. \tag{9.5-10}$$

Again, a single 1-dimensional array will be sufficient for the computation of the quantities $D_j^{(n-j)}$, which are either numerator or denominator sums of the sequence transformation (9.5-4), if the $D_j^{(n-j)}$ are stored in a 1-dimensional array d according to the following rule:

$$D_\nu^{(n-\nu)} \to d(n-\nu), \qquad n \geq 0, \quad 0 \leq \nu \leq n. \tag{9.5-11}$$

With this convention the recurrence formula (9.5-10) can be reformulated in terms of the elements of the array d:

$$d(n-j) \leftarrow d(n-j+1) - d(n-j), \qquad n \geq 1, \quad 1 \leq j \leq n. \tag{9.5-12}$$

If we compare Drummond's sequence transformation $\mathscr{D}_k^{(n)}(s_n, \omega_n)$, eq. (9.5-4), with the analogous sequence transformations $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), and $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-6), we see that the numerator and denominator sums of these transformations contain additional n-dependent coefficients such as $(\beta+n+j)^{k-1}$, $(\beta+n+j)_{k-1}$, or $(-\gamma-n-j)_{k-1}$, which are all of order $O(n^{k-1})$.

Similarly, in the difference operator representation (9.5-3) for Drummond's sequence transformation Δ^k acts only upon s_n/ω_n and $1/\omega_n$ whereas in the analogous difference operator representations (7.1-6), (8.2-5), and (9.2-5) for the other transformations mentioned above weighted differences of s_n/ω_n and $1/\omega_n$ are formed.

This implies that in the case of Drummond's sequence transformation, eq. (9.5-4), the sequence elements $s_n, \ldots, s_{n+k}$ and the remainder estimates $\omega_n, \ldots, \omega_{n+k}$, which are needed for the computation of $\mathscr{D}_k^{(n)}(s_n, \omega_n)$, all contribute equally in the numerator and denominator sums.

This is not true in the case of the other sequence transformations mentioned above since they contain additional weights. Essentially, this means that the information contained in the sequence elements and remainder estimates with larger indices will be emphasized more strongly in the computation of these sequence transformations.

Since the later elements of a convergent sequence $\{\!\{s_n\}\!\}$ are usually closer to the limit s than the earlier elements, it seems plausible to expect that Drummond's sequence transformation, eq. (9.5-4), which does not give special weight to the sequence elements and remainder estimates with higher indices, should normally be somewhat less powerful than the transformations $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, and $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$. We shall see later that this assumption is indeed normally true.

10. Brezinski's theta algorithm and related topics

10.1. The derivation of Brezinski's theta algorithm

It is well known that Wynn's ϵ algorithm accelerates linear convergence quite efficiently and that it is also able to sum even wildly divergent series. However, the ϵ algorithm is not able to accelerate logarithmic convergence. In the same way, Wynn's ρ algorithm is certainly one of the better accelerators for logarithmic convergence but fails to accelerate linear convergence and to sum divergent series. It would certainly be desirable to modify either the ϵ or the ρ algorithm in such a way that the advantageous features of the ϵ and the ρ algorithm could be combined. For that purpose, let us consider a recursive scheme of the following type:

$$T_{-1}^{(n)} = 0, \qquad T_0^{(n)} = s_n, \tag{10.1-1a}$$

$$T_{k+1}^{(n)} = T_{k-1}^{(n+1)} + w_k D_k^{(n)}, \qquad k, n \in \mathbb{N}_0. \tag{10.1-1b}$$

$D_k^{(n)}$ is a quantity which depends upon one or several other elements $T_\kappa^{(\nu)}$ of the table of this transformation. It is assumed that the functional form of $D_k^{(n)}$ is known. The quantity w_k is for the moment unspecified. Later, we will try to derive an expression for w_k which will guarantee that the above recursive scheme will lead to an acceleration of convergence.

The recursive scheme (10.1-1) contains the ϵ and the ρ algorithm as special cases. If we choose $w_k = 1$ together with

$$D_k^{(n)} = \frac{1}{T_k^{(n+1)} - T_k^{(n)}}, \qquad k, n \in \mathbb{N}_0, \tag{10.1-2}$$

the recursive scheme (10.1-1) corresponds to Wynn's ϵ algorithm, eq. (4.2-1), and if we choose $w_k = 1$ together with

$$D_k^{(n)} = \frac{x_{n+k+1} - x_n}{T_k^{(n+1)} - T_k^{(n)}}, \qquad k, n \in \mathbb{N}_0, \tag{10.1-3}$$

the recursive scheme (10.1-1) corresponds to Wynn's ρ algorithm, eq. (6.2-2).

We now want to analyze how the quantity w_k has to be chosen in order to guarantee that the sequence transformation $T_k^{(n)}$ will lead to an acceleration of convergence.

In Wynn's ϵ or ρ algorithm, only the transforms with even subscripts are used as approximations to the limit. The transforms with odd subscripts are only auxiliary quantities which diverge if the whole process converges. Since either the ϵ or the ρ algorithm will be our starting point for the construction of a new sequence transformation, we assume that $T_k^{(n)}$ behaves in the same way. This means that only the transforms with even subscripts will be used as approximations to the limit whereas the transforms with odd subscripts are only auxiliary quantities which diverge if the transforms with even subscripts converge.

Brezinski [26] argued that the exact numerical values of the transforms with odd subscripts do not really matter as long as they diverge if the whole process converges. Consequently, the most convenient choice for w_{2k} in eq. (10.1-1b) would be to proceed as in Wynn's ϵ or ρ algorithm, i.e.,

$$w_{2k} = 1, \qquad k \in \mathbb{N}_0. \tag{10.1-4}$$

The parameters w_{2k+1} can be determined by requiring that for fixed $k \in \mathbb{N}_0$ the sequence $\{\!\{T_{2k+2}^{(n)}\}\!\}$ should converge more rapidly than the sequence $\{\!\{T_{2k}^{(n+1)}\}\!\}$ in the following sense:

$$\lim_{n \to \infty} \frac{\Delta T_{2k+2}^{(n)}}{\Delta T_{2k}^{(n+1)}} = 0, \qquad k \in \mathbb{N}_0. \tag{10.1-5}$$

If we form in eq. (10.1-1b) the first difference with respect to n, we see that condition (10.1-5) is automatically fulfilled if we choose

$$w_{2k+1} = - \lim_{n \to \infty} \frac{\Delta T_{2k}^{(n+1)}}{\Delta D_{2k+1}^{(n)}}, \qquad k \in \mathbb{N}_0. \tag{10.1-6}$$

Unfortunately, in situations of practical interest it will normally not be possible to compute this limit $n \to \infty$. As a manageable alternative, Brezinski [26] suggested to use instead:

$$w_{2k+1}^{(n)} = - \frac{\Delta T_{2k}^{(n+1)}}{\Delta D_{2k+1}^{(n)}}, \qquad k, n \in \mathbb{N}_0. \tag{10.1-7}$$

This choice together with eq. (10.1-4) leads to the following recursive scheme for the sequence transformation $T_k^{(n)}$:

$$T_{-1}^{(n)} = 0, \qquad T_0^{(n)} = s_n, \tag{10.1-8a}$$

$$T_{2k+1}^{(n)} = T_{2k-1}^{(n+1)} + D_{2k}^{(n)}, \tag{10.1-8b}$$

$$T_{2k+2}^{(n)} = T_{2k}^{(n+1)} - \frac{\Delta T_{2k}^{(n+1)}}{\Delta D_{2k+1}^{(n)}} D_{2k+1}^{(n)}, \quad k, n \in \mathbb{N}_0. \tag{10.1-8c}$$

If we choose in this recursive scheme $D_k^{(n)}$ according to eq. (10.1-2) – which corresponds to Wynn's ϵ algorithm – we obtain Brezinski's ϑ algorithm [26]:

$$\vartheta_{-1}^{(n)} = 0, \qquad \vartheta_0^{(n)} = s_n, \tag{10.1-9a}$$

$$\vartheta_{2k+1}^{(n)} = \vartheta_{2k-1}^{(n+1)} + 1/\left[\Delta \vartheta_{2k}^{(n)}\right], \tag{10.1-9b}$$

$$\vartheta_{2k+2}^{(n)} = \vartheta_{2k}^{(n+1)} + \frac{\left[\Delta \vartheta_{2k}^{(n+1)}\right]\left[\Delta \vartheta_{2k+1}^{(n+1)}\right]}{\Delta^2 \vartheta_{2k+1}^{(n)}}, \quad k, n \in \mathbb{N}_0. \tag{10.1-9c}$$

As usual it is assumed that the difference operator Δ acts only upon the superscript n and not upon the subscript k.

Brezinski's derivation of his ϑ algorithm, which is based upon the the somewhat arbitrary choice (10.1-6), was purely experimental. However, it was certainly a very successful experiment. In numerical studies performed by Smith and Ford [29,30] it was demonstrated that Brezinski's ϑ algorithm is a very powerful as well as a very versatile sequence transformation since it is able to accelerate both linear and logarithmic convergence quite efficiently and to sum even wildly divergent series.

Unlike most of the other sequence transformations in this report, Brezinski's ϑ algorithm was not derived via a model sequence. In addition, the recursive scheme (10.1-9) is significantly more complicated than the recursive schemes of most other nonlinear sequence transformations. This should explain why only relatively little is known about the theoretical properties of Brezinski's ϑ algorithm. In his second book, Brezinski showed that the ϑ algorithm is invariant under translation according to eq. (3.1-4) (see p. 106 of ref. [20]). Also, Smith and Ford could prove that $\vartheta_2^{(n)}$ accelerates linear convergence (see pp. 225 – 226 of ref. [29]). Short discussions of the properties of the ϑ algorithm can be found in books by Brezinski [19,20] and Wimp [23].

Inspired by the success of his ϑ algorithm, Brezinski [27] suggested to use the approach, which led to the derivation of the ϑ algorithm, also in the case of other sequence transformations. Since Wynn's ϵ algorithm – which is the starting point for Brezinski's ϑ algorithm – and Wynn's ρ algorithm are structurally almost identical, it is a relatively obvious idea to try to use Brezinski's concept for the construction of a new sequence transformation which would be based upon Wynn's ρ algorithm. This can be accomplished quite easily. One only has to insert eq. (10.1-3) instead of eq. (10.1-2) into eqs. (10.1-8b) and (10.1-8c) to obtain the following recursive scheme:

$$\Theta_{-1}^{(n)} = 0, \qquad \Theta_0^{(n)} = s_n, \tag{10.1-10a}$$

$$\Theta_{2k+1}^{(n)} = \Theta_{2k-1}^{(n+1)} + \frac{x_{n+2k+1} - x_n}{\Delta \Theta_{2k}^{(n)}}, \tag{10.1-10b}$$

$$\Theta_{2k+2}^{(n)} = \Theta_{2k}^{(n+1)} - \frac{[x_{n+2k+2} - x_n][\Delta \Theta_{2k}^{(n+1)}][\Delta \Theta_{2k+1}^{(n+1)}]}{[x_{n+2k+2} - x_{n+1}][\Delta \Theta_{2k+1}^{(n)}] - [x_{n+2k+1} - x_n][\Delta \Theta_{2k+1}^{(n+1)}]},$$

$$k, n \in \mathbb{N}_0. \tag{10.1-10c}$$

Numerical tests showed that this sequence transformation $\Theta_k^{(n)}$ is more versatile than Wynn's ρ algorithm, from which it was derived, since it is able to accelerate linear convergence and to sum some divergent series. However, it is unfortunately much less efficient than Wynn's ρ algorithm in the case of logarithmic convergence, and it is also not particularly powerful in the case of linear convergence or divergence. This shows that in the case of Wynn's ρ algorithm Brezinski's experimental choice (10.1-6) does not lead to the same spectacular success as in the case of Wynn's ϵ algorithm since the resulting sequence transformation $\Theta_k^{(n)}$ – although clearly more versatile than Wynn's ρ algorithm – is not able to compete with other, more specialized sequence transformations.

10.2. Programming Brezinski's theta algorithm

A program for Brezinski's ϑ algorithm should have the same features as the other programs described in this report. This means it should read in the sequence elements $s_0, s_1, \ldots, s_m, \ldots$ successively starting with s_0. After the input of each new sequence element s_m as many new elements $\vartheta_k^{(n)}$ should be computed as is permitted by the recursive scheme (10.1-9). That new element $\vartheta_k^{(n)}$, which has the largest even subskript k, should be used as the new approximation to the limit of the sequence $\{s_n\}$.

Let us arrange the elements of the ϑ table in a rectangular scheme in such a way that the superscript n indicates the row and the subscript k the column of the 2-dimensional array:

$$\begin{array}{cccccc}
\vartheta_0^{(0)} & \vartheta_1^{(0)} & \vartheta_2^{(0)} & \ldots & \vartheta_n^{(0)} & \ldots \\
\vartheta_0^{(1)} & \vartheta_1^{(1)} & \vartheta_2^{(1)} & \ldots & \vartheta_n^{(1)} & \ldots \\
\vartheta_0^{(2)} & \vartheta_1^{(2)} & \vartheta_2^{(2)} & \ldots & \vartheta_n^{(2)} & \ldots \\
\vartheta_0^{(3)} & \vartheta_1^{(3)} & \vartheta_2^{(3)} & \ldots & \vartheta_n^{(3)} & \ldots \\
\vdots & \vdots & \vdots & \ddots & \vdots & \ddots \\
\vartheta_0^{(n)} & \vartheta_1^{(n)} & \vartheta_2^{(n)} & \ldots & \vartheta_n^{(n)} & \ldots \\
\vdots & \vdots & \vdots & \ddots & \vdots & \ddots
\end{array} \tag{10.2-1}$$

The entries in the first column of the array are the starting values $\vartheta_0^{(n)} = s_n$ of the recursion according to eq. (10.1-9a). The remaining elements of the ϑ table can be computed with the help of the recurrence formulas (10.1-9b) and (10.1-9c). The 4 elements, which are connected by the nonlinear recursion (10.1-9b), form the same pattern in the ϑ table as the 4 elements of the ϵ table which are connected by eq. (4.2-1b):

$$\begin{array}{ccc} & \vartheta_{2k}^{(n)} & \vartheta_{2k+1}^{(n)} \\ \vartheta_{2k-1}^{(n+1)} & \vartheta_{2k}^{(n+1)} & \end{array} \tag{10.2-2}$$

The 6 elements, which are connected by the nonlinear recursion (10.1-9c), form the following pattern in the ϑ table:

$$\begin{array}{ccc} & \vartheta_{2k+1}^{(n)} & \vartheta_{2k+2}^{(n)} \\ \vartheta_{2k}^{(n+1)} & \vartheta_{2k+1}^{(n+1)} & \\ \vartheta_{2k}^{(n+2)} & \vartheta_{2k+1}^{(n+2)} & \end{array} \tag{10.2-3}$$

These two patterns show that the recursions (10.1-9b) and (10.1-9c) have to proceed along a relatively complicated path in the ϑ table. Let us assume that the sequence elements s_0, $s_1, \ldots, s_{m-1}$ had been read in and as many elements of the ϑ table had been computed as it is permitted by the recursive scheme (10.1-9). After the input of the next sequence element s_m the string $\vartheta_j^{(m-[\![3j/2]\!])}$ with $0 \le j \le [\![(2m+1)/3]\!]$ can be computed. Again, $[\![x]\!]$ stands for the integral part of x, i.e., the largest integer ν satisfying $\nu \le x$. In this context, it is recommendable to rewrite the recursive scheme (10.1-9) in the following way:

$$\vartheta_0^{(n)} = s_n, \qquad n \ge 0, \tag{10.2-4a}$$

$$\vartheta_1^{(n-1)} = 1/\left[\vartheta_0^{(n)} - \vartheta_0^{(n-1)}\right], \qquad n \ge 1, \tag{10.2-4b}$$

$$\vartheta_{2j}^{(n-3j)} = \vartheta_{2j-2}^{(n-3j+1)} + \frac{\left[\Delta \vartheta_{2j-2}^{(n-3j+1)}\right]\left[\Delta \vartheta_{2j-1}^{(n-3j+1)}\right]}{\Delta^2 \vartheta_{2j-1}^{(n-3j)}}, \qquad n \ge 3,\ 1 \le j \le [\![n/3]\!], \tag{10.2-4c}$$

$$\vartheta_{2j+1}^{(n-3j-1)} = \vartheta_{2j-1}^{(n-3j)} + 1/\left[\Delta \vartheta_{2j}^{(n-3j-1)}\right], \qquad n \ge 4,\ j \le [\![(n-1)/3]\!]. \tag{10.2-4d}$$

It follows either from this recursive scheme or equivalently from the two geometric patterns (10.2-2) and (10.2-3) that for the computation of the transform $\vartheta_{2k}^{(n)}$ the sequence elements s_n, $s_{n+1}, \ldots, s_{n+3k}$ have to be known. Thus, $\vartheta_{2k}^{(n)}$ is a transformation of order $l = 3k$.

The approximation to the limit will depend upon the index m of the last sequence element s_m which was used in the recursive scheme (10.2-4). If m is a multiple of 3, $m = 3\mu$, our approximation to the limit will be the transformation

$$\{s_0, s_1, \ldots, s_{3\mu}\} \to \vartheta_{2\mu}^{(0)}, \tag{10.2-5}$$

if we have $m = 3\mu + 1$, our approximation to the limit will be

$$\{s_1, s_2, \ldots, s_{3\mu+1}\} \to \vartheta_{2\mu}^{(1)}, \tag{10.2-6}$$

and if we have $m = 3\mu + 2$, our approximation to the limit will be

$$\{s_2, s_3, \ldots, s_{3\mu+2}\} \to \vartheta_{2\mu}^{(2)}. \tag{10.2-7}$$

These three relationships can be combined into a single equation yielding

$$\{s_{m-3[\![m/3]\!]}, s_{m-3[\![m/3]\!]+1}, \ldots, s_m\} \to \vartheta_{2[\![m/3]\!]}^{(m-3[\![m/3]\!])}. \tag{10.2-8}$$

Because of the complicated structure of the nonlinear recursive scheme (10.2-4) a single 1-dimensional array will not suffice for the computation of the new string $\vartheta_j^{(n-[\![3j/2]\!])}$ with $0 \le j \le [\![(2n+1)/3]\!]$ after the input of the last sequence element s_n. Two 1-dimensional arrays A and B together with three auxiliary variables will be needed.

We shall use the convention that if the index n of the last sequence element s_n, which was read in, is even, $n = 2m$, the new string will be stored in the array A according to the rule

$$\vartheta_j^{(2m-[\![3j/2]\!])} \to A(j), \qquad 0 \le j \le [\![(4m+1)/3]\!], \tag{10.2-9}$$

and if the index n of the last sequence element s_n is odd, $n = 2m+1$, the new string will be stored in the array B according to the rule

$$\vartheta_j^{(2m-[\![3j/2]\!]+1)} \to B(j), \qquad 0 \le j \le [\![(4m+3)/3]\!]. \tag{10.2-10}$$

Let us now assume that the index n of the last sequence element s_n, which was read in, is even, i.e., $n = 2m$. Before the computation of the new string (10.2-9), which is to be stored in A, the array B contains the elements $\vartheta_j^{(2m-[\![3j/2]\!]-1)}$ with $0 \le j \le [\![(4m-1)/3]\!]$, whereas in A the elements $\vartheta_j^{(2m-[\![3j/2]\!]-2)}$ with $0 \le j \le [\![(4m-3)/3]\!]$ are stored. The recursive scheme (10.2-4) can then be expressed in terms of the elements of the arrays A and B in the following way:

$$A(0) \leftarrow s_{2m}, \qquad m \in \mathbb{N}_0, \tag{10.2-11a}$$

$$A(1) \leftarrow 1/[A(0) - B(0)], \tag{10.2-11b}$$

$$A(2j) \leftarrow A'(2j-2) + \frac{[B(2j-2) - A'(2j-2)][A(2j-1) - B(2j-1)]}{A(2j-1) - 2B(2j-1) + A'(2j-1)},$$

$$j \le [\![2m/3]\!], \tag{10.2-11c}$$

$$A(2j+1) \leftarrow A'(2j-1) + 1/[B(2j) - A(2j)], \qquad j \le [\![(2m-1)/3]\!]. \tag{10.2-11d}$$

The primed array elements $A'(2j-2)$ and $A'(2j-1)$ refer to the occupation of A after the computation of the string $\vartheta_j^{(2m-[\![3j/2]\!]-2)}$ with $0 \le j \le [\![(4m-3)/3]\!]$. Since these elements are overwritten during the computation of the new string, they have to be stored in auxiliary variables.

Let us now assume that the index n of the last sequence element s_n, which was read in, is odd, i.e., $n = 2m+1$. This implies that the rôle of the two arrays A and B has to be interchanged.

Before the computation of the new string (10.2-10), which is to be stored in B, the array A contains the elements $\vartheta_j^{(2m-[\![3j/2]\!])}$ with $0 \le j \le [\![(4m+1)/3]\!]$, and in B the elements $\vartheta_j^{(2m-[\![3j/2]\!]-1)}$ with $0 \le j \le [\![(4m-1)/3]\!]$ are stored. The recursive scheme (10.2-4) can then be expressed in terms of the elements of the arrays A and B in the following way:

$$B(0) \leftarrow s_{2m+1}, \qquad m \in \mathbb{N}_0, \tag{10.2-12a}$$

$$B(1) \leftarrow 1/[B(0) - A(0)], \tag{10.2-12b}$$

$$B(2j) \leftarrow B'(2j-2) + \frac{[A(2j-2) - B'(2j-2)][B(2j-1) - A(2j-1)]}{B(2j-1) - 2A(2j-1) + B'(2j-1)},$$

$$j \le [\![(2m+1)/3]\!], \tag{10.2-12c}$$

$$B(2j+1) \leftarrow B'(2j-1) + 1/[A(2j) - B(2j)], \qquad j \le [\![2m/3]\!]. \tag{10.2-12d}$$

The primed array elements $B''(2j-2)$ and $B'(2j-1)$ refer to the occupation of B after the computation of the string $\vartheta_j^{(2m-[\![3j/2]\!]-1)}$ with $0 \le j \le [\![(4m-1)/3]\!]$. Since these elements are overwritten during the computation of the new string, they have to be stored in auxiliary variables.

The following FORTRAN 77 subroutine THETA performs the recursive computation of Brezinski's ϑ algorithm in two 1-dimensional arrays A and B using the two recursive schemes (10.2-11) and (10.2-12). It is safeguarded against an exact or approximate vanishing of the denominators $\Delta^2 \vartheta_{2j-1}^{(n-3j)}$ and $\Delta \vartheta_{2j-2}^{(n-3j+1)}$ in eqs. (10.2-4c) and (10.2-4d) by using two variables HUGE and TINY. The elements s_n with $n = 0, 1, 2, \ldots$ of the sequence to be transformed have to be computed in a DO loop in the calling program. Whenever a new sequence element s_n is computed in the outer DO loop, this subroutine THETA has to be called again and a new string $\vartheta_j^{(n-[\![3j/2]\!])}$ with $0 \le j \le [\![(2n+1)/3]\!]$ will be calculated. The new sequence element s_n is read in via the variable SOFN and the approximation to the limit is returned via the variable ESTLIM.

It is important to note that this subroutine THETA only calculates the approximations to the limit according to eqs. (10.2-5)–(10.2-7). The convergence of the whole process has to be analyzed in the calling program.

On pp. 368–370 of Brezinski's second book [20] the listing of a FORTRAN IV program, which computes Brezinski's ϑ algorithm using three 1-dimensional arrays, can be found.

```
SUBROUTINE THETA(SOFN,N,A,B,LENGA,LENGB,ESTLIM)
DIMENSION A(0:LENGA), B(0:LENGB)
PARAMETER ( HUGE = 1.E+60 , TINY = 1.E-60 )
PARAMETER ( ZERO = 0.E0 , ONE  = 1.E0 , TWO = 2.E0 )
JMAX  = (2 * N + 1) / 3
NMOD2 = MOD(N,2)
IF (N.EQ.0) THEN
  A(0)   = SOFN
  ESTLIM = SOFN
  RETURN
END IF
```

```
      IF ( NMOD2 .EQ. 0) THEN
        AUX2    = ZERO
        AUX1    = A(0)
        A(0)    = SOFN
        DO 10 J = 1, JMAX
          AUX3  = AUX2
          AUX2  = AUX1
          IF ( J .LT. JMAX ) THEN
            AUX1  = A(J)
          END IF
          IF ( MOD(J,2) .EQ. 0 ) THEN
            DENOM = A(J-1) - TWO * B(J-1) + AUX2
            IF ( ABS(DENOM) .LT. TINY ) THEN
              A(J) = HUGE
            ELSE
              A(J) = AUX3 + (B(J-2) - AUX3 ) * ( A(J-1) - B(J-1) )
     1               / DENOM
            END IF
          ELSE
            DIFF  = A(J-1) - B(J-1)
            IF ( ABS(DIFF) .LT. TINY ) THEN
              A(J) = HUGE
            ELSE
              A(J) = AUX3 + ONE / DIFF
            END IF
          END IF
10      CONTINUE
        IF ( MOD(JMAX,2) .EQ. 0 ) THEN
          ESTLIM = A(JMAX)
        ELSE
          ESTLIM = A(JMAX-1)
        END IF
      ELSE
        AUX2    = ZERO
        AUX1    = B(0)
        B(0)    = SOFN
        DO 20 J = 1, JMAX
          AUX3  = AUX2
          AUX2  = AUX1
          IF ( J .LT. JMAX ) THEN
            AUX1  = B(J)
          END IF
          IF ( MOD(J,2) .EQ. 0 ) THEN
            DENOM = B(J-1) - TWO * A(J-1) + AUX2
            IF ( ABS(DENOM) .LT. TINY ) THEN
              B(J) = HUGE
            ELSE
              B(J) = AUX3 + ( A(J-2) - AUX3 ) * ( B(J-1) - A(J-1) )
     1               / DENOM
            END IF
          ELSE
            DIFF  = B(J-1) - A(J-1)
```

```
            IF ( ABS(DIFF) .LT. TINY ) THEN
              B(J) = HUGE
            ELSE
              B(J) = AUX3 + ONE / DIFF
            END IF
          END IF
20      CONTINUE
        IF ( MOD(JMAX,2) .EQ. 0 ) THEN
          ESTLIM = B(JMAX)
        ELSE
          ESTLIM = B(JMAX-1)
        END IF
      END IF
      RETURN
      END
```

10.3. The iteration of $\vartheta_2^{(n)}$

In section 5 it was shown how Aitken's Δ^2 algorithm, which according to eq. (5.1-5) is identical with $\epsilon_2^{(n)}$, can be iterated to give the sequence transformation $\mathscr{A}_k^{(n)}$. In the same way, it was shown in section 6.3 that $\rho_2^{(n)}$ can be iterated to give the sequence transformation $\mathscr{W}_k^{(n)}$. In this section, we want to analyze how the transform $\vartheta_2^{(n)}$ can be iterated. From the recursive scheme (10.1-9) we obtain the following expression:

$$\vartheta_2^{(n)} = s_{n+1} - \frac{[\Delta s_n][\Delta s_{n+1}][\Delta^2 s_{n+1}]}{[\Delta s_{n+2}][\Delta^2 s_n] - [\Delta s_n][\Delta^2 s_{n+1}]}, \qquad n \in \mathbb{N}_0. \tag{10.3-1}$$

It follows from this relationship that $\vartheta_2^{(n)}$, which is a kind of weighted Δ^3 process, is identical with Lubkin's W transformation [40]. Many other representations for $\vartheta_2^{(n)}$ can be derived by suitable manipulations of eq. (10.3-1). Examples are:

$$\vartheta_2^{(n)} = \frac{s_{n+1}[\Delta s_{n+2}][\Delta^2 s_n] - s_{n+2}[\Delta s_n][\Delta^2 s_{n+1}]}{[\Delta s_{n+2}][\Delta^2 s_n] - [\Delta s_n][\Delta^2 s_{n+1}]} \tag{10.3-2}$$

$$= \frac{\Delta^2[s_{n+1}/\Delta s_n]}{\Delta^2[1/\Delta s_n]}. \tag{10.3-3}$$

Comparison of eq. (10.3-3) with eqs. (2.4-8), (7.3-5), and (9.5-4) shows that $\vartheta_2^{(n)}$ is also a special case of Levin's u transformation or Drummonds sequence transformation with $\omega_{n+1} = \Delta s_n$:

$$\vartheta_2^{(n)} = u_2^{(n+1)}(\beta, s_{n+1}) \tag{10.3-4}$$

$$= \mathscr{D}_2^{(n+1)}(s_{n+1}, \Delta s_n). \tag{10.3-5}$$

In addition, a comparison of eqs. (5.1-12) and (10.3-3) shows that $\vartheta_2^{(n)}$ may also be considered to be a generalization of Aitken's Δ^2 process.

If we want to iterate eq. (10.3-1), we have to use the following recursive scheme:

$$\mathscr{J}_0^{(n)} = s_n, \qquad n \in \mathbb{N}_0, \tag{10.3-6a}$$

$$\mathscr{J}_{k+1}^{(n)} = \mathscr{J}_k^{(n+1)} - \frac{[\Delta \mathscr{J}_k^{(n)}][\Delta \mathscr{J}_k^{(n+1)}][\Delta^2 \mathscr{J}_k^{(n+1)}]}{[\Delta \mathscr{J}_k^{(n+2)}][\Delta^2 \mathscr{J}_k^{(n)}] - [\Delta \mathscr{J}_k^{(n)}][\Delta^2 \mathscr{J}_k^{(n+1)}]}, \qquad k, n \in \mathbb{N}_0. \tag{10.3-6b}$$

As usual, the difference operator Δ acts only upon the superscript n and not upon the subscript k. It follows from this recursive scheme that for the computation of $\mathscr{J}_k^{(n)}$ the sequence elements $s_n, s_{n+1}, \ldots, s_{n+3k}$ have to be known. Consequently, $\mathscr{J}_k^{(n)}$ is a transformation of order $3k$. In that respect, $\mathscr{J}_k^{(n)}$ is equivalent to $\vartheta_{2k}^{(n)}$ which needs the same set $s_n, s_{n+1}, \ldots, s_{n+3k}$ of sequence elements for its computation.

In sections 13 and 14, we shall see that $\mathscr{J}_k^{(n)}$ is a powerful sequence transformation which has similiar properties as $\vartheta_{2k}^{(n)}$, i.e., it is able to accelerate linear and logarithmic convergence and is also able to sum even wildly divergent series. This may be considered to be an indirect confirmation of the validity of Brezinski's choice (10.1-4) which is based upon the assumption that the exact numerical values of the transforms $\vartheta_{2k+1}^{(n)}$ do not really matter as long as they diverge if the transforms $\vartheta_{2k}^{(n)}$ converge. If we would replace the 4-term recursion (10.1-9b) by the 3-term recursion

$$\vartheta_{2k+1}^{(n)} = 1/[\Delta \vartheta_{2k}^{(n)}], \qquad k, n \in \mathbb{N}_0, \tag{10.3-7}$$

then with this modified ϑ algorithm we would obtain

$$\vartheta_{2k}^{(n)} = \mathscr{J}_k^{(n)}. \tag{10.3-8}$$

10.4. Programming the iterated theta algorithm

A program, which computes the sequence transformation $\mathscr{J}_k^{(n)}$, should have the same features as the other programs in this report. This means it should read in the sequence elements $s_0, s_1, \ldots, s_m, \ldots$ successively starting with s_0. After the input of each new sequence element s_m as many new elements $\mathscr{J}_k^{(n)}$ should be computed as is permitted by the recursive scheme (10.3-6). That new element $\mathscr{J}_k^{(n)}$, which has the largest subskript k, should be used as the new approximation to the limit of the sequence $\{s_n\}$.

Let us arrange the elements $\mathscr{J}_k^{(n)}$ in a rectangular scheme in such a way that the superscript n indicates the row and the subscript k the column of the 2-dimensional array:

$$\begin{array}{cccccc}
\mathscr{J}_0^{(0)} & \mathscr{J}_1^{(0)} & \mathscr{J}_2^{(0)} & \cdots & \mathscr{J}_n^{(0)} & \cdots \\
\mathscr{J}_0^{(1)} & \mathscr{J}_1^{(1)} & \mathscr{J}_2^{(1)} & \cdots & \mathscr{J}_n^{(1)} & \cdots \\
\mathscr{J}_0^{(2)} & \mathscr{J}_1^{(2)} & \mathscr{J}_2^{(2)} & \cdots & \mathscr{J}_n^{(2)} & \cdots \\
\mathscr{J}_0^{(3)} & \mathscr{J}_1^{(3)} & \mathscr{J}_2^{(3)} & \cdots & \mathscr{J}_n^{(3)} & \cdots \\
\vdots & \vdots & \vdots & \ddots & \vdots & \ddots \\
\mathscr{J}_0^{(n)} & \mathscr{J}_1^{(n)} & \mathscr{J}_2^{(n)} & \cdots & \mathscr{J}_n^{(n)} & \cdots \\
\vdots & \vdots & \vdots & \ddots & \vdots & \ddots
\end{array} \tag{10.4-1}$$

The entries in the first column of the array are the starting values $\mathscr{J}_0^{(n)} = s_n$ of the recursion according to eq. (10.3-6a). The remaining elements of the $\mathscr{J}$ table can be computed with the help of the recurrence formula (10.3-6b). The 5 elements, which are connected by the nonlinear recursion (10.3-6b), form the following pattern:

$$\begin{array}{ll} \mathscr{J}_k^{(n)} & \mathscr{J}_{k+1}^{(n)} \\ \mathscr{J}_k^{(n+1)} & \\ \mathscr{J}_k^{(n+2)} & \\ \mathscr{J}_k^{(n+3)} & \end{array} \tag{10.4-2}$$

It was remarked earlier that $\vartheta_2^{(n)}$ may be considered to be a generalization of Aitken's Δ^2 process. In the same way, the computational algorithm for the sequence transformation $\mathscr{J}_k^{(n)}$ is essentially a generalization of the computational scheme for Aitken's iterated Δ^2 process $\mathscr{A}_k^{(n)}$ which was discussed in section 5.2. For that purpose, the recursive scheme (10.3-6) is rewritten in the following way:

$$\mathscr{J}_0^{(n)} = s_n, \qquad n \geq 0, \tag{10.4-3a}$$

$$\mathscr{J}_l^{(n-3l)} = \mathscr{J}_{l-1}^{(n-3l+1)} - \frac{\left[\Delta \mathscr{J}_{l-1}^{(n-3l)}\right]\left[\Delta \mathscr{J}_{l-1}^{(n-3l+1)}\right]\left[\Delta^2 \mathscr{J}_{l-1}^{(n-3l+1)}\right]}{\left[\Delta \mathscr{J}_{l-1}^{(n-3l+2)}\right]\left[\Delta^2 \mathscr{J}_{l-1}^{(n-3l)}\right] - \left[\Delta \mathscr{J}_{l-1}^{(n-3l)}\right]\left[\Delta^2 \mathscr{J}_{l-1}^{(n-3l+1)}\right]},$$

$$n \geq 3 \qquad 1 \leq l \leq [\![n/3]\!]. \tag{10.4-3b}$$

As usual, $[\![x]\!]$ denotes the integral part of x, i.e., the largest integer ν satisfying $\nu \leq x$. It follows either from the geometric pattern (10.4-2) or from this recursive scheme that after the input of a new sequence element s_m the string $\mathscr{J}_\mu^{(m-3\mu)}$ with $0 \leq \mu \leq [\![m/3]\!]$ can be computed.

Again, the approximation to the limit of the sequence to be transformed depends upon the index m of the last sequence element s_m which was read in. If m is a multiple of 3, $m = 3\mu$, our approximation to the limit will be the transformation

$$\{s_0, s_1, \ldots, s_{3\mu}\} \to \mathscr{J}_\mu^{(0)}, \tag{10.4-4}$$

if we have $m = 3\mu + 1$, our approximation to the limit will be

$$\{s_1, s_2, \ldots, s_{3\mu+1}\} \to \mathscr{J}_\mu^{(1)}, \tag{10.4-5}$$

and if we have $m = 3\mu + 2$, our approximation to the limit will be

$$\{s_2, s_3, \ldots, s_{3\mu+2}\} \to \mathscr{J}_\mu^{(2)}. \tag{10.4-6}$$

With the help of the notation $[\![x]\!]$ for the integral part of x these three relationships can be combined into a single equation yielding

$$\{s_{m-3[\![m/3]\!]}, s_{m-3[\![m/3]\!]+1}, \ldots, s_m\} \to \mathscr{J}_{[\![m/3]\!]}^{(m-3[\![m/3]\!])}. \tag{10.4-7}$$

The recursive scheme (10.4-3) – or equivalently the geometric pattern (10.4-2) – looks relatively complicated. But nevertheless, it is possible to perform the computation of the string $\mathscr{J}_l^{(n-3l)}$ with $0 \le l \le [\![n/3]\!]$ in a single 1-dimensional array J if the elements of the table of this sequence transformation are stored according to the following rule:

$$\mathscr{J}_{[\![\nu/3]\!]}^{(n-\nu)} \to J(n-\nu), \qquad n \ge 0, \quad 0 \le \nu \le n. \tag{10.4-8}$$

With this convention, the recursive scheme (10.4-3) can be reformulated in terms of the elements of the array J:

$$J(n) \leftarrow s_n, \qquad n \ge 0, \tag{10.4-9a}$$

$$J(m) \leftarrow J(m+1) - \frac{[\Delta J(m)][\Delta J(m+1)][\Delta^2 J(m+1)]}{[\Delta J(m+2)][\Delta^2 J(m)] - [\Delta J(m)][\Delta^2 J(m+1)]},$$

$$m = n - 3l, \qquad n \ge 3, \qquad 1 \le l \le [\![n/3]\!]. \tag{10.4-9b}$$

The following FORTRAN 77 subroutine THETIT performs the recursive computation of the iterated ϑ_2 algorithm in a 1-dimensional array J using the recursive scheme (10.4-3). THETIT is safeguarded against an exact or approximate vanishing of the denominator in eq. (10.4-3c) by using two variables HUGE and TINY. The elements s_n with $n = 0, 1, 2, \ldots$ of the sequence to be transformed have to be computed in a DO loop in the calling program. Whenever a new sequence element s_n is computed in the outer DO loop this subroutine THETIT has to be called again and a new string $\mathscr{J}_l^{(n-3l)}$ with $0 \le l \le [\![n/3]\!]$ will be calculated. The new sequence element s_n is read in via the variable SOFN and the approximation to the limit is returned via the variable ESTLIM.

It is important to note that THETIT only calculates the approximations to the limit according to eqs. (10.4-4)–(10.4-6). The convergence of the whole acceleration or summation process has to be analyzed in the calling program.

```
      SUBROUTINE THETIT(SOFN,N,ARJ,LARRAY,ESTLIM)
      DIMENSION ARJ(0:LARRAY)
      PARAMETER ( HUGE = 1.E+60 , TINY = 1.E-60 )
      ARJ(N) = SOFN
      IF (N.LT.3) THEN
        ESTLIM  = SOFN
      ELSE
        LMAX = N/3
        M    = N
        DO 10 L = 1,LMAX
          M     = M - 3
          DIFFO = ARJ(M+1)  - ARJ(M)
          DIFF1 = ARJ(M+2)  - ARJ(M+1)
          DIFF2 = ARJ(M+3)  - ARJ(M+2)
          DENOM = DIFF2 * (DIFF1  - DIFFO)  - DIFFO * (DIFF2  - DIFF1)
          IF ( ABS(DENOM).LT.TINY ) THEN
            ARJ(M) = HUGE
```

```
            ELSE
              ARJ(M) = ARJ(M+1) - DIFF0 * DIFF1 * (DIFF2 - DIFF1) / DENOM
            END IF
10        CONTINUE
          ESTLIM = ARJ(MOD(N,3))
        END IF
        RETURN
        END
```

11. On the derivation of theta-type algorithms

11.1. New sequence transformations based upon Aitken's iterated Δ^2 process

In section 10.1 it was discussed how Brezinski's ϑ algorithm, eq. (10.1-9), can be derived by modifying Wynn's ϵ algorithm, eq. (4.2-1). The ϑ algorithm is a very powerful sequence transformation. It accelerates linear convergence and sums divergent series approximately as efficiently as the ϵ algorithm. However, unlike the ϵ algorithm the ϑ algorithm is also able to accelerate many logarithmically convergent sequences. Consequently, it is frequently emphasized in the literature that Brezinski's ϑ algorithm combines the advantageous features of both Wynn's ϵ and Wynn's ρ algorithm.

Brezinski [27] suggested to use his approach, which led to the ϑ algorithm, also in the case of other sequence transformations. This will be done in this section. However, one should not expect that Brezinski's approach will automatically lead to new sequence transformations that are more useful than the transformations from which they were derived. For instance, in the case of Wynn's ρ algorithm, eq. (6.2-2), which is formally almost identical with the ϵ algorithm, Brezinski's approach led to the sequence transformation $\Theta_k^{(n)}$, eq. (10.1-10), which is much less efficient than Wynn's ρ algorithm in the case of logarithmic convergence and which is also not very powerful in the case of linear convergence or divergence. Consequently, the sequence transformation $\Theta_k^{(n)}$ is practically useless although it is certainly more versatile than the ρ algorithm from which it was derived.

But even if Brezinski's ϑ concept does not automatically lead to practically useful new sequence transformations, it should nevertheless be worthwhile to investigate in which cases new sequence transformations can be obtained that are at least in some sense better than the transformations from which they were derived.

It follows from eqs. (10.1-1) and (10.1-8) that the essential step in the derivation of the ϑ algorithm consists in replacing a recursion of the general type

$$f_n = a_n + b_n \tag{11.1-1}$$

by a more complicated modified recursion

$$f_n = a_n - \frac{\Delta a_n}{\Delta b_n} b_n . \tag{11.1-2}$$

Thus, a new sequence transformation can be constructed by replacing a recursion of the type of eq. (11.1-1) by a recursion of the type of eq. (11.1-2) in the recursive scheme which defines a given sequence transformation. The new sequence transformation will have a more nonlinear structure than the original transformation. In addition, such a modification will normally increase the order of the transformation by one.

The probably closest relative of Wynn's ϵ algorithm is Aitken's iterated Δ^2 process, eq. (5.1-15). This follows from the fact that because of eq. (5.1-5) the ϵ algorithm may also be considered to be a generalization of Aitken's Δ^2 process, eq. (5.1-4). Aitken's iterated Δ^2 process and Wynn's ϵ algorithm have similar properties since they are both able to accelerate linear convergence and to sum many divergent series but are unable to accelerate logarithmic convergence. Consequently, it would be interesting to see how Brezinski's ϑ concept works in the case of Aitken's iterated Δ^2 process.

Let us now assume that a sequence transformation is defined by the following recursive scheme:

$$T_0^{(n)} = s_n, \qquad n \in \mathbb{N}_0, \tag{11.1-3a}$$

$$T_{k+1}^{(n)} = T_k^{(n)} + D_k^{(n)}, \qquad k, n \in \mathbb{N}_0. \tag{11.1-3b}$$

As in section 10.1 it is assumed that $D_k^{(n)}$ is a quantity which depends upon one or several elements of the table of this transformation. If this recursive scheme is to be modified along the lines of Brezinski's ϑ algorithm, a comparison with eqs. (11.1-1) and (11.1-2) shows that it must be changed in the following way:

$$T_0^{(n)} = s_n, \qquad n \in \mathbb{N}_0, \tag{11.1-4a}$$

$$T_{k+1}^{(n)} = T_k^{(n)} - \frac{\Delta T_k^{(n)}}{\Delta D_k^{(n)}} D_k^{(n)}, \qquad k, n \in \mathbb{N}_0. \tag{11.1-4b}$$

Aitken's iterated Δ^2 process, eq. (5.1-15), is of the form of eq. (11.1-3). Thus, if Aitken's iterated Δ^2 process is modified according to eq. (11.1-4), a new sequence transformation $\mathscr{B}_k^{(n)}$ results which is defined by the following nonlinear recursive scheme:

$$\mathscr{B}_0^{(n)} = s_n, \qquad n \in \mathbb{N}_0, \tag{11.1-5a}$$

$$\mathscr{B}_{k+1}^{(n)} = \mathscr{B}_k^{(n)} + \frac{\left[\Delta \mathscr{B}_k^{(n)}\right]^3 \left[\Delta^2 \mathscr{B}_k^{(n+1)}\right]}{\left[\Delta \mathscr{B}_k^{(n)}\right]^2 \left[\Delta^2 \mathscr{B}_k^{(n+1)}\right] - \left[\Delta \mathscr{B}_k^{(n+1)}\right]^2 \left[\Delta^2 \mathscr{B}_k^{(n)}\right]}, \qquad k, n \in \mathbb{N}_0. \tag{11.1-5b}$$

Again, it is assumed that the difference operator Δ acts only upon the superscript n and not upon the subscript k.

However, this is not the only possibility of modifying Aitken's iterated Δ^2 process. It follows from eq. (5.1-6) that the recursive scheme for Aitken's iterated Δ^2 process can also be written in the following way:

$$\mathscr{A}_0^{(n)} = s_n, \qquad n \in \mathbb{N}_0, \tag{11.1-6a}$$

$$\mathscr{A}_{k+1}^{(n)} = \mathscr{A}_k^{(n+1)} - \frac{\left[\Delta \mathscr{A}_k^{(n)}\right]\left[\Delta \mathscr{A}_k^{(n+1)}\right]}{\Delta^2 \mathscr{A}_k^{(n)}}, \qquad k, n \in \mathbb{N}_0. \tag{11.1-6b}$$

This version of Aitken's iterated Δ^2 process is a recursive scheme of the following type:

$$T_0^{(n)} = s_n, \qquad n \in \mathbb{N}_0, \tag{11.1-7a}$$

$$T_{k+1}^{(n)} = T_k^{(n+1)} + D_k^{(n)}, \qquad k, n \in \mathbb{N}_0. \tag{11.1-7b}$$

If we compare this recursive scheme with eqs. (11.1-1) and (11.1-2) we see that it has to be modified in the following way:

$$T_0^{(n)} = s_n, \qquad n \in \mathbb{N}_0, \tag{11.1-8a}$$

$$T_{k+1}^{(n)} = T_k^{(n+1)} - \frac{\Delta T_k^{(n+1)}}{\Delta D_k^{(n)}} D_k^{(n)}, \qquad k, n \in \mathbb{N}_0. \tag{11.1-8b}$$

However, a modification of the second version of Aitken's iterated Δ^2 process, eq. (11.1-6), according to eq. (11.1-8) does not produce a new sequence transformation since we obtain the recursive scheme for the sequence transformation $\mathscr{J}_k^{(n)}$, eq. (10.3-6), which was derived by iterating the expression for $\vartheta_2^{(n)}$, eq. (10.3-1).

It follows from eq. (5.1-7) that there is another possibility of rewriting the recursive scheme for Aitken's Δ^2 process in a way which would be suited for our purposes:

$$\mathscr{A}_0^{(n)} = s_n, \qquad n \in \mathbb{N}_0, \tag{11.1-9a}$$

$$\mathscr{A}_{k+1}^{(n)} = \mathscr{A}_k^{(n+2)} - \frac{\left[\Delta \mathscr{A}_k^{(n+1)}\right]^2}{\Delta^2 \mathscr{A}_k^{(n)}}, \qquad k, n \in \mathbb{N}_0. \tag{11.1-9b}$$

This version of Aitken's iterated Δ^2 process is a recursive scheme with the following general structure,

$$T_0^{(n)} = s_n, \qquad n \in \mathbb{N}_0, \tag{11.1-10a}$$

$$T_{k+1}^{(n)} = T_k^{(n+2)} + D_k^{(n)}, \qquad k, n \in \mathbb{N}_0, \tag{11.1-10b}$$

which according to eqs. (11.1-1) and (11.1-2) has to be modified in the following way:

$$T_0^{(n)} = s_n, \qquad n \in \mathbb{N}_0, \tag{11.1-11a}$$

$$T_{k+1}^{(n)} = T_k^{(n+2)} - \frac{\Delta T_k^{(n+2)}}{\Delta D_k^{(n)}} D_k^{(n)}, \qquad k, n \in \mathbb{N}_0. \tag{11.1-11b}$$

If the third version of Aitken's iterated Δ^2 process, eq. (11.1-9), is modified according to eq. (11.1-11), we obtain a new sequence transformation $\mathscr{C}_k^{(n)}$ which is defined by the following nonlinear recursive scheme:

$$\mathscr{C}_0^{(n)} = s_n, \qquad n \in \mathbb{N}_0, \tag{11.1-12a}$$

$$\mathscr{C}_{k+1}^{(n)} = \mathscr{C}_k^{(n+2)} + \frac{\left[\Delta\mathscr{C}_k^{(n+1)}\right]^2\left[\Delta\mathscr{C}_k^{(n+2)}\right]\left[\Delta^2\mathscr{C}_k^{(n+1)}\right]}{\left[\Delta\mathscr{C}_k^{(n+1)}\right]^2\left[\Delta^2\mathscr{C}_k^{(n+1)}\right] - \left[\Delta\mathscr{C}_k^{(n+2)}\right]^2\left[\Delta^2\mathscr{C}_k^{(n)}\right]}, \quad k, n \in \mathbb{N}_0. \tag{11.1-12b}$$

The recursive schemes (11.1-5) and (11.1-12) for the two new sequence transformations $\mathscr{B}_k^{(n)}$ and $\mathscr{C}_k^{(n)}$ have the same structure as the recursive scheme (10.3-6) for $\mathscr{J}_k^{(n)}$ since these transformations are all weighted Δ^3 methods. Consequently, the two new sequence transformations can be computed in the same way as $\mathscr{J}_k^{(n)}$. In this context it is recommendable to rewrite the recursive scheme (11.1-5) in the following way:

$$\mathscr{B}_0^{(n)} = s_n, \qquad n \geq 0, \tag{11.1-13a}$$

$$\mathscr{B}_j^{(n-3j)} = \mathscr{B}_{j-1}^{(n-3j)} + \frac{\left[\Delta\mathscr{B}_{j-1}^{(n-3j)}\right]^3\left[\Delta^2\mathscr{B}_{j-1}^{(n-3j+1)}\right]}{\left[\Delta\mathscr{B}_{j-1}^{(n-3j)}\right]^2\left[\Delta^2\mathscr{B}_{j-1}^{(n-3j+1)}\right] - \left[\Delta\mathscr{B}_{j-1}^{(n-3j+1)}\right]^2\left[\Delta^2\mathscr{B}_{j-1}^{(n-3j)}\right]},$$

$$n \geq 3, \qquad 1 \leq j \leq [\![n/3]\!]. \tag{11.1-13b}$$

As usual, $[\![x]\!]$ denotes the integral part of x, i.e., the largest integer ν satisfying $\nu \leq x$. The recursive scheme (11.1-12) should also be rewritten in the same way:

$$\mathscr{C}_0^{(n)} = s_n, \qquad n \geq 0, \tag{11.1-14a}$$

$$\mathscr{C}_j^{(n-3j)} = \mathscr{C}_{j-1}^{(n-3j+2)} + \frac{\left[\Delta\mathscr{C}_{j-1}^{(n-3j+1)}\right]^2\left[\Delta\mathscr{C}_{j-1}^{(n-3j+2)}\right]\left[\Delta^2\mathscr{C}_{j-1}^{(n-3j+1)}\right]}{\left[\Delta\mathscr{C}_{j-1}^{(n-3j+1)}\right]^2\left[\Delta^2\mathscr{C}_{j-1}^{(n-3j+1)}\right] - \left[\Delta\mathscr{C}_{j-1}^{(n-3j+2)}\right]^2\left[\Delta^2\mathscr{C}_{j-1}^{(n-3j)}\right]},$$

$$n \geq 3, \qquad 1 \leq j \leq [\![n/3]\!]. \tag{11.1-14b}$$

It follows from eqs. (11.1-13) and (11.1-14) that after the input of a new sequence element s_m the strings $\mathscr{B}_\mu^{(m-3\mu)}$ and $\mathscr{C}_\mu^{(m-3\mu)}$ with $0 \leq \mu \leq [\![m/3]\!]$ can be computed.

The approximations to the limit of the sequence to be transformed depend upon the index m of the last sequence element s_m which was read in. Let $\mathscr{X}_k^{(n)}$ stand for either $\mathscr{B}_k^{(n)}$ or $\mathscr{C}_k^{(n)}$. Then, if m is a multiple of 3, $m = 3\mu$, our approximation to the limit will be the transformation

$$\{s_0, s_1, \ldots, s_{3\mu}\} \to \mathscr{X}_\mu^{(0)}, \tag{11.1-15}$$

if we have $m = 3\mu + 1$, our approximation to the limit will be

$$\{s_1, s_2, \ldots, s_{3\mu+1}\} \to \mathscr{X}_\mu^{(1)}, \tag{11.1-16}$$

and if we have $m = 3\mu + 2$, our approximation to the limit will be

$$\{s_2, s_3, \ldots, s_{3\mu+2}\} \to \mathscr{X}_{\mu}^{(2)}. \tag{11.1-17}$$

These three relationships can be combined into a single equation yielding

$$\{s_{m-3[\![m/3]\!]}, s_{m-3[\![m/3]\!]+1}, \ldots, s_m\} \to \mathscr{X}_{[\![m/3]\!]}^{(m-3[\![m/3]\!])}. \tag{11.1-18}$$

As in the case of the iterated ϑ_2 algorithm $\mathscr{J}_k^{(n)}$, eq. (10.3-6), the recursive computation of the strings $\mathscr{B}_j^{(n-3j)}$ and $\mathscr{C}_j^{(n-3j)}$ with $0 \le j \le [\![n/3]\!]$ can be done in 1-dimensional arrays B and C, respectively, if the elements of the tables of these sequence transformations are stored according to the following rule:

$$\mathscr{X}_{[\![\nu/3]\!]}^{(n-\nu)} \to X(n-\nu), \qquad n \ge 0, \quad 0 \le \nu \le n. \tag{11.1-19}$$

Again, $\mathscr{X}_k^{(n)}$ stands for either $\mathscr{B}_k^{(n)}$ or $\mathscr{C}_k^{(n)}$, and X stands for the corresponding 1-dimensional array B or C.

With this convention, the recursive scheme (11.1-13) can be reformulated in terms of the elements of the array B:

$$B(n) \leftarrow s_n, \qquad n \ge 0, \tag{11.1-20a}$$

$$B(m) \leftarrow B(m) + \frac{[\Delta B(m)]^3 [\Delta^2 B(m+1)]}{[\Delta B(m)]^2 [\Delta^2 B(m+1)] - [\Delta B(m+1)]^2 [\Delta^2 B(m)]},$$

$$m = n - 3j, \qquad n \ge 3, \qquad 1 \le j \le [\![n/3]\!]. \tag{11.1-20b}$$

Similarly, the recursive scheme (11.1-14) can be reformulated in terms of the elements of the array C:

$$C(n) \leftarrow s_n, \qquad n \ge 0, \tag{11.1-21a}$$

$$C(m) \leftarrow C(m+2) + \frac{[\Delta C(m+1)]^2 [\Delta C(m+2)] [\Delta^2 C(m+1)]}{[\Delta C(m+1)]^2 [\Delta^2 C(m+1)] - [\Delta C(m+2)]^2 [\Delta^2 C(m)]},$$

$$m = n - 3j, \qquad n \ge 3, \qquad 1 \le j \le [\![n/3]\!]. \tag{11.1-21b}$$

It follows from their recursive schemes that $\mathscr{J}_k^{(n)}$, eq. (10.3-6), $\mathscr{B}_k^{(n)}$, eq. (11.1-5), and $\mathscr{C}_k^{(n)}$, eq. (11.1-12), are all transformations of order $l = 3k$. In that respect, they are equivalent to $\vartheta_{2k}^{(n)}$, eq. (10.1-9), which needs the same set $s_n, \ldots, s_{n+3k}$ of sequence elements for its computation.

Numerical tests showed that the sequence transformations $\mathscr{J}_k^{(n)}$, eq. (10.3-6), $\mathscr{B}_k^{(n)}$, eq (11.1-5), and $\mathscr{C}_k^{(n)}$, eq. (11.1-12), are clearly more versatile than Aitken's iterated Δ^2 process from which they were derived since they are also able to accelerate logarithmic convergence. The sequence transformation $\mathscr{B}_k^{(n)}$ is in general less powerful than the other two transformations, and $\mathscr{J}_k^{(n)}$ is normally the most powerful transformation, being roughly comparable with Brezinski's ϑ algorithm.

11.2. New nonlinear sequence transformations obtained from linear transformations

In this section we want to construct new sequence transformations by modifying the recursive schemes of the linear transformations $\Lambda_k^{(n)}(\beta, s_n)$, eq. (7.3-20), $\mathscr{F}_k^{(n)}(\alpha, s_n)$, eq. (8.4-11), and $\mathscr{P}_k^{(n)}(\zeta, s_n)$, eq. (9.4-11), according to eqs. (11.1-1) and (11.1-2).

The recursive schemes (7.3-21), (8.4-12), and (9.4-12) for the linear sequence transformations $\Lambda_k^{(n)}(\beta, s_n)$, $\mathscr{F}_k^{(n)}(\alpha, s_n)$, and $\mathscr{P}_k^{(n)}(\zeta, s_n)$ are all of the form of eq. (11.1-7). Consequently, these recursive schemes will be modified according to eq. (11.1-8). Hence, in the case of the recursive scheme (7.3-21) for the sequence transformation $\Lambda_k^{(n)}(\beta, s_n)$ we obtain a new nonlinear sequence transformation $\lambda_k^{(n)}$ which is defined by the following recursive scheme:

$$\lambda_0^{(n)} = s_n, \qquad n \in \mathbb{N}_0, \tag{11.2-1a}$$

$$\lambda_{k+1}^{(n)} = \lambda_k^{(n+1)} - \frac{(\beta+n)[\Delta\lambda_k^{(n)}][\Delta\lambda_k^{(n+1)}]}{(\beta+n+1)[\Delta\lambda_k^{(n+1)}] - (\beta+n)[\Delta\lambda_k^{(n)}]}, \quad k, n \in \mathbb{N}_0. \tag{11.2-1b}$$

As usual, it is assumed here that the difference operator Δ acts only upon the superscript n and not upon the subscript k.

The nonlinear sequence transformation $\sigma_k^{(n)}$ is obtained by modifying the recursive scheme (8.4-12) for $\mathscr{F}_k^{(n)}(\alpha, s_n)$:

$$\sigma_0^{(n)} = s_n, \qquad n \in \mathbb{N}_0, \tag{11.2-2a}$$

$$\sigma_{k+1}^{(n)} = \sigma_k^{(n+1)} - \frac{(\alpha+n+k)[\Delta\sigma_k^{(n)}][\Delta\sigma_k^{(n+1)}]}{(\alpha+n+k+1)[\Delta\sigma_k^{(n+1)}] - (\alpha+n+k)[\Delta\sigma_k^{(n)}]}, \quad k,n \in \mathbb{N}_0. \tag{11.2-2b}$$

Finally, a modification of the recursive scheme (9.4-12) for $\mathscr{P}_k^{(n)}(\zeta, s_n)$ yields the nonlinear sequence transformation $\mu_k^{(n)}$:

$$\mu_0^{(n)} = s_n, \qquad n \in \mathbb{N}_0, \tag{11.2-3a}$$

$$\mu_{k+1}^{(n)} = \mu_k^{(n+1)} - \frac{(\zeta+n-k)[\Delta\mu_k^{(n)}][\Delta\mu_k^{(n+1)}]}{(\zeta+n-k+1)[\Delta\mu_k^{(n+1)}] - (\zeta+n-k)[\Delta\mu_k^{(n)}]}, \quad k, n \in \mathbb{N}_0. \tag{11.2-3b}$$

These new sequence transformations $\lambda_k^{(n)}$, $\sigma_k^{(n)}$, and $\mu_k^{(n)}$ are all weighted Δ^2 methods, i.e., modifications of Aitken's iterated Δ^2 process, eq. (5.1-15). In this context it may be of interest that Aitken's iterated Δ^2 process can also be derived by modifying a linear sequence transformation along the lines of Brezinski's ϑ algorithm. Let us consider the following recursive scheme:

$$\mathscr{I}_0^{(n)} = s_n, \qquad n \in \mathbb{N}_0, \tag{11.2-4a}$$

$$\mathscr{I}_{k+1}^{(n)} = \mathscr{I}_k^{(n)} + \Delta\mathscr{I}_k^{(n)}, \quad k, n \in \mathbb{N}_0. \tag{11.2-4b}$$

Obviously, the sequence transformation $\mathscr{I}_k^{(n)}$, which is defined by a recursive scheme of the type of eq. (11.1-3), can also be written in the following way:

$$\mathscr{I}_k^{(n)} = E^k s_n. \tag{11.2-5}$$

Here, E denotes the shift operator which is defined by eq. (2.2-4). If the recursive scheme (11.2-4) is changed according to eq. (11.1-4), we obtain Aitken's iterated Δ^2 process, eq. (5.1-15).

The similarity of $\lambda_k^{(n)}$, $\sigma_k^{(n)}$, and $\mu_k^{(n)}$ with Aitken's iterated Δ^2 process implies that these sequence transformations can be computed in the same way as Aitken's iterated Δ^2 process. For that purpose, it is recommendable to rewrite the recursive scheme (11.2-1) for $\lambda_k^{(n)}$ in the following way:

$$\lambda_0^{(n)} = s_n, \qquad n \geq 0, \tag{11.2-6a}$$

$$\lambda_j^{(n-2j)} = \lambda_{j-1}^{(n-2j+1)} - \frac{(\beta+n-2j)\left[\Delta\lambda_{j-1}^{(n-2j)}\right]\left[\Delta\lambda_{j-1}^{(n-2j+1)}\right]}{(\beta+n-2j+1)\left[\Delta\lambda_{j-1}^{(n-2j+1)}\right] - (\beta+n-2j)\left[\Delta\lambda_{j-1}^{(n-2j)}\right]},$$

$$n \geq 2, \qquad 1 \leq j \leq [\![n/2]\!]. \tag{11.2-6b}$$

The recursive schemes (11.2-2) and (11.2-3) should also be rewritten in the same way:

$$\sigma_0^{(n)} = s_n, \qquad n \geq 0, \tag{11.2-7a}$$

$$\sigma_j^{(n-2j)} = \sigma_{j-1}^{(n-2j+1)} - \frac{(\alpha+n-j)\left[\Delta\sigma_{j-1}^{(n-2j)}\right]\left[\Delta\sigma_{j-1}^{(n-2j+1)}\right]}{(\alpha+n-j+1)\left[\Delta\sigma_{j-1}^{(n-2j+1)}\right] - (\alpha+n-j)\left[\Delta\sigma_{j-1}^{(n-2j)}\right]},$$

$$n \geq 2, \qquad 1 \leq j \leq [\![n/2]\!]. \tag{11.2-7b}$$

$$\mu_0^{(n)} = s_n, \qquad n \geq 0, \tag{11.2-8a}$$

$$\mu_j^{(n-2j)} = \mu_{j-1}^{(n-2j+1)} - \frac{(\zeta+n-3j)\left[\Delta\mu_{j-1}^{(n-2j)}\right]\left[\Delta\mu_{j-1}^{(n-2j+1)}\right]}{(\zeta+n-3j+1)\left[\Delta\mu_{j-1}^{(n-2j+1)}\right] - (\zeta+n-3j)\left[\Delta\mu_{j-1}^{(n-2j)}\right]},$$

$$n \geq 2, \qquad 1 \leq j \leq [\![n/2]\!]. \tag{11.2-8b}$$

As in the case of Aitken's iterated Δ^2 process, the approximations to the limit of these transformations depend upon the index m of the last sequence element s_m which was used in these recursions. Let $\mathscr{X}_k^{(n)}$ now stand for any of the sequence transformations $\lambda_k^{(n)}$, $\sigma_k^{(n)}$, or $\mu_k^{(n)}$. Then, if m is even, $m = 2\mu$, our approximations to the limit of the sequence are the transformations

$$\{s_0, s_1, \ldots, s_{2\mu}\} \to \mathscr{X}_\mu^{(0)}, \tag{11.2-9}$$

and if m is odd, $m = 2\mu + 1$, our approximations to the limit will be

$$\{s_1, s_2, \ldots, s_{2\mu+1}\} \to \mathscr{X}_{\mu}^{(1)}. \tag{11.2-10}$$

These two relationships can be combined into a single equation yielding

$$\{s_{m-2[\![m/2]\!]}, s_{m-2[\![m/2]\!]+1}, \ldots, s_m\} \to \mathscr{X}_{[\![m/2]\!]}^{(m-2[\![m/2]\!])}. \tag{11.2-11}$$

Only 1-dimensional arrays $\tilde{\lambda}$, $\tilde{\sigma}$, and $\tilde{\mu}$ will be needed for the computation of the sequence transformations $\lambda_k^{(n)}$, $\sigma_k^{(n)}$, and $\mu_k^{(n)}$ if the elements of the tables of these transformations are stored in the same way as the elements of the Aitken table according to eq. (5.2-6), i.e.,

$$\mathscr{X}_{[\![\nu/2]\!]}^{(n-\nu)} \to X(n-\nu), \qquad n \geq 0, \quad 0 \leq \nu \leq n. \tag{11.2-12}$$

Here, $\mathscr{X}_k^{(n)}$ stands for any of the three sequence transformations $\lambda_k^{(n)}$, $\sigma_k^{(n)}$, and $\mu_k^{(n)}$, and X stands for the corresponding 1-dimensional array, i.e., either for $\tilde{\lambda}$, $\tilde{\sigma}$, or for $\tilde{\mu}$. With this convention, the recursive scheme (11.2-6) for $\lambda_k^{(n)}$ can be reformulated in terms of the elements of the 1-dimensional array $\tilde{\lambda}$:

$$\tilde{\lambda}(n) = s_n, \qquad n \geq 0, \tag{11.2-13a}$$

$$\tilde{\lambda}(m) = \tilde{\lambda}(m+1) - \frac{(\beta+m)[\Delta\tilde{\lambda}(m)][\Delta\tilde{\lambda}(m+1)]}{(\beta+m+1)[\Delta\tilde{\lambda}(m+1)] - (\beta+m)[\Delta\tilde{\lambda}(m)]},$$

$$m = n - 2j, \qquad n \geq 2, \qquad 1 \leq j \leq [\![n/2]\!]. \tag{11.2-13b}$$

With the help of convention (11.2-12) the recursive schemes (11.2-7) for $\sigma_k^{(n)}$ and (11.2-8) for $\mu_k^{(n)}$ can also be reformulated in terms of the elements of the arrays $\tilde{\sigma}$ and $\tilde{\mu}$ yielding similiar expressions.

It follows from their recursive schemes that $\lambda_k^{(n)}$, eq. (11.2-1), $\sigma_k^{(n)}$, eq. (11.2-2), and $\mu_k^{(n)}$, eq. (11.2-3), are all transformations of order $l = 2k$. In that respect, they are equivalent to $\mathscr{A}_k^{(n)}$, eq. (5.1-15), and $\epsilon_{2k}^{(n)}$, eq. (4.2-1), which need the same set $s_n, \ldots, s_{n+2k}$ of sequence elements for their computation.

We shall see later that the nonlinear sequence transformations $\lambda_k^{(n)}$, $\sigma_k^{(n)}$, and $\mu_k^{(n)}$ are more versatile than the linear sequence transformations $\Lambda_k^{(n)}(\beta, s_n)$, eq. (7.3-20), $\mathscr{F}_k^{(n)}(\alpha, s_n)$, eq. (8.4-11), and $\mathscr{P}_k^{(n)}(\zeta, s_n)$, eq. (9.4-12), from which they were derived, since they are not only able to accelerate logarithmic convergence but also linear convergence. In addition, they can also sum many divergent series. Numerical tests showed that $\lambda_k^{(n)}$, eq. (11.2-1), is normally a more powerful sequence transformation than $\sigma_k^{(n)}$, eq. (11.2-2), or $\mu_k^{(n)}$, eq. (11.2-3).

12. A theoretical analysis of sequence transformations

12.1. Germain-Bonne's formal theory of convergence acceleration

The properties of linear and nonlinear sequence transformations are in some sense complementary. Theoretically, linear sequence transformations are now very well understood (see for instance refs. [4] and [7–11]) but their power as well as their practical usefulness is very limited. Nonlinear sequence transformations are often able to achieve spectacular results, but theoretically, only relatively little is known.

Any theory of nonlinear sequence transformations has to say something about the two fundamental questions which arise in connection with acceleration of convergence. Firstly, is the transformation under consideration regular, i.e., will the transformed sequence $\{s'_n\}$ converge to the same limit as the original sequence $\{s_n\}$. Secondly, will the transformed sequence converge more rapidly than the original sequence.

The first attempt to develop a general theory of the regularity and the acceleration properties of nonlinear sequence transformations is due to Germain-Bonne [33] who considered sequence transformations G_k with $k \in \mathbb{N}_0$ that are functions defined on vectors $\boldsymbol{x} = (x_1, x_2, \ldots, x_{k+1})$, i.e., functions of the type G_k: $\mathbb{R}^{k+1} \to \mathbb{R}$. Germain-Bonne postulated that these sequence transformations G_k possess some very general properties such as continuity, homogeneity and translativity. On the basis of these postulates Germain-Bonne could formulate some conditions which guarantee the regularity of such a sequence transformation. In addition, Germain-Bonne succeeded in formulating a general criterion which decides whether a sequence transformation G_k accelerates linear convergence or not. A good treatment of Germain-Bonne's formal theory of convergence acceleration [33] can also be found in Wimp's book (see pp. 101–105 of ref. [23]).

The applicability of Germain-Bonne's theory in its original version is quite limited. The reason is that Germain-Bonne treats sequence transformations which depend upon n only implicitly via the $k+1$ sequence elements $s_n, s_{n+1}, \ldots, s_{n+k}$, on which they act, but not explicitly. Consequently, Germain-Bonne's theory is limited to sequence transformations as for instance Wynn's ϵ algorithm, eq. (4.2-1), Aitken's iterated Δ^2 process, eq. (5.1-15), or Brezinski's ϑ algorithm, eq. (10.1-9), which are all defined by recursive schemes that do not depend explicitly upon n. It cannot be applied in the case of a sequence transformation such as $\lambda_k^{(n)}$, eq. (11.2-1), although it is a close relative of Aitken's iterated Δ^2 process, since its recursive scheme depends explicitly on n. Consequently, in this section Germain-Bonne's theory will be modified in such a way that it can be applied to the sequence transformations of this report which mostly depend explicitly upon n.

Let us therefore assume that for fixed $k \in \mathbb{N}_0$ a sequence transformation $G_k^{(n)}$ is a function which is defined on vectors $\boldsymbol{x} = (x_1, x_2, \ldots, x_{k+1}) \in \mathbb{R}^{k+1}$ and which may depend explicitly on $n \in \mathbb{N}_0$. In addition, we assume that such a sequence transformation $G_k^{(n)}$: $\mathbb{R}^{k+1} \to \mathbb{R}$ possesses for fixed $k \in \mathbb{N}_0$ and for all $n \in \mathbb{N}_0$ the following properties:

(H-0): *$G_k^{(n)}$ is defined and continuous on a subset $\boldsymbol{X}^{(n)}$ of $\mathbb{R}^{k+1}$.*

(H-1): *$G_k^{(n)}$ is a homogeneous function of degree one. This means that $G_k^{(n)}$ satisfies for arbitrary vectors $\boldsymbol{x} \in \boldsymbol{X}^{(n)}$ and for all $\lambda \in \mathbb{R}$ such that $\lambda \boldsymbol{x}$ is still an element of $\boldsymbol{X}^{(n)}$*

$$G_k^{(n)}(\lambda x_1, \lambda x_2, \ldots, \lambda x_{k+1}) = \lambda G_k^{(n)}(x_1, x_2, \ldots, x_{k+1}). \tag{12.1-1}$$

(H-2): *$G_k^{(n)}$ is invariant under translation in the sense of eq.* (3.1-4). *Consequently, for arbitrary $t \in \mathbb{R}$ and for arbitrary vectors $\boldsymbol{x} \in \boldsymbol{X}^{(n)}$ we have*

$$G_k^{(n)}(x_1 + t, x_2 + t, \ldots, x_{k+1} + t) = G_k^{(n)}(x_1, x_2, \ldots, x_{k+1}) + t. \tag{12.1-2}$$

(H-3): *A subset $\boldsymbol{X}^{(\infty)}$ of $\mathbb{R}^{k+1}$ exists such that for every vector $\boldsymbol{x} = (x_1, x_2, \ldots, x_{k+1})$ belonging to this subset the limiting transformation*

$$G_k^{(\infty)}(x_1, x_2, \ldots, x_{k+1}) = \lim_{n\to\infty} G_k^{(n)}(x_1, x_2, \ldots, x_{k+1}) \tag{12.1-3}$$

is uniquely defined and continuous. In addition, it is assumed that the limiting transformation $G_k^{(\infty)}$ is also homogeneous and invariant under translation according to (H-1) *and* (H-2).

The first three postulates (H-0)–(H-2) are essentially identical with the analogous postulates made by Germain-Bonne [33]. The main difference is that here a sequence transformation $G_k^{(n)}$ may depend explicitly upon n and not only implicitly via the $k+1$ sequence elements s_n, s_{n+1}, $\ldots$, s_{n+k} on which it acts. Some of Germain-Bonne's results are based on the limiting behaviour of a sequence transformation $G_k(s_n, s_{n+1}, \ldots, s_{n+k})$ as $n \to \infty$. If we want to formulate analogous results for a sequence transformation $G_k^{(n)}$, which may depend explicitly on n, we have to require that the limit $n \to \infty$ can be performed in the expression defining $G_k^{(n)}$ and that a unique limiting transformation $G_k^{(\infty)}$ exists which is defined and continuous on a suitable subset $\boldsymbol{X}^{(\infty)}$ of $\mathbb{R}^{k+1}$ and which is also homogeneous and translative according to (H-1) and (H-2).

Conditions for the existence of the limiting transformation $G_k^{(\infty)}$ were discussed by Smith and Ford (see p. 226 of ref. [29]). Their analysis was based upon that version of the Moore-Smith theorem [93] which can be found in Gleason's book (see p. 256 of ref. [94]). However, for our purposes it is probably simpler to postulate the validity of (H-3), since in all cases, in which we shall have to do such a limit $n \to \infty$, the existence of a limiting transformation $G_k^{(\infty)}$ with the required properties will always be quite obvious.

In (H-1) the restriction, that $\lambda \boldsymbol{x}$ has to be an element of $\boldsymbol{X}^{(n)}$, is necessary. The reason is that nonlinear sequence transformations are frequently not defined for constant sequences. In such a case, $\lambda = 0$ has to be excluded because $G_k^{(n)}$ would not be defined for the vector $\boldsymbol{x} = (0, 0, \ldots, 0)$.

Concerning (H-3) it should be noted that $G_k^{(n)}$ with n being finite and its limiting transformation $G_k^{(\infty)}$ are not necessarily defined and continuous on the same subset of $\mathbb{R}^{k+1}$, i.e., in general $\boldsymbol{X}^{(n)} \neq \boldsymbol{X}^{(\infty)}$. For instance, if in the recursive scheme (11.2-1) for $\lambda_k^{(n)}$ the limit $n \to \infty$ is performed according to eq. (12.1-3), we obtain the second version of Aitken's iterated Δ^2 process, eq. (11.1-6). If we compare the subsets of $\mathbb{R}^3$, for which the explicit expressions for $\mathscr{A}_1^{(n)}$ and $\lambda_1^{(n)}$ are continuous, we find that Aitken's Δ^2 process is defined for vectors $\boldsymbol{x} = (x_1, x_2, x_3)$ which satisfy $x_1 - 2x_2 + x_3 \neq 0$. A different restriction is necessary in the case of $\lambda_1^{(n)}$ as long as n is finite.

It was remarked above that a given sequence transformation $G_k^{(n)}$ will normally not be continuous for arbitrary vectors $\boldsymbol{x} \in \mathbb{R}^{k+1}$. If, however, a sequence transformation $G_k^{(n)}$ is defined and continuous for all vectors $\boldsymbol{x} \in \mathbb{R}^{k+1}$, then the regularity of this transformation can be proved quite easily.

Theorem 12-1: *If a sequence transformation $G_k^{(n)}$ is a continuous function on $\mathbb{R}^{k+1}$ for all $n \in \mathbb{N}_0$ and if its limiting transformation $G_k^{(\infty)}$ is also continuous on $\mathbb{R}^{k+1}$ and satisfies* (H-1) *and* (H-2), *then $G_k^{(n)}$ is regular, i.e., it preserves the limit s of every convergent sequence $\{\!\{s_n\}\!\}$.*

Proof: Since $G_k^{(n)}$ is continuous on $\mathbb{R}^{k+1}$ for all $n \in \mathbb{N}_0$ and since $\{s_n\}$ converges to some limit s, we have by continuity

$$\lim_{n \to \infty} G_k^{(n)}(s_n, s_{n+1}, \ldots, s_{n+k}) = G_k^{(\infty)}(s, s, \ldots, s). \tag{12.1-4}$$

Since (H-2) remains valid as $n \to \infty$, we may conclude that for all constant sequences $s, s, s, \ldots$ we have

$$G_k^{(\infty)}(s, s, \ldots, s) = s + G_k^{(\infty)}(0, 0, \ldots, 0). \tag{12.1-5}$$

In the same way, (H-1) remains valid as $n \to \infty$. Thus, we may conclude

$$G_k^{(\infty)}(0, 0, \ldots, 0) = 0. \tag{12.1-6}$$

Hence, it follows from eqs. (12.1-4)–(12.1-6) that a sequence transformation $G_k^{(n)}$, which satisfies the above assumptions, is regular.

Unfortunately, theorem 12-1 will be of little use in the case of nonlinear sequence transformations which are in general nonregular. A nonlinear sequence transformation $G_k^{(n)}$ is normally a rational function of the $k+1$ sequence elements $s_n, s_{n+1}, \ldots, s_{n+k}$ which are used for its computation. Since rational functions have poles, we cannot expect that nonlinear sequence transformations will be continuous on $\mathbb{R}^{k+1}$. Consequently, in the case of a nonlinear sequence transformation the convergence of an arbitrary sequence $\{s_n\}$ to some limit s does not imply that the transformed sequence converges at all, let alone to the same limit. In addition, eqs. (12.1-5) and (12.1-6) need not be valid since nonlinear sequence transformations are not necessarily defined for constant sequences.

Theorem 12-2: *Let $\mathbb{D}^{k+1}$ be the set of vectors $\boldsymbol{x} \in \mathbb{R}^{k+1}$ with distinct components. Every sequence transformation $G_k^{(n)}$, which is a continuous function for all vectors $\boldsymbol{x} \in \mathbb{D}^{k+1}$ and which also satisfies (H-1) and (H-2), can be expressed in the following way:*

$$G_k^{(n)}(x_1, x_2, \ldots, x_{k+1}) = x_1 + (x_2 - x_1) g_k^{(n)}\left(\frac{x_3 - x_2}{x_2 - x_1}, \ldots, \frac{x_{k+1} - x_k}{x_k - x_{k-1}} \right). \tag{12.1-7}$$

The associated transformation $g_k^{(n)}$ which is defined and continuous on a subset of $\mathbb{F}^{k-1}$, the set of vectors $\boldsymbol{y} \in \mathbb{R}^{k-1}$ with nonzero components, is given by

$$g_k^{(n)}\left(\frac{x_3 - x_2}{x_2 - x_1}, \ldots, \frac{x_{k+1} - x_k}{x_k - x_{k-1}} \right) = G_k^{(n)}\left(0, 1, 1 + \frac{x_3 - x_2}{x_2 - x_1}, \ldots, \sum_{j=0}^{k-1} \prod_{i=0}^{j-1} \frac{x_{i+3} - x_{i+2}}{x_{i+2} - x_{i+1}} \right). \tag{12.1-8}$$

Proof: According to (H-2) we can subtract x_1 from $G_k^{(n)}$, and according to (H-1) we can divide the $k+1$ arguments of $G_k^{(n)}$ by $x_2 - x_1$. This yields:

$$G_k^{(n)}(x_1, x_2, \ldots, x_{k+1}) = x_1 + (x_2 - x_1) G_k^{(n)}\left(0, 1, \frac{x_3 - x_1}{x_2 - x_1}, \ldots, \frac{x_{k+1} - x_1}{x_2 - x_1} \right). \tag{12.1-9}$$

We now need the following relationship which can be proved quite easily:

$$\frac{x_{\nu+1} - x_1}{x_2 - x_1} = \sum_{j=0}^{\nu-1} \prod_{i=0}^{j-1} \frac{x_{i+3} - x_{i+2}}{x_{i+2} - x_{i+1}}, \qquad \nu \in \mathbb{N}_0. \tag{12.1-10}$$

If we insert eq. (12.1-10) into eq. (12.1-9) we obtain eqs. (12.1-7) and (12.1-8). The continuity of $g_k^{(n)}$ on the subset of vectors $\boldsymbol{y} = (y_1, y_2, \ldots, y_{k-1}) \in \mathbb{F}^{k-1}$, which are generated from vectors $\boldsymbol{x} = (x_1, x_2, \ldots, x_{k+1}) \in \mathbb{D}^{k+1}$ via the relationship

$$y_\mu = (x_{\mu+2} - x_{\mu+1})/(x_{\mu+1} - x_\mu), \qquad 1 \leq \mu \leq k-1, \tag{12.1-11}$$

follows via eq. (12.1-7) from the continuity of $G_k^{(n)}$ on $\mathbb{D}^{k+1}$. However, an associated transformation $g_k^{(n)}$ in the sense of eqs. (12.1-7) and (12.1-8) is not defined for all vectors $\boldsymbol{y} \in \mathbb{F}^{k-1}$. For instance, if we choose $\boldsymbol{y} = (1, -1, 1, -1, \ldots)$, which clearly belongs to $\mathbb{F}^{k-1}$, we are not able to find a vector $\boldsymbol{x} = (x_1, x_2, \ldots, x_{k+1}) \in \mathbb{D}^{k+1}$ which satisfies eq. (12.1-11).

With the help of theorem 12-2 it can be analyzed for which convergent sequences $\{\!\{s_n\}\!\}$ a sequence transformation $G_k^{(n)}$ satisfying (H-0)–(H-3) will be regular.

Theorem 12-3: *Let $G_k^{(n)}$ be a sequence transformation which satisfies theorem* 12-2. *This means that according to eqs.* (12.1-7) *and* (12.1-8) *an associated transformation $g_k^{(n)}$ exists which is defined and continuous on that subset of vectors $\boldsymbol{y} = (y_1, y_2, \ldots, y_{k-1}) \in \mathbb{F}^{k-1}$ which is generated by all vectors $\boldsymbol{x} = (x_1, x_2, \ldots, x_{k+1}) \in \mathbb{D}^{k+1}$ according to eq.* (12.1-11). *If the limiting associated transformation, which is defined by*

$$g_k^{(\infty)}(y_1, y_2, \ldots, y_{k-1}) = \lim_{n \to \infty} g_k^{(n)}(y_1, y_2, \ldots, y_{k-1}), \tag{12.1-12}$$

is also defined and continuous on the same subset of $\mathbb{F}^{k-1}$ as $g_k^{(n)}$, then $G_k^{(n)}$ will preserve the limit s of every convergent sequence $\{\!\{s_n\}\!\}$ having the following properties:

(*i*): *For sufficiently large values of $n \in \mathbb{N}_0$ the sequence elements s_n are all distinct.*

(*ii*): *For sufficiently large values of $n \in \mathbb{N}_0$ the ratios $\Delta s_{n+1}/\Delta s_n$ all satisfy the inequality*

$$c \leq |\Delta s_{n+1}/\Delta s_n| \leq c', \qquad 0 < c < c' < \infty. \tag{12.1-13}$$

Proof: It follows from (i) that $\Delta s_n \neq 0$ for sufficiently large values of n. Consequently, $G_k^{(n)}$ can be rewritten for sufficiently large values of n according to eq. (12.1-7) yielding

$$G_k^{(n)}(s_n, s_{n+1}, \ldots, s_{n+k}) = s_n + \Delta s_n g_k^{(n)}\left(\frac{\Delta s_{n+1}}{\Delta s_n}, \ldots, \frac{\Delta s_{n+k-1}}{\Delta s_{n+k-2}}\right). \tag{12.1-14}$$

According to eq. (12.1-8) the associated transformation is given by

$$g_k^{(n)}\left(\frac{\Delta s_{n+1}}{\Delta s_n}, \ldots, \frac{\Delta s_{n+k-1}}{\Delta s_{n+k-2}}\right) = G_k^{(n)}\left(0, 1, 1 + \frac{\Delta s_{n+1}}{\Delta s_n}, \ldots, \sum_{j=0}^{k-1} \prod_{i=0}^{j-1} \frac{\Delta s_{n+i+1}}{\Delta s_{n+i}}\right). \tag{12.1-15}$$

$G_k^{(n)}$ will preserve the limit of $\{\!\{s_n\}\!\}$ if the second term on the right-hand side of eq. (12.1-14) vanishes as $n \to \infty$. Since $\{\!\{s_n\}\!\}$ converges, Δs_n vanishes as $n \to \infty$ and we only have to show that the associated transformation $g_k^{(n)}$ remains bounded as $n \to \infty$. The limiting associated transformation $g_k^{(\infty)}$ is by assumption defined and continuous on that subset of $\mathbb{F}^{k-1}$ which is generated from vectors $\boldsymbol{x} \in \mathbb{D}^{k+1}$ according to eq. (12.1-11). Continuity of a given function on a certain subset implies the boundedness of this function for all bounded elements belonging to this subset. It follows from (i) that for sufficiently large values of n every string $s_n, s_{n+1}, \ldots, s_{n+k}$ belongs to $\mathbb{D}^{k+1}$, and it follows from (ii) that for sufficiently large values of n the $k-1$ arguments of $g_k^{(n)}$ in eq. (12.1-15) are all nonzero and bounded. This implies that $g_k^{(n)}$ will remain bounded as $n \to \infty$. This concludes the proof of theorem 12-3.

Next, a criterion will be formulated which decides whether a sequence transformation $G_k^{(n)}$, which may depend explicitly upon n, is able to accelerate linear convergence or not. This is probably the most important result of our adaptation of Germain-Bonne's formal theory of convergence acceleration [33].

Theorem 12-4: *Let us assume that a sequence $\{\!\{s_n\}\!\}$ converges linearly to some limit s, i.e.,*

$$\lim_{n\to\infty} \frac{s_{n+1}-s}{s_n - s} = \rho, \qquad 0 < |\rho| < 1. \tag{12.1-16}$$

Then, a necessary and sufficient condition that a sequence transformation $G_k^{(n)}$: $\mathbb{R}^{k+1} \to \mathbb{R}$ accelerates the convergence of $\{\!\{s_n\}\!\}$ is that its associated transformation $g_k^{(n)}$ satisfies:

$$\lim_{n\to\infty} g_k^{(n)}(\rho_n, \rho_{n+1}, \ldots, \rho_{n+k-2}) = g_k^{(\infty)}(\rho, \rho, \ldots, \rho) = \frac{1}{1-\rho}. \tag{12.1-17}$$

Here, $\{\!\{\rho_n\}\!\}$ is an arbitrary sequence which converges to ρ. The above statement can also be formulated in terms of the limiting transformation $G_k^{(\infty)}$:

$$\lim_{n\to\infty} G_k^{(n)}\left(0, 1, 1+\rho_n, \ldots, \sum_{j=0}^{k-1} \prod_{i=0}^{j-1} \rho_{n+i}\right) = G_k^{(\infty)}\left(0, 1, 1+\rho, \ldots, \sum_{j=0}^{k-1} \rho^j\right) = \frac{1}{1-\rho},$$
$$0 < |\rho| < 1. \tag{12.1-18}$$

Proof: Since $\{\!\{s_n\}\!\}$ converges linearly, it follows from eqs. (2.6-3) and (2.6-4) that $\Delta s_n \sim \rho^n$ as $n \to \infty$. Thus, $G_k^{(n)}$ can for sufficiently large values of n be rewritten according to eq. (12.1-14). If we subtract s from both sides of eq. (12.1-14) and divide the resulting expression by $s_n - s$ we obtain:

$$\frac{G_k^{(n)}(s_n, s_{n+1}, \ldots, s_{n+k}) - s}{s_n - s} = 1 + \frac{\Delta s_n}{s_n - s} g_k^{(n)}\left(\frac{\Delta s_{n+1}}{\Delta s_n}, \ldots, \frac{\Delta s_{n+k-1}}{\Delta s_{n+k-2}}\right). \tag{12.1-19}$$

According to eq. (2.6-6) the convergence of a sequence $\{s_n\}$ to its limit s is accelerated by a sequence transformation $G_k^{(n)}$ if

$$\lim_{n\to\infty} \frac{G_k^{(n)}(s_n, s_{n+1}, \ldots, s_{n+k}) - s}{s_n - s} = 0. \tag{12.1-20}$$

Hence, if we perform the limit $n \to \infty$ in eq. (12.1-19), the left-hand side has to vanish if $G_k^{(n)}$ accelerates convergence. Thus, we only have to investigate under which conditions the right-hand side of eq. (12.1-19) also vanishes as $n \to \infty$. With the help of eq. (12.1-16) we find:

$$\lim_{n\to\infty} \frac{\Delta s_n}{s_n - s} = \lim_{n\to\infty} \frac{s_{n+1} - s}{s_n - s} - 1 = \rho - 1. \tag{12.1-21}$$

Next, we observe that because of the equivalence of eqs. (2.6-3) and (2.6-4) for sequences satisfying eq. (12.1-16) the arguments

$$\rho_n = \Delta s_{n+1}/\Delta s_n, \qquad n \in \mathbb{N}_0, \tag{12.1-22}$$

of the associated transformation $g_k^{(n)}$ in eq. (12.1-19) converge to ρ as $n \to \infty$. Now, the right-hand side of eq. (12.1-19) can only vanish as $n \to \infty$ if the limiting associated transformation $g_k^{(\infty)}$ satisfies eq. (12.1-17). In the same way, if eq. (12.1-17) is satisfied, then because of eq. (12.1-21) the right-hand side of eq. (12.1-19) vanishes. Consequently, the validity of eq. (12.1-17) is equivalent to the statement that $G_k^{(n)}$ accelerates linear convergence.

The sequence 0, 1, $1 + \rho$, ..., on which the limiting transformation $G_k^{(\infty)}$ acts in eq. (12.1-18), is apart from its first element and apart from a shift of the indices identical with the sequence of partial sums of the geometric series, eq. (2.6-2). This is best seen by rewriting the elements of this sequence in the following way:

$$\sigma_n(\rho) = \frac{1 - \rho^n}{1 - \rho} = \sum_{\nu=0}^{n-1} \rho^\nu, \qquad 0 < |\rho| < 1, \qquad n \in \mathbb{N}_0. \tag{12.1-23}$$

Obviously, this shifted sequence also converges to $1/(1 - \rho)$ as $n \to \infty$. Hence, we see that theorem 12-4, which decides whether a given sequence transformation $G_k^{(n)}$ accelerates linear convergence or not, requires that the limiting transformation $G_k^{(\infty)}$ is exact for the shifted sequence $\{\sigma_n(\rho)\}$ of partial sums of the geometric series. This fact indicates that there is a close connection between the exactness of a sequence transformation for the partial sums of the geometric series and its ability of accelerating linear convergence.

12.2. Applications of Germain-Bonne's theory

In this section, the properties of certain sequence transformations will be analyzed with the help of Germain-Bonne's formal theory of convergence acceleration. The most interesting feature of Germain-Bonne's theory is its treatment of the acceleration of linear convergence. Consequently, theorem 12-4 and related questions such as the exactness of a sequence transformation for the geometric series will be emphasized in this section.

First, we want to investigate the sequence transformations $\lambda_k^{(n)}$, eq. (11.2-1), $\sigma_k^{(n)}$, eq. (11.2-2), and $\mu_k^{(n)}$, eq. (11.2-3), which are close relatives of Aitken's iterated Δ^2 process, eq. (5.1-15). However, it follows from their recursive schemes (11.2-1)–(11.2-3) that unlike Aitken's Δ^2 process $\mathscr{A}_1^{(n)}$, eq. (5.1-4), which is by construction exact for the geometric series, $\lambda_1^{(n)}$, $\sigma_1^{(n)}$, and $\mu_1^{(n)}$ are not exact for the geometric series. Consequently, it is not obvious whether $\lambda_1^{(n)}$, $\sigma_1^{(n)}$, and $\mu_1^{(n)}$, accelerate linear convergence.

Theorem 12-5: *The sequence transformations* $\lambda_k^{(n)}$, *eq.* (11.2-1), $\sigma_k^{(n)}$, *eq.* (11.2-2), *and* $\mu_k^{(n)}$, *eq.* (11.2-3), *accelerate linear convergence if and only if Aitken's iterated* Δ^2 *process* $\mathscr{A}_k^{(n)}$, *eq.* (5.1-15), *accelerates linear convergence.*

Proof: According to theorem 12-4 a sequence transformation $G_k^{(n)}$ accelerates linear convergence if it possesses a limiting transformation $G_k^{(\infty)}$ which satisfies eq. (12.1-18). If we perform the limit $n \to \infty$ in the recursive scheme (11.2-1) for $\lambda_k^{(n)}$ according to eq. (12.1-3), we find that the limiting transformation of $\lambda_k^{(n)}$ is defined by a recursive scheme which is identical with the second version of Aitken's iterated Δ^2 process, eq. (11.1-6).

Hence, it follows from eq. (12.1-18) that $\lambda_k^{(n)}$ accelerates linear convergence if the elements of the sequence $\{\!\{\sigma_n(\rho)\}\!\}$, eq. (12.1-23) are transformed into $1/(1-\rho)$ by the recursive scheme (11.1-6).

However, since Aitken's iterated Δ^2 process does not depend explicitly upon n, this is at the same time the condition which determines whether Aitken's iterated Δ^2 process is able to accelerate linear convergence.

If we perform the limit $n \to \infty$ in the recursive schemes (11.2-2) for $\sigma_k^{(n)}$ and (11.2-3) for $\mu_k^{(n)}$ in the sense of eq. (12.1-3), we find that they have the same limiting transformation which is again defined by the recursive scheme for Aitken's iterated Δ^2 process, eq. (11.1-6). This concludes the proof of theorem 12-5.

Aitken's iterated Δ^2 process is defined by a recursive scheme and no explicit expression is known. Consequently, a general proof, that $\mathscr{A}_k^{(n)}$ accelerates linear convergence for all $k \geq 1$, does not seem possible. Instead, one can only construct an explicit rational expression for some special $\mathscr{A}_k^{(n)}$ with k, n being fixed integers. It can then be checked whether this explicit expression produces $1/(1-\rho)$ if it is applied to the elements of the sequence $\{\!\{\sigma_n(\rho)\}\!\}$, eq. (12.1-23). Unfortunately, these rational expressions become very complicated for larger values of the subscript k. Therefore, it is recommendable to consider only the simplest case.

Theorem 12-6: *Aitken's* Δ^2 *process* $\mathscr{A}_1^{(n)}$, *eq.* (5.1-4), *accelerates linear convergence for all* $n \in \mathbb{N}_0$.

Proof: According to theorem 12-4 we have to show that Aitken's Δ^2 process $\mathscr{A}_1^{(0)}$, eq. (5.1-4), produces $1/(1-\rho)$ if it acts upon the sequence elements $\sigma_0(\rho)$, $\sigma_1(\rho)$, and $\sigma_2(\rho)$ which are defined by eq. (12.1-23). Straightforward computation shows:

$$\mathscr{A}_1^{(0)} = \sigma_0(\rho) - \frac{[\sigma_1(\rho) - \sigma_0(\rho)]^2}{\sigma_2(\rho) - 2\sigma_1(\rho) + \sigma_0(\rho)} = \frac{1}{1-\rho}. \tag{12.2-1}$$

It follows from theorems 12-5 and 12-6 that $\lambda_1^{(n)}$, $\sigma_1^{(n)}$, and $\mu_1^{(n)}$ also accelerate linear convergence for all $n \in \mathbb{N}_0$.

Here, it must be emphasized that theorem 12-6 does not imply that $\mathscr{A}_k^{(n)}$ with $k > 1$ will also accelerate linear convergence. This has to be checked separately and independently for every $k > 1$. For instance, it follows from eq. (5.1-5) and theorem 12-6 that $\epsilon_2^{(n)}$ also accelerates linear convergence. However, in Wimp's book it is shown that in the case $\epsilon_4^{(n)}$ the limited associated transformation according to eqs. (12.1-7) and (12.1-8) cannot satisfy eq. (12.1-17) since it is unbounded in the vicinty of any $(\rho, \rho, \rho) \in \mathbb{R}^3$ (see pp. 127–128 of ref. [23]).

Next, the sequence transformations $\mathscr{J}_k^{(n)}$, eq. (10.3-6), $\mathscr{B}_k^{(n)}$, eq. (11.1-5), and $\mathscr{C}_k^{(n)}$, eq. (11.1-12), will be analyzed. Again, no explicit expressions are known for these transformations which are all defined by recursive schemes. Hence, as in the case of Aitken's iterated Δ^2 process, eq. (5.1-15), only special cases can be considered and it is again recommendable to consider only the simplest case.

A simple calculation shows that $\mathscr{J}_1^{(n)}$, $\mathscr{B}_1^{(n)}$, and $\mathscr{C}_1^{(n)}$ are exact for the partial sums (2.6-2) of the geometric series. Consequently, it is not surprising that these transformations accelerate linear convergence.

Theorem 12-7: *The sequence transformations* $\mathscr{J}_1^{(n)}$, *eq.* (10.3-6), $\mathscr{B}_1^{(n)}$, *eq.* (11.1-5), *and* $\mathscr{C}_1^{(n)}$, *eq.* (11.1-12), *accelerate linear convergence for all* $n \in \mathbb{N}_0$.

Proof: The recursive schemes, which define these sequence transformations, do not depend explicitly on n. Consequently, one only has to show that the explicit expressions for $\mathscr{J}_1^{(0)}$, $\mathscr{B}_1^{(0)}$, and $\mathscr{B}_1^{(0)}$ produce $1/(1-\rho)$ if they are applied to the first four elements of the sequence $\{\sigma_n(\rho)\}$, eq. (12.1-23). Straightforward computation shows that this and consequently theorem 12-7 is indeed true.

Since $\mathscr{J}_1^{(n)}$ is identical with $\vartheta_2^{(n)}$, theorem 12-7 also implies that $\vartheta_2^{(n)}$ accelerates linear convergence.

Next, we shall analyze those variants of the sequence transformations $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-6), and $\mathscr{D}_k^{(n)}(s_n, \omega_n)$, eq. (9.5-4), which are based upon Levin's [28] explicit remainder estimates (7.3-4), (7.3-6), and (7.3-10), as well as Smith and Ford's [29] modified remainder estimate (7.3-8). First, it will be shown that these sequence transformations are exact for the geometric series. For that purpose it is recommendable to modify a theorem, which was originally used by Smith and Ford (see p. 227 of ref. [29]) to prove that $u_k^{(n)}(\beta, s_n)$, eq. (7.3-5), $d_k^{(n)}(\beta, s_n)$, eq. (7.3-9), and $v_k^{(n)}(\beta, s_n)$, eq. (7.3-11), are exact for the geometric series, in such a way that it can also be applied in the case of the analogous variants of $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, and $\mathscr{D}_k^{(n)}(s_n, \omega_n)$.

Theorem 12-8: *Assume that a sequence transformation* $T_k^{(n)}(s_n, \omega_n)$ *is defined in the following way*:

$$T_k^{(n)}(s_n, \omega_n) = \frac{\Delta^k [P_{k-1}(n) s_n/\omega_n]}{\Delta^k [P_{k-1}(n)/\omega_n]}, \qquad k,\, n \in \mathbb{N}_0. \tag{12.2-2}$$

Here, $P_{k-1}(n)$ *is a polynomial of degree* $\leq k-1$ *in* n. *The sequence transformation* $T_k^{(n)}$ *is defined if the elements of* $\{s_n\}$ *are bounded in magnitude and if the sequence* $\{\omega_n\}$ *of remainder estimates is chosen in such a way that the denominator in eq.* (12.2-2) *does not vanish.*

Let us assume that the sequence $\{s_n\}$ *converges to some limit* s. *Then, for* $k \geq 1$ *and for* $n \geq 0$ *this sequence transformation* $T_k^{(n)}(s_n, \omega_n)$ *is exact for the sequence* $\{s_n\}$ *if the sequence* $\{\omega_n\}$ *of remainder estimate can be chosen in such a way that the denominator in eq.* (12.2-2) *does not vanish and that* ω_n *is proportional to* $s_n - s$, *i.e.*,

$$\omega_n = c(s_n - s), \qquad c \neq 0, \quad n \in \mathbb{N}_0. \tag{12.2-3}$$

Proof: Since this sequence transformation $T_k^{(n)}(s_n, \omega_n)$ is obviously invariant under translation in the sense of eq. (3.1-4), we can write

$$T_k^{(n)}(s_n, \omega_n) = s + \frac{\Delta^k[P_{k-1}(n)(s_n - s)/\omega_n]}{\Delta^k[P_{k-1}(n)/\omega_n]}, \qquad k, n \in \mathbb{N}_0. \tag{12.2-4}$$

If ω_n is proportional to $s_n - s$ according to eq. (12.2-3), the difference operator Δ^k in the numerator on the right-hand side acts only on $P_{k-1}(n)$ which is a annihilated because it is a polynomial of degree $\leq k-1$ in n. Since the denominator on the right-hand side of eq. (12.2-4) does not vanish by assumption, we see that $T_k^{(n)}$ is exact for $\{s_n\}$.

It is immediately obvious that the sequence transformations $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-6), and $\mathscr{D}_k^{(n)}(s_n, \omega_n)$, eq. (9.5-4), are all of the form of eq. (12.2-2). Hence, according to theorem 12-8 these sequence transformations are exact for the partial sums (2.6-2) of the geometric series if the remainder estimates ω_n satisfy

$$\omega_n = c\frac{z^{n+1}}{1-z}, \qquad c \neq 0, \quad n \in \mathbb{N}_0. \tag{12.2-5}$$

Since $z^\alpha/(1-z)$ is for fixed z and α also a constant, an equivalent condition for the exactness would be:

$$\omega_n = c' z^{n-\alpha+1}, \qquad c' \neq 0, \quad n \in \mathbb{N}_0, \quad \alpha \in \mathbb{R}. \tag{12.2-6}$$

Theorem 12-9: *The sequence transformations* $u_k^{(n)}(\beta, s_n)$, *eq.* (7.3-5), $y_k^{(n)}(\beta, s_n)$, *eq.* (8.4-2), *and* $Y_k^n(\gamma, s_n)$, *eq.* (9.4-2), *are all exact for the geometric series for* $k \geq 2$ *and* $n \geq 0$, *whereas* $t_k^{(n)}(\beta, s_n)$, *eq.* (7.3-7), $d_k^{(n)}(\beta, s_n)$, *eq.* (7.3-9), $v_k^{(n)}(\beta, s_n)$, *eq.* (7.3-11), $\tau_k^{(n)}(\beta, s_n)$, *eq.* (8.4-3), $\delta_k^{(n)}(\beta, s_n)$, *eq.* (8.4-4), $\varphi_k^{(n)}(\beta, s_n)$, *eq.* (8.4-5), $T_k^{(n)}(\gamma, s_n)$, *eq.* (9.4-3), $\Delta_k^{(n)}(\gamma, s_n)$, *eq.* (9.4-4), *and* $\Phi_k^{(n)}(\gamma, s_n)$, *eq.* (9.4-5), *are all exact for the geometric series for* $k \geq 1$ *and* $n \geq 0$. *Drummond's sequence transformation* $\mathscr{D}_k^{(n)}(s_n, \omega_n)$, *eq.* (9.5-4), *is also exact for* $k \geq 1$ *and* $n \geq 0$ *if the remainder estimates* (7.3-6), (7.3-8), *and* (7.3-10) *are used.*

Proof: We only have to prove that in the case of the partial sums (2.6-2) of the geometric series the remainder estimates, which define these variants of the sequence transformations $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-6), and $\mathscr{D}_k^{(n)}(s_n, \omega_n)$, eq. (9.5-4), lead to sequence transformations of the type of eq. (12.2-2) with remainder estimates ω_n that satisfy either eq. (12.2-5) or (12.2-6).

The remainder estimate (7.3-4) leads to $\omega_n = (\beta + n)z^n$. In the case of $u_k^{(n)}(\beta, s_n)$, eq. (7.3-5), and $y_k^{(n)}(\beta, s_n)$, eq. (8.4-2), the factor $\beta + n$ can be absorbed in eqs. (7.1-6) and (8.2-6), respectively, leading for $k \geq 2$ to new sequence transformations which are of the type of eq. (12.2-2). This proves the exactness of $u_k^{(n)}(\beta, s_n)$ and $y_k^{(n)}(\beta, s_n)$. The exactness of $Y_k^{(n)}(\gamma, s_n)$, eq. (9.4-2), can be proved in the same way because the remainder estimate (9.4-1) yields $\omega_n = -(\gamma + n) z^n$ and because $-\gamma - n$ can for $k \geq 2$ be absorbed in eq. (9.2-5) yielding a new sequence transformation satisfying eq. (12.2-2). However, in the case of Drummond's sequence transformation $\mathscr{D}_k^{(n)}(s_n, \omega_n)$, eq. (9.5-4), we would not obtain a sequence transformation of the type of eq. (12.2-2) if we absorb either $\beta + n$ or $-\gamma - n$. Thus, with neither of the two remainder estimates (7.3-4) or (9.4-1) Drummond's sequence transformation is exact for the geometric series.

The remainder estimate (7.3-6) leads to $\omega_n = z^n$. Because of eq. (12.2-6) this proves the exactness of $t_k^{(n)}(\beta, s_n)$, eq. (7.3-7), $\tau_k^{(n)}(\beta, s_n)$, eq. (8.4-3), and $T_k^{(n)}(\gamma, s_n)$, eq. (9.4-3), for $k \geq 1$. With this remainder estimate $\mathscr{D}_k^{(n)}(s_n, \omega_n)$ is also exact for $k \geq 1$.

The remainder estimate (7.3-8) leads to $\omega_n = z^{n+1}$. Because of eq. (12.2-6) this proves the exactness of $d_k^{(n)}(\beta, s_n)$, eq. (7.3-9), $\delta_k^{(n)}(\beta, s_n)$, eq. (8.4-4), and $\Delta_k^{(n)}(\gamma, s_n)$, eq. (9.4-4), for $k \geq 1$. With this remainder estimate $\mathscr{D}_k^{(n)}(s_n, \omega_n)$ is also exact for $k \geq 1$.

The remainder estimate (7.3-10) leads to $\omega_n = z^{n+1}/(1-z)$. Because of eq. (12.2-5) this proves the exactness of $v_k^{(n)}(\beta, s_n)$, eq. (7.3-11), $\varphi_k^{(n)}(\beta, s_n)$, eq. (8.4-5), and $\Phi_k^{(n)}(\gamma, s_n)$, eq. (9.4-5), for $k \geq 1$. With this remainder estimate $\mathscr{D}_k^{(n)}(s_n, \omega_n)$ is also exact for $k \geq 1$.

Since the sequence transformations, which are listed in theorem 12-9, are all exact for the geometric series, it is not surprising that they are also able to accelerate linear convergence.

Theorem 12-10: *The sequence transformations $u_k^{(n)}(\beta, s_n)$, eq. (7.3-5), $t_k^{(n)}(\beta, s_n)$, eq. (7.3-7), $y_k^{(n)}(\beta, s_n)$, eq. (8.4-2), $\tau_k^{(n)}(\beta, s_n)$, eq. (8.4-3), $Y_k^{(n)}(\gamma, s_n)$, eq. (9.4-2), and $T_k^{(n)}(\gamma, s_n)$, eq. (9.4-3), accelerate linear convergence if and only if Drummond's sequence transformation $\mathscr{D}_k^{(n)}(s_n, \omega_n)$, eq. (9.5-4), with $\omega_n = a_n$ accelerates linear convergence.*

The sequence transformations $d_k^{(n)}(\beta, s_n)$, eq. (7.3-9), $\delta_k^{(n)}(\beta, s_n)$, eq. (8.4-4), and $\Delta_k^{(n)}(\gamma, s_n)$, eq. (9.4-4), accelerate linear convergence if and only if Drummond's sequence transformation $\mathscr{D}_k^{(n)}(s_n, \omega_n)$, eq. (9.5-4), with $\omega_n = a_{n+1}$ accelerates linear convergence.

The sequence transformations $v_k^{(n)}(\beta, s_n)$, eq. (7.3-11), $\varphi_k^{(n)}(\beta, s_n)$, eq. (8.4-5), and $\Phi_k^{(n)}(\gamma, s_n)$, eq. (9.4-5), accelerate linear convergence if and only if Drummond's sequence transformation $\mathscr{D}_k^{(n)}(s_n, \omega_n)$, eq. (9.5-4), with $\omega_n = a_n a_{n+1}/(a_n - a_{n+1})$ accelerates linear convergence.

Proof: According to theorem 12-4 the sequence transformations mentioned above accelerate linear convergence if they possess limiting transformations which satisfy eq. (12.1-18). If we perform the limit $n \to \infty$ in the explicit expressions for these sequence transformations according to eq. (12.1-3), we find that their limiting transformations $G_k^{(\infty)}$ are Drummond's sequence transformation with $\omega_n = a_n$, $\omega_n = a_{n+1}$, or $\omega_n = a_n\, a_{n+1}/(a_n - a_{n+1})$, respectively.

Since Drummond's sequence transformation does not explicitly depend on n, this is at the same time the condition which determines whether $\mathscr{D}_k^{(n)}(s_n, \omega_n)$ with either $\omega_n = a_n$, $\omega_n = a_{n+1}$, or $\omega_n = a_n\, a_{n+1}/(a_n - a_{n+1})$ accelerates linear convergence or not.

Theorem 12-11: *Drummond's sequence transformation $\mathscr{D}_k^{(n)}(s_n, \omega_n)$, eq. (9.5-4), with either $\omega_n = a_n$, $\omega_n = a_{n+1}$, or $\omega_n = a_n a_{n+1}/(a_n - a_{n+1})$ accelerates linear convergence for $k \geq 1$ and $n \geq 0$.*

Proof: According to theorem 12-4 Drummond's sequence transformation $\mathscr{D}_k^{(n)}(s_n, \omega_n)$ accelerates linear convergence if $\mathscr{D}_k(s_0, \omega_0)$ is exact for the sequence $\{\sigma_n(\rho)\}$, eq. (12.1-23). The remainder estimate $\omega_n = a_n$ leads to $\omega_n = \rho^{n-1}$, the remainder estimate $\omega_n = a_{n+1}$ leads to $\omega_n = \rho^n$, and the remainder estimate $\omega_n = a_n a_{n+1}/(a_n - a_{n+1})$ leads to $\omega_n = \rho^n/(1-\rho)$. Since these remainder estimates are all of the form of either eq. (12.2-5) or (12.2-6), it follows from theorem 12-8 that for $k \geq 1$ $\mathscr{D}_k^{(0)}(s_0, \omega_0)$ with either $\omega_n = a_n$, $\omega_n = a_{n+1}$, or $\omega_n = a_n a_{n+1}/(a_n - a_{n+1})$ is exact for the sequence $\{\sigma_n(\rho)\}$, eq. (12.1-23). This completes the proof of theorem 12-11.

Similar results as in theorems 12-9 and 12-10 can also be derived for the analogous variants of the generalized transformations $\mathscr{L}_{k,l}^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-8), $\mathscr{S}_{k,l}^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-8), and $\mathscr{M}_{k,l}^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-7), with $l \geq 1$.

12.3. A modification of Germain-Bonne's theory for sequence transformations involving remainder estimates

In the last section Germain-Bonne's formal theory of convergence acceleration was applied to those variants of the sequence transformations $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-6), and $\mathscr{D}_k^{(n)}(s_n, \omega_n)$, eq. (9.5-4), which are based upon Levin's [28] explicit remainder estimates (7.3-4), (7.3-6), and (7.3-10), as well as Smith and Ford's [29] modified remainder estimate (7.3-8).

The sequence $\{\omega_n\}$ of remainder estimates plays a central rôle in the sequence transformations mentioned above since its choice will ultimately determine success or failure. Experience shows that simple remainder estimates (7.3-4), (7.3-6), (7.3-8), and (7.3-10) often work remarkably well in a variety of situations. However, one cannot expect that these simple remainder estimates will always lead to satisfactory results and in some cases it may well be much more efficient to use other remainder estimates $\{\omega_n\}$.

It is often possible to obtain explicit expressions for the remainders $\{r_n\}$ of a sequence $\{s_n\}$. Unfortunately, expressions of that kind are in most cases practically useless since they are normally too complicated. In some cases, however, it may be possible to derive with the help of simplifying assumptions, which are valid in the limit of large indices n, simple explicit expressions which can be used as remainder estimates $\{\omega_n\}$. If such an explicit remainder estimate ω_n does not depend explicitly upon one or several elements of $\{s_n\}$, the sequence transformations $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, and $\mathscr{D}_k^{(n)}(s_n, \omega_n)$ are linear sequence transformations. In addition, they are also defined and exact for constant sequences.

In such a case, one would of course like to know how the two sequences $\{s_n\}$ and $\{\omega_n\}$ have to be related in order to guarantee at least the regularity of the transformation, and under which circumstances the convergence of $\{s_n\}$ will be accelerated. These questions can at least partially be answered by a suitable modification of Germain-Bonne's theory.

Our approach is inspired by a modification of Germain-Bonne's theory which can be found in Brezinski's first book (see pp. 126–132 of ref. [19]). Brezinski considered sequence transformations which simultaneously act upon $k+1$ consecutive elements of the sequence $\{s_n\}$ to be transformed and on $k+1$ consecutive elements of an auxiliary sequence $\{x_n\}$. Brezinski's modification of Germain-Bonne's theory is suited for algorithms which remain well-defined if some or all elements of the auxiliary sequence $\{x_n\}$ are zero. This cannot be assumed here.

It follows from eqs. (7.1-6), (8.2-6), (9.2-5), and (9.5-3) that the sequence transformations $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, and $\mathscr{D}_k^{(n)}(s_n, \omega_n)$ are all of the form of eq. (12.2-2). If we assume that the remainder estimates $\{\!\{\omega_n\}\!\}$ do not explicitly depend upon the elements of $\{\!\{s_n\}\!\}$, then it is a typical feature of the sequence transformations mentioned above that they are linear functions of their first $k+1$ variables. Consequently, these transformations are all continuous in their first $k+1$ variables provided that the elements of $\{\!\{s_n\}\!\}$ are bounded. Much more critical is the continuity of these sequence transformations with respect to their second $k+1$ variables, the remainder estimates $\omega_n, \omega_{n+1}, \ldots, \omega_{n+k}$. Since the remainder estimates always occur in denominators, the elements of $\{\!\{\omega_n\}\!\}$ must not be zero for all finite values of n. In addition, the remainder estimates have to be chosen in such a way that the denominator of such a sequence transformation, which is the k-th difference of $P_{k-1}(n)$, a polynomial of degree $\leq k-1$ in n, divided by ω_n, will not vanish. Hence, a necessary but unfortunately not sufficient condition, which an admissible sequence $\{\!\{\omega_n\}\!\}$ of remainder estimates would have to satisfy, is that its elements are nonzero and distinct for all finite values of n. This implies that for every finite value of n a substring $\omega_n, \omega_{n+1}, \ldots, \omega_{n+k}$ has to belong to a suitable subset of $\mathbb{H}^{k+1}$, the intersection of $\mathbb{F}^{k+1}$ and $\mathbb{D}^{k+1}$.

In this section $\Gamma_k^{(n)}$ stands for a sequence transformation which acts upon $k+1$ consecutive elements of a convergent sequence $\{\!\{s_n\}\!\}$ and which also requires $k+1$ consecutive elements of a sequence $\{\!\{\omega_n\}\!\}$ of remainder estimates. The superscript n indicates that $\Gamma_k^{(n)}$ may depend explicitly upon n.

Hence, for fixed $k \in \mathbb{N}_0$ a sequence transformation $\Gamma_k^{(n)}$ is a function which may depend explicitly on $n \in \mathbb{N}_0$ and which is defined on vectors $\boldsymbol{x} = (x_1, x_2, \ldots, x_{k+1}) \in \mathbb{R}^{k+1}$ and $\boldsymbol{z} = (z_1, z_2, \ldots, z_{k+1}) \in \mathbb{H}^{k+1}$. In addition, we assume that such a sequence transformation $\Gamma_k^{(n)}$: $\mathbb{R}^{k+1} \times \mathbb{H}^{k+1} \to \mathbb{R}$ possesses for fixed $k \in \mathbb{N}_0$ and for all $n \in \mathbb{N}_0$ the following properties:

(A-0): *$\Gamma_k^{(n)}$ is defined and continuous on a subset of $\mathbb{R}^{k+1} \times \mathbb{H}^{k+1}$.*

(A-1): *$\Gamma_k^{(n)}$ is a homogeneous function of degree one in its first $k+1$ variables and a homogeneous function of degree zero in its second $k+1$ variables. This means that for all vectors $\boldsymbol{x} \in \mathbb{R}^{k+1}$ and $\boldsymbol{z} \in \mathbb{H}^{k+1}$, for which $\Gamma_k^{(n)}$ is defined and continuous, and for all $\lambda, \mu \in \mathbb{R}$ with $\mu \neq 0$ we have*

$$\begin{aligned} &\Gamma_k^{(n)}(\lambda x_1, \lambda x_2, \ldots, \lambda x_{k+1} \,|\, z_1, z_2, \ldots, z_{k+1}) \\ &\quad = \lambda \Gamma_k^{(n)}(x_1, x_2, \ldots, x_{k+1} \,|\, z_1, z_2, \ldots, z_{k+1}), \end{aligned} \tag{12.3-1a}$$

$$\begin{aligned} &\Gamma_k^{(n)}(x_1, x_2, \ldots, x_{k+1} \,|\, \mu z_1, \mu z_2, \ldots, \mu z_{k+1}) \\ &\quad = \Gamma_k^{(n)}(x_1, x_2, \ldots, x_{k+1} \,|\, z_1, z_2, \ldots, z_{k+1}). \end{aligned} \tag{12.3-1b}$$

(A-2): *$\Gamma_k^{(n)}$ is linear in its first $k+1$ variables. Consequently, for all vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^{k+1}$ and $\boldsymbol{z} \in \mathbb{H}^{k+1}$, for which is $\Gamma_k^{(n)}$ defined and continuous, we have*

$$\begin{aligned} &\Gamma_k^{(n)}(x_1 + y_1, x_2 + y_2, \ldots, x_{k+1} + y_{k+1} \,|\, z_1, z_2, \ldots, z_{k+1}) \\ &\quad = \Gamma_k^{(n)}(x_1, x_2, \ldots, x_{k+1} \,|\, z_1, z_2, \ldots, z_{k+1}) \\ &\qquad + \Gamma_k^{(n)}(y_1, y_2, \ldots, y_{k+1} \,|\, z_1, z_2, \ldots, z_{k+1}). \end{aligned} \tag{12.3-2}$$

(A-3): *Let* $\boldsymbol{c} = (c, c, \ldots, c) \in \mathbb{R}^{k+1}$ *be a vector with constant components and let* $\boldsymbol{z}$ *belong to the subset of* $\mathbb{H}^{k+1}$ *for which* $\Gamma_k^{(n)}$ *is defined and continuous. Then,* $\Gamma_k^{(n)}$ *is exact, i.e.,*

$$\Gamma_k^{(n)}(c, c, \ldots, c \mid z_1, z_2, \ldots, z_{k+1}) = c. \tag{12.3-3}$$

(A-4): *There exists a subset of* $\mathbb{H}^{k+1}$ *such that for all bounded vectors* $\boldsymbol{x} \in \mathbb{R}^{k+1}$ *and for all vectors* $\boldsymbol{z}$ *belonging to this subset the limiting transformation*

$$\Gamma_k^{(\infty)}(x_1, x_2, \ldots, x_{k+1} \mid z_1, z_2, \ldots, z_{k+1})$$
$$= \lim_{n \to \infty} \Gamma_k^{(n)}(x_1, x_2, \ldots, x_{k+1} \mid z_1, z_2, \ldots, z_{k+1}) \tag{12.3-4}$$

is uniquely defined and continuous on this subset of $\mathbb{R}^{k+1} \times \mathbb{H}^{k+1}$. *In addition, it is assumed that the limiting transformation* $\Gamma_k^{(\infty)}$ *is also homogeneous and linear according to* (A-1) *and* (A-2).

Similarly as in the case of the sequence transformations $G_k^{(n)}$ it cannot be assumed that a sequence transformation $\Gamma_k^{(n)}$ and its limiting transformation $\Gamma_k^{(\infty)}$ will be defined and continuous on the same subset of $\mathbb{R}^{k+1} \times \mathbb{H}^{k+1}$.

It follows from their explicit expressions that the sequence transformations $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-6), and $\mathscr{D}_k^{(n)}(s_n, \omega_n)$, eq. (9.5-4), satisfy (A-0)–(A-4) if suitable sequences $\{\omega_n\}$ of remainder estimates are used. Since these transformations are linear functions of the $k+1$ sequence elements $s_n, s_{n+1}, \ldots, s_{n+k}$ if the elements of $\{\omega_n\}$ do not depend explicitly upon the elements of $\{s_n\}$, they are defined and continuous for arbitrary sequences $\{s_n\}$ if the elements of $\{\omega_n\}$ are nonzero for all finite values of n and if the remainder estimates are chosen in such a way that the denominators of these transformations do not vanish. The denominators of $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, and $\mathscr{D}_k^{(n)}(s_n, \omega_n)$, which are all of the type of eq. (12.2-2), will be nonzero for all $k, n \in \mathbb{N}_0$ if the remainder estimates satisfy $\Delta^k\,(P_{k-1}(n)/\omega_n) \neq 0$. In the case of $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, we have according to eq. (7.1-6) $P_{k-1}(n) = (n+\beta)^{k-1}$, in the case of $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$ we have according to eq. (8.2-6) $P_{k-1}(n) = (n+\beta)_{k-1}$, in the case of $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$ we have according to eq. (9.2-5) $P_{k-1}(n) = (-n-\gamma)_{k-1}$, and in the case of $\mathscr{D}_k^{(n)}(s_n, \omega_n)$ we have according to eq. (9.5-3) $P_{k-1}(n) = 1$.

We are now in a position to formulate an analogue of theorem 12-2 for sequence transformations of the type $\Gamma_k^{(n)}$: $\mathbb{R}^{k+1} \times \mathbb{H}^{k+1} \to \mathbb{R}$.

Theorem 12-12: *Every sequence transformation* $\Gamma_k^{(n)}$, *which is defined and continuous for all vectors* $\boldsymbol{x} \in \mathbb{R}^{k+1}$ *and for all* $\boldsymbol{z}$ *belonging to a suitable subset of* $\mathbb{H}^{k+1}$ *and which also satisfies* (A-0)–(A-4), *can be expressed in the following way*:

$$\Gamma_k^{(n)}(x_1, x_2, \ldots, x_{k+1} \mid z_1, z_2, \ldots, z_{k+1})$$
$$= x_1 + z_1 \gamma_k^{(n)}\left(\frac{x_2 - x_1}{z_1}, \ldots, \frac{x_{k+1} - x_k}{z_k} \,\middle|\, \frac{z_2}{z_1}, \ldots, \frac{z_{k+1}}{z_k} \right). \tag{12.3-5}$$

The associated transformation $\gamma_k^{(n)}$, *which is defined and continuous on a suitable subset of* $\mathbb{R}^k \times \mathbb{F}^k$, *is given by*

$$\gamma_k^{(n)}\left(\frac{x_2 - x_1}{z_1}, \ldots, \frac{x_{k+1} - x_k}{z_k} \,\middle|\, \frac{z_2}{z_1}, \ldots, \frac{z_{k+1}}{z_k} \right)$$

$$= \Gamma_k^{(n)}\left(0, \frac{x_2 - x_1}{z_1}, \ldots, \sum_{j=0}^{k-1} \frac{x_{j+2} - x_{j+1}}{z_{j+1}} \prod_{i=0}^{j-1} \frac{z_{i+2}}{z_{i+1}} \,\middle|\, 1, \frac{z_2}{z_1}, \ldots, \prod_{j=0}^{k-1} \frac{z_{j+2}}{z_{j+1}} \right). \tag{12.3-6}$$

Proof: It follows from (A-2) and (A-3) that $\Gamma_k^{(n)}$ is invariant under translation in the sense of eq. (3.1-4). Consequently, we can subtract x_1 from $\Gamma_k^{(n)}$. Since by assumption $z_1 \neq 0$, it follows from (A-1) that we can divide the $2k+2$ arguments of $\Gamma_k^{(n)}$ by z_1. This yields:

$$\Gamma_k^{(n)}(x_1, x_2, \ldots, x_{k+1} \mid z_1, z_2, \ldots, z_{k+1})$$

$$= x_1 + z_1 \Gamma_k^{(n)}\left(0, \frac{x_2 - x_1}{z_1}, \ldots, \frac{x_{k+1} - x_1}{z_1} \,\middle|\, 1, \frac{z_2}{z_1}, \ldots, \frac{z_{k+1}}{z_1} \right). \tag{12.3-7}$$

We now need the following two relationships which can be proved quite easily:

$$\frac{z_{\nu+1}}{z_1} = \prod_{i=0}^{\nu-1} \frac{z_{i+2}}{z_{i+1}}, \qquad \nu \in \mathbb{N}_0, \tag{12.3-8}$$

$$\frac{x_{\nu+1} - x_1}{z_1} = \sum_{j=0}^{\nu-1} \frac{x_{j+2} - x_{j+1}}{z_{j+1}} \prod_{i=0}^{j-1} \frac{z_{i+2}}{z_{i+1}}, \qquad \nu \in \mathbb{N}_0. \tag{12.3-9}$$

If we insert eqs. (12.3-8) and (12.3-9) into eq. (12.3-7) we obtain eqs. (12.3-5) and (12.3-6). The continuity of $\gamma_k^{(n)}$ on a suitable subset of $\mathbb{R}^k \times \mathbb{F}^k$ follows from the continuity of $\Gamma_k^{(n)}$ on a subset of $\mathbb{R}^{k+1} \times \mathbb{H}^{k+1}$ via eq. (12.3-5).

With the help of theorem 12-12 it can be analyzed for which convergent sequences $\{s_n\}$ and for which sequences $\{\omega_n\}$ of remainder estimates a sequence transformation $\Gamma_k^{(n)}$ satisfying (A-0)–(A-4) will be regular.

Theorem 12-13: *Let* $\Gamma_k^{(n)}$ *be a sequence transformation which satisfies theorem* 12-12. *This means that according to eqs.* (12.3-5) *and* (12.3-6) *an associated transformation* $\gamma_k^{(n)}$ *exists which is continuous on a suitable subset of* $\mathbb{R}^k \times \mathbb{F}^k$. *Let us assume that a sequence* $\{s_n\}$ *converges to some limit* s, *and that the elements of a sequence* $\{\omega_n\}$ *of remainder estimates – although they are different from zero for all finite values of* n *– approach zero as* $n \to \infty$. *Then,* $\Gamma_k^{(n)}$ *is regular if the elements of* $\{s_n\}$ *and* $\{\omega_n\}$ *satisfy:*

(*i*): *For sufficiently large values of* $n \in \mathbb{N}_0$ *the ratios* $\Delta s_n / \omega_n$ *are all bounded, i.e.,*

$$|\Delta s_n / \omega_n| \leq c, \qquad 0 \leq c < \infty. \tag{12.3-10}$$

(*ii*): *For all bounded vectors* $\boldsymbol{y} = (y_1, y_2, \ldots, y_k) \in \mathbb{R}^k$ *the associated transformation* $\gamma_k^{(n)}$ *remains bounded as* $n \to \infty$:

$$\lim_{n \to \infty} \left| \gamma_k^{(n)}\left(y_1, \ldots, y_k \mid \omega_{n+1}/\omega_n, \ \ldots, \ \omega_{n+k}/\omega_{n+k-1}\right) \right| \leq M, \qquad 0 < M < \infty. \tag{12.3-11}$$

Proof: It follows from eq. (12.3-5) that $\Gamma_k^{(n)}$ can be written in the following way:

$$\begin{aligned} &\Gamma_k^{(n)}\left(s_n, \ s_{n+1}, \ldots, s_{n+k} \mid \omega_n, \ \omega_{n+1}, \ldots, \omega_{n+k}\right) \\ &= s_n + \omega_n \gamma_k^{(n)}\left(\frac{\Delta s_n}{\omega_n}, \ldots, \frac{\Delta s_{n+k-1}}{\omega_{n+k-1}} \,\middle|\, \frac{\omega_{n+1}}{\omega_n}, \ldots, \frac{\omega_{n+k}}{\omega_{n+k-1}} \right). \end{aligned} \tag{12.3-12}$$

According to eq. (12.3-6) the associated transformation is given by

$$\begin{aligned} &\gamma_k^{(n)}\left(\frac{\Delta s_n}{\omega_n}, \ldots, \frac{\Delta s_{n+k-1}}{\omega_{n+k-1}} \,\middle|\, \frac{\omega_{n+1}}{\omega_n}, \ldots, \frac{\omega_{n+k}}{\omega_{n+k-1}} \right) \\ &= \Gamma_k^{(n)}\left(0, \ \frac{\Delta s_{n+1}}{\omega_n}, \ \ldots, \ \sum_{j=0}^{k-1} \frac{\Delta s_{n+j}}{\omega_{n+j}} \prod_{i=0}^{j-1} \frac{\omega_{n+i+1}}{\omega_{n+i}} \,\middle|\, 1, \ \frac{\omega_{n+1}}{\omega_n}, \ldots, \prod_{j=0}^{k-1} \frac{\omega_{n+j+1}}{\omega_{n+j}} \right). \end{aligned} \tag{12.3-13}$$

$\Gamma_k^{(n)}$ preserves the convergence of $\{s_n\}$ to its limit s if the second term on the right-hand side of eq. (12.3-12) vanishes as $n \to \infty$. Since $\{\omega_n\}$ approaches zero as $n \to \infty$, we only have to show that the associated transformation $\gamma_k^{(n)}$ remains bounded as $n \to \infty$. Since $\Gamma_k^{(n)}$ is according to (A-2) linear in its first $k+1$ components, we may conclude from eq. (12.3-13) that $\gamma_k^{(\infty)}$ is bounded for all $n \in \mathbb{N}_0$ if its $2k$ arguments remain bounded as $n \to \infty$, and if it remains continuous in its second k variables as $n \to \infty$. It follows from (i) and (ii) that this is indeed the case which proves theorem 12-13.

If we compare theorem 12-13 with the analogous theorem 12-3, which formulates criteria for the regularity of sequence transformations $G_k^{(n)}$: $\mathbb{H}^{k+1} \to \mathbb{R}$, we see that theorem 12-13 is quite liberal with respect to the set of admissible sequences $\{s_n\}$ since only convergence to some limit is assumed. However, given a convergent sequence $\{s_n\}$, the criteria, which have to be satisfied by an admissible sequence $\{\omega_n\}$ of remainder estimates, are quite restrictive.

The next theorem deals with the acceleration of linear convergence by sequence transformations $\Gamma_k^{(n)}$. The following theorem is virtually identical with the analogous theorem 12-4 which deals with sequence transformations $G_k^{(n)}$. In both cases the decisive criterion is that the limiting transformations $\Gamma_k^{(\infty)}$ and $G_k^{(\infty)}$ have to be exact for a shifted sequence of partial sums of the geometric series. This again emphasizes the importance of the geometric series for a theoretical analysis of the acceleration of linear convergence.

Theorem 12-14: *Let us assume that the elements of the sequences $\{s_n\}$ and $\{\omega_n\}$ satisfy*:

$$(i): \qquad \lim_{n \to \infty} s_n = s, \tag{12.3-14}$$

$$(ii): \qquad \lim_{n \to \infty} \frac{s_n - s}{\omega_n} = c, \qquad c \neq 0, \tag{12.3-15}$$

$$(iii): \qquad \lim_{n \to \infty} \frac{\omega_{n+1}}{\omega_n} = \rho, \qquad 0 < |\rho| < 1. \tag{12.3-16}$$

Then, a necessary and sufficient condition that a sequence transformation $\Gamma_k^{(n)}$ accelerates the convergence of the sequence $\{s_n\}$ is that its associated transformation $\gamma_k^{(n)}$ satisfies:

$$\lim_{n \to \infty} \gamma_k^{(n)}(y_n, y_{n+1}, \ldots, y_{n+k-1} \,|\, z_n, z_{n+1}, \ldots, z_{n+k-1})$$
$$= \gamma_k^{(\infty)}(y, y, \ldots, y \,|\, z, z, \ldots, z) = \frac{y}{1-z}. \tag{12.3-17}$$

Here, $\{y_n\}$ and $\{z_n\}$ are essentially arbitrary sequences which converge to y and z, respectively. The above statement can also be formulated in terms of the limiting sequence transformation $\Gamma_k^{(\infty)}$:

$$\lim_{n \to \infty} \Gamma_k^{(n)}\left(0, y_n, \ldots, \sum_{j=0}^{k-1} y_{n+j} \prod_{i=0}^{j-1} z_{n+i} \,\middle|\, 1, z_n, \ldots, \prod_{j=0}^{k-1} z_{n+j}\right)$$
$$= \Gamma_k^{(\infty)}\left(0, y, \ldots, y \sum_{j=0}^{k-1} z^j \,\middle|\, 1, z, \ldots, z^k\right) = \frac{y}{1-z}, \qquad 0 < |z| < 1. \tag{12.3-18}$$

Proof: If we subtract s from both sides of eq. (12.3-12) and divide the resulting expression by $s_n - s$ we obtain:

$$\frac{\Gamma_k^{(n)}(s_n, s_{n+1}, \ldots, s_{n+k} \,|\, \omega_n, \omega_{n+1}, \ldots, \omega_{n+k}) - s}{s_n - s}$$
$$= 1 + \frac{\omega_n}{s_n - s} \gamma_k^{(n)}\left(\frac{\Delta s_n}{\omega_n}, \ldots, \frac{\Delta s_{n+k-1}}{\omega_{n+k-1}} \,\middle|\, \frac{\omega_{n+1}}{\omega_n}, \ldots, \frac{\omega_{n+k}}{\omega_{n+k-1}}\right). \tag{12.3-19}$$

According to eq. (2.6-6) the convergence of a sequence $\{s_n\}$ to its limit s is accelerated by a sequence transformation $\Gamma_k^{(n)}$ if

$$\lim_{n \to \infty} \frac{\Gamma_k^{(n)}(s_n, s_{n+1}, \ldots, s_{n+k} \,|\, \omega_n, \omega_{n+1}, \ldots, \omega_{n+k}) - s}{s_n - s} = 0. \tag{12.3-20}$$

Hence, if we perform the limit $n \to \infty$ in eq. (12.3-19), the left-hand side has to vanish if $\Gamma_k^{(n)}$ accelerates convergence. Thus, we only have to investigate under which conditions the right-hand side of eq. (12.3-19) also vanishes as $n \to \infty$. It follows from eq. (12.3-15) that $\omega_n/(s_n - s) \to 1/c$ as $n \to \infty$. In addition, it follows from eq. (12.3-16) that the second k arguments of the associated transformation $\gamma_k^{(n)}$ all approach ρ as $n \to \infty$, and with the help of eqs. (12.3-15) and (12.3-16) we find that the first k arguments of $\gamma_k^{(n)}$ all satisfy:

$$\lim_{n\to\infty} \frac{\Delta s_n}{\omega_n} = \lim_{n\to\infty} \left\{ \frac{\omega_{n+1}}{\omega_n} \frac{s_{n+1} - s}{\omega_{n+1}} - \frac{s_n - s}{\omega_n} \right\} = c(\rho - 1). \tag{12.3-21}$$

Hence, if we perform the limit $n \to \infty$ in eq. (12.3-19) we find:

$$1 + \frac{1}{c} \gamma_k^{(\infty)}(c(\rho-1), c(\rho-1), \ldots, c(\rho-1) \mid \rho, \rho, \ldots, \rho) = 0. \tag{12.3-22}$$

Now we only have to set $c(\rho - 1) = y$ and $\rho = z$ in order to see that if a sequence transformation $\Gamma_k^{(n)}$ accelerates the convergence of $\{\!\{s_n\}\!\}$, then its associated transformation has to satisfy eq. (12.3-17). In the same way, if eq. (12.3-17) is satisfied by the associated transformation of a sequence transformation $\Gamma_k^{(n)}$, then it follows from eqs. (12.3-15), (12.3-16), and (12.3-19) that $\Gamma_k^{(n)}$ accelerates the convergence of $\{\!\{s_n\}\!\}$ according to eq. (12.3-20). This proves theorem 12-14.

First, it should be remarked that condition (ii), eq. (12.3-15), is identical with eq. (7.3-1). This is another confirmation that the elements of a sequence of remainder estimates should be chosen in such a way that ω_n is proportional to the leading term of an asymptotic expansion of $s_n - s$ as $n \to \infty$.

The first $k + 1$ arguments of $\Gamma_k^{(\infty)}$ in eq. (12.3-18) are apart from the factor y identical with the elements of the sequence $\{\!\{\sigma_n(z)\}\!\}$ which are defined by eq. (12.1-23). Consequently, the right-hand side of eq. (12.3-18) can be rewritten in the following way:

$$\Gamma_k^{(\infty)}\left(0, y, \ldots, y \frac{1 - z^k}{1 - z} \;\middle|\; 1, z, \ldots, z^k\right) = \frac{y}{1 - z}, \qquad 0 < |z| < 1. \tag{12.3-23}$$

Next, theorem 12-14 will be used to prove that the sequence transformations $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-6), and $\mathscr{D}_k^{(n)}(s_n, \omega_n)$, eq. (9.5-4), are able to accelerate linear convergence if the remainder estimates are chosen in such a way that ω_n is proportional to the leading term on an asymptotic expansion of $s_n - s$ as $n \to \infty$.

Theorem 12-15: *We assume that a sequence transformation $T_k^{(n)}(s_n, \omega_n)$ can be written in the following way*:

$$T_k^{(n)}(s_n, \omega_n) = \frac{\sum_{j=0}^{k} (-1)^j \binom{k}{j} f_k(n+j) \frac{s_{n+t}}{\omega_{n+j}}}{\sum_{j=0}^{k} (-1)^j \binom{k}{j} f_k(n+j) \frac{1}{\omega_{n+j}}}, \qquad k, n \in \mathbb{N}_0. \tag{12.3-24}$$

If the sequences $\{s_n\}$ *and* $\{\omega_n\}$ *are as in theorem* 12-14 *and if the coefficients* $f_k(n)$ *satisfy*

$$\lim_{n \to \infty} f_k(n) = 1, \qquad k \in \mathbb{N}_0, \tag{12.3-25}$$

then $T_k^{(n)}(s_n, \omega_n)$ *accelerates the convergence of* $\{s_n\}$ *for* $k \geq 1$.

Proof: According to theorem 12-14 we have to show that the limiting transformation $T_k^{(\infty)}$ satisfies either eq. (12.3-18) or (12.3-23). However, the limiting transformation of all sequence transformations $T_k^{(n)}(s_n, \omega_n)$ satisfying eqs. (12.3-24) and (12.3-25) is Drummond's sequence transformation $\mathscr{D}_k^{(n)}(s_n, \omega_n)$, eq. (9.5-4). Consequently, it is sufficient to show that Drummond's sequence transformation is exact for the sequence $y\{\sigma_n(z)\}$, with $\{\sigma_n(z)\}$ defined by eq. (12.1-23), if the remainder estimates $\omega_n = z^n$ are used, i.e.,

$$\frac{\sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{y}{z^j} \frac{1 - z^j}{1 - z}}{\sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{1}{z^j}} = \frac{y}{1 - z}. \tag{12.3-26}$$

The numerator sum in eq. (12.3-26) may be rewritten in the following way:

$$\sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{y}{z^j} \frac{1 - z^j}{1 - z} = \frac{y}{1 - z} \sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{1}{z^j} - \frac{y}{1 - z} \sum_{j=0}^{k} (-1)^j \binom{k}{j}. \tag{12.3-27}$$

It follows from eq. (2.4-8) that the second sum on the right-hand side is zero for $k \geq 1$. This shows that eq. (12.3-26) is correct and proves theorem 12-15.

It can be deduced directly from their explicit representations that $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-6), and $\mathscr{D}_k^{(n)}(s_n, \omega_n)$, eq. (9.5-4), with $l \geq 1$ satisfy eqs. (12.3-24) and (12.3-25). In addition, it can be shown that the generalized transformations $\mathscr{L}_{k,l}^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-8), $\mathscr{S}_{k,l}^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-8), and $\mathscr{M}_{k,l}^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-7), also satisfy eqs. (12.3-24) and (12.3-25) for sufficiently large values of k. Consequently, these sequence transformations accelerate the convergence of a linearly convergent sequence $\{s_n\}$ if the remainder estimates $\{\omega_n\}$ are chosen in such a way that conditions (i)–(iii) of theorem 12-14 are fulfilled.

It is in fact by no means trivial that the generalized sequence transformations $\mathscr{L}_{k,l}^{(n)}(\beta, s_n, \omega_n)$, $\mathscr{S}_{k,l}^{(n)}(\beta, s_n, \omega_n)$, $\mathscr{M}_{k,l}^{(n)}(\gamma, s_n, \omega_n)$ with $l \geq 1$ also accelerate convergence if the sequences $\{s_n\}$ and $\{\omega_n\}$ satisfy conditions (i)–(iii) of theorem 12-14. It follows from the model sequences (7.1-10), (8.2-10), and (9.2-9) that these transformations were derived assuming that $s_n - s \sim P_l(n)\omega_n$ as $n \to \infty$ with $P_l(n)$ being a polynomial of degree l in n. However, in theorem 12-14 it is assumed that $s_n - s \sim \omega_n$ as $n \to \infty$. Hence, we see that sequence transformations $T_k^{(n)}(s_n, \omega)$ satisfying eqs. (12.3-24) and (12.3-25) accelerate linear convergence even if instead of the "right" sequence $\{\omega_n\}$ of remainder estimates a "wrong" sequence $\{\omega_n'\}$ of remainder estimates with $\omega_n' = P_l(n)\omega_n$ with $l \in \mathbb{N}_0$ is used.

This behaviour is quite typical of all sequence transformations of the type of eq. (12.2-2) which are defined as ratios of finite differences.

With the help of the following theorem it can also be shown that the sequence transformations $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-6), and $\mathscr{D}_k^{(n)}(s_n, \omega_n)$, eq. (9.5-4), as well as the generalized transformations $\mathscr{L}_{k,l}^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-8), $\mathscr{S}_{k,l}^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-8), and $\mathscr{M}_{k,l}^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-7), with $l \geq 1$ are all exact for the geometric series.

Theorem 12-16: *We assume that a sequence transformation $T_k^{(n)}(s_n, \omega_n)$ can be written in the following way:*

$$T_k^{(n)}(s_n, \omega_n) = \frac{\sum_{j=0}^{k} (-1)^j \binom{k}{j} \varphi_{k-1}(n+j) \frac{s_{n+j}}{\omega_{n+j}}}{\sum_{j=0}^{k} (-1)^j \binom{k}{j} \varphi_{k-1}(n+j) \frac{1}{\omega_{n+j}}}, \qquad k, n \in \mathbb{N}_0. \tag{12.3-28}$$

If for sufficiently large values of k the coefficients $\varphi_{k-1}(n)$ are polynomials of degree $\leq k-1$ in n, then for sufficiently large value of k such a sequence transformation $T_k^{(n)}(s_n, \omega_n)$ is exact for the partial sums of the geometric series, eq. (2.6-2), if the remainder estimates are chosen according to $\omega_n = z^{n+\alpha}$ with $\alpha \in \mathbb{R}$.

Proof: The numerator polynomial in eq. (12.3-28) can then be rewritten in the following way:

$$\begin{aligned} &\frac{1}{1-z} \sum_{j=0}^{k} (-1)^j \binom{k}{j} \varphi_{k-1}(n+j) \frac{1-z^{n+j}}{z^{n+j+\alpha}} \\ &= \frac{1}{1-z} \sum_{j=0}^{k} (-1)^j \binom{k}{j} \varphi_{k-1}(n+j) \frac{1}{z^{n+j+\alpha}} \\ &- \frac{1}{z^\alpha(1-z)} \sum_{j=0}^{k} (-1)^j \binom{k}{j} \varphi_{k-1}(n+j). \end{aligned} \tag{12.3-29}$$

Let us now assume that k is large enough such that $\varphi_{k-1}(n)$ is a polynomial of degree $\leq k-1$ in n. Then it follows from eq. (2.4-8) that the second sum in eq. (12.3-29) vanishes. This proves theorem 12-16.

Obviously, all sequence transformations mentioned above satisfy the prerequisites of theorem 12-16.

12.4. A critical assessment of Germain-Bonne's theory

With the help of either the original version of Germain-Bonne's formal theory of convergence acceleration or its modifications it can be decided whether a sequence transformation is regular, i.e., whether the transformed sequence $\{s_n'\}$ converges to the same limit as the original sequence $\{s_n\}$. In addition, a necessary and sufficient condition could be formulated by means of which it can be decided whether a sequence transformation is able to accelerate linear convergence or not. Theoretically, these results are certainly remarkable achievements, in particular since for their derivation only some very general properties of the sequence transformation such as continuity, homogeneity, and translativity had to be assumed. Also, concerning the sequences $\{s_n\}$, which are to be transformed, only relatively little has to be assumed. In most cases it is sufficient that the sequences converge or – if the acceleration of linear convergence is analyzed – that they converge linearly.

However, it must not be overlooked that Germain-Bonne's formal theory of convergence acceleration has some serious shortcomings which definitely limit its practical usefulness, although it certainly is a beautiful mathematical theory. For instance, the generality of Germain-Bonne's theory and its modifications – although highly desirable from a theoretical point of view – is at the same time a major weakness since it implies that the results of this theory are quite general and cannot be as specific as one would like them to be.

Germain-Bonne's theory is only able to make statements as for instance that a sequence transformation is regular or that it is able to accelerate linear convergence. However, from a practical point of view the statement that a given sequence transformation is able to accelerate linear convergence is just as useful – or as useless – as the statement that a given series converges.

Germain-Bonne's theory is essentially asymptotic in nature because only the sequence elements s_n and the transforms $G_k^{(n)}$ or $\Gamma_k^{(n)}$ with large values of n matter. This asymptotic nature is essential because it makes a theoretical analysis possible. However, it refers to a situation – sequence elements s_n or transforms s_n' with large indices n – which one would like to avoid by using sequence transformations. In addition, the predictive value of an asymptotic theory is often quite limited. A given numerical technique need not be particularly powerful for moderately large values of n, let alone for small values of n even if it is guaranteed that this technique will work well in the limit $n \to \infty$.

In actual computations only a relatively small number of sequence elements will normally be known, say $s_n, s_{n+1}, \ldots, s_{n+k}$, and one would like to know how the information, which is contained in these sequences elements, can be extracted and utilized in an optimal way. Here, Germain-Bonne's theory or its variants cannot help at all since it does not discriminate among sequence transformations which all have the same properties in the limit $n \to \infty$. For instance, according to theorem 12-10 certain variants of the sequence transformations $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), and $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-6), are able to accelerate linear convergence if and only if the analogous variants of Drummond's sequence transformation $\mathscr{D}_k^{(n)}(s_n, \omega_n)$, eq. (9.5-4), are able to accelerate linear convergence. Consequently, Drummond's sequence transformation plays a very important rôle in our modification of Germain-Bonne's formal theory of convergence acceleration since it allows a unified treatment of a large class of sequence transformations. But it is wrong to assume that Drummond's sequence transformation will be particularly useful in actual computations. In fact, we shall see later that Drummond's sequence transformation is normally significantly less powerful than the other sequence transformations mentioned above.

Germain-Bonne's theory is essentially a successful theory of the acceleration of linear convergence and it does not help at all if for instance logarithmic convergence is to be accelerated. This is quite deplorable since the acceleration of logarithmic convergence is a much more annoying problem than the acceleration of linear convergence – both theoretically and computationally.

Another problem of great practical relevance is the determination of the antilimit s of a divergent sequence $\{s_n\}$ by employing a suitable sequence transformation $G_k^{(n)}$ or $\Gamma_k^{(n)}$. In such a case a formal theory, which involves a limit $n \to \infty$, makes no sense. Instead, any theoretical analysis of such a summation process would have to say something about the convergence of a sequence of transforms $G_k^{(n)}$ or $\Gamma_k^{(n)}$ to the antilimit s and how this convergence is affected if the subscript k is increased while the superscript n is held fixed. Again, Germain-Bonne's theory and its modifications cannot contribute anything.

Hence, Germain-Bonne's theory of convergence acceleration is not able to treat several problems of great practical relevance and has to be supplemented by other theoretical approaches. However, it is not likely that Germain-Bonne's theory can be improved significantly without making much more detailed assumptions about both the sequence transformations, which are to be analyzed, and the sequences, which are to be accelerated or summed.

13. Summation and convergence acceleration of Stieltjes series

13.1. Stieltjes series and Stieltjes functions

This section deals with the summation of divergent asymptotic series, as they for instance occur in the theory of special functions or in quantum mechanical perturbation theory. It is well known that a given function $f(z)$ admits at most one asymptotic power series. The converse, however, is not true, i.e., it may happen that there are several different functions which all admit the same asymptotic power series. In a theoretical analysis of summation processes, the possible nonuniqueness of asymptotic expansions is certainly quite annoying and the set of admissible asymptotic series should be suitably restricted in order to avoid these complications.

These nonuniqueness problems can be avoided in the case of Stieltjes series which assume an exceptional position among divergent series. For Stieltjes series there exists a highly developed convergence and representation theory (see refs. [18,22,79,95,96]). For instance, it can be shown that Padé approximants are able to sum even wildly divergent Stieltjes series if the terms a_n of these series do not grow faster in magnitude than $c^{n+1}(2n)!$ as $n \to \infty$ with c being a suitable positive constant (see theorems 1.2 and 1.3 of ref. [87]). This implies that Padé approximants are able to sum the divergent Euler series, eq. (1.1-7), which is, as we shall see later, a Stieltjes series.

Stieltjes series are also of considerable physical interest since many quantum mechanical perturbation expansions are Stieltjes series. For instance, it could be proved rigorously that the Rayleigh-Schrödinger perturbation expansions for the energy eigenvalues of the quartic anharmonic oscillator are Stieltjes series [86,97]. It also follows from the asymptotic behaviour (1.1-5) of the perturbation series coefficients that Padé approximants are able to sum the divergent perturbation series (1.1-4) for the ground state energy of the quartic anharmonic oscillator.

In this section, the summation of divergent Stieltjes series by means of nonlinear sequence transformations will be investigated both theoretically and numerically. It will be one of the main results of this section that the sequence transformations $\mathscr{L}_k^{(n)}$ $(\beta,\ s_n,\ \omega_n)$, eq. (7.1–7), $\mathscr{S}_k^{(n)}(\beta,\ s_n,\ \omega_n)$, eq. (8.2–7), and $\mathscr{M}_k^{(n)}(\gamma,\ s_n,\ \omega_n)$, eq. (9.2–6), which all require a sequence $\{\omega_n\}$ of remainder estimates for their computation, sum divergent Stieltjes series much more efficiently than for instance Padé approximants, which may be computed with the help of Wynn's ϵ algorithm, eq. (4.2–1), or also Brezinski's ϑ algorithm, eq. (10.1–9). This is probably due to the fact that in the case of a Stieltjes series it is comparatively easy to find simple remainder estimates $\{\omega_n\}$, which may be used in the sequence transformations mentioned above and which in spite of their simplicity yield rigorous and tight upper bounds for the remainders of truncated Stieltjes series.

In order to make this section more self-contained, first those properties of Stieltjes series and Stieltjes functions, which are of particular importance for our purposes, will be reviewed.

A function $f(z)$ with $z \in \mathbb{C}$ will be called *Stieltjes function* if it can be expressed as a *Stieltjes integral*,

$$f(z) = \int_0^\infty \frac{\mathrm{d}\psi(t)}{1+zt}, \qquad |\arg(z)| < \pi. \tag{13.1-1}$$

Here, $\psi(t)$ is a positive measure on $0 \le t < \infty$ which has for all $m \in \mathbb{N}_0$ finite and positive moments μ_m defined by

$$\mu_m = \int_0^\infty t^m \,\mathrm{d}\psi(t), \qquad m \in \mathbb{N}_0. \tag{13.1-2}$$

A formal series expansion of the following type, which need not be convergent,

$$f(z) = \sum_{m=0}^{\infty} (-1)^m \mu_m z^m, \tag{13.1-3}$$

is called a *Stieltjes series* if its coefficients μ_m are moments of a positive measure $\psi(t)$ on $0 \le t < \infty$ according to eq. (13.1-2), i.e.,

$$f(z) = \sum_{m=0}^{\infty} (-1)^m z^m \int_0^\infty t^m \,\mathrm{d}\psi(t). \tag{13.1-4}$$

A good example for a Stieltjes function with a wildly divergent Stieltjes series is the so-called Euler integral, eq. (1.1-6), and its associated asymptotic series, the so-called Euler series, eq. (1.1-7).

Theorem 13-1: *Every Stieltjes function $f(z)$ can be written in the following way*:

$$f(z) = \sum_{m=0}^{n} (-1)^m \mu_m z^m + (-z)^{n+1} \int_0^\infty \frac{t^{n+1}\,\mathrm{d}\psi(t)}{1+zt}, \qquad |\arg(z)| < \pi. \tag{13.1-5}$$

Proof: We only have to insert the relationship

$$\sum_{m=0}^{n} x^m = \frac{1 - x^{n+1}}{1 - x} \tag{13.1-6}$$

with $x = -zt$ into eq. (13.1-1) and do the moment integrals according to eq. (13.1-2).

Whether a Stieltjes series converges or diverges depends upon the behaviour of the remainder integral on the right-hand side of eq. (13.1-5). The next theorem shows that this remainder integral is bounded by the first term of the power series (13.1-3) which was not included in the partial sum in eq. (13.1-5). This bound will also help us to find a simple and – as we shall see later – efficient sequence of remainder estimates for the sequence transformations $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-6), and $\mathscr{D}_k^{(n)}(s_n, \omega_n)$, eq. (9.5-4).

Theorem 13-2: *The remainder term in theorem* 13-1,

$$R_n(z) = (-z)^{n+1} \int_0^\infty \frac{t^{n+1}\, \mathrm{d}\psi(t)}{1 + zt}, \tag{13.1-7}$$

satisfies depending upon the value of $\vartheta = \arg(z)$ *the following inequalities*:

$$|R_n(z)| \le \mu_{n+1} |z^{n+1}|, \qquad |\vartheta| \le \pi/2, \tag{13.1-8a}$$

$$|R_n(z)| \le \mu_{n+1} |z^{n+1} \operatorname{cosec} \vartheta|, \quad \pi/2 < |\vartheta| < \pi. \tag{13.1-8b}$$

Proof: Setting $z = r\, \mathrm{e}^{\mathrm{i}\vartheta}$ gives

$$|1 + zt| = \left[1 + 2rt \cos \vartheta + r^2 t^2\right]^{1/2}. \tag{13.1-9}$$

Next, one has to look for the value of t with $0 \le t < \infty$ for which $|1 + zt|$ assumes its minimal value. Differentiation with respect to t gives the extremal condition

$$t = -\frac{\cos \vartheta}{r}. \tag{13.1-10}$$

Now, two different cases have to be distinguished:

(i): $|\vartheta| \le \pi/2$. Then, $\cos \vartheta \ge 0$. Consequently, $|1 + zt|$ assumes its minimal value for $t = 0$ and we obtain the estimate (13.1-8a).

(ii): $\pi/2 < |\vartheta| < \pi$. Then, $\cos \vartheta < 0$. Combination of eqs. (13.1-9) and (13.1-10) yields

$$|\sin \vartheta| \le |1 + zt|. \tag{13.1-11}$$

If this inequality is used in eq. (13.1-7), estimate (13.1-8b) follows. This shows that theorem 13-2 is correct.

It also follows from theorem 13-2 that every Stieltjes function possesses an asymptotic series valid uniformly in every sector $|\arg(z)| < \vartheta$ for any $\vartheta < \pi$ and that this asymptotic series is a

Stieltjes series (see p. 398 of ref. [86]). It can also be proved that for every Stieltjes series there exists at least one associated Stieltjes function. Since this possible nonuniqueness is very inconvenient in summation processes, a criterion would be needed which makes it possible to prove that there is a one-to-one correspondence at least between certain divergent Stieltjes series and certain Stieltjes functions. Thus, a condition would be needed which is stronger than the existence of an asymptotic power series of the type of eq. (13.1-3) but weaker than the existence of a convergent Stieltjes series.

On the basis of Carleman's theorem (see p. 39 of ref. [98]) a sufficient condition can be formulated which guarantees that there exists a one-to-one correspondence between a Stieltjes function and its associated asymptotic series.

A Stieltjes function $f(z)$, which is analytic in a sectorial region of the complex plane, is said to satisfy a *strong asymptotic condition* and its associated Stieltjes series is called a *strong asymptotic series* if suitable positive constants A and ξ can be found such that

$$\left| f(z) - \sum_{m=0}^{n} (-1)^m \mu_m z^m \right| \leq A \xi^{n+1} (n+1)! |z|^{n+1} \tag{13.1-12}$$

holds for all $n \in \mathbb{N}_0$ and for all z in this sectorial region.

The validity of such a strong asymptotic condition implies that a Stieltjes function $f(z)$ is uniquely determined by its asymptotic series (see p. 40 of ref. [98]). Such a strong asymptotic condition can only be valid if the Stieltjes moments μ_n, which are defined by eqs. (13.1-3) and (13.1-4), satisfy for all $n \in \mathbb{N}_0$ (see p. 43 of ref. [98])

$$\mu_n \leq A \xi^n n!. \tag{13.1-13}$$

The moments of the Euler series, eq. (1.1-7), satisfy this inequality. Hence, we may conclude that the Euler integral, eq. (1.1-6), is uniquely determined by its asymptotic series (1.1-7). In the same way, it follows from the asymptotic behaviour (1.1-5) of the series coefficients that the perturbation series (1.1-4) for the ground state energy of the quartic anharmonic oscillator is a strong asymptotic series. Consequently, the ground state energy of the quartic anharmonic oscillator is uniquely determined by its divergent perturbation series (see also p. 41 of ref. [98]).

However, there are Stieltjes series of considerable physical interest which have moments μ_n that behave like $(kn)!$ with $k > 1$ as $n \to \infty$. For instance, in the case of the perturbation expansions for the energy eigenvalues of the sextic or octic anharmonic oscillator there is ample numerical evidence that the coefficients of these series grow as $(2n)!$ or $(3n)!$, respectively, as $n \to \infty$ (see p. 43 of ref. [98]) Obviously, a strong asymptotic condition cannot be valid in such a case. However, it can be shown (see p. 43 of ref. [98]) that a function $f(z)$, which is analytic within a sectorial region of the complex plane, is also uniquely determined by its asymptotic series if $f(z)$ satisfies a *modified strong asymptotic condition of order k* and if its asymptotic series is a *modified strong asymptotic series of order k*. This means that suitable positive constants A and ξ can be found such that

$$\left| f(z) - \sum_{m=0}^{n} (-1)^m \mu_m z^m \right| \leq A \xi^{n+1} [k(n+1)]! |z|^{n+1} \tag{13.1-14}$$

holds for all $n \in \mathbb{N}_0$ and for all z in this sectorial region.

Again, such a modified strong asymptotic condition of order k can only be valid if the Stieltjes moments μ_n satisfy for all $n \in \mathbb{N}_0$ (see p. 406 of ref. [86])

$$\mu_n \leq A \xi^n (kn)!. \tag{13.1-15}$$

The bounds for the remainders $R_n(z)$ in theorem 13-2 are also of considerable importance for convergence acceleration and summation processes because it helps us to find simple and manageable remainder estimates $\{\omega_n\}$. Because of the specific structure of the sequence transformations $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, and $\mathscr{D}_k^{(n)}(s_n, \omega_n)$, all those quantities which are independent of n, do not have to be included in the remainder estimate ω_n. Consequently, it is not necessary to distinguish the two different cases $|\vartheta| \leq \pi$ and $\pi/2 < |\vartheta| < \pi$ in theorem 13-2 and for every sector $|\arg(z)| < \vartheta$ with $\vartheta < \pi$ a suitable estimate ω_n for the remainder $R_n(z)$ of a Stieltjes series would be

$$\omega_n = (-1)^{n+1} \mu_{n+1} z^{n+1}, \qquad n \in \mathbb{N}_0. \tag{13.1-16}$$

This choice is identical with the remainder estimate (7.3-8) of Smith and Ford [29]. Hence, for the summation of divergent Stieltjes series the most natural choices among the numerous variants of the sequence transformations $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, and $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$ would be the transformations $d_k^{(n)}(\beta, s_n)$, eq. (7.3-9), $\delta_k^{(n)}(\beta, s_n)$, eq. (8.4-4), and $\Delta_k^{(n)}(\gamma, s_n)$, eq. (9.4-4).

13.2. Theoretical error estimates

In this section theoretical error estimates for the summation of a divergent Stieltjes series by means of nonlinear sequence transformations will be derived. However, the error estimates of this section can also be applied if the convergence of sequences with strictly alternating remainders is accelerated.

There are only relatively few references in the literature in which the summation of divergent Stieltjes series by means of nonlinear sequence transformations is analyzed. In articles by Wynn [99], Common [100], Allen, Chui, Madych, Narcowich, and Smith [101], and Karlsson and Sydow [102] the Padé summation of Stieltjes series was analyzed. Then, there is an article by Sidi [103] on the summation of certain wildly divergent series by Levin's u and t transformations, eqs. (7.3-5) and (7.3-7). Sidi could show that if the divergent series satisfies certain conditions, Levin's u and t transformation produce sequences of approximants which converge to the Borel sum of the divergent series. Other sequence transformations were apparently not yet treated in the literature. This is not too surprising since many nonlinear sequence transformations as for instance Brezinski's ϑ algorithm, eq. (10.1-9), are defined by relatively complicated recursive schemes and otherwise only very little is known about these transformations. Currently, a detailed theoretical analysis of the efficiency of such a sequence transformation in convergence acceleration and summation processes seems to be more or less impossible.

However, sequence transformations of the type of eq. (12.2-2) can be analyzed relatively easily if suitable assumptions concerning the sequences $\{s_n\}$ and $\{\omega_n\}$ are made. Consequently, in this section the emphasis will be on the sequence transformations $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7),

$\mathcal{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), $\mathcal{M}_k^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-6), and $\mathcal{D}_k^{(n)}(s_n, \omega_n)$, eq. (9.5-4), as well as on the mild generalizations $\mathcal{L}_{k,l}^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-8), $\mathcal{S}_{k,l}^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-8), and $\mathcal{M}_{k,l}^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-7). The following theorem will be the basis of our analysis.

Theorem 13-3: *Let us assume that a sequence transformation $G_k^{(n)}(s_n, s_{n+1}, \ldots, s_{n+k})$ with k, $n \in \mathbb{N}_0$ is invariant under translation according to eq.* (3.1-4). *Then a necessary and sufficient condition that this sequence transformation is able to sum a divergent sequence $\{\!\{s_n\}\!\}$ to its antilimit s on a path $\mathcal{P} = \{\!\{(n_j, k_j)\}\!\}$ with $j \in \mathbb{N}_0$ is that*

$$\lim_{j\to\infty} G_{k_j}^{(n_j)}(s_{n_j} - s, s_{n_j+1} - s, \ldots, s_{n_j+k_j} - s) = 0. \tag{13.2-1}$$

Proof: Since $G_k^{(n)}$ is by assumption invariant under translation according to eq. (3.1-4), we have for arbitrary integers n_j and k_j:

$$G_{k_j}^{(n_j)}(s_{n_j}, s_{n_j+1}, \ldots, s_{n_j+k_j}) = s + G_{k_j}^{(n_j)}(s_{n_j} - s, s_{n_j+1} - s, \ldots, s_{n_j+k_j} - s). \tag{13.2-2}$$

Performing the limit $j \to \infty$ shows that theorem 13-3 is correct.

It can be proved quite easily by a typical 2ε proof that if the antilimit s of a divergent series $\{\!\{s_n\}\!\}$ exists on a given path $\mathcal{P}$, then it is uniquely determined on this path. For different paths, however, no general statement concerning the uniqueness of the antilimit s can be made. In summation processes, one is normally only interested in horizontal paths, i.e., in paths in which n_j is ultimately constant and in which only k_j is increased. Of course, theorem 13-3 can be reformulated in such a way that it applies to convergence acceleration processes.

Theorem 13-4: *Let us assume that a sequence transformation $G_k^{(n)}(s_n, s_{n+1}, \ldots, s_{n+k})$ with k, $n \in \mathbb{N}_0$ is invariant under translation according to eq.* (3.1-4). *Then a necessary and sufficient condition that this sequence transformation preserves the limit s of a convergent sequence $\{\!\{s_n\}\!\}$ on a path $\mathcal{P} = \{\!\{(n_j, k_j)\}\!\}$ with $j \in \mathbb{N}_0$ is that*

$$\lim_{j\to\infty} G_{k_j}^{(n_j)}(s_{n_j} - s, s_{n_j+1} - s, \ldots, s_{n_j+k_j} - s) = 0. \tag{13.2-3}$$

What is gained if summation and convergence acceleration processes are analyzed with the help of theorems 13-3 and 13-4. Since the limit or antilimit s of a sequence $\{\!\{s_n\}\!\}$ is normally not known, it is in most cases very hard or even impossible to estimate, how close $G_k^{(n)}(s_n, s_{n+1}, \ldots, s_{n+k})$ and s are. However, it will become clear in the sequel that it is frequently comparatively easy to obtain a theoretical estimate for the magnitude of the error term $G_k^{(n)}(s_n - s, s_{n+1} - s, \ldots, s_{n+k} - s)$ and its dependence upon k and n.

Theorems 13-3 and 13-4 remain of course valid if the sequence transformation $G_k^{(n)}$, which only depends upon $k+1$ sequence elements $s_n, s_{n+1}, \ldots, s_{n+k}$, is replaced by a sequence transformation $\Gamma_k^{(n)}$ which in addition to the $k+1$ sequence elements $s_n, s_{n+1}, \ldots, s_{n+k}$ also depends upon $k+1$ remainder estimates $\omega_n, \omega_{n+1}, \ldots, \omega_{n+k}$.

In this section we shall try to make some quantitative predictions about the magnitude of the summation error if a given sequence transformation, which is of the type of eq. (12.2-2), is used for the summation of a divergent Stieltjes series. Unfortunately, it seems that such an error analysis cannot be done in the case of a completely arbitrary Stieltjes series. However, if we apply sequence transformations of the type of eq. (12.2-2) to some suitably chosen model sequences, valuable insight into the mechanism as well as the power of these sequence transformations can be gained.

Our error analysis will be based upon theorems 13-3 and 13-4, i.e., we shall try to estimate the magnitude of the error term $G_k^{(n)}(s_n - s, s_{n+1} - s, \ldots, s_{n+k} - s)$ and its dependence upon k and n. In addition, we assume that the sequences $\{s_n\}$ and $\{\omega_n\}$ possess the following properties:

(S-0): *The elements of $\{s_n\}$ are the partial sums of an infinite series which either converges to some limit s or in the case of divergence can be summed to give s.*

(S-1): *The elements of the sequence $\{\omega_n\}$ of remainder estimates for $\{s_n\}$ are strictly alternating in sign.*

(S-2): *For all $n \in \mathbb{N}_0$ the ratio $(s_n - s)/\omega_n$ can be expressed as a factorial series, i.e.,*

$$\frac{s_n - s}{\omega_n} = \sum_{j=0}^{\infty} \frac{c_j}{(\beta + n)_j}, \qquad \beta \in \mathbb{R}_+, \quad n \in \mathbb{N}_0. \tag{13.2-4}$$

On the basis of these assumptions the summation of divergent Stieltjes series as well as the acceleration of the convergence of certain alternating series can be analyzed.

Concerning (S-1) it should be remarked that if we chose the remainder estimates $\{\omega_n\}$ according to eq. (13.1-16) then the positivity of the Stieltjes moments μ_n according to eq. (13.1-2) implies that we are restricted to power series with positive arguments z. If z would be an arbitrary complex number, it could not be guaranteed that our remainder estimates ω_n will have strictly alternating signs if they are chosen according to eq. (13.1-16).

The requirement that $(s_n - s)/\omega_n$ can be expressed as a factorial series according to eq. (13.2-4) may appear to be somewhat restrictive. However, this is not necessarily more restrictive than the analogous requirement that $(s_n - s)/\omega_n$ can be expressed as a series in inverse powers of $\beta + n$,

$$\frac{s_n - s}{\omega_n} = \sum_{j=0}^{\infty} \frac{c'_j}{(\beta + n)^j}, \qquad \beta \in \mathbb{R}_+, \quad n \in \mathbb{N}_0. \tag{13.2-5}$$

In Nielsen's book it is described how inverse power series and factorial series can be transformed into each other (see pp. 272–282 of ref. [77]). Assumptions (S-0)–(S-2) will now be used to obtain quantitative error estimates in summation and convergence acceleration processes.

Theorem 13-5: *Let us assume that the sequences $\{s_n\}$ and $\{\omega_n\}$ satisfy (S-0)–(S-2) and that the sequence transformation $\mathcal{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), is used for the transformation of $\{s_n\}$. Then we obtain for fixed $k \in \mathbb{N}$ and for all $n \in \mathbb{N}_0$ the following estimate for the error term*:

$$|\mathcal{S}_k^{(n)}(\beta, s_n, \omega_n) - s| \le \left| \frac{\omega_n}{(\beta + n)_{2k}} \sum_{j=0}^{\infty} \frac{c_{k+j}(j+1)_k}{(\beta + n + 2k)_j} \right|. \tag{13.2-6}$$

This implies for fixed $k \in \mathbb{N}$ and for large values of n the following order estimate:

$$\frac{\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n) - s}{s_n - s} = O(n^{-2k}), \qquad n \to \infty. \tag{13.2-7}$$

Proof: First, we observe that $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), is invariant under translation according to eq. (3.1-4). This implies that the magnitude of the transformation error and its dependence upon k and n can be analyzed by estimating the magnitude of $\mathscr{S}_k^{(n)}(\beta, s_n - s, \omega_n)$. The starting point of our analysis is eq. (8.2-6) which is rewritten in the following way:

$$\mathscr{S}_k^{(n)}(\beta, s_n - s, \omega_n) = \frac{\Delta^k\{(\beta+n)_{k-1}(s_n - s)/\omega_n\}}{\Delta^k\{(\beta+n)_{k-1}/\omega_n\}}. \tag{13.2-8}$$

In the numerator in eq. (13.2-8) $(s_n - s)/\omega_n$ is replaced by the factorial series (13.2-4) yielding

$$\Delta^k \frac{(\beta+n)_{k-1}(s_n - s)}{\omega_n} = \Delta^k (\beta+n)_{k-1} \sum_{j=0}^{\infty} \frac{c_j}{(\beta+n)_j} \tag{13.2-9}$$

$$= \Delta^k \sum_{j=0}^{k-1} c_j (\beta+n+j)_{k-j-1} + \Delta^k \sum_{j=0}^{\infty} \frac{c_{k+j}}{(\beta+n+k-1)_{j+1}}. \tag{13.2-10}$$

Next, we derive from eq. (8.4-10) the following two relationships:

$$\Delta^k (a+n)_m = (-1)^k (-m)_k (a+n+k)_{m-k}, \tag{13.2-11}$$

$$\Delta^k [1/(b+n)_m] = (-1)^k (m)_k / (b+n)_{k+m}. \tag{13.2-12}$$

If these two relationships are used in eq. (13.2-10), we find that the first sum vanishes since it is a polynomial of degree k_1 in n and we obtain for the numerator in eq. (13.2-8):

$$\Delta^k \frac{(\beta+n)_{k-1}(s_n - s)}{\omega_n} = \frac{\Gamma(\beta+n+k-1)}{\Gamma(\beta+n+2k)} \sum_{j=0}^{\infty} \frac{c_{k+j}(j+1)_k}{(\beta+n+2k)_j}. \tag{13.2-13}$$

With the help of eq. (2.4-8) we obtain for the denominator in eq. (13.2-8):

$$\Delta^k \frac{(\beta+n)_{k-1}}{\omega_n} = (-1)^k \sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{(\beta+n+j)_{k-1}}{\omega_{n+j}}. \tag{13.2-14}$$

According to (S-1) the elements of $\{\omega_n\}$ strictly alternate in sign. This gives us immediately the following estimate:

$$|(\beta+n)_{k-1}/\omega_n| \le |\Delta^k\{(\beta+n)_{k-1}/\omega_n\}|. \tag{13.2-15}$$

Combination of eqs. (13.3-8), (13.2-13) and (13.2-15) gives eq. (13.2-6). The order estimate (13.2-7) follows from the fact that according to eq. (13.2-4)

$$(s_n - s)/\omega_n = c_0\left[1 + \mathrm{O}(n^{-1})\right], \qquad n \to \infty, \tag{13.2-16}$$

and that $(\beta + n)_{2k} = \mathrm{O}(n^{2k})$ as $n \to \infty$. This proves theorem 13-5.

Since the denominator sum (13.2-14), which consists of $k+1$ terms, is estimated by a single term according to eq. (13.2-15), the error estimate (13.2-6) is quite conservative.

It is a typical feature of the error estimate (13.2-6) and also of some analogous error estimates for other sequence transformations, which will be derived later in this section, that the error estimate is directly proportional to ω_n. Consequently, no distinction between convergent and divergent sequences $\{s_n\}$ of partial sums has to be made. It also follows from the error estimate (13.2-6) that $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), is able to sum a divergent series satisfying (S-0)–(S-2) on a horizontal path if the coefficients c_j of the factorial series (13.2-4) do not grow too fast in magnitude as $j \to \infty$.

The next theorem, which can be proved in essentially the same way as theorem 13-5, shows that our error analysis is able to distinguish between the sequence transformation $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), and its mild generalization $\mathscr{S}_{k,l}^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-8), with $l \geq 1$.

Theorem 13-6: *Let us assume that the sequences $\{s_n\}$ and $\{\omega_n\}$ satisfy (S-0), (S-1), and*

$$\frac{s_n - s}{\omega_n} = \sum_{j=0}^{\infty} \frac{d_j}{(\beta + n + l)_j}, \qquad \beta \in \mathbb{R}_+, \qquad l \in \mathbb{N}, \qquad n \in \mathbb{N}_0, \tag{13.2-17}$$

and that the sequence transformation $\mathscr{S}_{k,l}^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-8), with $l \geq 1$ is used for the transformation of $\{s_n\}$. Then we obtain for fixed k, $l \in \mathbb{N}$ with $k \geq l+1$ and for all $n \in \mathbb{N}_0$ the following estimate for the error term:

$$\left|\mathscr{S}_{k,l}^{(n)}(\beta, s_n, \omega_n) - s\right| \leq \left| \frac{\omega_n}{(\beta + n + l)_{2k-l}} \sum_{j=0}^{\infty} \frac{d_{k-l+j}(j+1)_k}{(\beta + n + 2k)_j} \right|. \tag{13.2-18}$$

This implies for fixed k, $l \in \mathbb{N}$ with $k \geq l+1$ and for large values of n the following order estimate:

$$\frac{\mathscr{S}_{k,l}^{(n)}(\beta, s_n, \omega_n) - s}{s_n - s} = \mathrm{O}(n^{l-2k}), \qquad n \to \infty. \tag{13.2-19}$$

If we compare eqs. (13.2-4) and (13.2-17), it seems that in theorems 13-5 and 13-6 the existence of two different factorial series expansions for the ratio $(s_n - s)/\omega_n$ are assumed. However, the factorial series (13.2-4) and (13.2-17) are not independent. In Nielsen's book it is shown how a factorial series of the type of eq. (13.2-4) can be transformed into a factorial series of the type of eq. (13.2-17) (see pp. 252–253 of ref. [77]).

Theorems 13-5 and 13-6 indicate that $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), should normally be more efficient than its generalization $\mathscr{S}_{k,l}^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-8), with $l \geq 1$.

In the following theorem, the efficiency of Drummond's sequence transformation $\mathscr{D}_k^{(n)}(s_n, \omega_n)$, eq. (9.5-4), is analyzed.

Theorem 13-7: *Let us assume that the sequences* $\{s_n\}$ *and* $\{\omega_n\}$ *satisfy* (*S*-0)–(*S*-2) *and that Drummond's sequence transformation* $\mathscr{D}_k^{(n)}(s_n, \omega_n)$, *eq.* (9.5-4), *is used for the transformation of* $\{s_n\}$. *Then we obtain for fixed* $k \in \mathbb{N}$ *and for all* $n \in \mathbb{N}_0$ *the following estimate for the error term*:

$$|\mathscr{D}_k^{(n)}(\beta, s_n, \omega_n) - s| \le \frac{\omega_n}{(\beta+n)_{k+1}} \sum_{j=0}^{\infty} \frac{c_{j+1}(j+1)_k}{(\beta+n+k+1)_j}. \tag{13.2-20}$$

This implies for fixed $k \in \mathbb{N}$ *and for large values of n the following order estimate*:

$$\frac{\mathscr{D}_k^{(n)}(s_n, \omega_n) - s}{s_n - s} = \mathrm{O}(n^{-k-1}), \qquad n \to \infty. \tag{13.2-21}$$

This theorem, which can be proved in exactly the same way as theorem 13-5, indicates that in particular for larger values of k Drummond's sequence transformation $\mathscr{D}_k^{(n)}(s_n, \omega_n)$, eq. (9.5-4), should be significantly less powerful than $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2–7), or its mild generalization $\mathscr{S}_{k,l}^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2–8), with $l \ge 1$.

It would be interesting to do the same kind of error analysis also for the sequence transformation $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-6), and its mild generalization $\mathscr{M}_{k,l}^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-7). However, if we would try to estimate the error term of this transformation for arbitrary sequences $\{s_n\}$ and $\{\omega_n\}$ satisfying (S-0)–(S-2) we would in general end up with very complicated formulas which would contribute little to our understanding. This is due to the fact that for arbitrary β and γ we would have to use Leibniz' theorem for finite differences (see p. 35 of ref. [72]),

$$\Delta^k[f(n)g(n)] = \sum_{j=0}^{k} \binom{k}{j} [\Delta^j f(n)][\Delta^{k-j} g(n)]. \tag{13.2-22}$$

Much more revealing and enlightening is, however, the following observation:

Theorem 13-8: *Assume that* $\gamma = \beta + k - 2$ *holds. Then,*

$$\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n) = \mathscr{S}_k^{(n)}(\beta, s_n, \omega_n). \tag{13.2-23}$$

Proof: If we use the following relationship for Pochhammer symbols (see eq. (I.5) on p. 239 of ref. [104]),

$$(a-m)_m = (-1)^m (1-a)_m, \tag{13.2-24}$$

we obtain

$$(-\gamma-n)_{k-1} = (-1)^{k-1}(n+\gamma-k+2)_{k-1}. \tag{13.2-25}$$

If we insert this relationship into eq. (9.2-6) and use $\gamma = \beta + k - 2$, we obtain eq. (8.2-7).

Theorem 13-8 does not imply that the two strings $\mathscr{M}_j^{(n)}(\gamma, s_n, \omega_n)$ and $\mathscr{S}_j^{(n)}(\beta, s_n, \omega_n)$ with $0 \leq j \leq k$ are identical if $\gamma = \beta + k - 2$ holds. Only the last elements of the two strings are guaranteed to be identical, but not the others.

In the case of Levin's sequence transformation $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), an analysis of the magnitude of the error term as in theorem 13-5 would again be very complicated and would lead to lengthy and messy expressions. This is due to the fact that Levin's sequence transformation is constructed on the basis of the model sequence (7.1-1), which is merely a truncation of the inverse power series (13.2-5) after k terms, and that in the calculus of finite differences Pochhammer symbols and not powers are the most simple functions. However, at least some order estimates can be obtained relatively easily in the case of the Levin transformation.

Theorem 13-9: *Let us assume that the sequences* $\{s_n\}$ *and* $\{\omega_n\}$ *satisfy* $(S\text{-}0)$, $(S\text{-}1)$, *and eq.* (13.2-5) *and that Levin's sequence transformation* $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, *eq.* (7.1-7), *is used for the transformation of* $\{s_n\}$. *Then we obtain for large values of n and for fixed* $k \in \mathbb{N}$ *the following order estimate*:

$$\frac{\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n) - s}{s_n - s} = \mathrm{O}(n^{-2k}), \qquad n \to \infty. \tag{13.2-26}$$

Proof: Obviously, $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$ is invariant under translation according to eq. (3.1-4). This implies that the magnitude of the error term and its dependence upon k and n can be analyzed by estimating the magnitude of $\mathscr{L}_k^{(n)}(\beta, s_n - s, \omega_n)$. The starting point of our analysis is eq. (7.1-6) which is rewritten in the following way:

$$\mathscr{L}_k^{(n)}(\beta, s_n - s, \omega_n) = \frac{\Delta^k\{(\beta + n)^{k-1}(s_n - s)/\omega_n\}}{\Delta^k\{(\beta + n)^{k-1}/\omega_n\}}. \tag{13.2-27}$$

In the numerator in eq. (13.2-27) $(s_n - s)/\omega_n$ is replaced by the power series (13.2-5) yielding

$$\Delta^k \frac{(\beta + n)^{k-1}(s_n - s)}{\omega_n} = \Delta^k (\beta + n)^{k-1} \sum_{j=0}^{\infty} \frac{c_j'}{(\beta + n)^j} \tag{13.2-28}$$

$$= \Delta^k \sum_{j=0}^{k-1} c_j'(\beta + n)^{k-j-1} + \Delta^k \sum_{j=0}^{\infty} \frac{c_{k+j}'}{(\beta + n)^{j+1}}. \tag{13.2-29}$$

The first sum on the right-hand side of eq. (13.2-29) is annihilated by Δ^k since it is a polynomial of degree $k - 1$ in n. The large n behaviour of the second sum can be estimated with the help of the relationship

$$\Delta^m n^{-\alpha} = \mathrm{O}(n^{-\alpha - m}), \qquad \alpha > 0, \tag{13.2-30}$$

to give

$$\Delta^k \frac{(\beta + n)^{k-1}(s_n - s)}{\omega_n} = \mathrm{O}(n^{-k-1}), \qquad n \to \infty. \tag{13.2-31}$$

Since the remainder estimates $\{\omega_n\}$ are strictly alternating in sign according to (S-1), we obtain the following estimate for the denominator in eq. (13.2-27),

$$\left|(\beta+n)^{k-1}/\omega_n\right| \leq \left|\Delta^k\left\{(\beta+n)^{k-1}/\omega_n\right\}\right|. \tag{13.2-32}$$

If we combine eqs. (13.2-27), (13.2-31), and (13.2-32) and take into account that $(\beta+n)^{k-1} = O(n^{k-1})$ as $n \to \infty$, we see that theorem 13-9 is correct.

In the following theorem, which can be proved in essentially the same way as theorem 13-9, an order estimate for Levin's generalized sequence transformation $\mathscr{L}_{k,l}^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1–8), with $l \geq 1$ is derived.

Theorem 13-10: *Let us assume that the sequences* $\{s_n\}$ *and* $\{\omega_n\}$ *satisfy* (*S-0*), (*S-1*), *and eq.* (13.2-5) *and that Levin's generalized sequence transformation* $\mathscr{L}_{k,l}^{(n)}(\beta, s_n, \omega_n)$, *eq.* (7.1–8), *with* $l \geq 1$ *is used for the transformation of* $\{s_n\}$. *Then we obtain for fixed* k, $l \in \mathbb{N}$ *with* $k \geq l+1$ *and for large values of* n *the following order estimate*:

$$\frac{\mathscr{L}_{k,l}^{(n)}(\beta, s_n, \omega_n) - s}{s_n - s} = O(n^{l-2k}), \qquad n \to \infty. \tag{13.2-33}$$

A comparison of the order estimates (13.2-7), (13.2-19), (13.2-26), and (13.2-33) shows that Levin's sequence transformation $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), should be roughly comparable with $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), and that for fixed $l \geq 1$ Levin's generalized transformation $\mathscr{L}_{k,l}^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-8), should be roughly comparable with $\mathscr{S}_{k,l}^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-8). In addition, a comparison with the order estimate (13.2-21) shows that these sequence transformations should all be significantly more powerful than Drummond's sequence transformation $\mathscr{D}_k^{(n)}(s_n, \omega_n)$, eq. (9.5-4). A more detailed comparison cannot be made here since this would require additional knowledge about the sequence $\{s_n\}$ and the remainder estimates $\{\omega_n\}$.

The error analysis of this section is restricted to convergent or divergent sequences $\{s_n\}$ with strictly alternating remainder estimates $\{\omega_n\}$. This restriction is essential because otherwise the denominator sums of the pertaining sequence transformations cannot be estimated by a single term as it was for instance done in eqs. (13.2-15) and (13.2-32). If we want to analyze the transformation of sequences with nonalternating remainders, additional assumptions about the behaviour of the remainders have to be made. For instance, in the case of logarithmic convergence we could assume something like

$$\frac{(\beta+n)_{k-1}}{\omega_n} = \sum_{j=0}^{\infty} c_j \frac{\Gamma(\beta+n+k-1)}{\Gamma(\delta+n-j)}, \qquad \beta, \delta \in \mathbb{R}_+, \quad n \in \mathbb{N}_0, \tag{13.2-34}$$

because then the denominator of eq. (13.2-8) could be computed with the help of eq. (8.4-10) and we would obtain an explicit expression for the transformation error $\mathscr{S}_k^{(n)}(\beta, s_n - s, \omega_n)$. With the help of similar assumptions Sidi [56,105] could derive various error estimates for Levin's sequence transformation $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), in convergence acceleration and summation processes.

13.3. Summation of the Euler series

Padé approximants are generally accepted to be valuable numerical tools for the treatment of scientific problems. Therefore, it is certainly interesting to compare Padé approximants with the other sequence transformations of this report.

Unfortunately, the theoretical error estimates for the Padé summation of a Stieltjes series, which can be found in the literature, are not directly comparable with the error estimates of section 13.2, in which the error is always directly proportional to the remainder estimate ω_n. For instance, in the articles by Allen, Chui, Madych, Narcowich, and Smith [101], and by Karlsson and Sydow [102] the summation error is expressed in terms of polynomials which are orthogonal with respect to the measure $\psi(t)$ in the Stieltjes integral (13.1-1).

Consequently, we first would have to derive something like theorem 13-5 for Padé approximants before we could compare Padé approximants and the sequence transformations $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-6), and $\mathscr{D}_k^{(n)}(s_n, \omega_n)$, eq. (9.5-4), and their mild generalizations $\mathscr{L}_{k,l}^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-8), $\mathscr{S}_{k,l}^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-8), and $\mathscr{M}_{k,l}^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-7), with respect to their ability of summing divergent Stieltjes series or accelerating the convergence of some alternating series. Unfortunately, no such theorem could be derived which treats the Padé summation of an arbitrary Stieltjes series.

However, there is a notable exception. In the case of the Euler integral, eq. (1.1-6), and its associated asymptotic series, the so-called Euler series, eq. (1.1-7), Sidi [84] could show that their Padé approximants can be expressed in closed form via Drummond's sequence transformation $\mathscr{D}_k^{(n)}(s_n, \omega_n)$, eq. (9.5-4), with $\omega_n = a_{n+1}$. Sidi's proof is based upon the well-known fact that Wynn's ϵ algorithm, eq. (4.2-1), which according to eq. (4.2-10) is able to compute the Padé approximants $[n+k/k]$, is exact for the model sequence (4.1-3). In the case of the Euler series, eq. (1.1-7), we have

$$s_n = \sum_{\nu=0}^{n} (-1)^\nu \nu! z^\nu, \tag{13.3-1}$$

$$a_n = (-1)^n n! z^n. \tag{13.3-2}$$

If we insert these relationships into the model sequence (4.1-3), we obtain

$$s_n = s + \sum_{j=0}^{k-1} c_j (-1)^{n+j+1} (n+j+1)! z^{n+j+1}. \tag{13.3-3}$$

This model sequence for the Padé approximants $[n+k/k]$ of the Euler series can be rewritten in the following way:

$$\frac{s_n - s}{(-1)^{n+1}(n+1)! z^{n+1}} = \sum_{j=0}^{k-1} c_j (-1)^j z^j (n+2)_j. \tag{13.3-4}$$

The sum on the right-hand side of eq. (13.3-4) is a polynomial of degree $k-1$ in n. Consequently, it can be annihilated by the difference operator Δ^k. Hence, it follows from eq. (9.5-2) that the Padé approximants $[n+k/k]$ for the Euler series can be expressed in closed form in terms of Drummond's sequence transformation, eq. (9.5-4),

$$[n+k/k] = \epsilon_{2k}^{(n)} = \mathscr{D}_k^{(n)}(s_n,\ a_{n+1}), \qquad k,\ n \in \mathbb{N}_0. \tag{13.3-5}$$

From this relationship we may conclude that in the case of the Euler series Drummond's sequence transformation is much more efficient than Wynn's ϵ algorithm. If the diagonal Padé approximant $[n/n]$, which according to eq. (4.1-8) is a ratio of two polynomials $p_n(z)$ and $q_n(z)$ of degree n in z, is computed with the help of Wynn's ϵ algorithm as $\epsilon_{2n}^{(0)}$, the partial sums $s_0, s_1, \ldots, s_{2n}$ of the Euler series will be needed. If the same diagonal Padé approximant $[n/n]$ is computed as $\mathscr{D}_n^{(0)}(s_0,\ a_1)$, then according to eq. (9.5-4) only the partial sums $s_0, s_1, \ldots, s_{n+1}$ will be needed. Consequently, in the case of the Euler series, Drummond's sequence transformation is approximately twice as efficient as Wynn's ϵ algorithm.

How can this behaviour be explained? If Drummond's sequence transformation is applied to a sequence of partial sums of the Euler series, then we find that $\mathscr{D}_n^{(0)}(s_0,\ a_1)$ is the ratio of two polynomials $p_n(z)$ and $q_n(z)$ of degree n in z. However, it follows from eq. (9.5-4) that the $2n+2$ coefficients of the two polynomials are not all independent. In fact, these two polynomials $p_n(z)$ and $q_n(z)$ are completely determined by the $n+2$ terms $a_0, a_1, \ldots, a_{n+1}$. If the same ratio $p_n(z)/q_n(z)$ is computed via Wynn's ϵ algorithm, it is implicitly assumed that the $2n+2$ coefficients of the two polynomials are independent apart from a common normalization condition. This implies that Wynn's ϵ algorithm needs $2n+1$ independent conditions – in this case the $2n+1$ partial sums $s_0, s_1, \ldots, s_{2n}$ – for the construction of the ratio $p_n(z)/q_n(z)$.

Actually, it is a typical feature of all sequence transformation $T_k^{(n)}(s_n,\ \omega_n)$ of the type of eq. (12.2-2) that the coefficients of the numerator and denominator sums are not independent.

It is a natural idea to try to apply theorem 13-7, which gives an error estimate for Drummond's sequence transformation, also for of the Padé summation of the Euler series. Assumptions (S-0) and (S-1) are obviously satisfied. However, it is not clear whether and how a sequence $\{\omega_n\}$ of remainder estimates can be found such that assumption (S-2), which requires that $(s_n - s)/\omega_n$ can be represented as a factorial series according to eq. (13.2-4), is valid. No explicit proof for the existence of such a factorial series could be found in the case of the Euler series, if the remainder estimates were chosen according to

$$\omega_n = (-1)^{n+1}(n+1)!z^{n+1}, \qquad n \in \mathbb{N}_0. \tag{13.3-6}$$

Consequently, it can only be investigated numerically whether the error analysis of theorem 13-7 provides an adequate description of the Padé summation of the Euler series.

A close relative of the Euler integral, eq. (1.1-6), is the so-called exponential integral

$$E_1(z) = \int_z^\infty \frac{\mathrm{e}^{-x}}{x}\mathrm{d}x. \tag{13.3-7}$$

By means of some elementary operations we find:

$$z\, e^z\, E_1(z) = \int_0^\infty \frac{e^{-t}\, dt}{1 + t/z}. \tag{13.3-8}$$

If we compare this relationship with eq. (1.1-6) and also use eq. (1.1-7), we see that the associated Stieltjes series of the integral in eq. (13.3-8) is the Euler series with argument $1/z$,

$$z\, e^z\, E_1(z) \sim \sum_{m=0}^{\infty} (-1)^m m! z^{-m} = {}_2F_0(1, 1;\, -1/z), \qquad z \to \infty. \tag{13.3-9}$$

The radius of convergence of the hypergeometric series ${}_2F_0$ on the right-hand side of eq. (13.3-9) is zero, i.e., the series diverges quite rapidly for all finite values of z. Since reliable programs for the exponential integral $E_1(z)$ with $z \in \mathbb{R}_+$ are available, eq. (13.3-9) is well suited to test the ability of a sequence transformation of summing even wildly divergent series. In this report, the exponential integral $E_1(z)$ will be computed with the help of the routine S13AAF of the NAG FORTRAN library [106]. This function computes an approximation for the exponential integral in DOUBLE PRECISION (15–16 decimal digits) using appropriate Chebyshev expansions.

In table 13-1 the effect of Aitken's iterated Δ^2 process, eq. (5.1-15), of Drummond's sequence transformation, eq. (9.5-4), and of Wynn's ϵ algorithm, eq. (4.2-1), on the partial sums

$$s_n = \sum_{m=0}^{n} (-1)^m m! z^{-m}, \qquad n \in \mathbb{N}_0 \tag{13.3-10}$$

of the divergent series ${}_2F_0$ in eq. (13.3-9) with $z = 3$ is compared. In Drummond's sequence transformation, eq. (9.5-4), the remainder estimates are chosen according to eq. (13.1-16) which in this case means

$$\omega_n = a_{n+1} = (-1)^{n+1}(n+1)! z^{-n-1}, \qquad n \in \mathbb{N}_0. \tag{13.3-11}$$

The partial sums and the three different transforms in table 13-1 were computed in QUADRUPLE PRECISION (31–32 decimal digits). When these computations were repeated in DOUBLE PRECISION (15–16 digits) in order to study the numerical stability of the pertaining numerical processes, it turned out that the two computations agreed at least up to 12 decimal digits.

In all cases, the approximants were chosen in such a way that the information, which is contained in the finite string $s_0, s_1, \ldots, s_n$ of partial sums, is exploited optimally. This means that in the case of Aitken's iterated Δ^2 process, eq. (5.1-15), the approximants were chosen according to eq. (5.2-6), and in the case of Wynn's ϵ algorithm, eq. (4.2-1), they were chosen according to eq. (4.3-6).

A comparison of these three sequence transformations is quite interesting. Aitken's iterated Δ^2 process and Wynn's ϵ algorithm are closely related since they are both generalizations of Aitken's Δ^2 process, eq. (5.1-4), and one would like to know which one of these two generaliza-

Table 13-1
Summation of the asymptotic series ${}_2F_0(1, 1; -1/z) = z\,\mathrm{e}^z E_1(z)$ for $z = 3$

n	partial sum s_n eq. (13.1-10)	$\mathscr{A}_{[\![n/2]\!]}^{(n-2[\![n/2]\!])}$ eq. (5.1-15)	$\mathscr{D}_n^{(0)}(s_0, a_1)$ eq. (9.5-4)	$\epsilon_{2[\![n/2]\!]}^{(n-2[\![n/2]\!])}$ eq. (4.2-1)
10	$0.4831550069\times 10^{02}$	0.78625130019479	0.78625125348502	0.78626367674141
11	$-0.1770160037\times 10^{03}$	0.78625114835779	0.78625123263883	0.78624220653206
12	$0.7243100137\times 10^{03}$	0.78625122394910	0.78625122525386	0.78625447790898
13	$-0.3181436062\times 10^{04}$	0.78625121766831	0.78625122252501	0.78624881508686
14	$0.1504537896\times 10^{05}$	0.78625122089403	0.78625122147819	0.78625215335611
15	$-0.7608869613\times 10^{05}$	0.78625122063943	0.78625122106292	0.78625052018310
16	$0.4099597043\times 10^{06}$	0.78625122077179	0.78625122089311	0.78625150842397
17	$-0.2344314565\times 10^{07}$	0.78625122076057	0.78625122082175	0.78625100153477
18	$0.1418133105\times 10^{08}$	0.78625122076626	0.78625122079099	0.78625131522011
19	$-0.9048109119\times 10^{08}$	0.78625122076568	0.78625122077742	0.78625114787954
20	$0.6072683904\times 10^{09}$	0.78625122076597	0.78625122077131	0.78625125348502
21	$-0.4276977981\times 10^{10}$	0.78625122076594	0.78625122076850	0.78625119524201
22	$0.3154082874\times 10^{11}$	0.78625122076595	0.78625122076718	0.78625123263883
23	$-0.2430623561\times 10^{12}$	0.78625122076596	0.78625122076656	0.78625121141456
24	$0.1953763123\times 10^{13}$	0.78625122076595	0.78625122076626	0.78625122525386
25	$-0.1635311587\times 10^{14}$	0.78625122076596	0.78625122076611	0.78625121720071
26	$0.1423065021\times 10^{15}$	0.78625122076596	0.78625122076603	0.78625122252501
27	$-0.1285630059\times 10^{16}$	0.78625122076596	0.78625122076600	0.78625121935772
28	$0.1204177785\times 10^{17}$	0.78625122076596	0.78625122076598	0.78625122147819
29	$-0.1167898319\times 10^{18}$	0.78625122076596	0.78625122076597	0.78625122019177
30	$0.1171526266\times 10^{19}$	0.78625122076596	0.78625122076596	0.78625122106292
NAG function S13AAF		0.78625122076594	0.78625122076594	0.78625122076594

tions fares better. In addition, since the series (13.3-9) is the Euler series with argument $1/z$, the validity of eq. (13.3-5) can be checked numerically by comparing the results for Wynn's ϵ algorithm and Drummond's sequence transformation.

The clear winner in table 13-1 is Aitken's iterated Δ^2 process which produces 14 decimal digits after $n = 23$ (there is strong independent evidence that the last digit produced by the NAG function S13AAF in table 13-1 is incorrect and that Aitken's iterated Δ^2 process and Drummond's sequence transformation produce the correct result). It is followed by Drummond's sequence transformation, which reaches an accuracy of 14 decimal digits after $n = 30$, and the clear loser is Wynn's ϵ algorithm.

The results in table 13-1 show that eq. (13.3-5) is obviously valid in the case of the divergent series ${}_2F_0(1, 1; -1/z)$ because we find

$$\mathscr{D}_n^{(0)}(s_0, a_1) = \epsilon_{2n}^{(0)}. \tag{13.3-12}$$

Since $\mathscr{D}_{30}^{(0)}(s_0, a_1)$ is able to produce an accuracy of 14 decimal digits, it follows from eq. (13.3-12) that Wynn's ϵ algorithm will need the partial sums $s_0, s_1, \ldots, s_{60}$ of the asymptotic series in eq. (13.3-9) to produce the same accuracy.

In table 13-2 the same divergent series ${}_2F_0$ in eq. (13.3-9) with $z = 3$ is summed by the sequence transformations $d_n^{(0)}(\beta, s_0)$, eq. (7.3-9), and $\delta_n^{(0)}(\beta, s_0)$, eq. (8.4-4), with $\beta = 1$ and

Table 13-2
Summation of the asymptotic series ${}_2F_0(1, 1; -1/z) = z\,e^z E_1(z)$ for $z = 3$

n	partial sum s_n eq. (13.1-10)	$d_n^{(0)}(1, s_0)$ eq. (7.3-9)	$\delta_n^{(0)}(1, s_0)$ eq. (8.4-4)	$\Delta_n^{(0)}(17, s_0)$ eq. (9.4-4)
3	$0.6666666667\times10^{00}$	0.78709677419355	0.78672985781991	0.78633660627852
4	$0.9629629630\times10^{00}$	0.78607714016933	0.78622197922362	0.78625813355638
5	$0.4691358025\times10^{00}$	0.78628225839245	0.78625036724446	0.78625167667778
6	$0.1456790123\times10^{01}$	0.78624675493384	0.78625141640628	0.78625123654802
7	$-0.8477366255\times10^{00}$	0.78625162955159	0.78625123162756	0.78625121997903
8	$0.5297668038\times10^{01}$	0.78625123599599	0.78625121903376	0.78625122068020
9	$-0.1313854595\times10^{02}$	0.78625120523222	0.78625122051031	0.78625122077447
10	$0.4831550069\times10^{02}$	0.78625122396512	0.78625122077239	0.78625122076641
11	$-0.1770160037\times10^{03}$	0.78625122056582	0.78625122077131	0.78625122076576
12	$0.7243100137\times10^{03}$	0.78625122068924	0.78625122076646	0.78625122076598
13	$-0.3181436062\times10^{04}$	0.78625122079175	0.78625122076590	0.78625122076596
14	$0.1504537896\times10^{05}$	0.78625122076354	0.78625122076593	0.78625122076595
15	$-0.7608869613\times10^{05}$	0.78625122076528	0.78625122076595	0.78625122076596
16	$0.4099597043\times10^{06}$	0.78625122076622	0.78625122076596	0.78625122076596
17	$-0.2344314565\times10^{07}$	0.78625122076593	0.78625122076596	0.78625122076596
18	$0.1418133105\times10^{08}$	0.78625122076595	0.78625122076596	0.78625122076596
NAG function S13AAF		0.78625122076594	0.78625122076594	0.78625122076594

$\Delta_n^{(0)}(\gamma, s_0)$, eq. (9.4-4), with $\gamma = 17$. These three sequence transformations use the same remainder estimate (13.3-11) as $\mathscr{D}_n^{(0)}(s_0, a_1)$ in table 13-1.

Table 13-2 was also produced in QUADRUPLE PRECISION. When this computation was repeated in DOUBLE PRECISION, it turned out that in the last two columns all 14 digits agreed. Only in the case of the Levin transformation $d_n^{(0)}(\beta, s_0)$, eq. (7.3-9), it occasionally happened that the last digit disagreed. Thus, numerical instabilities are no problem here.

A comparison of the results in tables 13-1 and 13-2 confirms the error analysis in section 13.2, which indicates that Drummond's sequence transformation $\mathscr{D}_k^{(n)}(s_n, \omega_n)$, eq. (9.5-4), should be significantly less powerful than the sequence transformations $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), and $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-6), if the same remainder estimates $\{\omega_n\}$ are used. Even $d_n^{(0)}(\beta, s_0)$, which is somewhat weaker than the other two transformations in table 13-2, is clearly more powerful than the transformations in table 13-1, and both $\Delta_n^{(0)}(\gamma, s_0)$ and $\delta_n^{(0)}(\beta, s_0)$ are approximately twice as efficient as $\mathscr{D}_n^{(0)}(s_0, a_1)$. This observation is at least qualitatively in agreement with the order estimates (13.2-7) and (13.2-21).

In view of its slow convergence a Padé summation of the divergent series ${}_2F_0$ in eq. (13.3-9) does not seem feasible if its argument is significantly smaller than $z = 3$ as it was chosen in table 13-1. If, however, variants of the sequence transformations $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), and $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-6), are used, the summation of the divergent series in eq. (13.3-9) can be done even for relatively small arguments. Table 13-3 shows that the sequence transformations $d_n^{(0)}(\beta, s_0)$, eq. (7.3-9), and $\delta_n^{(0)}(\beta, s_0)$, eq. (8.4-4), with $\beta = 1$ and $\Delta_n^{(0)}(\gamma, s_0)$, eq. (9.4-4), with $\gamma = 29$ are able to sum the divergent series ${}_2F_0$ in eq. (13.3-9) with an accuracy of 14 decimal digits even if the argument of the series is as small as $z = 1/2$.

Table 13-3
Summation of the asymptotic series ${}_2F_0(1, 1; -1/z) = z\,e^z E_1(z)$ for $z = 1/2$

n	partial sum s_n eq. (13.1-10)	$d_n^{(0)}(1, s_0)$ eq. (7.3-9)	$\delta_n^{(0)}(1, s_0)$ eq. (8.4-4)	$\Delta_n^{(0)}(29, s_0)$ eq. (9.4-4)
15	$-0.4147067254\times10^{17}$	0.46145531715043	0.46145531958535	0.46145595366489
16	$0.1329725286\times10^{19}$	0.46145530923846	0.46145531701552	0.46145551453546
17	$-0.4529093729\times10^{20}$	0.46145531613431	0.46145531625982	0.46145536941468
18	$0.1633052915\times10^{22}$	0.46145531735759	0.46145531613493	0.46145532757622
19	$-0.6214401349\times10^{23}$	0.46145531627646	0.46145531616450	0.46145531778365
20	$0.2488938643\times10^{25}$	0.46145531605612	0.46145531620445	0.46145531622965
21	$-0.1046565329\times10^{27}$	0.46145531622971	0.46145531622787	0.46145531618769
22	$0.4609744216\times10^{28}$	0.46145531627375	0.46145531623807	0.46145531623838
23	$-0.2122526902\times10^{30}$	0.46145531624564	0.46145531624153	0.46145531624494
24	$0.1019714416\times10^{32}$	0.46145531623631	0.46145531624231	0.46145531624191
25	$-0.5102726985\times10^{33}$	0.46145531624080	0.46145531624227	0.46145531624156
26	$0.2655415912\times10^{35}$	0.46145531624283	0.46145531624210	0.46145531624194
27	$-0.1434925159\times10^{37}$	0.46145531624214	0.46145531624197	0.46145531624189
28	$0.8040791666\times10^{38}$	0.46145531624170	0.46145531624191	0.46145531624184
29	$-0.4666476909\times10^{40}$	0.46145531624180	0.46145531624188	0.46145531624188
30	$0.2801466126\times10^{42}$	0.46145531624189	0.46145531624187	0.46145531624187
NAG function S13AAF		0.46145531624187	0.46145531624187	0.46145531624187

In the case of table 13-3 it is essential to use QUADRUPLE PRECISION. In DOUBLE PRECISION, a heavy loss of significant digits occurs. The best results in DOUBLE PRECISION are obtained by $\Delta_n^{(0)}(\gamma, s_0)$ for $n = 20$ (10 decimal digits). For larger values of n, the accuracy of the results deteriorates rapidly due to numerical instabilities, leading to nonsensical results for the Levin transformation $d_{30}^{(0)}(\beta, s_0)$ and to only 3 digits accuracy for $\delta_{30}^{(0)}(\beta, s_0)$ and $\Delta_{30}^{(0)}(\gamma, s_0)$.

If the other variants of the sequence transformations $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), and $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-6), which are based upon the remainder estimates (7.3-4), (7.3-6), (7.3-10), and (9.4-1), are used for the summation of the divergent series ${}_2F_0$ in eq. (13.3-9), it turns out that these transformations are roughly comparable with $d_n^{(0)}(\beta, s_0)$, $\delta_n^{(0)}(\beta, s_0)$, or $\Delta_n^{(0)}(\gamma, s_0)$, which were used in tables 13-2 and 13-3.

The other sequence transformations of this report do not sum the divergent series ${}_2F_0$ in eq. (13.3-9) as efficiently as the transformations mentioned above. For instance, Brezinski's ϑ algorithm, eq. (10.1-9), or other transformations, which are based upon the ϑ algorithm, as for instance $\mathscr{J}_k^{(n)}$, eq. (10.3-6), $\mathscr{B}_k^{(n)}$, eq. (11.1-5), and $\mathscr{C}_k^{(n)}$, eq. (11.1-12), all rank between Aitken's iterated Δ^2 algorithm, eq. (5.1-15), and Drummond's sequence transformation, eq. (9.5-4), with respect to their ability of summing the divergent series in eq. (13.3-9). The sequence transformations $\lambda_k^{(n)}$, eq. (11.2-1), $\sigma_k^{(n)}$, eq. (11.2-2), and $\mu_k^{(n)}$, eq. (11.2-3), sum the divergent series ${}_2F_0$ in eq. (13.3-9) slightly less efficient than Drummond's sequence transformation.

It is a remarkable fact that compared with $d_n^{(0)}(\beta, s_0)$, eq. (7.3-9), $\delta_n^{(0)}(\beta, s_0)$, eq. (8.4-4), or $\Delta_n^{(0)}(\gamma, s_0)$, eq. (9.4-4), which were used in tables 13-2 and 13-3, the Padé summation of the divergent series ${}_2F_0$ in eq. (13.3-9) is hopelessly inefficient, even if the Padé approximants are computed via Drummond's sequence transformation according to eq. (13.3-5) and not via Wynn's ϵ algorithm.

How can this inferiority of Padé approximants be explained? It was remarked earlier, that the Padé approximants $[n+k/k]$ for the Euler series can be constructed on the basis of the model sequence (13.3-5). The remainder r_n of this model sequence is of order $O(z^{n+k}n^{n+k})$ as $n \to \infty$. However, it follows from theorem 13-2 that the remainder integral $R_n(z)$ of the Euler series (1.1-7) with $z \in \mathbb{R}_+$ is rigorously bounded by $(n+1)!z^{n+1}$ which is only of order $O(z^{n+1}n^{n+1})$ as $n \to \infty$. Hence, we see that the remainder of the model sequence (13.3-3) for Padé approximants $[n+k/k]$ yields unrealistically large estimates in the case of a wildly divergent series such as the Euler series. Consequently, it is to be expected that in the case of such a wildly divergent series Padé approximants will be less efficient than sequence transformations which use tighter remainder estimates as for instance suitable variants of $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), and $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-6).

There is considerable numerical evidence that the inferiority of Padé approximants in summation processes is not restricted to the Euler series. For instance, the numerical tests performed by Smith and Ford [29,30], who also considered the summation of several alternating divergent series, showed that Levin's u transformation, eq. (7.3-5), is significantly more powerful than Wynn's ϵ algorithm.

The inferiority of Padé approximants in summation processes also becomes quite obvious in the case of the following class of auxiliary functions,

$$F_m(z) = \int_0^1 u^{2m} \, e^{-zu^2} \, du, \qquad m \in \mathbb{N}_0, \quad z \in \mathbb{R}_+. \tag{13.3-13}$$

These auxiliaray functions $F_m(z)$ are of considerable importance in molecular *ab initio* calculations with Gaussian-type basis functions since the nuclear attraction and interelectronic repulsion integrals are ultimately expressed in terms of these functions. In molecular calculations these auxiliary functions have to be computed over a wide range of parameters m and arguments z so frequently that it amounts to a significant part of the whole integral evaluation time. In the case of larger arguments z, it is recommendable to compute this auxiliary function via its asymptotic expansion,

$$F_m(z) \sim \frac{\Gamma(m+1/2)}{2z^{m+1/2}} - \frac{e^{-z}}{2z} \, {}_2F_0(1, 1/2-m; -1/z), \qquad z \to \infty. \tag{13.3-14}$$

In ref. [63] it was shown that Levin's d transformation, eq. (7.3-9), sums this divergent series much more efficiently than Wynn's ϵ algorithm. Later, in ref. [107] the effect of the sequence transformations $u_k^{(n)}(\beta, s_n)$, eq. (7.3-5), $y_k^{(n)}(\beta, s_n)$, eq. (8.4-2), and $Y_k^{(n)}(\gamma, s_n)$, eq. (9.4-2), on the divergent series ${}_2F_0$ in eq. (13.3-14) was compared. Similarly as in the case of the divergent series ${}_2F_0$ in eq. (13.3-9) it was found that Levin's u transformation is slightly less efficient than the analogous new sequence transformations $y_k^{(n)}(\beta, s_n)$ and $Y_k^{(n)}(\gamma, s_n)$.

13.4. A Stieltjes series with a finite radius of convergence

Let us consider the following integral representation for the logarithm which is defined for all

z belonging to the cut complex plane which is cut along $-\infty < z \leq -1$,

$$\frac{1}{z}\ln(1+z) = \int_0^1 \frac{\mathrm{d}t}{1+zt}. \tag{13.4-1}$$

The integral in eq. (13.4-1) is a Stieltjes integral as the one in eq. (13.1-1). To see this we only have to set $\psi(t) = t$ for $0 \leq t \leq 1$ and $\psi(t) = 1$ for $1 < t < \infty$ in eq. (13.1-1). The moments μ_m of this positive measure $\psi(t)$ are given by

$$\int_0^\infty t^m \,\mathrm{d}\psi(t) = \int_0^1 t^m \,\mathrm{d}t = \frac{1}{m+1}, \qquad m \in \mathbb{N}_0. \tag{13.4-2}$$

If we use these moments μ_m in eq. (13.1-3), we obtain the following power series for the logarithm which is by construction a Stieltjes series:

$$\ln(1+z) = \sum_{m=0}^{\infty} \frac{(-1)^m z^{m+1}}{m+1} = z\,{}_2F_1(1, 1; 2, -z). \tag{13.4-3}$$

The power series in eq. (13.4-3) converges absolutely for all $z \in \mathbb{C}$ with $|z| < 1$, for $z = 1$ the series converges conditionally, and all for $z \in \mathbb{C}$ with $|z| > 1$ the series diverges. However, as long as the argument $z \in \mathbb{C}$ does not lie on the cut, the divergent series can at least in principle be summed.

It may be interesting to note that the infinite series (13.4-3) for ln(2) occurs also in solid state physics since it gives the Madelung constant of a 1-dimensional lattice of oppositely charged ions (see pp. 74–75 of ref. [108]). According to Killingbeck the infinite series (13.4-3) occurs also if correlation effects in atoms are treated via perturbation theory (see p. 969 of ref. [109]).

In the last section, it was demonstrated both theoretically and numerically that the sequence transformations $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), and $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-6), sum the wildly divergent series ${}_2F_0$ in eq. (13.3-9), which is essentially the Euler series, significantly more efficiently than Padé approximants. In addition, some arguments were presented which indicate that this inferiority of Padé approximants is not restricted to the Euler series (1.1-7) and will occur also in the case of other wildly divergent series. For $1 < z < \infty$, the sequence of partial sums of the Stieltjes series in eq. (13.4-3),

$$s_n = \sum_{m=0}^{n} \frac{(-1)^m z^{m+1}}{m+1}, \qquad n \in \mathbb{N}_0, \tag{13.4-4}$$

obviously diverges but not as wildly as the partial sums of the divergent series ${}_2F_0$ in eq. (13.3-9). Consequently, it should be interesting to investigate whether the striking superiority of the sequence transformations $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, and $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, over Padé approximants is also observed in the case of the Stieltjes series (13.4-3).

In this context it would of course be helpful to have some theoretical summation error estimates. In the case of the sequence transformations $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), and $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-6), this poses no problems. If the argument z of the Stieltjes series in eq. (13.4-3) is positive, the error analysis of section 13-2 can be used since the remainders of the power series are then strictly alternating.

In the case of Padé approximants the error analysis of section 13-2 cannot be applied because only the Padé approximants of the Euler series, eq. (1.1-7), can be computed via Drummond's sequence transformation. In Wimp's book [23] the effect of Wynn's ϵ algorithm, eq. (4.2-2), on the following model sequence is studied. The elements of this model sequence are defined by Poincaré-type asymptotic expansions in inverse powers of n,

$$s_n \sim s + \lambda^n n^{\vartheta} \sum_{j=0}^{\infty} c_j/n^j, \qquad c_0 \neq 0, \quad n \to \infty. \tag{13.4-5}$$

A sequence of the type of eq. (13.4-5) should be a reasonably good model for the behaviour of the partial sums s_n of the series (13.4-3) as $n \to \infty$. The sequence (13.4-5) obviously converges linearly if $|\lambda| < 1$ and it diverges if $|\lambda| > 1$. Assuming $\lambda \neq 1$ and $\vartheta \neq 0, 1, \ldots, k-1$ in eq. (13.4-5), Wimp obtained for fixed $k \in \mathbb{N}$ the following order estimate (see p. 127 of ref. [23]):

$$\epsilon_{2k}^{(n)} - s = \frac{c_0 \lambda^{n+2k} n^{\vartheta - 2k} k! (-\vartheta)_k}{(\lambda - 1)^{2k}} \left[1 + O\left(\frac{1}{n} \right) \right], \qquad n \to \infty. \tag{13.4-6}$$

Since $s_n - s \sim \lambda^n n^{\vartheta}$ as $n \to \infty$, we obtain from eq. (13.4-6) the following order estimate:

$$\frac{\epsilon_{2k}^{(n)} - s}{s_n - s} \sim O(n^{-2k}), \qquad n \to \infty. \tag{13.4-7}$$

The error estimate (13.4-6) shows quite clearly that for $|\lambda| < 1$ Wynn's ϵ algorithm accelerates the convergence of the linearly convergent sequence. It also follows from the error estimate (13.4-6) that the limit s can be determined more easily if λ is negative which is well in agreement with experience.

However, if we compare the order estimate (13.4-7) with the order estimates (13.2-7) and (13.2-26) for $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), and $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), respectively, we find that Wynn's ϵ algorithm should be significantly less efficient than the other two sequence transformations mentioned above. This follows from the fact that for the computation of $\epsilon_{2k}^{(n)}$, which according to eq. (13.4-7) gives an order estimate of order $O(n^{-2k})$, $2k+1$ sequence elements $s_n, s_{n+1}, \ldots, s_{n+2k}$ will be needed, whereas for the computation of $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, and $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, which also give an error estimate of order $O(n^{-2k})$, only $k+1$ sequence elements $s_n, s_{n+1}, \ldots, s_{n+k}$ will be needed.

In table 13-4 the effect of the sequence transformations $t_k^{(n)}(\beta, s_n)$, eq. (7.3-7), $\tau_k^{(n)}(\beta, s_n)$, eq. (8.4-3), with $\beta = 1$ and Wynn's ϵ algorithm, eq. (4.2-1), on the partial sums of the divergent Stieltjes series in eq. (13.4-3) with $z = 5$ are compared.

The results in table 13-4 are another striking example for the inferiority of Padé approximants in summation processes. The results also indicate that our conclusions concerning the efficiency of Padé approximants in summation processes, which were based upon a comparison of the order estimates in theorems 13-5 and 13-9 and in eq. (13.4-6), should at least be qualitatively correct.

Table 13-4
Summation of the divergent series $z\,{}_2F_1(1, 1; 2; -z) = \ln(1+z)$ for $z = 5$

n	partial sum s_n eq. (13.4-4)	$\epsilon_{2[\![n/2]\!]}^{(n-2[\![n/2]\!])}$ eq. (4.2-1)	$t_n^{(0)}(1, s_0)$ eq. (7.3-7)	$\tau_n^{(0)}(1, s_0)$ eq. (8.4-3)
10	$0.3639603183\times10^{07}$	1.79198007997771	1.79175951159974	1.79175959220168
11	$-0.1670544890\times10^{08}$	1.79159768463775	1.79175946864530	1.79175949178480
12	$0.7719479148\times10^{08}$	1.79179758764032	1.79175946794412	1.79175947333854
13	$-0.3587706103\times10^{09}$	1.79173348919423	1.79175946933559	1.79175946997338
14	$0.1675734598\times10^{10}$	1.79176609278102	1.79175946926071	1.79175946936268
15	$-0.7861008566\times10^{10}$	1.79175520193427	1.79175946922241	1.79175946925230
16	$0.3701778279\times10^{11}$	1.79176062438322	1.79175946922743	1.79175946923241
17	$-0.1749098431\times10^{12}$	1.79175875744767	1.79175946922828	1.79175946922884
18	$0.8289578584\times10^{12}$	1.79175967119854	1.79175946922806	1.79175946922819
19	$-0.3939413724\times10^{13}$	1.79175934919749	1.79175946922805	1.79175946922808
20	$0.1876711762\times10^{14}$	1.79175950460547	1.79175946922806	1.79175946922806
21	$-0.8960496379\times10^{14}$	1.79175944882296	1.79175946922806	1.79175946922806
22	$0.4286962951\times10^{15}$	1.79175947543322	1.79175946922805	1.79175946922806
23	$-0.2054830571\times10^{16}$	1.79175946573795	1.79175946922805	1.79175946922806
24	$0.9866098385\times10^{16}$	1.79175947031756	1.79175946922806	1.79175946922806
25	$-0.4744606005\times10^{17}$	1.79175946862827	1.79175946922806	1.79175946922806
FORTRAN function QLOG		1.79175946922806	1.79175946922806	1.79175946922806

Table 13-4 was produced in QUADRUPLE PRECISION. When the same computation was repeated in DOUBLE PRECISION, some loss of accuracy was observed. The best results were produced by $t_n^{(0)}(1, s_0)$ and $\tau_n^{(0)}(1, s_0)$ for n between 15 and 18 (approximately 11 decimal digits). For larger values of n the accuracy deteriorates.

The Stieltjes series (13.4-3) is not only suited to test the efficiency of a sequence transformation in summation processes. A very popular test case, which is frequently found in the literature, is the conditionally convergent series (13.4-3) for ln(2) which converges quite slowly. According to Bender and Orszag (see p. 372 of ref. [2]) about 7000 terms of the series in eq. (13.4-3) with $z = 1$ will be needed to compute ln(2) with a relative accuracy of 0.01 percent. The same sequence transformations as in table 13-4 accelerate the convergence of the series for ln(2) also in table 13-5. This time, the ϵ algorithm is comparatively successful since it only needs the partial sums (13.3-4) up to $n = 18$ to produce an accuracy of 14 decimal digits. However, the other two sequence transformations in table 13-5 are still significantly more powerful.

Table 13-5 was again produced in QUADRUPLE PRECISION. When this computation was repeated in DOUBLE PRECISION, no loss of accuracy was observed.

The potential of the Stieltjes series (13.4-3) for $\ln(1 + z)$ to test the performance of sequence transformations is not yet exhausted. If the argument z of the power series satisfies $-1 < z < 0$, all of its terms have the same sign. The convergence of this series will become quite bad if z approaches -1 because for $z = -1$ it becomes the series (1.1-2) for $\zeta(1)$ which diverges. Consequently, it should be interesting to find out whether and how well the convergence of the Stieltjes series (13.4-3) can be accelerated if its argument z is close to -1.

In this context, it would again be helpful to have some theoretical error estimates. In the case of Padé approximants this poses no problems. If we assume that the elements of the model

Table 13-5
Acceleration of the conditionally convergent series $z\,{}_2F_1(1, 1; 2; -z) = \ln(1+z)$ for $z = 1$

n	partial sum s_n eq. (13.4-4)	$\epsilon_{2[\![n/2]\!]}^{(n-2[\![n/2]\!])}$ eq. (4.2-1)	$t_n^{(0)}(1, s_0)$ eq. (7.3-7)	$\tau_n^{(0)}(1, s_0)$ eq. (8.4-3)
3	0.58333333333333	0.69047619047619	0.69313725490196	0.69321533923304
4	0.78333333333333	0.69333333333333	0.69314393939394	0.69314971751412
5	0.61666666666667	0.69308943089431	0.69314740192831	0.69314726571364
6	0.75952380952381	0.69315245478036	0.69314717779003	0.69314718328808
7	0.63452380952381	0.69314574314574	0.69314718001500	0.69314718064517
8	0.74563492063492	0.69314733235438	0.69314718060123	0.69314718056257
9	0.64563492063492	0.69314714248772	0.69314718055924	0.69314718056003
10	0.73654401154401	0.69314718496213	0.69314718055985	0.69314718055995
11	0.65321067821068	0.69314717951778	0.69314718055995	0.69314718055995
12	0.73013375513376	0.69314718068816	0.69314718055995	0.69314718055995
13	0.65870518370518	0.69314718053085	0.69314718055995	0.69314718055995
14	0.72537185037185	0.69314718056369	0.69314718055995	0.69314718055995
15	0.66287185037185	0.69314718055912	0.69314718055995	0.69314718055995
16	0.72169537978362	0.69314718056005	0.69314718055995	0.69314718055995
17	0.66613982422806	0.69314718055992	0.69314718055995	0.69314718055995
18	0.71877140317543	0.69314718055995	0.69314718055995	0.69314718055995
FORTRAN function QLOG		0.69314718055995	0.69314718055995	0.69314718055995

sequence (13.4-5) is still a good model for the behaviour of the partial sums (13.4-4) as $n \to \infty$, we may conclude from eq. (13.4-6) that Wynn's ϵ algorithm will accelerate the convergence of the sequence (13.4-5).

In the case of the sequence transformations $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), and $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-6), the situation is more complicated since the error analysis of section 13-2, which rests upon the assumption that the remainder estimates strictly alternate in sign, cannot be applied here. Consequently, we have to find estimates of the type of eq. (13.4-6) for the other sequence transformations mentioned above.

Theorem 13-11: *Let us assume that the elements of the sequence* $\{s_n\}$ *satisfy*

$$s_n = s + \lambda^n n^{\vartheta}\left[c_0 + O(n^{-1})\right], \qquad c_0 \neq 0, \quad \lambda \neq 0, 1, \quad n \to \infty, \tag{13.4-8}$$

that the elements of the sequence of remainder estimates $\{\omega_n\}$ *can be chosen in such a way that*

$$\omega_n = \lambda^n n^{\vartheta}\left[d_0 + O(n^{-1})\right], \qquad d_0 \neq 0, \quad n \to \infty, \tag{13.4-9}$$

and that the ratio $(s_n - s)/\omega_n$ *can for all* $n \in \mathbb{N}_0$ *be expressed as a factorial series,*

$$\frac{s_n - s}{\omega_n} = \sum_{j=0}^{\infty} \frac{\gamma_j}{(\beta+n)_j}, \qquad \beta \in \mathbb{R}_+. \tag{13.4-10}$$

If the sequence transformation $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, *eq.* (8.2-7), *is used for the transformation of* $\{s_n\}$, *we obtain for fixed* $k \in \mathbb{N}$ *the following order estimate*:

$$\frac{\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n) - s}{s_n - s} = \mathrm{O}(n^{-2k}), \qquad n \to \infty. \tag{13.4-11}$$

Proof: We can proceed as in theorem 13-5, i.e., the starting point for the proof of theorem 13-11 is eq. (13.2-8). Since eqs. (13.2-4) and (13.4-10) are identical we find that the numerator of this expression is also given by eq. (13.2-13) which is obviously of order $\mathrm{O}(n^{-k-1})$ as $n \to \infty$.

In order to obtain an estimate for the denominator $\Delta^k[(\beta+n)_{k-1}/\omega_n]$ we use (see eq. (41) on p. 21 of ref. [23])

$$\Delta^k[z^n n^\alpha] \sim z^n (z-1)^k n^\alpha, \qquad z \neq 1, \quad n \to \infty. \tag{13.4-12}$$

This relationship gives us the following asymptotic estimate for the denominator in eq. (13.2-8):

$$\Delta^k[(\beta+n)_{k-1}/\omega_n] \sim [1-\lambda]^k \lambda^{-n-k} n^{k-\vartheta-1}, \qquad n \to \infty. \tag{13.4-13}$$

If we combine the expressions for the numerator and the denominator and take into account that $s_n - s \sim \lambda^n n^\vartheta$ as $n \to \infty$, we obtain eq. (13.4-11).

In the next theorem, which can be proved in essentially the same way as theorem 13-11, it is shown that Levin's sequence transformation $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), also leads to an error estimate of order $\mathrm{O}(n^{-2k})$.

Theorem 13-12: *Let us assume that the elements of the sequence* $\{s_n\}$ *satisfy*

$$s_n = s + \lambda^n n^\vartheta [c_0 + \mathrm{O}(n^{-1})], \qquad c_0 \neq 0, \quad \lambda \neq 0, 1, \quad n \to \infty, \tag{13.4-14}$$

that the elements of the sequence of remainder estimates $\{\omega_n\}$ *can be chosen in such a way that*

$$\omega_n = \lambda^n n^\vartheta [d_0 + \mathrm{O}(n^{-1})], \qquad d_0 \neq 0, \quad n \to \infty, \tag{13.4-15}$$

and that the ratio $(s_n - s)/\omega_n$ *can for all* $n \in \mathbb{N}_0$ *be expressed as a power series of the following type*,

$$\frac{s_n - s}{\omega_n} = \sum_{j=0}^{\infty} \frac{\gamma_j'}{(\beta+n)^j}, \qquad \beta \in \mathbb{R}_+. \tag{13.4-16}$$

If the sequence transformation $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, *eq.* (7.1-7), *is used for the transformation of* $\{s_n\}$, *we obtain for fixed* $k \in \mathbb{N}$ *the following order estimate*:

$$\frac{\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n) - s}{s_n - s} = \mathrm{O}(n^{-2k}), \qquad n \to \infty. \tag{13.4-17}$$

Table 13-6
Acceleration of the absolutely convergent series $z\,{}_2F_1(1, 1; 2; -z) = \ln(1+z)$ for $z = -0.9$

n	partial sum s_n eq. (13.4-4)	$\epsilon_{2[\![n/2]\!]}^{(n-2[\![n/2]\!])}$ eq. (4.2-1)	$t_n^{(0)}(1, s_0)$ eq. (7.3-7)	$\tau_n^{(0)}(1, s_0)$ eq. (8.4-3)
15	−2.23201245730299	−2.30249119271252	−2.30258308878949	−2.30258507564758
16	−2.24182256418514	−2.30253886165435	−2.30258429564850	−2.30258508829481
17	−2.25016115503498	−2.30255980521704	−2.30258477584394	−2.30258509172157
18	−2.25727090091747	−2.30257257511841	−2.30258496686622	−2.30258509264961
19	−2.26334973364700	−2.30257828055989	−2.30258504284117	−2.30258509290084
20	−2.26856016170088	−2.30258170574611	−2.30258507305389	−2.30258509296883
21	−2.27303639307444	−2.30258325724815	−2.30258508506691	−2.30258509298723
22	−2.27688984443081	−2.30258417685642	−2.30258508984292	−2.30258509299220
23	−2.28021344622568	−2.30258459820761	−2.30258509174154	−2.30258509299355
24	−2.28308503817645	−2.30258484529651	−2.30258509249623	−2.30258509299391
25	−2.28557006967230	−2.30258495961133	−2.30258509279620	−2.30258509299401
26	−2.28772376363538	−2.30258502604192	−2.30258509291542	−2.30258509299404
27	−2.28959286232476	−2.30258505703207	−2.30258509296280	−2.30258509299404
28	−2.29121704463416	−2.30258507490093	−2.30258509298163	−2.30258509299404
29	−2.29263008324333	−2.30258508329701	−2.30258509298911	−2.30258509299405
30	−2.29386079429003	−2.30258508810542	−2.30258509299209	−2.30258509299405
FORTRAN function QLOG		−2.30258509299405	−2.30258509299405	−2.30258509299405

The error estimate (13.4-17) for Levin's sequence transformation has in principle already been derived by Sidi (see eq. (3.14) on p. 840 of ref. [105]).

In the same way, it can be proved that for fixed $k \in \mathbb{N}$ and for large values of n the sequence transformations $\mathscr{L}_{k,l}^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-8), and $\mathscr{S}_{k,l}^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-8), with $l \in \mathbb{N}$ lead to error estimates of order $O(n^{l-2k})$ and that Drummond's sequence transformation $\mathscr{D}_k^{(n)}(s_n, \omega_n)$, eq. (9.5-4), leads to an error estimate of order $O(n^{-k-1})$.

On the basis of these order estimates it is to be expected that in the case of negative arguments z the different variants of $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), and $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), should accelerate the convergence of the Stieltjes series in eq. (13.4-3) more efficiently than the analogous variants of the generalized transformations $\mathscr{L}_{k,l}^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-8), and $\mathscr{S}_{k,l}^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-8), with $l \geq 1$. Also, Drummond's sequence transformation, eq. (9.5-4), should be roughly as efficient as Wynn's ϵ algorithm, eq. (4.2-1).

In table 13-6 the convergence of the absolutely convergent Stieltjes series (13.4-3) with $z = -0.9$ is accelerated by the same sequence transformations as in tables 13-4 and 13-5. The inferiority of Wynn's ϵ algorithm is again evident.

Table 13-6 was produced in QUADRUPLE PRECISION. When the same computation was repeated in DOUBLE PRECISION, a heavy loss of accuracy was observed. No transformation was able to produce an accuracy of more than 8 decimal digits. These 8 digits were produced by $\tau_n^{(0)}(1, s_0)$ for $n = 17$, by $t_n^{(0)}(1, s_0)$ for $n = 21$, and by Wynn's ϵ algorithm for $n = 30$. For larger values of n the accuracy deteriorated again. However, it seems that Wynn's ϵ algorithm is not as much affected by numerical instablities as the other two sequence transformations in table 13-6.

More extensive numerical tests showed that in the case of the convergent or divergent Stieltjes series in eq. (13.4-3) $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), had a slight plus over $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq.

(7.1-7). A comparison of the numerous variants of the sequence transformations $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$ and $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$ showed that they differ in their ability of summing or accelerating the Stieltjes series in eq. (13.4-3). Those variants, which are based upon the remainder estimates (7.3-4) and (7.3-8), are approximately as efficient as $t_k^{(n)}(\beta, s_n)$, eq. (7.3-7), and $\tau_k^{(n)}(\beta, s_n)$, eq. (7.3-7). However, $v_k^{(n)}(\beta, s_n)$, eq. (7.3-11), and $\varphi_k^{(n)}(\beta, s_n)$, eq. (8.4-5), which are both based upon the remainder estimate (7.3-10), were significantly less efficient. Also, $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-6), and its variants were somewhat less efficient than the analogous variants of $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$ and $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$.

Almost as efficient as the sequence transformations mentioned above were $\mathscr{J}_k^{(n)}$, eq. (10.3-6), and $\mathscr{C}_k^{(n)}$, eq. (11.1-12), followed by Brezinski's ϑ algorithm, eq. (10.1-9). Next came Aitken's iterated Δ^2 process, eq. (5.1-15), which was again more efficient than Wynn's ϵ algorithm, eq. (4.2-1). The transformations $\mathscr{B}_k^{(n)}$, eq. (11.1-5), $\lambda_k^{(n)}$, eq. (11.2-1), $\sigma_k^{(n)}$, eq. (11.2-2), and $\mu_k^{(n)}$, eq. (11.2-3), were weaker than the ϵ algorithm.

Of all transformations tested Drummond's sequence transformation, eq. (9.5-4), was least efficient. It was able to sum or accelerate the Stieltjes series (13.4-3) moderately well for positive arguments z, i.e., as long as the terms of the power series for $\ln(1+z)$ had alternating signs. However, it failed completely if the argument z approached -1. For instance, for $z = -0.9$ Drummond's sequence transformation produced a sequence of transforms $\mathscr{D}_n^{(0)}(s_0, a_1)$ which rapidly diverged with increasing n. This example shows that asymptotic order estimates – although undeniably quite helpful for the classification of sequence transformations – do not necessarily tell the whole truth about the capability of a sequence transformation.

14. The acceleration of logarithmic convergence

14.1. Properties of logarithmically convergent sequences and series

It is tempting to believe that the most formidable task for a nonlinear sequence transformation is the summation of a wildly divergent series such as the Euler series, eq. (1.1-7), and that convergence acceleration should not be overly troublesome. In the case of alternating series or sequences with strictly alternating remainders, this is indeed normally true. However, it will become clear later in this section that the acceleration of the convergence of a monotonic sequence or a series with terms, that all have the same sign, can be a more formidable computational problem than the summation of an alternating divergent series.

The numerical examples presented in sections 13.3 and 13.4 showed that several sequence transformations are able to sum efficiently divergent series with alternating terms. In addition, it was shown that it is frequently possible to sum alternating divergent series with an accuracy that is close to or identical with machine accuracy. Particularly efficient and also remarkably reliable were variants of Levin's sequence transformation $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), and partly even more so variants of the new sequence transformations $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), and $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-6).

The situation is much less satisfactory if logarithmic convergence has to be accelerated. Many series with positive terms are known which converge so slowly that the evaluation of such a series by successively adding up the terms would be hopeless. In such a case, the use of a convergence acceleration method is indispensable. However, there is a considerable amount of theoretical and numerical evidence which indicates that convergence acceleration methods are generally less efficient and also more susceptible to rounding errors in the case of series with positive terms than in the case of alternating series. Consequently, it is often easier to sum an alternating divergent series, even if it diverges quite wildly, than to accelerate the convergence of a slowly convergent series with terms that all have the same sign.

A good example of a very slowly convergent series with positive terms is the series (1.1-2) for the Riemann zeta function. It is well known that this series converges for all $z \in \mathbb{C}$ with $\mathrm{Re}(z) > 1$. However, it follows from eqs. (7.3-12) and (7.3-14) that the remainder r_n of the series (1.1-2), which is defined by

$$r_n = \sum_{m=n+1}^{\infty} (m+1)^{-z}, \tag{14.1-1}$$

is of order $O(n^{1-z})$ as $n \to \infty$. Consequently, the computation of $\zeta(z)$ with the help of the series (1.1-2) would only be feasible if $\mathrm{Re}(z)$ is relatively large. But even then, the use of convergence acceleration techniques would be recommendable.

In order to make this section, in which the acceleration of logarithmic convergence by means of nonlinear sequence transformations will be treated, more selfcontained, first some properties of logarithmically convergent sequences and series are reviewed. A sequence $\{s_n\}$, which converges to some limit s, is said to converge logarithmically if

$$\lim_{n\to\infty} \frac{s_{n+1} - s}{s_n - s} = 1. \tag{14.1-2}$$

This definition of logarithmic convergence is inconvenient since it involves the limit s of the sequence $\{s_n\}$ which is normally not known. Thus, it would be advantageous to have an alternative criterion for logarithmic convergence which only involves the differences Δs_n. Such an equivalent criterion can be formulated if it is assumed that the elements of the sequence $\{s_n\}$ are partial sums of an infinite series with real terms a_m that all have the same sign. Then, Clark, Gray, and Adams could show (see theorem 2 on p. 26 of ref. [35]) that the sequence of partial sums converges logarithmically according to eq. (14.1-2) if the following condition holds:

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \frac{\Delta s_n}{\Delta s_{n-1}} = 1. \tag{14.1-3}$$

Eq. (14.1-3) implies that for larger indices n the terms a_n of a logarithmically convergent series differ only slightly. This fact is not only responsible for the often prohibitively slow convergence of logarithmically convergent series but also affects convergence acceleration processes in a very unpleasant way.

A sequence transformation can only beat the conventional process of successively adding up the terms of a series if it does not only use the numerical values of the terms. It also has to extract from the terms of the series some additional information about the behaviour of the partial sums s_n as n increases.

All sequence transformations of this report retrieve and utilize this additional information by computing rational expressions of weighted differences of partial sums. In the case of alternating series the computation of these weighted differences normally does not lead to a serious loss of significant digits. If, however, the terms a_n of a series all have the same sign and do not differ much in magnitude, this additional information, which has to be retrieved, is hidden somewhere in the later digits. Consequently, in the case of logarithmic convergence the computation of weighted differences is likely to lead to a cancellation of significant digits and ultimately, i.e., in the case of large transformation orders, completely nonsensical results are to be expected. This explains why rounding errors are more or less inevitable if logarithmic convergence is accelerated and why the acceleration or summation of alternating series is frequently remarkably stable. A good discussion of these stability problems can be found in an article by Longman [58].

A theoretical analysis of the acceleration of logarithmic convergence is also far from simple. For instance, it would certainly be quite helpful if an analogue of Germain-Bonne's theory of the acceleration of linear convergence [33] could be developed because then a decision based on some theoretical criteria could be made whether a given sequence transformation is able to accelerate logarithmic convergence or not. In the case of linear convergence this question can be decided on the basis of theorems 12-4 and 12-14. It only has to be shown that the sequence transformation under consideration is exact for a sequence which apart from a shift of indices consists of the partial sums of the geometric series, and it is guaranteed that linear convergence will be accelerated.

Smith and Ford had speculated whether it might be possible to develop an analogue of Germain-Bonne's theory also in the case of logarithmic convergence, i.e., whether some special sequence could be found such that the exactness of a sequence transformation for this sequence would imply that all logarithmically convergent sequences will be accelerated. They also had presented some potential candidates which in their opinion might possibly be suited to serve as this special sequence (see p. 238 of ref. [29]). In the meantime, this question has been answered by Delahaye and Germain-Bonne [110], but unfortunately the answer is negative. Delahaye and Germain-Bonne [110] showed that no algorithm exists which would be able to accelerate the convergence of every logarithmically convergent sequence. Consequently, a general theory in the spirit of Germain-Bonne's theory, which would cover the acceleration of all logarithmically convergent sequences, cannot exist. Such an analogue of Germain-Bonn's theory can exist at most for suitably restricted subsets of the set of logarithmically convergent sequences. But it seems that even this has not yet been accomplished so far.

14.2. Exactness results and error estimates

As mentioned in the last section, a theoretical analysis of the acceleration of logarithmic convergence is far from simple. Particularly hard is the analysis of the acceleration properties of those sequence transformations which are defined by a complicated nonlinear recursive scheme as for instance Brezinski's ϑ algorithm, eq. (10.1-9). In those cases, apart from the defining recursive scheme only very little else is normally known. However, at least for Levin's sequence transformation $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), some exactness results and asymptotic error estimates can be derived quite easily if suitable model sequences are considered.

Similar exactness results and error estimates as for Levin's sequence transformation can also be derived for the sequence transformations $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), and $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-6). However, they will not be considered here. The reason is that numerical tests showed that those variants of these sequence transformations, which should be able to accelerate logarithmic convergence, are significantly less efficient than the analogous variants of Levin's sequence transformation. For the moment, no completely satisfactory explanation can be given why these otherwise very powerful sequence transformations perform so weakly in the case of logarithmic convergence. It is at least conceivable that in the case of logarithmically convergent sequences inverse powers of n are better suited for a description of the n-dependence of the ratios $(s_n - s)/\omega_n$ than Pochhammer symbols, which are the basis of the sequence transformations $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), and $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-6). However, this is only speculation.

The explicit expressions of Levin's u transformation, eq. (7.3-5), and t transformation, eq. (7.3-7), are very similar. Also, with respect to the acceleration of linear convergence or the summation of alternating divergent series these two sequence transformations have virtually identical properties. However, in the literature on convergence acceleration it is always emphasized that Levin's u transformation is one of the best accelerators of logarithmic convergence, whereas Levin's t transformation completely fails to accelerate logarithmic convergence. In view of the otherwise close similarity of these two sequence transformations this different behaviour with respect to the acceleration of logarithmic convergence is certainly quite puzzling. It will now be shown that the different properties of Levin's u and t transformation can be understood on the basis of the different exactness properties of these two sequence transformations.

Quite common in practical applications are logarithmically convergent sequences $\{s_n\}$ with remainders $\{r_n\}$ that are of order $O(n^{-\alpha})$ with $\alpha \in \mathbb{R}_+$ as $n \to \infty$. Hence, for large values of n the elements of such a logarithmically convergent sequence can be characterized in the following way:

$$s_n = s + n^{-\alpha}\left[c + O(n^{-1})\right], \qquad c \neq 0, \quad \alpha \in \mathbb{R}_+, \quad n \to \infty. \tag{14.2-1}$$

The explicit expressions for $u_k^{(n)}(\beta, s_n)$, eq. (7.3-5), and $t_k^{(n)}(\beta, s_n)$, eq. (7.3-7), contain the terms a_n of the series which is to be transformed. If these two transformations are to be applied to sequences of the type of eq. (14.2-1), the terms a_n in the explicit expressions have to be replaced by the differences Δs_{n-1}. If we compute these differences and apply some simplifying assumptions, which are permitted if n is large, we see that sequences of the type of eq. (14.2-1) satisfy:

$$(s_n - s)/\Delta s_{n-1} = O(n), \qquad n \to \infty. \tag{14.2-2}$$

This relationship is quite typical of logarithmically convergent sequences of the type of eq. (14.2-1). Also, eq. (14.2-2) is essentially identical with the remainder estimate (7.3-4), which is the basis of Levin's u transformation.

In the following theorem the exactness properties of Levin's u and t transformation for a special class of logarithmically convergent model sequences are analyzed. Model sequences belonging to this special class have the same behaviour as $n \to \infty$ as the dominant term of sequences of the type of eq. (14.2-1). Consequently, the following theorem makes it plausible why Levin's u transformation accelerates the convergence of sequences of the type of eq. (14.2-1), and why Levin's t transformation fails to accelerate the convergence of these sequences.

Theorem 14-1: *Assume that a sequence transformation $\mathscr{T}_k^{(n)}(s_n)$ is defined in the following way:*

$$\mathscr{T}_k^{(n)}(s_n) = \frac{\Delta^k [P_{k-1}(n) s_n / \Delta s_{n-1}]}{\Delta^k [P_{k-1}(n) / \Delta s_{n-1}]}, \qquad k, n \in \mathbb{N}_0. \tag{14.2-3}$$

$P_{k-1}(n)$ is for sufficiently large values of k a polynomial of degree $\leq k-1$ in n. Obviously, $\mathscr{T}_k^{(n)}(s_n)$ is defined as long as Δ^k does not annihilate $P_{k-1}(n)/\Delta s_{n-1}$, i.e., as long as $P_{k-1}(n)/\Delta s_{n-1}$ is not a polynomial of degree $\leq k-1$ in n.

Let us assume that the sequence $\{s_n\}$, which converges to some limit s, belongs to the domain of the sequence transformation $\mathscr{T}_k^{(n)}$ and that its elements satisfy for all $n \in \mathbb{N}_0$

$$(s_n - s)/\Delta s_{n-1} = \gamma n + \delta, \qquad \gamma, \delta \in \mathbb{R}, \quad \gamma \neq 0. \tag{14.2-4}$$

If for sufficiently large values of k the degree of $P_{k-1}(n)$ is exactly $k-1$, $\mathscr{T}_k^{(n)}(s_n)$ does not accelerate the convergence of $\{s_n\}$, and if $P_{k-1}(n)$ is a polynomial of degree $\leq k-2$ in n, $\mathscr{T}_k^{(n)}(s_n)$ is exact for $\{s_n\}$.

Proof: Since $\mathscr{T}_k^{(n)}(s_n)$ is obviously invariant under translation in the sense of eq. (3.1-4), we can write

$$\mathscr{T}_k^{(n)}(s_n) = s + \frac{\Delta^k [P_{k-1}(n)(s_n - s)/\Delta s_{n-1}]}{\Delta^k [P_{k-1}(n)/\Delta s_{n-1}]}, \qquad k, n \in \mathbb{N}_0. \tag{14.2-5}$$

Next, the ratio $(s_n - s)/\Delta s_{n-1}$ in eq. (14.2-5) is replaced by $\gamma n + \delta$ according to eq. (14.2-4). This yields

$$\mathscr{T}_k^{(n)}(s_n) = s + \frac{\Delta^k [P_{k-1}(n)(\gamma n + \delta)]}{\Delta^k [P_{k-1}(n)/\Delta s_{n-1}]}, \qquad k, n \in \mathbb{N}_0, \quad \gamma \neq 0. \tag{14.2-6}$$

Let us now assume that k is large enough such that $P_{k-1}(n)$ is a polynomial of degree $\leq k-1$ in n. If the degree of $P_{k-1}(n)$ in eq. (14.2-6) is exactly $k-1$, the product $(\gamma n + \delta) P_{k-1}(n)$ is a polynomial of degree k in n. Consequently, this product will not be annihilated by Δ^k and $\mathscr{T}_k^{(n)}(s_n)$ will not accelerate $\{s_n\}$. If, however, $P_{k-1}(n)$ is of degree $\leq k-2$ in n, the product $(\gamma n + \delta) P_{k-1}(n)$ is a polynomial of degree $\leq k-1$ in n. Consequently, this product will be annihilated by Δ^k and $\mathscr{T}_k^{(n)}(s_n)$ is exact for $\{s_n\}$. This proves theorem 14-1.

In the case of Levin's u transformation, eq. (7.3-5), the polynomial $P_{k-1}(n)$ is given by $(\beta + n)^{k-2}$, i.e., it is a polynomial of degree $k-2$ in n. Consequently, $u_k^{(n)}(\beta, s_n)$ will be exact for sequences $\{s_n\}$ satisfying eq. (14.2-4). In the case of Levin's t transformation, eq. (7.3-7), the polynomial $P_{k-1}(n)$ is given by $(\beta + n)^{k-1}$, i.e., it is a polynomial of degree $k-1$ in n. This implies that $t_k^{(n)}(\beta, s_n)$ will not accelerate the convergence of a sequence satisfying eq. (14.2-4). It also follows from theorem 14-1 that for sufficiently large values of k Levin's generalized sequence transformation $\mathscr{L}_{k,l}^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-8), with $l \geq 2$ and $\omega_n = \Delta s_{n-1}$ is also exact for every sequence $\{s_n\}$ satisfying eq. (14.2-4).

A simple example of a logarithmically convergent sequence, which for all $n \in \mathbb{N}_0$ satisfies eq. (14.2-4), would be

$$s_n = s + \frac{(a)_{n+1}}{(b)_{n+1}}, \qquad a, b \in \mathbb{R}_+, \quad a < b. \tag{14.2-7}$$

In sections 13.2 and 13.4 error estimates for the summation of divergent Stieltjes series and the acceleration of the convergence of Stieltjes series by means of sequence transformations as for instance $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), or $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), were derived. It could be shown that the application of these sequence transformations to the partial sums of convergent or divergent Stieltjes series lead to asymptotic error estimates which were of order $O(n^{-2k})$ as $n \to \infty$. In the next theorem a similar asymptotic error analysis is done for a large class of logarithmically convergent sequences.

Theorem 14-2: *Let us assume that the elements of the sequence $\{s_n\}$, which converges logarithmically to some limit s, satisfy*

$$s_n = s + n^{-\alpha}\left[b_0 + O(n^{-1})\right], \qquad b_0 \neq 0, \quad \alpha \in \mathbb{R}_+, \quad n \to \infty. \tag{14.2-8}$$

Let us also assume that the elements of a sequence of remainder estimates $\{\omega_n\}$ can be chosen in such a way that

$$\omega_n = n^{-\alpha}\left[d_0 + O(n^{-1})\right], \qquad d_0 \neq 0, \quad n \to \infty, \tag{14.2-9}$$

and that the ratio $(s_n - s)/(\omega_n)$ can for all $n \in \mathbb{N}_0$ be expanded in a power series of the following type,

$$\frac{s_n - s}{\omega_n} = \sum_{j=0}^{\infty} \frac{c_j}{(\beta + n)^j}, \qquad \beta \in \mathbb{R}_+. \tag{14.2-10}$$

If the sequence transformation $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), is used for the acceleration of the convergence of $\{s_n\}$, we obtain for fixed $k \in \mathbb{N}$ and for $n \to \infty$ the following order estimate:

$$\frac{\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n) - s}{s_n - s} = O(n^{-k}), \qquad n \to \infty. \tag{14.2-11}$$

Proof: We can proceed as in theorem 13-9, i.e., the starting point for the proof of theorem 14-2 is the representation of the transformation error $\mathscr{L}_k^{(n)}(\beta, s_n - s, \omega_n)$ as in eq. (13.2-27). Since the series expansions (13.2-4) and (14.2-10) are structurally identical, we can conclude that according to eq. (13.2-31) the numerator of the transformation error is also of order $O(n^{-k-1})$ as $n \to \infty$.

In order to obtain an estimate for the denominator $\Delta^k[(\beta+n)^{k-1}/\omega_n]$ we take into account that according to eq. (14.2-9) the remainder estimate ω_n is of order $O(n^{-\alpha})$ as $n \to \infty$. This implies that $(\beta+n)^{k-1}/\omega_n$ is of order $O(n^{k+\alpha-1})$ as $n \to \infty$. Hence, with the help of eq. (13.2-30) we obtain the following order estimate for the denominator of the transformation error:

$$\Delta^k\left[(\beta+n)^{k-1}/\omega_n\right] = O(n^{\alpha-1}), \qquad n \to \infty. \tag{14.2-12}$$

If we combine this relationship with the expression for the numerator of the transformation error according to eq. (13.2-21), which is of order $O(n^{-k-1})$ as $n \to \infty$, we find that the transformation error $\mathscr{L}_k^{(n)}(\beta, s_n - s, \omega_n)$ is of order $O(n^{-\alpha-k})$ as $n \to \infty$. If we next divide the transformation error by $s_n - s$ and use eq. (14.2-8), we obtain eq. (14.2-11) which proves theorem 14-2.

An essentially identical result as theorem 14-2 was derived previously by Sidi (see theorem 4.2 on p. 320 of ref. [56]). A comparison of the asymptotic order estimate (14.2-11) with the analogous order estimates (13.2-26) and (13.4-17) in theorems 13-9 and 13-12, respectively, which are all of order $O(n^{-2k})$ as $n \to \infty$, shows that the acceleration of logarithmic convergence is indeed a much more formidable task than the acceleration of linear convergence, and it can be even harder than the summation of wildly divergent series.

In extensive numerical studies performed by Smith and Ford [29,30] it was demonstrated that Levin's u and v transformations, eqs. (7.3-5) and (7.3-11), respectively, are among the best accelerators for logarithmic convergence. Hence, we have to conclude that the relative inefficiency of Levin's sequence transformation according to theorem 14-2 is entirely due to the complicated nature of logarithmically convergent sequences and that it cannot be attributed to an intrinsic weakness of Levin's sequence transformation.

The next theorem, which can be proved in the same way as theorem 14-2, shows that Levin's generalized sequence transformation $\mathscr{L}_{k,l}^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-8), with $l \geq 1$, is also able to accelerate logarithmically convergent sequences of the type of eq. (14.2-8). However, with increasing $l \in \mathbb{N}$ the efficiency of the acceleration process deteriorates.

Theorem 14-3: *Let us assume that the sequences $\{s_n\}$ and $\{\omega_n\}$ are chosen as in theorem* 14-2. *If the sequence transformation $\mathscr{L}_{k,l}^{(n)}(\beta, s_n, \omega_n)$, eq.* (7.1-8), *is used for the acceleration of the convergence of $\{s_n\}$, we obtain for fixed k, $l \in \mathbb{N}$ with $k \geq l+1$ and for $n \to \infty$ the following order estimate*:

$$\frac{\mathscr{L}_{k,l}^{(n)}(\beta, s_n, \omega_n) - s}{s_n - s} = O(n^{l-k}), \qquad n \to \infty. \tag{14.2-13}$$

14.3. Some numerical test series

Due to the lack of theoretical criteria, by means of which it can be decided whether a given sequence transformation is able to accelerate logarithmic convergence or not, numerical testing will be of particular importance. In the literature on convergence acceleration the partial sums of the series expansion (1.1-2) for $\zeta(2)$,

$$\zeta(2) = \sum_{m=0}^{\infty} (m+1)^{-2} = \pi^2/6, \tag{14.3-1}$$

are frequently used to test the ability of a sequence transformation of accelerating logarithmic convergence. It follows from eqs. (7.3-12) and (7.3-14) that the remainders r_n of this series are of order $O(n^{-1})$ as $n \to \infty$, i.e., the convergence of the sequence of partial sums of this series is prohibitively slow and the series for $\zeta(2)$ indeed appears to be a good test problem for logarithmic convergence. However, the next theorem shows that it can happen that a sequence transformation is able to accelerate the convergence of a sequence if the remainders behave like an integral power of $1/n$ as $n \to \infty$ but fails to accelerate convergence if the remainders behave like a nonintegral power of $1/n$. Consequently, the partial sums of the series (14.3-1) for $\zeta(2)$ are not suited to test the ability of a sequence transformation of accelerating the convergence of a large class of logarithmically convergent sequences.

Theorem 14-4: *Let us assume that the linear sequence transformation $\Lambda_k^{(n)}(\beta, s_n)$, eq. (7.3-20), which corresponds to the special case $x_n = 1/(\beta + n)$ of the Richardson extrapolation scheme , eq. (6.1-5), is applied to the following logarithmically convergent model sequence:*

$$s_n = s + \sum_{j=0}^{\infty} \frac{c_j}{(\beta+n)^{\alpha+j}}, \qquad n \in \mathbb{N}_0, \quad \alpha, \beta \in \mathbb{R}_+, \quad c_0 \neq 0. \tag{14.3-2}$$

If α is a positive integer, i.e., $\alpha = 1, 2, \ldots$, we obtain for fixed $k \in \mathbb{N}$ and $k \geq \alpha$ and for large values of n the following asymptotic order estimate

$$\Lambda_k^{(n)}(\beta, s_n) - s = O(n^{-k-1}), \tag{14.3-3}$$

which shows that $\Lambda_k^{(n)}(\beta, s_n)$ accelerates the convergence of this sequence according to eq. (2.7-7) for sufficiently large values of k.

If, however, α is not a positive integer, $\Lambda_k^{(n)}(\beta, s_n)$ does not accelerate the convergence of the sequence (14.3-2).

Proof: Obviously, $\Lambda_k^{(n)}(\beta, s_n)$, eq. (7.3-20), is invariant with respect to translation according to eq. (3.1-4). Hence, with the help of eqs. (7.3-17) and (7.3-19) we can write:

$$\Lambda_k^{(n)}(\beta, s_n) - s = \frac{\Delta^k\left[(\beta+n)^k (s_n - s)\right]}{k!}. \tag{14.3-4}$$

Let us now assume that α is a positive integer, i.e., $\alpha = m$ with $m \in \mathbb{N}$, and that $k \geq m$. Then, with the help of eq. (14.3-2) we obtain for the numerator of the right-hand side of eq. (14.3-4):

$$\Delta^k\left[(\beta+n)^k (s_n - s)\right] = \Delta^k \sum_{j=0}^{k-m} c_j (\beta+n)^{k-m-j} + \Delta^k \sum_{j=0}^{\infty} c_{k-m+j+1} (\beta+n)^{-j-1}. \tag{14.3-5}$$

The first sum on the right-hand side is a polynomial of degree $k-m$ in n, i.e., it is annihilated by Δ^k, and according to eq. (13.2-30) the second sum will produce a result which is of order $O(n^{-k-1})$ as $n \to \infty$. This proves eq. (14.3-3).

Let us now assume that α is not a positive integer. Then, with the help of eqs. (13.2-30) and (14.3-2) we obtain for the numerator of the right-hand side of eq. (14.3-4) the following asymptotic estimate

$$\Delta^k\left[(\beta+n)^k(s_n-s)\right] = \Delta^k \sum_{j=0}^{\infty} c_j(\beta+n)^{k-\alpha-j} = O(n^{-\alpha}), \qquad n \to \infty, \tag{14.3-6}$$

which proves the second part of theorem 14-4.

Hence, if we want to use the Richardson extrapolation scheme, eq. (6.1-5), for the acceleration of the logarithmically convergent sequence (14.3-2) if α is not a positive integer, we cannot use the interpolation points $x_n = 1/(\beta+n)$ and we would at least have to find a different set of interpolation points $\{x_n\}$. There is considerable numerical evidence that the situation is quite analogous in the case of Wynn's ρ algorithm, eq. (6.2-2). The standard form of Wynn's ρ algorithm, eq. (6.2-4), corresponds to the choice $x_n = \beta + n$ for the interpolation points. It is together with its iteration $\mathscr{W}_k^{(n)}$, eq. (6.3-4), probably the best accelerator for the partial sums of the series (14.3-1) for $\zeta(2)$, but is apparently not able to accelerate the convergence of a sequence with remainders that behave like $n^{-1/2}$ as $n \to \infty$. However, it will be shown later that the general form of Wynn's ρ algorithm, eq. (6.2-2), is able to accelerate the convergence of sequences with remainders $r_n \sim n^{-1/2}$ as $n \to \infty$ if the interpolation points $\{x_n\}$ are chosen according to $x_n = (\beta+n)^{1/2}$ with $\beta > 0$.

We shall see later that for instance Brezinski's ϑ algorithm, eq. (10.1-9), or its iteration $\mathscr{J}_k^{(n)}$, eq. (10.3-6), are able to accelerate the convergence of sequences, whose remainders behave like $n^{-1/2}$ as $n \to \infty$. Consequently, $\vartheta_k^{(n)}$ and $\mathscr{J}_k^{(n)}$ are more flexible and versatile than the Richardson extrapolation scheme or Wynn's ρ algorithm since these sequence transformations only work if appropriate interpolation points $\{x_n\}$ are used.

This implies that because of theorem 14-4 the infinite series (14.3-1) for $\zeta(2)$ is not suited to test the ability of a sequence transformation of accelerating logarithmic convergence. Instead, one should use test problems with remainders that behave like a nonintegral power of $1/n$ as $n \to \infty$. A simple idea would be to use the infinite series (1.1-2) for $\zeta(z)$ for nonintegral arguments and not for $z = 2$. However, only if z is an even positive integer, a simple explicit expression for $\zeta(z)$ is known (see p. 19 of ref. [34]). Therefore, the use of the infinite series (1.1-2) for $\zeta(z)$ with $z \notin \mathbb{N}$ as a test problem for logarithmic convergence would be somewhat inconvenient and the emphasis in this report will be on other test problems.

Well suited for our purposes is the following series expansion (see p. 14 of ref. [111]):

$$A = \sum_{m=0}^{\infty} \frac{(2m-1)!!}{(2m)!!} \frac{1}{4m+1}. \tag{14.3-7}$$

Here, A stands for the so-called lemniscate constant which can be expressed in terms of the complete elliptic integral K (see pp. 358–359 of ref. [34]),

$$A = \int_0^1 \frac{\mathrm{d}t}{[1-t^4]^{1/2}} = 2^{-1/2}K(1/2) = \frac{[\Gamma(1/4)]^2}{4(2\pi)^{1/2}}. \tag{14.3-8}$$

If we use the following expression for the ratio of two gamma functions which holds for $|z| \to \infty$ (see p. 12 of ref. [34]),

$$\frac{\Gamma(z+\alpha)}{\Gamma(z+\beta)} = z^{\alpha-\beta}\left[1 + O(z^{-1})\right], \qquad |\arg(z)| < \pi, \tag{14.3-9}$$

we find that the terms of the series (14.3-7) behave like $n^{-3/2}$ as $n \to \infty$. Consequently, it follows from eqs. (7.3-12) and (7.3-14) that the remainders of this series behave like $n^{-1/2}$ as $n \to \infty$.

When Smith and Ford [29] investigated numerically the performance of sequence transformations in convergence acceleration and summation processes, they used the series (14.3-7) for the lemniscate constant A to test the ability of a sequence transformation of accelerating logarithmic convergence. Smith and Ford observed that the standard version of Wynn's ρ algorithm, eq. (6.2-4), was not able to accelerate the convergence of this series (see p. 235 of ref. [29]).

Another test problem, which is well suited for our purposes, is the following series expansion for $1/z$ in terms of the so-called reduced Bessel functions,

$$1/z = \sum_{m=0}^{\infty} \hat{k}_{m-1/2}(z)/[2^m m!], \qquad z \in \mathbb{R}_+. \tag{14.3-10}$$

This series expansion was derived and used in connection with explicit expressions for certain molecular multicenter integrals of exponentially declining basis functions (see eq. (6.5) of ref. [112]). In table I of ref. [64] it was shown that this series converges extremely slowly. For $z = 1$ the infinite series (14.3-10) produced an accuracy of only three decimal digits after 1 000 000 terms.

The so-called reduced Bessel function $\hat{k}_\nu(z)$ of real or complex order ν, which was introduced by Steinborn and Filter (see eqs. (3.1) and (3.2) of ref. [113]) as an exponentially declining basis function in electronic structure calculations, is defined by

$$\hat{k}_\nu(z) = (2/\pi)^{1/2} z^\nu K_\nu(z). \tag{14.3-11}$$

Here, $K_\nu(z)$ is a modified Bessel function of the second kind (see p. 66 of ref. [34]). If the order ν of the reduced Bessel function is half-integral and positive, $\nu = n + 1/2$ with $n \in \mathbb{N}_0$, a reduced Bessel function can be represented as an exponential multiplied by a terminating confluent hypergeometric series ${}_1F_1$ (see eq. (3.7) of ref. [114]),

$$\hat{k}_{n+1/2}(z) = 2^n (1/2)_n \, e^{-z} \, {}_1F_1(-n;\ -2n;\ 2z). \tag{14.3-12}$$

The polynomial part of these reduced Bessel functions with half-integral orders $\nu = n + 1/2$ with $n \in \mathbb{N}_0$ has also been investigated independently in the mathematical literature. There, the following notation is used (see p. 34 of ref. [115]):

$$\vartheta_n(z) = e^z \hat{k}_{n+1/2}(z) = 2^n (1/2)_n \, {}_1F_1(-n;\ -2n;\ 2z), \qquad n \in \mathbb{N}_0. \tag{14.3-13}$$

Together with some other, closely related polynomials, which are denoted by $y_n(z)$, the polynomials $\vartheta_n(z)$ are called Bessel polynomials. In Grosswald's book on Bessel polynomials [115] numerous applications of these polynomials in vastly differing fields are described. For instance, they are applied in number theory, in statistics, or for the analysis of complex electrical networks.

In the context of convergence acceleration and summation it may be interesting to note that Bessel polynomials occur also in the theory of Padé approximants. In the book by Baker and Graves-Morris it is shown that the Padé approximants $[l/m]$ for e^z are given by (see eq. (2.12) of part I of ref. [22])

$$[l/m] = \frac{{}_1F_1(-l; -l-m; z)}{{}_1F_1(-m; -l-m; -z)}, \qquad l, m \in \mathbb{N}_0. \tag{14.3-14}$$

Comparison of eqs. (14.3-13) and (14.3-14) shows that the diagonal elements $[n/n]$ of the Padé table for e^z can be expressed as ratios of Bessel polynomials,

$$[n/n] = \frac{\vartheta_n(z/2)}{\vartheta_n(-z/2)}, \qquad n \in \mathbb{N}_0. \tag{14.3-15}$$

With the help of some well known monotonicity properties of the modified Bessel function of the second kind, $K_\nu(z)$, it can be shown that the reduced Bessel functions $\hat{k}_\nu(z)$ are positive and bounded by their values at the origin provided that $\nu > 0$ and $z \geq 0$ (see eq. (3.1) of ref. [114]). In the case of reduced Bessel functions with half-integral orders this implies:

$$0 < \hat{k}_{n+1/2}(z) \leq \hat{k}_{n+1/2}(0) = 2^n(1/2)_n, \qquad 0 \leq z < \infty, \quad n \in \mathbb{N}_0. \tag{14.3-16}$$

Grosswald's book [115] also contains a chapter on the asymptotic properties of Bessel polynomials. There, it is shown that for fixed and finite argument z the Bessel polynomials $\vartheta_n(z)$ satisfy (see p. 125 of ref. [115])

$$\vartheta_n(z) \sim \frac{(2n)!}{2^n n!} e^z, \qquad n \to \infty. \tag{14.3-17}$$

If we combine eqs. (14.3-13) and (14.3-17) we find that the dominant term of the Poincaré-type asymptotic expansion in inverse powers of n of a reduced Bessel function $\hat{k}_{n+1/2}(z)$ with fixed and finite argument z corresponds to its value at the origin,

$$\hat{k}_{n+1/2}(z) = 2^n(1/2)_n[1 + O(n^{-1})] = \hat{k}_{n+1/2}(0)[1 + O(n^{-1})], \qquad n \to \infty. \tag{14.3-18}$$

Higher terms of the asymptotic expansion of a reduced Bessel function $\hat{k}_{n+1/2}(z)$ in inverse powers of n can in principle be obtained from related expansions for Bessel polynomials $y_n(z)$. In Grosswald's book the coefficients for terms up to an order $O(n^{-3})$ can be found (see p. 130 of ref. [115]) and in an article by Salzer [116] the coefficients for the terms up to an order $O(n^{-4})$.

Starting from eq. (14.3-18) it can be proved quite easily with the help of eq. (14.3-9) that the terms of the infinte series (14.3-10) behave like $n^{-3/2}$ as $n \to \infty$ (see p. 3709 of ref. [64]). In view of eqs. (7.3-12) and (7.3-14) this implies that the remainders of the partial sums of the series (14.3-10) behave like $n^{-1/2}$ as $n \to \infty$.

From the series expansion (14.3-10) in terms of reduced Bessel functions another series of the same convergence type can be derived. The new series is essentially the z-independent part of the infinite series (14.3-10). If we take into account that $\hat{k}_{-1/2}(z) = e^{-z}/z$, we can conclude from eq. (14.3-12) that only the first term of the infinite series (14.3-10) is singular at the origin. Consequently, the following limit exists,

$$\lim_{z \to 0} \left[1/z - \hat{k}_{-1/2}(z)\right] = \lim_{z \to 0} \sum_{m=1}^{\infty} \hat{k}_{m-1/2}(z)/[2^m m!] = 1, \tag{14.3-19}$$

and we obtain with the help of eq. (14.3-16):

$$\sum_{m=0}^{\infty} \frac{(2m-1)!!}{(2m+2)!!} = \frac{1}{2} \sum_{m=0}^{\infty} \frac{(1/2)_m}{(m+1)!} = 1. \tag{14.3-20}$$

Again, it follows from eq. (14.3-9) that the terms of this series behave like $n^{-3/2}$ as $n \to \infty$, which according to eqs. (7.3-12) and (7.3-14) implies that the remainders of this series behave like $n^{-1/2}$.

Several sequence transformations are exact for the infinite series (14.3-20) since its partial sums satisfy the prerequisites of theorem 14-1. With the help of a summation theorem by Gauss for a hypergeometric series ${}_2F_1$ with unit argument (see p. 40 of ref. [34]) it can be proved quite easily that the remainders of the infinite series (14.3-20) satisfy

$$\sum_{m=n+1}^{\infty} \frac{(2m-1)!!}{(2m+2)!!} = \frac{(1/2)_{n+1}}{(n+1)!}, \tag{14.3-21}$$

which shows that the partial sums of the infinite series (14.3-20) are of the type of the sequence (14.2-7). Hence, it follows from theorem 14-1 that Levin's u transformation, eq. (7.3-5), will only need the partial sums s_{n-1}, s_n, s_{n+1}, and s_{n+2} to be exact for the infinite series (14.3-20). Also, from eq. (10.3-4) it follows that Brezinski's ϑ algorithm, eq. (10.1-9), and its iteration $\mathcal{J}_k^{(n)}$, eq. (10.3-6), only need the partial sums s_n, s_{n+1}, s_{n+2}, and s_{n+3} to be exact for the infinite series (14.3-20).

Levin's v transformation, eq. (7.3-11), is even more efficient than the sequence transformations mentioned above because only the partial sums s_{n-1}, s_n, and s_{n+1} are needed to sum the infinite series (14.3-20) exactly. This follows from the fact that in the case of the infinite series (14.3-20) the remainder estimate (7.3-10), which is the basis for the v transformation,

$$\frac{a_{n+1} a_n}{a_{n+1} - a_n} = -\frac{1}{3} \frac{(1/2)_{n+1}}{(n+1)!}, \tag{14.3-22}$$

is proportional to the remainder (14.3-21) of the infinite series (14.3-20). Consequently, in this case the ratio $(s_n - s)/\omega_n$ is independent of n which implies that $(\beta + n)^{k-1}(s_n - s)/\omega_n$ will be annihilated by Δ^k for $k \geq 1$.

The terms of the three test series (14.3-7), (14.3-10), and (14.3-20) all behave like $n^{-3/2}$ as $n \to \infty$. Consequently, we may expect that these three series should be roughly comparable with respect to their rates of convergence as well as in convergence acceleration processes. However, at least in convergence acceleration processes these three test series are definitely not equivalent. The acceleration of the convergence of the series expansion (14.3-10) in terms of reduced Bessel functions is a much more formidable task, in particular for larger values of z, than the acceleration of the convergence of the other two series (14.3-7) and (14.3-20). In fact, for sufficiently large values of z it is virtually impossible to accelerate the convergence of the infinite series (14.3-10). This is probably a consequence of the exponential decline of the terms and also of the partial sums of this series. Numerical tests showed that for larger values of $z \in \mathbb{R}_+$ a reduced Bessel function is approximated by its value at the origin,

$$\hat{k}_{n+1/2}(z) \approx \hat{k}_{n+1/2}(0) = 2^n (1/2)_n, \qquad n \in \mathbb{N}_0, \tag{14.3-23}$$

with reasonable accuracy only if n is very large. For instance, if we require that eq. (14.3-23) should be accurate to one percent for $z = 8$ then we would need $n \geq 1400$, and for $z = 4$ we would still need $n \geq 400$. Consequently, in particular for larger values of z a partial sum s_n of the series (14.3-10) is essentially a linear combination of some quantities which decline exponentially, and only for relatively large values of n it can actually be observed that s_n behaves like $n^{-1/2}$ as $n \to \infty$. Thus, it is likely that for a sequence transformation, which tries to extract and utilize some regularity in the behaviour of the partial sums, the partial sums of the infinite series (14.3-10) appear for larger values of z to be much more irregular than the partial sums of the other two infinite series (14.3-7) and (14.3-20).

14.4. Numerical examples

In this section, the acceleration of logarithmically convergent sequences will be studied numerically. The emphasis will be on the test series (14.3-7), (14.3-10), and (14.3-20), which should converge approximately as slowly as the infinite series (1.1-2) for $\zeta(3/2)$. But since the infinite series (14.3-1) for $\zeta(2)$ is the most popular test problem for logarithmic convergence in the literature, it is of interest to see how in particular the new sequence transformations fare if they are applied to the partial sums

$$s_n = \sum_{m=0}^{n} (m+1)^{-2}, \qquad n \in \mathbb{N}_0, \tag{14.4-1}$$

of the infinite series (14.3-1) for $\zeta(2)$.

In table 14-1 the partial sums (14.4-1) are transformed by the standard form of $\mathscr{W}_k^{(n)}$, eq. (6.3-4), by $\Lambda_k^{(n)}(\beta, s_n)$, eq. (7.3-20), with $\beta = 1$, which corresponds to the special case $x_n = 1/(\beta + n)$ of the Richardson extrapolation scheme, eq. (6.1-5), and by $\mathscr{J}_k^{(n)}$, eq. (10.3-6). In all cases, the approximants were chosen in such a way that the information, which is contained in the finite string $s_0, s_1, \ldots, s_n$ of partial sums, is exploited in an optimal way. This means that in the case of $\mathscr{W}_k^{(n)}$, eq. (6.3-4), the approximations to ζ (2) were chosen according to eq. (6.3-9), and in the case of $\mathscr{J}_k^{(n)}$, they were chosen according to eq. (10.4-7).

Table 14-1
Acceleration of the series (14.3-1) for $\zeta(2) = \pi^2/6$

n	partial sum s_n eq. (14.4-1)	$\mathcal{W}_{[\![n/2]\!]}^{(n-2[\![n/2]\!])}$ eq. (6.3-4)	$\Lambda_n^{(0)}(1, s_0)$ eq. (7.3-20)	$\mathcal{J}_{[\![n/3]\!]}^{(n-3[\![n/3]\!])}$ eq. (10.3-6)
0	1.0000000000000	1.0000000000000	1.0000000000000	1.0000000000000
1	1.2500000000000	1.2500000000000	1.5000000000000	1.2500000000000
2	1.3611111111111	1.6500000000000	1.6250000000000	1.3611111111111
3	1.4236111111111	1.6468253968254	1.6435185185185	1.6388888888889
4	1.4636111111111	1.6449013949014	1.6449652777778	1.6423611111111
5	1.4913888888889	1.6449244489889	1.6449513888889	1.6436111111111
6	1.5117970521542	1.6449342449874	1.6449351851852	1.6449225865209
7	1.5274220521542	1.6449341126465	1.6449339434186	1.6449297924298
8	1.5397677311665	1.6449340660297	1.6449340411698	1.6449321959755
9	1.5497677311665	1.6449340666548	1.6449340662475	1.6449340557022
10	1.5580321939765	1.6449340668515	1.6449340671488	1.6449340629267
11	1.5649766384209	1.6449340668489	1.6449340668835	1.6449340652730
12	1.5708937981842	1.6449340668482	1.6449340668472	1.6449340668410
13	1.5759958390005	1.6449340668482	1.6449340668476	1.6449340668458
14	1.5804402834450	1.6449340668482	1.6449340668482	1.6449340668473
15	1.5843465334450	1.6449340668482	1.6449340668482	1.6449340668482
$\pi^2/6$		1.6449340668482	1.6449340668482	1.6449340668482

All sequence transformations in table 14-1 accelerate the convergence of the infinite series (14.3-1) for $\zeta(2)$ quite efficiently. The winner in table 14-1 is $\mathcal{W}_k^{(n)}$, eq. (6.3-4), which together with the standard form of Wynn's ρ algorithm, eq. (6.2-4), is the best accelerator for the series for $\zeta(2)$. Somewhat less efficient are $\Lambda_k^{(n)}(\beta, s_n)$, eq. (7.3-20), which in the case of the partial sums (14.4-1) is identical with Levin's u transformation, eq. (7.3-5), and $\mathcal{J}_k^{(n)}$, eq. (10.3-6).

Other good accelerators for the partial sums (14.4-1) are Levin's v transformation, eq. (7.3-11), which is as efficient as Levin's u transformation, eq. (7.3-5), and Brezinski's ϑ algorithm, eq. (10.1-9), and $\sigma_k^{(n)}$, eq. (11.2-2), which are almost as efficient as $\mathcal{J}_k^{(n)}$, eq. (10.3-6).

The partial sums and the three transforms in table 14-1 were computed in QUADRUPLE PRECISION (31–32 decimal digits). When these computations were repeated in DOUBLE PRECISION (15–16 decimal digits) the loss of some significant digits was observed. This is not surprising since the acceleration of logarithmic convergence is – as emphasized previously – an inherently unstable process. Of the three sequence transformations in table 14-1 it was again $\mathcal{W}_k^{(n)}$, eq. (6.3-4), which turned out to be the numerically most stable transformation since it lost at most 3 significant digits in DOUBLE PRECISION. The other two transformations, which lost up to 5 decimal digits, are apparently more sensitive to rounding errors.

Next, we want to see how the convergence of the test series (14.3-7), (14.3-10), and (14.3-20) can be accelerated. Since the remainders r_n of these series behave like $n^{-1/2}$ as $n \to \infty$, we expect that these series will converge significantly more slowly than the series (14.3-1) for $\zeta(2)$, whose remainders are of order $O(n^{-1})$. Here, it must be emphasized that it is not clear how and to what extent the slower convergence of the test series (14.3-7), (14.3-10), and (14.3-20) will affect convergence acceleration processes.

Table 14-2
Acceleration of the series (14.3-7) for the lemniscate constant A

n	partial sum s_n eq. (14.4-2)	$\mathscr{L}_{n,2}^{(0)}(1, s_0, a_0)$ eq. (7.1-8)	$u_n^{(0)}(1, s_0)$ eq. (7.3-5)	$\mathscr{J}_{[\![n/3]\!]}^{(n-3[\![n/3]\!])}$ eq. (10.3-6)
3	1.1657051282051	1.2190476190476	1.3163120567376	1.3037037037037
4	1.1817896870287	1.3343421605717	1.3108727079053	1.3080867850099
5	1.1935084370287	1.3103293923028	1.3109952008776	1.3095200070979
6	1.2025318745287	1.3108082123295	1.3110289627926	1.3110119624014
7	1.2097550695718	1.3110277257117	1.3110291499078	1.3110229739315
8	1.2157059973061	1.3110318476640	1.3110287979182	1.3110263202535
9	1.2207187157130	1.3110290080304	1.3110287737803	1.3110287611468
10	1.2250162047862	1.3110287461269	1.3110287766205	1.3110287718416
11	1.2287537180105	1.3110287708982	1.3110287771522	1.3110287750875
12	1.2320431110268	1.3110287771312	1.3110287771540	1.3110287771349
13	1.2349672811610	1.3110287772480	1.3110287771466	1.3110287771425
14	1.2375891404731	1.3110287771553	1.3110287771460	1.3110287771447
15	1.2399574101139	1.3110287771452	1.3110287771460	1.3110287771461
16	1.2421104860230	1.3110287771458	1.3110287771461	1.3110287771461
17	1.2440790912340	1.3110287771461	1.3110287771461	1.3110287771461
18	1.2458881405432	1.3110287771461	1.3110287771461	1.3110287771461
$[\Gamma(1/4)]^2/[4(2\pi)^{1/2}]$		1.3110287771461	1.3110287771461	1.3110287771461

First, we shall accelerate the convergence of the sequence of partial sums of the infinite series (14.3-7) for the lemniscate constant A,

$$s_n = \sum_{m=0}^{n} \frac{(2m-1)!!}{(2m)!!} \frac{1}{4m+1}. \tag{14.4-2}$$

In table 14-2 we see the effect of Levin's generalized sequence transformation $\mathscr{L}_{k,l}^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-8), with $l = 2$, $\omega_n = a_n$, and $\beta = 1$, of Levin's u transformation, eq. (7.3-5), with $\beta = 1$, and of $\mathscr{J}_k^{(n)}$ on the partial sums (14.4-2). As usual, the approximants were chosen in such a way that the information, which is contained in the finite string $s_0, s_1, \ldots, s_n$ of partial sums, is exploited optimally. This means that in the case of the Levin transformations the approximations to the lemniscate constant A were chosen according to eq. (7.5-4).

If we compare the results in tables 14-1 and 14-2, we see that the significantly slower convergence of the infinite series (14.3-7) does not affect the efficiency of convergence acceleration too much. The convergence of the transforms in table 14-2 is almost as fast as in table 14-1. However, it seems that the slower convergence of the test series (14.3-7) has a detrimental effect on the numerical stability of the transformations. As usual, table 14-2 was produced in QUADRUPLE PRECISION. When the same computations were repeated in DOUBLE PRECISION, a larger number of significant digits were lost than in table 14-1. The best results were obtained by Levin's u transformation which achieved a relative accuracy of 11 decimal digits after $n = 11$. For larger values of n, the accuracy deteriorated again. For instance, for $n = 18$ the best results were obtained by $\mathscr{J}_{12}^{(0)}$ which achieved an accuracy of 8 decimal digits.

If we compare eqs. (7.1-8) and (7.3-5), we find that Levin's u transformation may also be considered to be a special case of Levin's generalized sequence transformation $\mathscr{L}_{k,l}^{(n)}(\beta, s_n, \omega_n)$,

$$u_k^{(n)}(\beta, s_n) = \mathscr{L}_{k,1}^{(n)}(\beta, s_n, a_n). \tag{14.4-3}$$

This relationship shows that the numerical data in table 14-2 are in agreement with theorem 14-3 which predicts that the efficiency of Levin's generalized sequence transformation $\mathscr{L}_{k,l}^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-8), for the acceleration of the convergence of sequences of the type of eq. (14.2-8) should decrease with increasing $l \in \mathbb{N}$. This behaviour is apparently quite typical since it was also observed when the partial sums of the test series (14.3-10) were accelerated by $u_k^{(n)}(\beta, s_n)$, eq. (7.3-5), and $\mathscr{L}_{k,l}^{(n)}(\beta, s_n, a_n)$, eq. (7.1-8), with $l \geq 2$.

Other good sequence transformations for the test series (14.3-7) are Levin's v transformation, eq. (7.3-11), which is as good or even slightly better than Levin's u transformation, eq. (7.3-5), and Brezinski's ϑ algorithm, eq. (10.1-9), which is as good as $\mathscr{L}_{k,2}^{(n)}(\beta, s_n, a_n)$.

Theorem 14-4 predicts that $\Lambda_k^{(n)}(\beta, s_n)$, eq. (7.3-20), which corresponds to the special case $x_n = 1/(\beta + n)$ of the Richardson extrapolation scheme, eq. (6.1-5), is not able to accelerate the convergence of the series (14.3-7), (14.3-10), and (14.3-20). Numerical tests confirmed this prediction. In addition, the standard forms of Wynn's ρ algorithm, eq. (6.2-4), and of its iteration $\mathscr{W}_k^{(n)}$, eq. (6.3-4), also fail to accelerate the convergence of these series. Since the remainders of the test series mentioned above all behave like $n^{-1/2}$ as $n \to \infty$, and since the Richardson extrapolation scheme $\mathscr{N}_k^{(n)}(s_n, x_n)$, eq. (6.1-5), is by construction exact for the model sequence (6.1-6), it is an obvious idea to choose an alternative set of interpolation points $\{\!\{x_n\}\!\}$ according to

$$x_n = (\beta + n)^{-1/2}, \qquad n \in \mathbb{N}_0, \quad \beta \in \mathbb{R}_+. \tag{14.4-4}$$

Practical experience has shown that if the Richardson extrapolation scheme, eq. (6.1-5), can successfully handle a certain problem if the interpolation points $\{\!\{x_n\}\!\}$ are used, then the general forms of Wynn's ρ algorithm, eq. (6.2-2), and of its iteration $\mathscr{W}_k^{(n)}$, eq. (6.3-3), are usually able to handle the same problem if the interpolation points $\{\!\{\xi_n\}\!\}$ with $\xi_n = 1/x_n$ are used. Hence, if we want to use these nonlinear sequence transformations for the acceleration of the convergence of the series (14.3-7), (14.3-10), and (14.3-20), we should choose the elements of the set $\{\!\{\xi_n\}\!\}$ of interpolation points according to

$$\xi_n = (\beta + n)^{1/2}, \qquad n \in \mathbb{N}_0, \quad \beta \in \mathbb{R}_+. \tag{14.4-5}$$

In table 14-3 the partial sums (14.4-2) are accelerated by the Richardson extrapolation scheme, eq. (6.1-5), and by the general forms of Wynn's ρ algorithm, eq. (6.2-2), and of $\mathscr{W}_k^{(n)}$, eq. (6.3-3). In the case of the Richardson extrapolation scheme the interpolation points $\{\!\{x_n\}\!\}$ were chosen according to eq. (14.4-4) with $\beta = 1$, and in the case of $\rho_k^{(n)}$ and $\mathscr{W}_k^{(n)}$ the interpolation points $\{\!\{\xi_n\}\!\}$ were chosen according to eq. (14.4-5) with $\beta = 1$.

The most efficient transformation in table 14-3 is the general form of Wynn's ρ algorithm, eq. (6.2-2), followed by the Richardson extrapolation scheme, eq. (6.1-5), and the general form of

Table 14-3
Acceleration of the series (14.3-7) for the lemniscate constant A

n	partial sum s_n eq. (14.4-2)	$\mathscr{N}_n^{(0)}(s_0, x_0)$ eq. (6.1-5) $x_n = (n+1)^{-1/2}$	$\rho_{2[\![n/2]\!]}^{(n-2[\![n/2]\!])}$ eq. (6.2-2) $\xi_n = (n+1)^{1/2}$	$\mathscr{W}_{[\![n/2]\!]}^{(n-2[\![n/2]\!])}$ eq. (6.3-3) $\xi_n = (n+1)^{1/2}$
10	1.2250162047862	1.3110281470344	1.3110287489950	1.3110289097428
11	1.2287537180105	1.3110291844571	1.3110287903217	1.3110290181070
12	1.2320431110268	1.3110287720479	1.3110287927338	1.3110289480394
13	1.2349672811610	1.3110287371888	1.3110287908112	1.3110286302938
14	1.2375891404731	1.3110287841474	1.3110287783445	1.3110287340891
15	1.2399574101139	1.3110287800479	1.3110287774952	1.3110287694697
16	1.2421104860230	1.3110287759641	1.3110287770923	1.3110287666673
17	1.2440790912340	1.3110287770396	1.3110287771290	1.3110287675098
18	1.2458881405432	1.3110287772843	1.3110287771523	1.3110287670755
19	1.2475580797723	1.3110287771349	1.3110287771467	1.3110287674918
20	1.2491058660392	1.3110287771336	1.3110287771459	1.3110287670448
21	1.2505456974656	1.3110287771492	1.3110287771460	1.3110287770066
22	1.2518895646423	1.3110287771468	1.3110287771461	1.3110287770883
23	1.2531476731141	1.3110287771456	1.3110287771461	1.3110287771050
24	1.2543287710802	1.3110287771461	1.3110287771461	1.3110287770977
25	1.2554404064530	1.3110287771461	1.3110287771461	1.3110287770960
$[\Gamma(1/4)]^2/[4(2\pi)^{1/2}]$		1.3110287771461	1.3110287771461	1.3110287771461

$\mathscr{W}_k^{(n)}$, eq. (6.3-3). However, a comparison of tables 14-2 and 14-3 shows that even if we choose the interpolation points according to eqs. (14.4-4) and (14.4-5), the transformations in table 14-3 are clearly less efficient than Levin's u transformation or $\mathscr{J}_k^{(n)}$.

As usual, table 14-3 was produced in QUADRUPLE PRECISION. When the same computations were repeated in DOUBLE PRECISION, it was observed that the Richardson extrapolation scheme is much more unstable than the other two transformations. The Richardson extrapolation scheme achieved a relative accuracy of 8 decimal digits for $n = 12$. For larger values of n the accuracy of the transforms deteriorated rapidly, yielding totally nonsensical results for $n \geq 22$. The other two transformations also did not accomplish more than a relative accuracy of 7 or 8 decimal digits. However, they maintained this relative accuracy throughout the whole range of n between $n = 10$ and $n = 25$.

The results in table 14-3 and similar results for the other two test series (14.3-10) and (14.3-20) indicate that the Richardson extrapolation scheme, eq. (6.1-5), is able to accelerate the convergence of a sequence, whose remainders r_n behave like $n^{-1/2}$ as $n \to \infty$, if the interpolation points $\{\!\{x_n\}\!\}$ are chosen according to eq. (14.4-4). Similary, the general forms of Wynn's ρ algorthm, eq. (6.2-2), or of $\mathscr{W}_k^{(n)}$, eq. (6.3-3), should be able to accelerate the convergence of such a sequence if the interpolation points $\{\!\{\xi_n\}\!\}$ are chosen according to eq. (14.4-5).

Unfortunately, in practical applications these observations are not necessarily very helpful. Let us assume that only the numerical values of a few sequence elements are known but nothing about the behaviour of the remainders. In such a situation, it will be very hard or even impossible to find out whether the remainders of this sequence behave like an integral or like a nonintegral power of $1/n$. If one wants to use the Richardson extrapolation scheme, eq. (6.1-5), in such an unfavourable situation, it may be a good idea to follow a recommendation by Beleznay [117]. He suggested to choose the interpolation points according to

$$x_n = (n+\beta)^{-\alpha}, \qquad n \in \mathbb{N}_0, \quad \alpha, \beta \in \mathbb{R}_+, \tag{14.4-6}$$

and to optimize the exponent α of the interpolation points in such a way that the error term $|\mathscr{N}_{n-1}^{(1)} - \mathscr{N}_{n-1}^{(0)}|$ becomes minimal. This technique was later used by Liegener, Beleznay, and Ladik [118] to extrapolate the results of Hartree-Fock calculations on periodic chains. A similar approach is of course also possible in the case of the general forms of Wynn's ρ algorithm, eq. (6.2-2), or of its iteration $\mathscr{W}_k^{(n)}$, eq. (6.3-3). In that case, the interpolation points would have to be chosen according to

$$\xi_n = (n+\beta)^{\alpha}, \qquad n \in \mathbb{N}_0, \quad \alpha, \beta \in \mathbb{R}_+. \tag{14.4-7}$$

There is still another possibility of modifying either the Richardson extrapolation scheme, eq. (6.1-5), or the general forms of the rational transformations $\rho_k^{(n)}$, eq. (6.2-2), and $\mathscr{W}_k^{(n)}$, eq. (6.3-3), in such a way that it will be unimportant whether the remainder of the sequence to be transformed behaves like an integral or a nonintegral power of $1/n$. This approach is inspired by a nonlinear variant of the Richardson extrapolation scheme, which in Wimp's book is called GBW (Germain-Bonne Wimp) transformation (see p. 106 of ref. [23]). This GBW transformation is obtained from the Richardson extrapolation scheme by choosing the interpolation points according to

$$x_n = \Delta s_n = a_{n+1}, \qquad n \in \mathbb{N}_0. \tag{14.4-8}$$

If the interpolation points $\{x_n\}$ are chosen according to eq. (14.4-8), then it follows from eq. (6.1-6) that the Richardson extrapolation scheme, eq. (6.1-5), is exact for the following model sequence:

$$s_n = s + \sum_{j=0}^{k-1} c_j [\Delta s_n]^{j+1}, \qquad k, n \in \mathbb{N}_0, \quad \beta \in \mathbb{R}_+. \tag{14.4-9}$$

According to Wimp this GBW transformation works quite well in the case of iteration sequences (see pp. 106–108 of ref. [23]). However, the GBW transformation is apparently not able to accelerate logarithmic convergence.

Let us now assume that the remainders r_n of a sequence behave like $n^{-\alpha}$ as $n \to \infty$. Then it follows from eq. (14.2-2) that the product $[n \Delta s_{n-1}]$ also behaves like $n^{-\alpha}$ as $n \to \infty$. Hence, if we choose the interpolation points $\{x_n\}$ for the Richardson extrapolation scheme according to

$$x_n = (\beta + n) \Delta s_{n-1}, \qquad n \in \mathbb{N}_0, \quad \beta \in \mathbb{R}_+, \tag{14.4-10}$$

it is at least guaranteed that the interpolation points x_n behave like the r_n as $n \to \infty$.

If the interpolation points $\{x_n\}$ are chosen according to eq. (14.4-10), then obviously the Richardson extrapolation scheme, eq. (6.1-5), is exact for the following model sequence:

$$s_n = s + \sum_{j=0}^{k-1} c_j [(\beta + n) \Delta s_{n-1}]^{j+1}, \qquad k, n \in \mathbb{N}_0, \quad \beta \in \mathbb{R}_+. \tag{14.4-11}$$

Table 14-4
Acceleration of the series (14.3-7) for the lemniscate constant A

n	partial sum s_n eq. (14.4-2)	$\mathcal{N}_n^{(0)}(s_0, x_0)$ eq. (6.1-5) $x_n = (n+1)a_n$	$\rho_{2[\![n/2]\!]}^{(n-2[\![n/2]\!])}$ eq. (6.2-2) $\xi_n = 1/[(n+1)a_n]$	$\mathcal{W}_{[\![n/2]\!]}^{(n-2[\![n/2]\!])}$ eq. (6.3-3) $\xi_n = 1/[(n+1)a_n]$
10	1.2250162047862	1.3106718871541	1.3108757199219	1.3110586974716
11	1.2287537180105	1.3109702421830	1.3110068620892	1.3110406715649
12	1.2320431110268	1.3110807598150	1.3110296350804	1.3110524519639
13	1.2349672811610	1.3110705535246	1.3110300492139	1.3110491470884
14	1.2375891404731	1.3110388651127	1.3110297665932	1.3110531590519
15	1.2399574101139	1.3110247934048	1.3110284817234	1.3110289425634
16	1.2421104860230	1.3110247919628	1.3110287561737	1.3110288767162
17	1.2440790912340	1.3110277629872	1.3110287810924	1.3110291518189
18	1.2458881405432	1.3110291063652	1.3110287803056	1.3110289945504
19	1.2475580797723	1.3110291052833	1.3110287815108	1.3110288722980
20	1.2491058660392	1.3110288500150	1.3110287770120	1.3110288696319
21	1.2505456974656	1.3110287479173	1.3110287771445	1.3110288718241
22	1.2518895646423	1.3110287537771	1.3110287771529	1.3110288775601
23	1.2531476731141	1.3110287734917	1.3110287771494	1.3110288714005
24	1.2543287710802	1.3110287796054	1.3110287771460	1.3110288733834
25	1.2554404064530	1.3110287785488	1.3110287771461	1.3110288715363
$[\Gamma(1/4)]^2/[4(2\pi)^{1/2}]$		1.3110287771461	1.3110287771461	1.3110287771461

It was remarked previously that if the Richardson extrapolation scheme, eq. (6.1-5), is able to handle a certain problem using the interpolation points $\{\!\{x_n\}\!\}$, then the extrapolation points $\{\!\{\xi_n\}\!\}$ with $\xi_n = 1/x_n$ should be used if the same problem is to be treated by the general forms of Wynn's ρ algorithm, eq. (6.2-2), and of its iteration $\mathcal{W}_k^{(n)}$, eq. (6.3-3). Hence, the appropriate interpolation points $\{\!\{\xi_n\}\!\}$ for $\rho_k^{(n)}$ and $\mathcal{W}_k^{(n)}$ would be

$$\xi_n = 1/\big[(\beta+n)\Delta s_{n-1}\big], \qquad n \in \mathbb{N}_0, \quad \beta \in \mathbb{R}_+. \tag{14.4-12}$$

In table 14-4 the partial sums (14.4-2) are accelerated by the Richardson extrapolation scheme, eq. (6.1-5), and by the general forms of Wynn's ρ algorithm, eq. (6.2-2), and of its iteration $\mathcal{W}_k^{(n)}$, eq. (6.3-3). In the case of the Richardson extrapolation scheme the interpolation points $\{\!\{x_n\}\!\}$ were chosen according to eq. (14.4-10) with $\beta = 1$, and in the case of $\rho_k^{(n)}$ and $\mathcal{W}_k^{(n)}$ the interpolation points $\{\!\{\xi_n\}\!\}$ were chosen according to eq. (14.4-12) with $\beta = 1$. If we compare tables 14-3 and 14-4, we find that the rate of convergence of the transforms is slower in table 14-4, but otherwise, the results are quite similar.

As usual, table 14-4 was produced in QUADRUPLE PRECISION. When the same computation was repeated in DOUBLE PRECISION, it was again observed that the Richardson extrapolation scheme is much more sensitive to rounding errors than the other two transformations. For $n = 15$ the Richardson extrapolation scheme achieved a relative accuracy of 6 decimal digits, and for larger values of n the accuracy deteriorated rapidly yielding nonsensical results for $n \geq 22$. Of the other two transformations in table 14-4, $\mathcal{W}_k^{(n)}$, eq. (6.3-3), was this time the numerically more stable transformation. Wynn's ρ algorithm, eq. (6.2-2), achieved for $n \geq 19$ a relative accuracy for 7 decimal digits, whereas $\mathscr{W}_k^{(n)}$, eq. (6.3-3), achieved for $n \geq 18$ a relative accuracy of 8 decimal digits.

Next, the acceleration of the convergence of the other two test series (14.3-10) and (14.3-20) will be considered. The infinite series (14.3-20) may be considered to be a special case of the infinite series (14.3-10) since it was derived from it by performing the limit $z \to 0$. Because of eq. (14.3-18) we expect that these two series (14.3-10) and (14.3-20) should have roughly the same convergence properties. However, in convergence acceleration processes these two series differ significantly. As remarked previously, Levin's u and v transformation, eqs. (7.3-5) and (7.3-11), respectively, Brezinski's ϑ algorithm, eq. (10.1-9), and $\mathscr{J}_k^{(n)}$, eq. (10.3-6), are all exact for the partial sums of the infinite series (14.3-20), whereas no sequence transformation is known which is exact for the sequence of partial sums

$$s_n = \sum_{m=0}^{n} \hat{k}_{m-1/2}(z)/[2^m m!], \qquad n \in \mathbb{N}_0 \quad z \in \mathbb{R}_+, \tag{14.4-13}$$

of the infinite series (14.3-10). In the case of those sequence transformations, which are not exact for the series (14.3-20), it was observed quite consistently that the series expansion (14.3-20) can be accelerated more easily than the series expansion (14.3-10). Consequently, we shall not consider explicitly the acceleration of the convergence of the series (14.3-20). Instead, we shall focus our attention on the acceleration of the convergence of the infinite series (14.3-10) which is much more interesting in this context. The acceleration of the convergence of the series expansion (14.3-10) is particularly hard for larger values of z. In fact, for sufficiently large values of z, every sequence transformation has so far been brought down to its knees. This strong dependence of the success of a convergence acceleration process on the the magnitude of the argument z make the series expansion (14.3-10) of $1/z$ in terms of reduced Bessel functions a very interesting test problem.

In table 14-5 the convergence of the partial sums (14.4-13) is accelerated by Brezinski's ϑ algorithm, eq. (10.1-9), by its iteration $\mathscr{J}_k^{(n)}$, eq. (10.3-6), and by $\lambda_k^{(n)}$, eq. (11.2-1), with $\beta = 1$. As in tables 14-1 and 14-2 $\mathscr{J}_k^{(n)}$ has a slight plus over Brezinski's ϑ algorithm. The third transformation in table 14-5, $\lambda_k^{(n)}$, is clearly less efficient than the other two.

As usual, table 14-5 was produced in QUADRUPLE PRECISION. When the same computations were repeated in DOUBLE PRECISION, the loss of some significant digits was again observed. Relatively insensitive to rounding errors was $\mathscr{J}_k^{(n)}$, which for $n = 22$ reproduced 10 decimal digits. For $n = 22$ Brezinski's ϑ algorithm reproduced 8 decimal digits, wheras $\lambda_k^{(n)}$ reproduced 7 decimal digits.

If we compare the results produced by $\mathscr{J}_k^{(n)}$, eq. (10.3-6), in tables 14-2 and 14-5 we see that the convergence of the series (14.3-7) for the lemniscate constant A can apparently be accelerated much more easily than the convergence of the series expansion (14.3-10) for $1/z$ in terms of reduced Bessel functions. Similar results were observed also in the case of other sequence transformations.

The third sequence transformation in table 14-5, $\lambda_k^{(n)}$, eq. (11.2-1), was derived by modifying the recursive scheme (7.3-21) for $\Lambda_k^{(n)}(\beta, s_n)$ along the lines of Brezinski's ϑ algorithm. There are some interesting differences between $\lambda_k^{(n)}$, which may be considered to be an iterated weighted Δ^2 process, and $\Lambda_k^{(n)}(\beta, s_n)$, which corresponds to the special case $x_n = 1/(\beta + n)$ of the Richard-

Table 14-5
Acceleration of the series expansion (14.3-10) for $z = 4/5$

n	partial sum s_n eq. (14.4-13)	$\vartheta_{2[\![n/3]\!]}^{(n-3[\![n/3]\!])}$ eq. (10.1-9)	$\mathscr{J}_{[\![n/3]\!]}^{(n-3[\![n/3]\!])}$ eq. (10.3-6)	$\lambda_{[\![n/2]\!]}^{(n-2[\![n/2]\!])}$ eq. (11.2-1)
7	1.0422312196170	1.2497381860187	1.2479931939358	1.2173888687023
8	1.0550056275790	1.2499224049805	1.2494113420204	1.2172621135266
9	1.0656857514131	1.2497936062022	1.2499952677367	1.2015462070872
10	1.0747865667307	1.2500104486053	1.2499875622609	1.2083962702181
11	1.0826618965033	1.2500117160875	1.2499952780892	1.2511499315352
12	1.0895638413456	1.2500026110013	1.2499738050830	1.2501603786137
13	1.0956774851981	1.2500122711095	1.2499996707633	1.2500076445011
14	1.1011421634246	1.2500120099841	1.2499999749148	1.2500005196325
15	1.1060650318428	1.2500122009821	1.2499999979586	1.2500007343758
16	1.1105300244656	1.2500168769473	1.2500000006228	1.2500006080295
17	1.1146039429560	1.2500002080018	1.2499999999894	1.2500010863749
18	1.1183407028756	1.2500032915285	1.2499999999664	1.2500008892374
19	1.1217843613599	1.2500000030342	1.2499999999776	1.2499999946191
20	1.1249713188304	1.2500000009966	1.2499999999625	1.2499999323380
21	1.1279319483516	1.2500000107858	1.2499999999720	1.2499999648641
22	1.1306918204772	1.2499999999866	1.2500000000029	1.2499999366280
exact		1.2500000000000	1.2500000000000	1.2500000000000

son extrapolation scheme, with respect to their ability of accelerating logarithmic convergence. The linear sequence transformation $\Lambda_k^{(n)}(\beta, s_n)$ is one of the best accelerators for the series (14.3-1) for $\zeta(2)$ but according to theorem 14-4 is not able to accelerate the convergence of the test series (14.3-7), (14.3-10), and (14.3-20). The nonlinear sequence transformation $\lambda_k^{(n)}$ is clearly less efficient than $\Lambda_k^{(n)}(\beta, s_n)$ in the case of the series for $\zeta(2)$, but is at least moderately powerful in the case of the test series (14.3-7), (14.3-10), and (14.3-20). This example shows once more that the modification of the recursive scheme of a sequence transformation along the lines of Brezinski's ϑ algorithm does not automatically lead to a sequence transformation which is able to outperform the transformation, from which it was derived, in all respects. However, it is quite likely that the new transformation will be more versatile than the transformation from which it was derived.

The greater flexibility of those sequence transformations, which are derived along the lines of Brezinski's ϑ algorithm, is probably responsible for their ability of accelerating the convergence of the test series (14.3-7), (14.3-10), and (14.3-20), whose remainders all behave like $n^{-1/2}$ as $n \to \infty$. A sequence transformation like the standard form of Wynn's ρ algorithm, eq. (6.2-4), is often able to achieve really spectacular results if the remainders of the sequence to be transformed behave like an integral power of $1/n$, but it fails completely if the remainders behave like a nonintegral powers of $1/n$. In such a case, the general form of Wynn's ρ algorithm, eq. (6.2-2), together with an appropriate set of interpolation points $\{x_n\}$ has to be used. However, one should not expect that it will always be an easy task to find an appropriate set of interpolation points $\{x_n\}$.

If the convergence of a given sequence $\{s_n\}$ is to be accelerated by Levin's sequence transformation $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), together with one of the explicit remainder estimates

(7.3-4), (7.3-6), (7.3-8), and (7.3-10), then the elements of the sequence $\{s_n\}$ also supply the remainder estimates $\{\omega_n\}$. We can expect that Levin's sequence transformation will produce good results if the the remainder estimates $\{\omega_n\}$ can be chosen in such a way that the ratio $(s_n - s)/\omega_n$ will depend upon n only quite weakly, i.e., if it is a constant apart from terms, that are at least of order $O(n^{-1})$ as $n \to \infty$ or smaller. However, in practical applications it may happen that unless n is very large, the explicit remainder estimates (7.3-4), (7.3-6), (7.3-8), and (7.3-10) yield only relatively bad approximations for the actual remainders $\{r_n\}$ of the sequence to be transformed. In such a case, it is to be expected that Levin's sequence transformation $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), will be a relatively weak sequence transformation if it uses one the the explicit remainder estimates (7.3-4), (7.3-6), (7.3-8), and (7.3-10).

According to eq. (14.3-16) the reduced Bessel functions $\hat{k}_{n+1/2}(z)$ with $z \in \mathbb{R}_+$ and $n \in \mathbb{N}_0$ are positive and bounded by their values at the origin. However, it was already remarked in section 14.3 that due to the exponential decline of the reduced Bessel functions $\hat{k}_{n+1/2}(0)$ is a good approximation for $\hat{k}_{n+1/2}(z)$ with a larger argument z only if n is relatively large. This has some unpleasant consequences if for instance Levin's u transformation, eq. (7.3-5), is to be used for an acceleration of the convergence of the series expansion (14.3-10) of $1/z$ in terms of reduced Bessel functions. For larger values of z the product $(\beta + n)\hat{k}_{n-1/2}(z)$ with $\beta \in \mathbb{R}_+$ will be a good approximation for the remainder

$$r_n = \sum_{m=n+1}^{\infty} \hat{k}_{m-1/2}(z)/[2^m m!], \qquad n \in \mathbb{N}_0 \quad z \in \mathbb{R}_+, \tag{14.4-14}$$

of the infinite series (14.3-10) only if n is very large. In such a situation, it should be worthwhile to look for other sets of remainder estimates $\{\omega_n\}$ even if Levin's sequence transformation $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), would then be a linear sequence transformation. One simple possibility would be to choose

$$\omega_n = (n+1)^{-1/2}, \qquad n \in \mathbb{N}_0. \tag{14.4-15}$$

Another possibilty, which would also lead to remainder estimates that behave like $n^{-1/2}$ as $n \to \infty$, would be to choose the remainder estimates according to eq. (14.3-21),

$$\omega_n = (2n-1)!!/(2n)!!, \qquad n \in \mathbb{N}_0. \tag{14.4-16}$$

In table 14-6 the partial sums (14.4-13) are accelerated by $u_k^{(n)}(\beta, s_n)$, eq. (7.3-5), and by $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), with either $\omega_n = (n+1)^{-1/2}$ or $\omega_n = (2n-1)!!/(2n)!!$. In all cases $\beta = 1/2$ was used which gives slightly better results than $\beta = 1$.

The results in table 14-6 show quite clearly that the remainder estimates (14.4-15) and (14.4-16) produce significantly better results than the remainder estimate (7.3-4) which is the basis of Levin's u transformation. There is indirect evidence that this improved convergence of the transforms is indeed due to the better approximation of the remainders (14.4-14) by the remainder estimates (14.4-15) and (14.4-16). If the convergence of the series (14.3-7) for the lemniscate constant A is accelerated by Levin's sequence transformation, eq. (7.1-7), with the remainder estimates being chosen according to eq. (14.4-15), then the results obtained in this way

Table 14-6
Acceleration of the series expansion (14.3-10) for $z = 4/5$

n	partial sum s_n eq. (14.4-13)	$u_n^{(0)}(1/2, s_0)$ eq. (7.3-5)	$\mathscr{L}_n^{(0)}(1/2, s_0, \omega_0)$ eq. (7.1-7) $\omega_n = (n+1)^{-1/2}$	$\mathscr{L}_n^{(0)}(1/2, s_0, \omega_0)$ eq. (7.1-7) $\omega_n = (2n-1)!!/(2n)!!$
7	1.0422312196170	1.2472807413200	1.2500112144531	1.2500305329006
8	1.0550056275790	1.2519888543148	1.2499669957586	1.2499700063760
9	1.0656857514131	1.2498098463228	1.2500047257200	1.2500032427111
10	1.0747865667307	1.2498185973113	1.2500012644944	1.2500013433652
11	1.0826618965033	1.2500735225247	1.2499994685087	1.2499995333305
12	1.0895638413456	1.2499966179905	1.2500000284158	1.2500000125919
13	1.0956774851981	1.2499942161928	1.2500000253626	1.2500000250149
14	1.1011421634246	1.2500019544458	1.2499999930747	1.2499999939416
15	1.1060650318428	1.2499999104639	1.2500000000686	1.2499999999124
16	1.1105300244656	1.2499998754938	1.2500000003615	1.2500000003511
17	1.1146039429560	1.2500000414919	1.2499999999164	1.2499999999261
18	1.1183407028756	1.2499999971441	1.2499999999999	1.2499999999984
19	1.1217843613599	1.2499999980000	1.2500000000043	1.2500000000042
20	1.1249713188304	1.2500000007366	1.2499999999990	1.2499999999991
21	1.1279319483516	1.2499999999235	1.2500000000000	1.2500000000000
22	1.1306918204772	1.2499999999760	1.2500000000000	1.2500000000000
exact		1.2500000000000	1.2500000000000	1.2500000000000

are as good or only marginally better than the results obtained by Levin's u transformation. Hence, in the case of the series (14.3-7) for the lemniscate constant A the remainder estimates (14.4-15) do not lead to a spectacular improvement.

As usual, table 14-6 was produced in QUADRUPLE PRECISION. When the same computation was repeated in DOUBLE PRECISION, Levin's u transformation produced for $n = 15$ a relative accuracy of 8 decimal digits. The other two transformations produced also for $n = 15$ a relative accuracy of 10 decimal digits. For larger values of n, the accuracy deteriorated again.

15. Synopsis

15.1. General considerations

In this report a large number of mainly nonlinear sequence transformations for the acceleration of convergence and the summation of divergent series were discussed. Some of those sequence transformations as for instance Wynn's ϵ algorithm are well established in the literature, while many others are new. The properties of these sequence transformations were analyzed and efficient algorithms for their evaluation were derived. In sections 13 and 14 the performance of these sequences transformations was tested by applying them to certain slowly convergent and divergent series, which are hopefully realistic models for a large part of the slowly convergent or divergent series that can occur in scientific problems and in applied mathematics.

It still has to be discussed how one should actually proceed if the convergence of a slowly convergent sequence or series has to be accelerated or if a divergent series has to be summed. In view of the numerous different types of sequences and series, which can occur in practical problems, and because of the large number of sequence transformations, which are known, the selection of an appropriate sequence transformation is certainly a nontrivial problem.

If the terms of the series, which is to be transformed, are known analytically or if it is known how the elements of the sequence $\{s_n\}$ of partial sums behave as $n \to \infty$, it is normally comparitively easy to find a suitable sequence transformation. Unfortunately, it can happen that only a few elements of a slowly convergent or divergent sequence $\{s_n\}$ are available and that the behaviour of the sequence elements s_n as $n \to \infty$ is not known. In such an unfavourable case, in which it is often not easy to decide whether $\{s_n\}$ converges at all, and if it does, whether it converges linearly or logarithmically, the choice of an appropriate sequence transformation is by no means simple and also of decisive importance for the success of the whole approach.

It is well known that the performance of a sequence transformation depends in most cases quite strongly upon the type of convergence of the sequence to which it is applied. Apparently, there is no sequence transformation which excells in every respect. Of all the sequence transformations in this report, only Levin's u transformation, eq. (7.3-5), to a somewhat lesser extent also Levin's v transformation, eq. (7.3-11), Brezinski's ϑ algorithm, eq. (10.1-9), and its iteration $\mathscr{J}_k^{(n)}$, eq. (10.3-6), are powerful accelerators for both linear and logarithmic convergence and are also able to sum efficiently even wildly divergent series. In all test cases considered they were among the better sequence transformations.

It is tempting to believe that it would be sufficient to use only the four sequence transformations mentioned above in situations, in which apart from the numerical values of a few sequence elements little is known. Since these sequence transformations are known to work well in a variety of different situations, it seems reasonable to expect that they will accomplish at least something. However, in many of the test problems of this report other, less versatile sequence transformations were actually more efficient. Hence, even if the four transformations mentioned above are successful, they do not necessarily give the best results, and it may well be worthwhile to look for other sequence transformations which are possibly more efficient, in particular if only relatively few sequence elements are available.

There are also some other aspects which should be taken into consideration. In this report, only the most common types of sequences and series were considered, i.e., either linearly and logarithmically convergent sequences and series or alternating divergent series. This does not exhaust all possibilities. Therefore, it is not certain whether the four sequence transformations mentioned above will also be able to handle successfully other types of convergence. For instance, Smith and Ford report that in the case of some slowly convergent series with terms having irregular sign patterns Wynn's ϵ algorithm, eq. (4.2-1), clearly outperformed Levin's u transformation, eq. (7.3-5), and Brezinski's ϑ algorithm, eq. (10.1-9), which both did not accomplish much (see table 5 on p. 484 of ref. [30]). Also, in a situation, in which apart from the numerical values of only a few sequence elements very little is known, it is often not clear whether and how well the whole process has already converged. Even if a sequence transformation produces a sequence of transforms which apparently converges to some limit it cannot be excluded that this convergence is an artifact. If two different sequence transformations converge to the same limit, an artifact still cannot be ruled out but it is much less likely. Consequently, in such a situation it should be worthwhile to use more than a single sequence transformation.

In the opinion of the author the approximate determination of the limit or antilimit s of a slowly convergent or divergent sequence $\{s_n\}$ is essentially an experimental problem which should be handled with utmost care. The numerical evidence supplied by a single sequence transformation is not necessarily sufficient, and it is usually a good idea to compare the results produced by several sequence transformations. In order to facilitate the task of selecting appropriate sequence transformations, short résumés of the properties of all sequence transformations, which occur in this report, will now be be given.

15.2. Wynn's epsilon algorithm and related transformations

Wynn's ϵ algorithm, eq. (4.2-1), and its close relative, Aitken's iterated Δ^2 process, eq. (5.1-15), are both able to accelerate linear convergence and to sum alternating divergent series even if they diverge as wildly as the Euler series, eq. (1.1-7), but they are not able to accelerate logarithmic convergence.

Practical experience and also some theoretical estimates indicate that Wynn's ϵ algorithm is only a moderately powerful sequence transformation, in particular if wildly divergent series must be summed. In the case of the two Stieltjes series (13.3-9) for the exponential integral and (13.4-3) for the logarithm, and also in other tests not discussed in this report, Wynn's ϵ algorithm was not only clearly outperformed by Levin's sequence transformation $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), and the new sequence transformations $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), and $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-6), but also frequently by Aitken's iterated Δ^2 algorithm, albeit to a lesser extent.

On the basis of these results it looks as if Wynn's ϵ algorithm should be dismissed. However, the real strength of Wynn's ϵ algorithm is not its efficiency but its robustness. The ϵ algorithm is remarkably insensitive to rounding errors and can also tolerate input data which either have a low relative accuracy or which behave in a comparatively irregular way. Due to its robustness, Wynn's ϵ algorithm is often able to produce meaningful and reliable results in situations in which other sequence transformations, which are in principle much more powerful, fail. For instance, in ref. [64] the convergence of some infinite series with very complicated terms was accelerated by Wynn's ϵ algorithm and by Levin's u transformation, eq. (7.3-5). Since these infinite series converge linearly, it was to be expected that Levin's u transformation would do better than Wynn's ϵ algorithm. However, it was found that the ϵ algorithm converged more rapidly than the u transformation. In addition, the ϵ algorithm was apparently not affected by numerical instabilities whereas in the case of the u transformation a dangerous accumulation of rounding errors was observed (see pp. 3716–3717 of ref. [64]).

Superficially, Aitken's iterated Δ^2 process appears to be a better sequence transformation than Wynn's ϵ algorithm. However, to a certain extent Aitken's iterated Δ^2 process combines the disadvantageous features of both Wynn's ϵ algorithm, which is only moderately powerful, and of the transformations $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), and $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-6), which are not very robust since they are powerful sequence transformations only if the remainder estimates $\{\omega_n\}$ are good approximations of the actual remainders $\{r_n\}$. Experience indicates that Aitken's iterated Δ^2 process is in general less efficient than $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, and $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, and that it is less robust and more susceptible to rounding errors than Wynn's ϵ algorithm.

15.3. Wynn's rho algorithm and related transformations

The properties of Wynn's ϵ algorithm, eq. (4.2-1), and of Wynn's ρ algorithm, eq. (6.2-2), are complementary. The ρ algorithm is often a good or even very good accelerator for logarithmic convergence but is unable to accelerate linear convergence or to sum divergent series. But with respect to robustness, the ϵ and the ρ algorithm are very similar. Experience indicates that the ρ algorithm is in general less sensitive to rounding errors than other sequence transformations which are also able to accelerate logarithmic convergence. This is certainly no mean accomplishment, in particular since the acceleration of logarithmic convergence is an inherently unstable process.

The power of Wynn's ρ algorithm, which is essentially an intelligent way of computing and extrapolating to infinity an interpolating rational function of the type of eq. (6.2-1), depends decisively upon an appropriate choice of the interpolation points $\{x_n\}$. In this respect, the ρ algorithm closely resembles the Richardson extrapolation scheme, eq. (6.1-5), which is essentially an efficient way of computing and extrapolating to zero an interpolating polynomial of the type of eq. (6.1-3). According to theorem 14-4 the linear sequence transformation $\Lambda_k^{(n)}(\beta, s_n)$, eq. (7.3-20), which corresponds to the special case $x_n = 1/(\beta + n)$ of the Richardson extrapolation scheme, eq. (6.1-5), is only able to accelerate the convergence of logarithmically convergent sequences if the remainders of these sequences behave like integral powers of $1/n$ as $n \to \infty$.

The standard form of Wynn's ρ algorithm, eq. (6.2-4), which corresponds to the choice $x_n = \beta + n$ for the interpolation points in eq. (6.2-2), is a very powerful accelerator for logarithmically convergent sequences with remainders that behave like integral powers of $1/n$ as $n \to \infty$. However, experience indicates that the standard form of the ρ algorithm is unable to accelerate convergence if the remainders of the sequence to be transformed behave like a nonintegral power of $1/n$. In section 14.4 it was shown that the general form of Wynn's ρ algorithm, eq. (6.2-2), is apparently able to accelerate the convergence of sequences with remainders that behave like $n^{-1/2}$ as $n \to \infty$ if the interpolation points $\{x_n\}$ are chosen in such a way that $x_n \sim n^{1/2}$ as $n \to \infty$.

The iterated sequence transformation $\mathscr{W}_k^{(n)}$, eq. (6.3-3), is also only able to accelerate logarithmic convergence. The power of $\mathscr{W}_k^{(n)}$ depends as in the case of Wynn's ρ algorithm, from which it was derived, strongly upon an appropriate choice of the interpolation points $\{x_n\}$. The numerical results presented in section 14.4 indicate that $\mathscr{W}_k^{(n)}$ has similar properties as Wynn's ρ algorithm. The standard form of $\mathscr{W}_k^{(n)}$, eq. (6.3-4), is apparently not able to accelerate the convergence of sequences whose remainders behave like nonintegral powers of $1/n$. However, the results in section 14.4 also show that the general form of $\mathscr{W}_k^{(n)}$, eq. (6.3-3), is apparently able to accelerate the convergence of sequences with remainders that behave like $n^{-1/2}$ as $n \to \infty$ if the interpolation points $\{x_n\}$ are chosen in such a way that $x_n \sim n^{1/2}$ as $n \to \infty$. It also seems that $\mathscr{W}_k^{(n)}$ is relatively insensitive to rounding errors. However, because of the limited experience with this transformation it seems to be too early for a definite assessment of its merits as well as its weaknesses. Further tests of this sequence transformation should therefore be of interest.

The fact, that Wynn's ρ algorithm and its iteration $\mathscr{W}_k^{(n)}$ are only successful if an appropriate set of interpolation points is used, severely limits the practical usefulness of these transformations in situations in which only the numerical values of a few sequence elements are known. In such a situation, it may be a good idea to use a modification of a suggestion by Beleznay [117]. In this approach, the interpolation points are chosen according to eq. (14.4-7) and the free parameter α is optimized after the input of every new sequence s_n. Another possibility would be to choose the interpolation points according to eq. (14.4-12). However, these two suggestions are not yet sufficiently tested and it seems to be too early for a definite assessment of their practical usefulness.

15.4. Levin's sequence transformation and related transformations

It is a typical feature of Levin's sequence transformations $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), and of the related transformations $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), and $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-6), that they not only require the sequence elements $s_n, s_{n+1}, \ldots, s_{n+k}$, but also the remainder estimates $\omega_n, \omega_{n+1}, \ldots, \omega_{n+k}$. This explicit incorporation of remainder estimates is both the strength as well as the weakness of these sequence transformations. If it is possible to find a sequence of remainder estimates $\{\omega_n\}$ that are good approximations of the remainders $\{r_n\}$ of the sequence to be transformed, then experience as well as some theoretical estimates indicate that such a sequence transformation is extremely powerful. If, however, a good sequence of remainder estimates cannot be found, such a sequence transformation will probably perform quite poorly.

Theoretical estimates as well as practical experience indicate that the remainder estimates $\{\omega_n\}$ should be chosen in such a way that the ratios $(s_n - s)/\omega_n$ depend on n only quite weakly and approach a constant as $n \to \infty$. In practical applications Levin's sequence transformation $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$ has so far exclusively been used in connection with the simple remainder estimates (7.3-4), (7.3-6), (7.3-8), and (7.3-10), which can also be used in the case of the new transformations $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$ and $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$. The remainder estimate (7.3-4) gives Levin's u transformation, eq. (7.3-5), which is certainly one of the most powerful and most versatile sequence transformations. It is a powerful accelerator for both linear and logarithmic convergence and is able to sum efficiently divergent alternating series. The remainder estimates (7.3-6) and (7.3-8) give Levin's t and d transformations, eqs. (7.3-7) and (7.3-9), respectively, which are powerful accelerators for linear convergence and are able to sum efficiently divergent alternating series. However, they are unable to accelerate logarithmic convergence. The remainder estimate (7.3-10) gives Levin's v transformation, eq. (7.3-11), which has similar properties as Levin's u transformation. Finally, the remainder estimate $\omega_n = 1/(\beta + n)$ gives the linear sequence transformation $\Lambda_k^{(n)}(\beta, s_n)$, eq. (7.3-20), which can also be obtained from the Richardson extrapolation scheme, eq. (6.1-5), by choosing $x_n = 1/(\beta + n)$. It is able to accelerate logarithmic convergence if the remainders of the sequence to be transformed behave like an integral power of $1/n$ as $n \to \infty$.

The simple remainder estimates (7.3-4), (7.3-6), (7.3-8), and (7.3-10) are essentially asymptotic in nature because they were derived using some simplifications which are valid for large values of n. However, in convergence acceleration or summation processes it is tried to approximate the limit or antilimit of a sequence $\{s_n\}$ using only the information stored in the first few sequence elements, say $s_0, s_1, \ldots, s_m$, with m being a relatively small number. Therefore, it is by no means clear whether the simple remainder estimates (7.3-4), (7.3-6), (7.3-8), and (7.3-10) lead to good approximations of the actual remainders for only moderately large or even small indices.

There is some evidence that the simple remainder estimates (7.3-4), (7.3-6), (7.3-8), and (7.3-10) lead to efficient sequence transformations if the terms a_n of the series, which is to be accelerated or summed, approach their asymptotic limits relatively fast. Let us for instance assume that a_n is the term of a power series in z and that a_n behaves like $n^{\alpha} z^n$ as $n \to \infty$. Then, the sequence transformations $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, and $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$ should work well in combination with one of the simple remainder estimates (7.3-4), (7.3-6), (7.3-8), and (7.3-10) if the leading term $n^{\alpha} z^n$ is a good approximation for a_n already for moderately large or even small values of n. If this is not the case, the simple remainder estimates (7.3-4), (7.3-6), (7.3-8), and (7.3-10) will probably not work particulary well.

The infinite series (14.3-10) is a good example for the complications which can occur in this context. It was remarked earlier that the reduced Bessel functions, eq. (14.3-11), approach their asymptotic limits according to eq. (14.3-18) quite slowly. This slow approach decreases the efficiency of Levin's u transformation, eq. (7.3-5), considerably. The results presented in table 14-6 show that in this case it is advantageous to use other, explicit remainder estimates, which are not obtained from the elements of the sequence $\{s_n\}$ to be transformed, even if Levin's sequence transformation is then a linear sequence transformation. Unfortunately, such an approach is only possible if the remainders are known analytically and if simple and yet good approximations for the remainders can be derived.

Also, under unfavourable circumstances the simple remainder estimates (7.3-4), (7.3-6), (7.3-8), and (7.3-10) may have a detrimental effect on the robustness of the sequence transformations $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, and $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$. The elements of the sequence $\{s_n\}$ are not only input data, but they are also used to compute the remainder estimates $\{\omega_n\}$. Consequently, the elements of $\{s_n\}$ induce two fundamentally different kinds of errors. More or less inevitable are the errors due to the limited accuracy of the input data. However, the elements of the sequence $\{s_n\}$ induce also potentially large errors among the remainder estimates $\{\omega_n\}$, either because they are not accurate enough or because they deviate too much from their asymptotic limits and therefore produce bad remainder estimates. The worst scenario, which can be imagined in this context, would be that the terms a_n of a series are not very accurate and that the terms approach their limiting expressions only quite slowly and in an irregular fashion. In such a situation the sequence transformations $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, and $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$ will be in trouble and it is likely that Wynn's ϵ algorithm, although in principle only moderately powerful, will produce better results.

Levin's sequence transformation $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), is based upon the assumption that the ratio $(s_n - s)/\omega_n$ can be approximated by a polynomial in $1/(\beta + n)$, whereas the new sequence transformations $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), and $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-6), were derived assuming that the ratio $(s_n - s)/\omega_n$ can be approximated by truncated factorial series or related expressions. Since power series and factorial series have different properties, it is not surprising that the new sequence transformations and Levin's transformation behave differently in convergence acceleration and summation processes. With respect to the acceleration of linear convergence the new sequence transformations $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$ and $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$ are approximately as efficient as Levin's sequence transformation. The new sequence transformations are particularly well suited to sum wildy divergent alternating series such as the Euler series, eq. (1.1-7). In that respect, they are usually at least as good as Levin's sequence transformation and often they are even clearly better. However, the new transformations perform quite poorly if logarithmic convergence is to be accelerated. Also, the linear transformations $\mathscr{F}_k^{(n)}(\alpha, s_n)$, eq. (8.4-11), and $\mathscr{P}_k^{(n)}(\gamma, s_n)$, eq. (9.4-11), are much less efficient than their analogue $\Lambda_k^{(n)}(\gamma, s_n)$, eq. (7.3-20).

The stability properties of the sequence transformations $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, and $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$ depend very much upon the sequence which is to be transformed and upon the remainder estimates being used. However, at least some statements of a more general nature, which are based upon experience, can be made. It seems that the transformation of both convergent and divergent alternating series is in general remarkably stable. Also, the acceleration of linear convergence usually poses no particular stability problems. The acceleration of logarithmic convergence is always a problem which may easily lead to a serious loss of accuracy. But it cannot be said that Levin's sequence transformation is more sensitive to rounding errors than most other sequence transformations. Also, it is probably safe to say that the sequence transformations $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, and $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$ are in general more efficient and at the same time less robust than Wynn's ϵ algorithm if linear convergence is accelerated or if divergent alternating series are summed. In the same way, Wynn's ρ algorithm is apparently more robust than the u and v transformation or the linear sequence transformation $\Lambda_k^{(n)}(\beta, s_n)$ in the case of logarithmic convergence.

Finally, there is Drummond's sequence transformation $\mathscr{D}_k^{(n)}(s_n, \omega_n)$, eq. (9.5-4), which is another relative of Levin's sequence transformation since it also uses a sequence of remainder estimates $\{\omega_n\}$. Drummond's sequence transformation is very important theoretically, in particular in connection with Germain-Bonne's formal theory of convergence accleration [33] and the explicit construction of Padé approximants for the Euler series, eq. (1.1-7). However, in practical applications Drummond's sequence transformations is at most moderately powerful. It is significantly less powerful than the sequence transformations $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), and $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-6), but has the same weaknesses as these transformations.

15.5. Brezinski's theta algorithm and related transformations

Brezinski's ϑ algorithm, eq. (10.1-9), and its iteration $\mathscr{J}_k^{(n)}$, eq. (10.3-6), combine many of the advantageous features of Wynn's ϵ algorithm, eq. (4.2-1), and of Wynn's ρ algorithm, eq. (6.2-2). They are able to accelerate linear convergence and to sum even wildly divergent alternating series, and they are also able to accelerate logarithmic convergence.

In those tests, in which linear convergence had to be accelerated or divergent alternating series had to be summed, Brezinski's ϑ algorithm and its iteration $\mathscr{J}_k^{(n)}$ were usually better than the ϵ algorithm, but less powerful than the sequence transformations $\mathscr{L}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (7.1-7), $\mathscr{S}_k^{(n)}(\beta, s_n, \omega_n)$, eq. (8.2-7), and $\mathscr{M}_k^{(n)}(\gamma, s_n, \omega_n)$, eq. (9.2-6).

With respect to the acceleration of logarithmic convergence Brezinski's ϑ algorithm and its iteration $\mathscr{J}_k^{(n)}$ are more reliable than the standard form of Wynn's ρ algorithm, eq. (6.2-4), since they are not restricted to sequences with remainders that behave like an integral power of $1/n$ as $n \to \infty$, and they are easier to use than the general form of Wynn's ρ algorithm, eq. (6.2-2), since no interpolation points are needed. In those tests, in which logarithmic convergence had to be accelerated, they were approximately as powerful as Levin's u and v transformations.

It also seems that with respect to numerical stability and robustness Brezinski's ϑ algorithm and its iteration $\mathcal{J}_k^{(n)}$ are less robust than the ϵ or the ρ algorithm and also more susceptible to rounding errors.

The other sequence transformations, which were also derived along the lines of Brezinski's ϑ algorithm are $\mathscr{B}_k^{(n)}$, eq. (11.1-5), $\mathscr{C}_k^{(n)}$, eq. (11.1-12), $\lambda_k^{(n)}$, eq. (11.2-1), $\sigma_k^{(n)}$, eq. (11.2-2), and $\mu_k^{(n)}$, eq. (11.2-3). It is a typical feature of these transformations that they are much more versatile than the transformations from which they were derived. This means they are all able to accelerate linear and logarithmic convergence and are also able to sum even wildly divergent alternating series.

Unfortunately, it is also a typical feature of these sequence transformations that their performance in the numerical tests described in sections 13 and 14 was quite inconsistent and more or less unpredictable. For instance, $\sigma_k^{(n)}$ turned out to be a very powerful accelerator for the infinite series (14.3-1) for $\zeta(2)$, a powerful accelerator for the infinite series (14.3-20), but a relatively weak accelerator for the infinite series (14.3-7) and (14.3-10). No explanation for this inconsistent behaviour can be given. At best, the sequence transformations listed above were as good as Brezinski's ϑ algorithm or its iteration $\mathcal{J}_k^{(n)}$, but in most cases they were significantly weaker. Also, it seems that the sequence transformations mentionend above are not more robust and less susceptible to rounding errors than the ϑ algorithm or $\mathcal{J}_k^{(n)}$. Hence, it seems that the most promising choices among all sequence transformations, which were derived along the lines of Brezinski's ϑ algorithm, are the ϑ algorithm, eq. (10.1-9), and $\mathcal{J}_k^{(n)}$, eq. (10.3-6).

Acknowledgement

I would like to thank Professor E.O. Steinborn for stimulating discussions, for his encouragement, for his constant support, and for the excellent working conditions at the Institut für Physikalische und Theoretische Chemie der Universität Regensburg.

The research, which ultimately led to this report, was begun during a stay at the Faculty of Mathematics of the University of Waterloo, Ontario, Canada. I would like to thank Professor J. Čížek and Professor J. Paldus for their invitation to work with them as a postdoctoral fellow in the Quantum Theory Group of the Department of Applied Mathematics. Their hospitality, their generosity, and the inspiring atmosphere which they provided is highly appreciated. Special thanks to Professor J. Čížek who aroused my interest in asymptotic techniques, the summation of divergent series, and the construction of rational approximants.

Many thanks also to H. Homeier, who patiently read and discussed the numerous preliminary versions of this manuscript, and to M. Middleton, the $\mathrm{T_EX}$ expert at the Universität Regensburg and the author of many useful macros, who helped me to typeset this manuscript in $\mathrm{R_ZT_EX}$, the local variant of $\mathrm{PCT_EX}$.

References

[1] S.E. Haywood and J.D. Morgan III, *Discrete basis-set approach for calculating Bethe logarithms*, Phys. Rev. A **32** (1985), 3179–3186.

[2] C.M. Bender and S. A. Orszag, *Advanced mathematical methods for scientists and engineers* (McGraw-Hill, New York, 1978).
[3] C.M. Bender and T.T. Wu, *Anharmonic oscillator. II. A study in perturbation theory in large order*, Phys. Rev. D 7 (1973), 1620–1636.
[4] K. Knopp, *Theorie and Anwendung der unendlichen Reihen* (Springer-Verlag, Berlin, 1964).
[5] J. Stirling, *Methodus differentialis sive tractatus de summatione et interpolatione serium infinitarum* (London, 1730). English translation by F. Holliday, *The differential method, or, a treatise concerning the summation and interpolation of infinite series* (London, 1749).
[6] L. Euler, *Institutiones calculi differentialis cum eius usu in analysi finitorum ac doctrina serium. Part II.1. De transformatione serium* (Academia Imperialis Scientiarum Petropolitana, 1755). This book was reprinted as Vol. X of *Leonardi Euleri Opera Omnia, Seria Prima* (Teubner, Leipzig and Berlin, 1913).
[7] G.H. Hardy, *Divergent series* (Oxford University Press, Oxford, 1949).
[8] G.M. Petersen, *Regular matrix transformations* (McGraw-Hill, London, 1966).
[9] A. Peyerimhoff, *Lectures on summability* (Springer-Verlag, Berlin, 1969).
[10] K. Zeller and W. Beekmann, *Theorie der Limitierungsverfahren* (Springer-Verlag, Berlin, 1970).
[11] R.E. Powell and S.M. Shah, *Summability theory and its applications* (Van Nostrand Reinhold, London, 1972).
[12] A.C. Aitken, *On Bernoulli's numerical solution of algebraic equations*, Proc. Roy. Soc. Edinburgh **46** (1926), 289–305.
[13] J. Todd, *Motivation for working in numerical analysis*, in J. Todd (ed.), *Survey of numerical analysis* (McGraw-Hill, New York, 1962), pp. 1–26.
[14] E.E. Kummer, *Eine neue Methode, die numerischen Summen langsam convergirender Reihen zu berechnen*, J. Reine Angew. Math. **16** (1837), 206–214.
[15] D. Shanks, *Non-linear transformations of divergent and slowly convergent sequences*, J. Math. and Phys. (Cambridge, Mass.) **34** (1955), 1–42.
[16] P. Wynn, *On a device for computing the $e_m(S_n)$ transformation*, Math. Tables Aids Comput. **10** (1956), 91–96.
[17] J.R. Schmidt, *On the numerical solution of linear simultaneous equations by an iterative method*, Philos. Mag. **32** (1941), 369–383.
[18] G.A. Baker, Jr., *Essentials of Padé approximants* (Academic Press, New York, 1975).
[19] C. Brezinski, *Accélération de la convergence en analyse numérique* (Springer-Verlag, Berlin, 1977).
[20] C. Brezinski, *Algorithmes d'accélération de la convergence – Étude numérique* (Editions Technip, Paris, 1978).
[21] C. Brezinski, *Padé-type approximation and general orthogonal polynomials* (Birkhäuser Verlag, Basel, 1980).
[22] G.A. Baker, Jr., and P. Graves-Morris, *Padé approximants. Part I: Basic theory. Part II: Extensions and applications* (Addison-Wesley, Reading, Mass., 1981).
[23] J. Wimp, *Sequence transformations and their applications* (Academic Press, New York, 1981).
[24] C. Brezinski, *Convergence acceleration methods: The past decade*, J. Comput. Appl. Math. **12 & 13** (1985), 19–36.
[25] P. Wynn, *On a Procrustean technique for the numerical transformation of slowly convergent sequences and series*, Proc. Camb. Phil. Soc. **52** (1956), 663–671.
[26] C. Brezinski, *Accélération de suites à convergence logarithmique*, C. R. Acad. Sc. Paris **273** (1971), 727–730.
[27] C. Brezinski, *Some new convergence acceleration methods*, Math. Comput. **39** (1982), 133–145.
[28] D. Levin, *Development of non-linear transformations for improving convergence of sequences*, Int. J. Comput. Math. B 3 (1973), 371–388.
[29] D.A. Smith and W.F. Ford, *Acceleration of linear and logarithmic convergence*, SIAM J. Numer. Anal. **16** (1979), 223–240.
[30] D.A. Smith and W.F. Ford, *Numerical comparison of nonlinear convergence accelerators*, Math. Comput. **38** (1982), 481–499.
[31] C. Brezinski, *A general extrapolation algorithm*, Numer. Math. **35** (1980), 175–180.
[32] T. Håvie, *Generalized Neville type extrapolation schemes*, BIT **19** (1979), 204–213.
[33] B. Germain-Bonne, *Transformations de suites*, Rev. Française Automat. Informat. Rech. Operat. **7** (R-1) (1973), 84–90.
[34] W. Magnus, F. Oberhettinger, and R.P. Soni, *Formulas and theorems for the special functions of mathematical physics* (Springer-Verlag, New York, 1966).

[35] W.D. Clark, H.L. Gray, and J.E. Adams, *A note on the T-transformation of Lubkin*, J. Res. Natl. Bur. Stand. B **73** (1969), 25–29.
[36] C. Brezinski, *Algorithm 585: A subroutine for the general interpolation and extrapolation problems*, ACM Trans. Math. Software **8** (1982), 290–301.
[37] P. Wynn, *On the convergence and the stability of the epsilon algorithm*, SIAM J. Numer. Anal. **3** (1966), 91–122.
[38] P. Wynn, *A note on programming repeated applications of the ϵ-algorithm*, R.F.T.I.-Chiffres **8** (1965), 23–62.
[39] P. Wynn, *Singular rules for certain non-linear algorithms*, BIT **3** (1963), 175–195.
[40] S. Lubkin, *A method of summing infinite series*, J. Res. Natl. Bur. Stand. **48** (1952), 228–254.
[41] R.R. Tucker, *The δ^2 process and related topics*, Pacif. J. Math. **22** (1967), 349–359.
[42] R.R. Tucker, *The δ^2 process and related topics II*, Pacif. J. Math. **28** (1969), 455–463.
[43] F. Cordellier, *Sur la régularité des procédés δ^2 d'Aitken et W de Lubkin*, in L. Wuytack (ed.), *Padé approximation and its applications* (Springer-Verlag, Berlin, 1979), pp. 20–35.
[44] G.E. Bell and G.M. Phillips, *Aitken acceleration of some alternating series*, BIT **24** (1984), 70–77.
[45] A.J. MacLeod, *Acceleration of vector sequences by multi-dimensional Δ^2-methods*, Commun. Appl. Numer. Meth. **2** (1986), 385–392.
[46] J.E. Drummond, *Summing a common type of slowly convergent series of positive terms*, J. Austral. Math. Soc. **B 19** (1976), 416–421.
[47] P. Bjørstad, G. Dahlquist, and E. Grosse, *Extrapolations of asymptotic expansions by a modified Aitken δ^2-formula*, BIT **21** (1981), 56–65.
[48] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling, *Numerical recipes* (Cambridge University Press, Cambridge, 1986).
[49] P. Hillion, *Méthode d'Aitken itérée pour les suites oscillantes d'approximations*, C. R. Acad. Sc. Paris A **280** (1975), 1701–1704.
[50] P.J. Davis, *Interpolation and approximation* (Dover, New York, 1975).
[51] D.C. Joyce, *Survey of extrapolation processes in numerical analysis*, SIAM Rev. **13** (1971), 435–490.
[52] E.H. Neville, *Iterative interpolation*, J. Indian Math. Soc. **20** (1934), 87–120.
[53] L. F. Richardson, *The deferred approach to the limit. I. Single lattice*, Phil. Trans. Roy. Soc. London A **226** (1927), 229–349.
[54] A. Cuyt and L. Wuytack, *Nonlinear methods in numerical analysis* (North-Holland, Amsterdam, 1987).
[55] T.N. Thiele, *Interpolationsrechnung* (Teubner, Leipzig, 1909).
[56] A. Sidi, *Convergence properties of some nonlinear sequence transformations*, Math. Comput. **33** (1979), 315–326.
[57] T. Fessler, W.F. Ford, and D.A. Smith, *HURRY: An acceleration algorithm for scalar sequences and series*, ACM Trans. Math. Software **9** (1983), 346–354.
[58] I.M. Longman, *Difficulties of convergence acceleration*, in M.G. de Bruin and H. van Rossum (eds.), *Padé approximation and its applications Amsterdam 1980* (Springer-Verlag, Berlin, 1981), pp. 273–289.
[59] Sister M.C. Fasenmyer, *Some generalized hypergeometric polynomials*, Bull. Amer. Math. Soc. **53** (1947), 806–812. A good discussion of Sister Celine's technique can also be found in chapter 14 of E. D. Rainville, *Special functions* (Chelsea, New York, 1960).
[60] A.J. Thakkar, *A Technique for increasing the utility of the Wigner-Kirkwood expansion for the second virial coefficient*, Mol. Phys. **36** (1978), 887–892.
[61] A.C. Tanner and A.J. Thakkar, *Discrete and continuum contributions to multipole polarizabilities and shielding factors of hydrogen*, Int. J. Quantum Chem. **24** (1983), 345–352.
[62] E.J. Weniger, J. Grotendorst, and E.O. Steinborn, *Some applications of nonlinear convergence accelerators*, Int. J. Quantum Chem. Symp. **19** (1986), 181–191.
[63] J. Grotendorst and E.O. Steinborn, *Use of nonlinear convergence accelerators for the efficient evaluation of GTO molecular integrals*, J. Chem. Phys. **84** (1986), 5617–5623.
[64] J. Grotendorst, E.J. Weniger, and E.O. Steinborn, *Efficient evaluation of infinite-series representations for overlap, two-center nuclear attraction, and Coulomb integrals using nonlinear convergence accelerators*, Phys. Rev. A **33** (1986), 3706–3726.
[65] I.S. Gradshteyn and I.M. Ryzhik, *Table of integrals, series, and products* (Academic Press, New York, 1980).
[66] H.E. Salzer, *A Simple method for summing certain slowly convergent series*, J. Math. and Phys. (Cambridge, Mass.) **33** (1954), 356–359.

[67] H.E. Salzer, *Formulas for the partial summation of series*, Math. Tables Aids Comput. **10** (1956), 149–156.
[68] H.E. Salzer and G.M. Kimbro, *Improved formulas for complete and partial summation of certain series*, Math. Comput. **15** (1961), 23–39.
[69] J. Wimp, *Some transformations of monotone sequences*, Math. Comput. **26** (1972), 251–254.
[70] A. Sidi, *An algorithm for a special case of a generalization of the Richardson extrapolation process*, Numer. Math. **38** (1982), 299–307.
[71] N.E. Nörlund, *Vorlesungen über Differenzenrechnung*, (Chelsea, New York, 1954).
[72] L.M. Milne-Thomson, *The calculus of finite differences* (Chelsea, New York, 1981).
[73] A. Sidi, *Some properties of a generalization of the Richardson extrapolation process*, J. Inst. Math. Appl. **24** (1979), 327–346.
[74] W.F. Ford and A. Sidi, *An algorithm for a generalization of the Richardson extrapolation process*, SIAM J. Numer. Anal. **24** (1987), 1212–1232.
[75] D. Levin and A. Sidi, *Two new classes of nonlinear transformations for accelerating the convergence of infinite integrals and series*, Appl. Math. Comput. **9** (1981), 175–215.
[76] A. Sidi and D. Levin, *Rational approximations from the d-transformation*, IMA J. Numer. Anal. **2** (1982), 153–167.
[77] N. Nielsen, *Die Gammafunktion* (Chelsea, New York, 1965).
[78] N.E. Nörlund, *Leçons sur les séries d'interpolation* (Gautier-Villars, Paris, 1926).
[79] E. Borel, *Leçons sur les séries divergentes* (Gautier-Villars, Paris, 1928). Reprinted by Éditions Jacques Gabay, Paris, 1988.
[80] W. Wasow, *Asymptotic expansions for ordinary differential equations* (Dover, New York, 1987).
[81] S. Iseki and Y. Iseki, *Asymptotic expansion for the remainder of a factorial series*, Mem. Natl. Defense Acad. Japan **20** (1980), 1–6.
[82] E. Landau, *Über die Grundlagen der Theorie der Fakultätenreihen*, Sitzungsb. Königl. Bay. Akad. Wissensch. München, math.-phys. Kl. **36** (1906), 151–218.
[83] G.N. Watson, *The transformation of an asymptotic series into a convergent series of inverse factorials*, Rend. Circ. Mat. Palermo **34** (1912), 41–88.
[84] A. Sidi, *A new method for deriving Padé approximants for some hypergeometric functions*, J. Comput. Appl. Math. **7** (1981), 37–40.
[85] C.M. Bender and T.T. Wu, *Anharmonic oscillator*, Phys. Rev. **184**, (1969), 1231–1260.
[86] B. Simon, *The anharmonic oscillator: A singular perturbation theory*, in D. Bessis (ed.), *Cargèse lectures in physics* (Gordon and Breach, New York, 1972), Vol. 5, pp. 383–414.
[87] B. Simon, *Large orders and summability of eigenvalue perturbation theory: A mathematical overview*, Int. J. Quantum Chem. **21** (1982), 3–25.
[88] J. Čížek and E.R. Vrscay, *Large order perturbation theory in the context of atomic and molecular physics – Interdisciplinary aspects*, Int. J. Quantum Chem. **21** (1982), 27–68.
[89] H.J. Silverstone, J.G. Harris, J. Čížek, and J. Paldus, *Asymptotics of high-order perturbation theory for the one-dimensional anharmonic oscillator by quasisemiclassical methods*, Phys. Rev. A **32** (1985), 1965–1980. See p. 1966, Eq. (1), p. 1977, Eq. (69), and p. 1979, Eq. (71).
[90] J. Čížek, R.J. Damburg, S. Graffi, V. Grecchi, E.M. Harrell II, J.G. Harris, S. Nakai, J. Paldus, R.Kh. Propin, and H.J. Silverstone, *$1/R$ expansion for H_2^+: Calculation of exponentially small terms and asymptotics*, Phys. Rev. A **33** (1986), 12–54. See p. 13, Eq. (2), p. 15, Eqs. (28) and (29), p. 36, Eq. (229), p. 37, Table IV, pp. 38-39, Eq. (232), p. 43, Eq. (236), and p. 45, Eq. (238).
[91] G. Alvarez, *Coupling-constant behavior of the cubic anharmonic oscillator*, Phys. Rev. A **37** (1988), 4079–4083. See p. 4079, Eq. (4), p. 4081, Eq. (24).
[92] J.E. Drummond, *A formula for accelerating the convergence of a general series*, Bull. Austral. Math. Soc. **6** (1972), 69–74.
[93] E.H. Moore and H.L. Smith, *A general theory of limits*, Amer. J. Math. **44** (1922), 102–121.
[94] A.M. Gleason, *Fundamentals of abstract analysis* (Addison-Wesley, Reading, Mass., 1966).
[95] O. Perron, *Die Lehre von den Kettenbrüchen, Band II: Analytisch-funktionentheoretische Kettenbrüche*, (Teubner, Stuttgart, 1957).
[96] H.S. Wall, *Analytic theory of continued fractions*, (Chelsea, New York, 1973).

[97] B. Simon, *Coupling constant analyticity for the anharmonic oscillator*, Ann. Phys. **58** (1970), 76–136.

[98] M. Reed and B. Simon, *Methods of modern mathematical physics IV: Analysis of operators* (Academic Press, New York, 1978).

[99] P. Wynn, *Upon the Padé table derived from a Stieltjes series*, SIAM J. Numer. Anal. **5** (1968), 805–834.

[100] A.K. Common, *Padé approximants and bounds to series of Stieltjes*, J. Math. Phys. **9** (1968), 32–38.

[101] G.D. Allen, C.K. Chui, W.R. Madych, F.J. Narcowich, and P.W. Smith, *Padé approximation of Stieltjes series*, J. Approx. Theor. **14** (1975), 302–316.

[102] J. Karlsson and B. von Sydow, *The convergence of Padé approximants to series of Stieltjes*, Ark. Matem. **14** (1976), 43–53.

[103] A. Sidi, *Borel summability and converging factors for some everywhere divergent series*, SIAM J. Math. Anal. **17** (1986), 1222–1231.

[104] L.J. Slater, *Generalized hypergeometric functions* (Cambridge University Press, Cambridge, 1966).

[105] A. Sidi, *Analysis of convergence of the T-transformation for power series*, Math. Comput. **35** (1980), 833–850.

[106] The NAG Library, Mark 5 (1975), Numerical Analysis Group, NAG Central Office, Oxford, UK.

[107] E.J. Weniger and E.O. Steinborn, *Nonlinear sequence transformations for the efficient evaluation of auxiliary functions for GTO molecular integrals*, in M. Defranceschi and J. Delhalle (eds.), *Numerical determination of the electronic structure of atoms, diatomic and polyatomic molecules* (Kluwer, Dordrecht, 1989), pp. 341–346.

[108] R.A. Levy, *Principles of solid state physics* (Academic Press, New York, 1968).

[109] J. Killingbeck, *Quantum-mechanical perturbation theory*, Rep. Prog. Phys. **40** (1977), 963–1031.

[110] J.P. Delahaye and B. Germain-Bonne, *The set of logarithmically convergent sequences cannot be accelerated*, SIAM J. Numer. Anal. **19** (1982), 840–844.

[111] J. Todd, *The lemniscate constants*, Commun. ACM **18** (1975), 14–19.

[112] E. Filter and E.O. Steinborn, *The three-dimensional convolution of reduced Bessel functions and other functions of physical interest*, J. Math. Phys. **19** (1978), 79–84.

[113] E.O. Steinborn and E. Filter, *Translations of fields represented by spherical-harmonic expansions for molecular calculations III. Translations of reduced Bessel functions, Slater-type s-orbitals, and other functions*, Theor. Chim. Acta **38** (1975), 273–281.

[114] E.J. Weniger and E.O. Steinborn, *Numerical properties of the convolution theorems of B functions*, Phys. Rev. A **28** (1983), 2026–2041.

[115] E. Grosswald, *Bessel polynomials* (Springer-Verlag, Berlin, 1978).

[116] H.E. Salzer, *Note on the Dočev-Grosswald asymptotic series for generalized Bessel polynomials*, J. Comput. Appl. Math. **9** (1983), 131–135.

[117] F. Beleznay, *Estimations for asymptotic series using a modified Romberg algorithm: I. Finite-size scaling calculations*, J. Phys. A **19** (1986), 551–562.

[118] C.-M. Liegener, F. Beleznay, and J. Ladik, *Application of a modified Romberg algorithm to Hartree-Fock calculations on periodic chains*, Phys. Lett. A **123** (1987), 399–401.