International Series of
Numerical Mathematics, Vol. 85

GENERALIZATIONS OF RICHARDSON EXTRAPOLATION WITH APPLICATIONS TO NUMERICAL INTEGRATION

Avram Sidi

TECHNION - Israel Institute of Technology, Department of Computer Science, Haifa 32000, Israel

1. INTRODUCTION

The application of extrapolation (or convergence acceleration) methods to the solution of problems in numerical analysis has become very widespread in recent years. One important area in which extrapolation methods have proved to be remarkably efficient is that of numerical integration. Of the methods that have proved to be useful in this area, the Richardson extrapolation process and some of its generalizations have been the subject of intense research. It is the purpose of this paper to survey briefly the recent developments relating to the generalizations of the Richardson extrapolation process with special emphasis on their application to the numerical evaluation of finite and infinite range integrals.

2. A GENERALIZED RICHARDSON EXTRAPOLATION

2.1 The Extrapolation Problem

Let $A(y)$ be a scalar function of a discrete or continuous variable y, defined for $0 < y \le b < \infty$. Let there exist constants A and α_k, $k=1,2,...$, and functions $\phi_k(y)$, $k=1,2,...$, which form an asymptotic sequence in the sense that

$$\phi_{k+1}(y) = o(\phi_k(y)) \quad \text{as } y \to 0+, \tag{2.1}$$

and assume that $A(y)$ has the asymptotic expansion

$$A(y) \sim A + \sum_{k=1}^{\infty} \alpha_k \phi_k(y) \quad \text{as } y \to 0+. \tag{2.2}$$

Here $A(y)$ and $\phi_k(y)$, $k=1,2,...$, are assumed to be known for $0 < y \le b$, but A and α_k, $k=1,2,...$, are unknown. The problem is to approximate A, which, in many cases is $\lim_{y\to 0+} A(y)$ when the latter exists. (When $\lim_{y\to 0+} A(y)$ does not exist A is said to be the antilimit of $A(y)$ as $y \to 0+$.)

2.2 Method of Approximation (Generalized Richardson Extrapolation)

Pick a decreasing sequence y_l, l=0,1,2,..., in $(0,b]$, such that $\lim_{l\to\infty} y_l = 0$. Then the solution for A_p^j of the system of linear equations

$$A(y_l) = A_p^j + \sum_{k=1}^{p} \bar{\alpha}_k \phi_k(y_l),\ j \le l \le j+p, \tag{2.3}$$

is taken to be an approximation to A. We note that the equations in (2.3) are obtained by truncating the infinite sum in (2.2) at the term $\alpha_p \phi_p(y)$, replacing A and $\alpha_1, \ldots, \alpha_p$ by A_p^j and $\bar{\alpha}_1, \ldots, \bar{\alpha}_p$ respectively, treating the latter as unknowns (p+1 in number), and collocating at the points $y_j, y_{j+1}, \ldots, y_{j+p}$ to obtain p+1 linear equations for these unknowns. The process described by (2.3) is known as a generalized Richardson extrapolation process.

For some cases of interest it can be shown that for $j \to \infty$ with p fixed $A_p^j - A = O(\phi_{p+1}(y_j))$, i.e., the effect of $\phi_1, \ldots, \phi_p$ is eliminated by the extrapolation process, and A_p^j tends to A more quickly than $A(y_{j+p})$ for $j \to \infty$. the best convergence properties, however, are possessed by the limiting process in which j is fixed and $p \to \infty$. This is observed in most cases of interest, and can be proved rigorously for some cases.

In the problem that is solved by the classical Richardson extrapolation process $\phi_i(y) = y^i$, i=1,2,..., and this problem arises, for example, from finite difference approximations of derivatives, Euler-Maclaurin expansions for the trapezoidal rule approximation of finite range integrals of smooth functions, etc. Bulirsch and Stoer (1964) consider the case in which $\phi_i(y) = y^{\gamma_i}$, i=1,2,..., with arbitrary real γ_i and $y_l = y_0 \rho^l$, $0 < \rho < 1$, and give a thorough convergence analysis for it.

The general setting of (2.1)-(2.3) above can be found in Hart et al. (1968, p.39). By Cramer's rule it can be shown that A_p^j has a determinant representation. Levin (1973) seems to be the first to point this out explicitly. This representation may be used effectively in deriving algorithms for computing the A_p^j, and also in analyzing them. As pointed out also by Brezinski (1980), many of the known convergence acceleration methods are directly or indirectly related to the general setting described above.

It should be mentioned, however, that unless the functions $\phi_k(y)$ in (2.2) are known precisely, the generalized Richardson extrapolation, as defined through (2.3), may be useless in the sense that if $A(y)$ does not satisfy (2.2) with those $\phi_k(y)$ that are used in (2.3), the approximations A_p^j to A will not be better than $A(y_j)$,...,$A(y_{j+p})$. Next, in order for the extrapolation method defined through (2.3) to be useful, the functions $\phi_k(y)$ have to be easily computable. Therefore, the first important task of the practitioner is to determine what kind of an asymptotic expansion $A(y)$ has for $y \to 0+$. While for some problems this may be available a priori, for others careful analysis may be necessary. As a rule, a good extrapolation method then is one which requires as little analysis as possible and can be applied to as large a family of functions as possible. Finally, the choice of the y_l in (2.3) is very crucial since it effects the numerical stability of A_p^j very strongly. We shall say more on these points in the next sections.

2.3 Recursive Algorithms for A_p^j

It seems that Schneider (1975) was the first to give a recursive algorithm for the implementation of the extrapolation process defined through (2.3) when the $\phi_k(y_l)$ have no particular structure. Using different techniques, the same algorithm was later derived by Håvie (1979), and then by Brezinski (1980), who designated it the E-algorithm. Håvie (1979) also gave simplified versions of the E-algorithm for some special cases. Brezinski (1980) proved some convergence results for A_p^j in the case p is fixed and $j \to \infty$, assuming that the $\phi_j(y_l)$ satisfy certain conditions.

Recently Ford and Sidi (1987) derived a different recursive algorithm for implementing the same extrapolation process, and their algorithm turns out to be computationally more economical than the E-algorithm. The steps of the Ford-Sidi algorithm are summarized below:
Denote an arbitrary sequence $b(l)$, l=0,1,2,..., by b. The sequence 1,1,1,..., will be denoted by I. Define the sequences g_k, k=1,2,..., and a by $g_k(l) = \phi_k(y_l)$, k=1,2,..., and $a(l) = A(y_l)$, l=0,1,2,... .

(1) For $b = a, I$, and $b = g_k$, k=2,3,..., set $\psi_0^j(b) = b(j)/g_1(j)$, j=0,1,... .

(2) For p = 1,2,..., let $D_p^j = \psi_{p-1}^{j+1}(g_{p+1}) - \psi_{p-1}^j(g_{p+1})$
and $\psi_p^j(b) = [\psi_{p-1}^{j+1}(b) - \psi_{p-1}^j(b)]/D_p^j$ with $b = a, I$, and $b = g_k$, $k \geq p+1$, j=0,1,... .

(3) Set $A_p^j = \psi_p^j(a)/\psi_p^j(I)$, all $j, p \geq 0$.

For details see Ford and Sidi (1987).

3. GREP: A FURTHER GENERALIZATION OF RICHARDSON EXTRAPOLATION

3.1 The Extrapolation Problem

Let us now assume that the function $A(y)$ is of the form

$$A(y) = A + \sum_{k=1}^{m} \phi_k(y)\beta_k(y). \tag{3.1}$$

As before, A is the limit or antilimit of $A(y)$ as $y \to 0+$, and the problem is to approximate it. The functions $\phi_k(y)$ are also as before, except that they are not required to satisfy (2.1). The functions $\beta_k(y)$ are defined for $0 \le y \le b$, and have the asymptotic expansions

$$\beta_k(y) \sim \sum_{i=0}^{\infty} \beta_{ki} y^{ir_k} \quad \text{as } y \to 0+, \tag{3.2}$$

where $r_k > 0$ are known constants. Again $A(y)$ and $\phi_k(y)$, $k=1,\dots,m$, are assumed to be known for $0 < y \le b$, but A and the β_{ki} are unknown. The class of such functions $A(y)$ is denoted $F^{(m)}$.

3.2 The Method of Approximation: GREP

Pick a decreasing sequence y_l, $l=0,1,2,\dots$, in $(0,b]$, such that $\lim_{l\to\infty} y_l = 0$. Then define $A_n^{(m,j)}$, the approximation to A, and the parameters $\bar{\beta}_{ki}$, through the system of linear equations

$$A(y_l) = A_n^{(m,j)} + \sum_{k=1}^{m} \phi_k(y_l) \sum_{i=0}^{n_k} \bar{\beta}_{ki} y_l^{ir_k}, \; j \le l \le j+N, \tag{3.3}$$

where n denotes the vector of integers $(n_1,\dots,n_m)$, and $N = \sum_{k=1}^{m}(n_k+1)$.

This extrapolation method, which is denoted GREP, was considered by Sidi (1979) (see also Davis and Rabinowitz (1983, pp. 46-47)), who also gave some error bounds and convergence results for it. In the same paper several examples of functions $A(y)$ in $F^{(m)}$ that arise naturally in practical numerical problems are given. Actually, GREP was proposed as a general framework to deal with these problems in a unified fashion.

A very efficient recursive technique for solving (3.3) for $A_n^{(m,j)}$ in the case $m=1$ was given by Sidi (1982b), and it was denoted the W-algorithm. The W-algorithm will be described in Section 7 in conjunction with a new extrapolation method for infinite oscillatory integrals.

Recently Ford and Sidi (1987) gave a very efficient recursive technique for solving (3.3) for $A_n^{(m,j)}$ when m is arbitrary and $r_k = r$, $k=1,\dots,m$, which is probably the most common form of GREP. This technique, denoted the $W^{(m)}$-algorithm, reduces to the W-algorithm when $m=1$. A FORTRAN 77 program that implements the $W^{(m)}$-algorithm can be found in Ford and Sidi (1987).

Finally, the expansion in (3.1) and (3.2) can be viewed as a generalization of that in (2.2) in the sense that (1) the unknown constants α_k in (2.2) are now being replaced by some unknown functions $\beta_k(y)$ of known form for $y \to 0+$ and (2) as opposed to $A(y)$ in Section 2, which is represented by one asymptotic expansion as $y \to 0+$, $A(y)$ in the present section is represented by a sum of m different asymptotic expansions as $y \to 0+$.

The functions $\beta_k(y)$ can be further generalized to have the asymptotic expansions

$$\beta_k(y) \sim \sum_{i=0}^{\infty} \beta_{ki}\,\omega_{ki}(y) \quad \text{as } y \to 0+, \tag{3.2'}$$

with the functions $\omega_{ki}(y)$, $i=0,1,\ldots$, forming an asymptotic sequence as $y \to 0+$ for each $k=1,\ldots,m$.

4. PRACTICAL REMARKS ON THE USE OF GREP

Let us write the linear system in (3.3) in the form $Qc = d$, where $c = (A_n^{(m,j)}, \bar{\beta}_{10}, \ldots, \bar{\beta}_{k,n_k})^T$, $d = (A(y_j), \ldots, A(y_{j+N}))^T$, and Q is the $(N+1)\times(N+1)$ matrix of this system with the vector $(1,1,\ldots,1)^T$ as its first column. The solution for $A_n^{(m,j)}$ can now be expressed as

$$A_n^{(m,j)} = \sum_{l=0}^{N} \gamma_l A(y_{j+l}), \tag{4.1}$$

where $(\gamma_0, \gamma_1, \ldots, \gamma_N)$ is the first row of the matrix Q^{-1}. Obviously, $\sum_{l=0}^{N} \gamma_l = 1$, thus $\sum_{l=0}^{N} |\gamma_l| \geq 1$. The theoretical results given in Sidi (1979) and numerical results suggest that best convergence is attained for the case in which $n_1, \ldots, n_m$ all tend to infinity simultaneously. (This can be achieved, for example, by taking $n_k = \nu$, $k=1,\ldots,m$, and by letting $\nu \to \infty$.) It is interesting to note that $\kappa_\infty(Q) \geq \sum_{l=0}^{N} |\gamma_l|$, where $\kappa_\infty(Q)$ denotes the condition number of Q in the l_∞-norm, although $\kappa_\infty(Q)$ may be much larger than $\sum_{l=0}^{N} |\gamma_l|$, in general. This ill-conditioning of Q does not seem to affect the numerical accuracy of $A_n^{(m,j)}$, however. What does affect the numerical accuracy of $A_n^{(m,j)}$ is in fact $\sum_{l=0}^{N} |\gamma_l|$, as will be described in the next paragraph. In addition, there is a strong connection between the rate of convergence of $A_n^{(m,j)}$ to A and the rate at which $\sum_{l=0}^{N} |\gamma_l|$ changes as $n_k \to \infty$, $k=1,\ldots,m$. In general, the following are observed to take place simultaneously:

(1) $A_n^{(m,j)} \to A$ quickly.

(2) $\sum_{l=0}^{N} |\gamma_l|$ is small, and if it increases, its rate of increase is slow.

In relation to (2) we note that if $\sum_{l=0}^{N} |\gamma_l|$ increases quickly, then the approximation $A_n^{(m,j)}$ becomes less stable numerically too. As a matter of fact, if ε denotes the maximum of the absolute errors in the values of $A(y_l)$, $j \leq l \leq j+N$, then an estimate of the error in the computed value of $A_n^{(m,j)}$ is $(\sum_{l=0}^{N} |\gamma_l|)\varepsilon$.

This estimate seems to be quite realistic in many cases of interest.

In practice, the growth of $\sum_{l=0}^{N} |\gamma_l|$ can be effectively controlled by a wise choice of the y_l that go into the definition of GREP. As a rule of thumb, $\sum_{l=0}^{N} |\gamma_l|$ has a small rate of increase or stays bounded if the sequences $\phi_k(y_l)$, $l=0,1,2,...$ are quickly varying, for each k. Quick variation can be achieved by picking the y_l such that, for instance, $\phi_k(y_l) = C_k(l)\exp(p_k(l))$, where $C_k(l)$ and $p_k(l)$ are slowly varying functions of l. (A function is slowly varying in l if, for example, it is monotonic and behaves like δl^{α} as $l \to \infty$, for some constants δ and α.) The most ideal situation is, of course, one in which $\sum_{l=0}^{N} |\gamma_l| = 1$, i.e., $\gamma_l \geq 0$ for all l. This can be achieved, again by a wise choice of the y_l, for a large class of infinite series and integrals of oscillatory nature.

5. APPLICATION OF GREP TO FINITE RANGE INTEGRALS

Through the appropriate Euler-Maclaurin expansions for the (ordinary or offset) trapezoidal rule approximations, GREP can be used to obtain approximations to one-dimensional finite range integrals and to N-dimensional integrals over hypercubes or hypersimplexes. The best known example of this is the classical Romberg integration for smooth integrands, see, for example, Davis and Rabinowitz (1983).

We now wish to mention briefly the recent developments regarding Euler-Maclaurin expansions for singular integrals.

Euler-Maclaurin expansions for singular integrals of the form $\int_0^1 w(x)f(x)dx$, where $w(x) = x^s(\log x)^{s'}$, $s > -1$, $s' = 0,1,...$, and $f(x)$ is smooth, were first given in two important works by Navot (1961, 1962). Using a different approach, Navot's results were later derived by Lyness and Ninham (1967). Lyness (1976) derived Euler-Maclaurin expansions for N-dimensional integrals over the hypercube of functions having a corner singularity, and Lyness and Monegato (1980) derived the Euler-Maclaurin expansions for integrals over the hypersimplex of functions having singularities at the vertices of the hypersimplex. Recently, Sidi (1983) derived Euler-Maclaurin expansions for integrals over the two-dimensional unit square or standard simples of functions having algebraic and/or logarithmic edge singularities. The techniques and results of Sidi (1983) can be extended to N dimensions.

For further references on Euler-Maclaurin expansions and their generalizations, see Davis and Rabinowitz (1983).

In all of the above mentioned works, the integral in question is being approximated by a trapezoidal rule first. Let us denote this approximation by $A(h)$, where h denotes the integration step size. If $I[f]$ denotes the value of the integral, then the Euler-Maclaurin expansion, roughly speaking, takes the form

$$A(h) \sim I[f] + \sum_{k=1}^{m} \phi_k(h) \sum_{i=0}^{\infty} \beta_{ki} h^{ir_k} \quad \text{as } h \to 0, \tag{5.1}$$

with $\phi_k(h)$ being of the form h^{μ_k} or $h^{\mu_k} \log h$, etc., depending on the nature of the singularities. Obviously, with the correspondence $y \leftrightarrow h$, (5.1) is of the form mentioned in Section 3, and hence by picking a suitable sequence $y_l \equiv h_l$, l=0,1,2,..., of integration step sizes, GREP is directly applicable. If the sequence h_l, l=0,1,2,..., is taken such that $h_l = h_0 \rho^l$, for some h_0 and some ρ, $0 < \rho < 1$, for example, the approximations $A^{(m,j)}_{(\nu,\nu,\ldots,\nu)}$ are very accurate and numerically stable with increasing ν, as $\sum_{l=0}^{N} |\gamma_l|$ stays small. This choice of the h_l may make GREP prohibitively expensive, as the number of function evaluations increases exponentially with ν. In fact, for multidimensional integrals this number becomes enormously large even for small values of ν. For these integrals it becomes more economical to take, for instance, $h_l = h_0/(l+1)$, l=1,2,..., for some h_0. In this case, $A^{(m,j)}_{(\nu,\nu,\ldots,\nu)}$ is prone to numerical instability with increasing ν, as $\sum_{l=0}^{N} |\gamma_l|$ increases rapidly with ν. Yet such a choice of the h_l may be very useful, provided one starts with more accuracy in the $A(h_l)$ than is required in the final answer.

Example 5.1. Consider the case $I[f] = \int_0^1 x^{-1/2} f(x)dx$ with $A(h) = h\sum_{i=1}^{n} x_i^{-1/2} f(x_i)$, where $x_i = (i-1/2)h$, $h = 1/n$, n=1,2,... . Then, from Navot (1961), we have

$$A(h) \sim I[f] + \sum_{i=1}^{\infty} \alpha_i h^{2i} + \sum_{i=1}^{\infty} \beta_i h^{i-1/2} \quad \text{as } h \to 0, \tag{5.2}$$

i.e., m=2, $\phi_1(h) = h^2$, $\phi_2(h) = h^{1/2}$, $r_1 = 2$, $r_2 = 1$. In Table 5.1 we display $\sum_{l=0}^{N} |\gamma_l|$ for $n_1 = n_2 = \nu = 0,1,\ldots,5$ with $h_l = 2^{-l}$ and $h_l = (l+1)^{-1}$.

$n_1=n_2=\nu$	$\sum_{l=0}^{N} \mid\gamma_l\mid$ with $h_l=2^{-l}$	$\sum_{l=0}^{N} \mid\gamma_l\mid$ with $h_l=(l+1)^{-1}$
0	9.71	1.74×10^1
1	23.05	3.23×10^2
2	33.40	6.35×10^3
3	40.91	1.29×10^5
4	44.78	2.69×10^6
5	46.82	5.69×10^7

Table 5.1 - $\sum_{l=0}^{N} \mid\gamma_l\mid$ for Example 5.1

6. APPLICATION OF GREP TO INFINITE RANGE INTEGRALS

One of the most interesting applications of extrapolation methods in general and GREP in particular is that of computing infinite integrals $I[f] = \int_a^\infty f(t)dt$ of functions $f(x)$ having different types of behavior at infinity. The main result that enables us to employ extrapolation methods in this case is Theorem 1 in Levin and Sidi (1981), see also Sidi (1979, Theorem 2.1). Roughly speaking, this theorem says the following:

Let $f(x)$ be defined for $x > a \geq 0$, be integrable at infinity, and satisfy a homogeneous linear differential equation of order m of the form

$$f(x) = \sum_{k=1}^{m} p_k(x) f^{(k)}(x), \tag{6.1}$$

where

$$p_k(x) \sim \sum_{j=0}^{\infty} p_{kj} x^{i_k - j} \quad \text{as } x \to \infty, \tag{6.2}$$

with i_k an integer $\leq k$, $k=1,\ldots,m$. Then, under mild conditions on $f(x)$ and $p_k(x)$,

$$\int_x^\infty f(t)dt \sim \sum_{k=0}^{m-1} x^{\rho_k} f^{(k)}(x) \sum_{i=0}^{\infty} \beta_{ki} x^{-i} \quad \text{as } x \to \infty, \tag{6.3}$$

for some integers ρ_k such that $\rho_k \leq k+1$, $k=0,\ldots,m-1$.

Based on the fact that $I[f] = \int_a^\infty f(t)dt = \int_a^x f(t)dt + \int_x^\infty f(t)dt$, and on the expansion in (6.3), Levin and Sidi (1981) define the D-transformation for approximating $I[f]$. This transformation is defined by the linear system of equations

$$\int_a^{x_l} f(t)\,dt = D_n^{(m,j)} + \sum_{k=0}^{m-1} x_l^{\rho_k} f^{(k)}(x_l) \sum_{i=0}^{n_k} \bar{\beta}_{ki} x_l^{-i}, \qquad j \le l \le j+N, \tag{6.4}$$

where $N = \sum_{k=0}^{m-1}(n_k+1)$, n denotes the vector $(n_0,n_1,...,n_{m-1})$, and the x_l are chosen so that $a < x_0 < x_1 <...,$ and $\lim_{l\to\infty} x_l = \infty$.

Obviously, the D-transformation is GREP with $y \equiv \frac{1}{x}$, $A(y) \equiv \int_a^x f(t)\,dt$, and $\phi_k(y) \equiv x^{\rho_{k-1}} f^{(k-1)}(x)$, $r_k = 1$, $k=1,...,m$. The P-transformation of Levin (1975) for Fourier transforms is a special case of the D-transformation with $m=1$ and $x_l = l+1$, $l=0,1,...$.

The D-transformation can be modified slightly by replacing ρ_k in (6.4) by $k+1$, i.e., writing $\phi_k(y) \equiv x^k f^{(k-1)}(x)$, $k=1,...,m$, thus making the knowledge of the ρ_k unnecessary. This makes the use of the D-transformation easier, as the only input required from the user now is the integer m (the order of the differential equation (6.1)-(6.2) satisfied by $f(x)$), and the choice of the x_l. When in doubt, one may begin with $m=1,2,...,$ etc. Levin and Sidi (1981) provide some practical means for determining m by inspection of $f(x)$.

The $W^{(m)}$-algorithm of Ford and Sidi (1987) can be used to implement the D-transformation very efficiently.

Example 6.1. $I[f] = \int_0^\infty \log(1+t)/(1+t^2)\,dt = 1.4603621167531193.$

The integrand $f(x) = \log(1+x)/(1+x^2)$ satisfies (6.1) and (6.2) with $m=2$. Replacing ρ_k by $k+1$ in (6.4), and picking $x_l = e^{0.4l}$, $l=0,1,...,$ in the D-transformation, we obtain the results shown in Table 6.1.

ν	Relative error in $D^{(2,0)}_{(\nu,\nu)}$
0	0.613 D0
2	0.501 D-3
4	0.481 D-6
6	0.182 D-11
8	0.745 D-14

Table 6.1 - D-transformation for Example 6.1

The integrals $F(x_l) = \int_0^{x_l} f(t)\,dt$, $l=0,1,...,2\nu+2$, are used in constructing $D^{(2,0)}_{(\nu,\nu)}$. Although $D^{(2,0)}_{(8,8)}$ has relative error of order 10^{-14}, $F(x_{19})$ has relative error of order 10^{-3}.

The example above has also been treated in Levin and Sidi (1981) with different parameters. The reason for the quick convergence shown in Table 6.1 is that $\sum_{l=0}^{N} |\gamma_l|$ stays small for increasing ν. If we choose $x_l = l+1$, $l=0,1,\ldots$, for example, $\sum_{l=0}^{N} |\gamma_l|$ increases rapidly with increasing ν, hence $D_{(\nu,\nu)}^{(2,0)}$ converges less rapidly than in Table 6.1, and is also prone to numerical instabilities.

We mention in passing that Levin and Sidi (1981) give a similar treatment to infinite series $\sum_{r=1}^{\infty} f_r$, which results in the development of the d-transformation.

To summarize, we mention that the D-transformation, as described above, can be successfully applied to a wide range of infinite integrals, with no more computational effort than required for the simplest extrapolation methods. This can be attributed to (1) the simplicity of the $\phi_k(y)$, which can all be expressed in a straightforward way in terms of $f(x)$ and its derivatives, and (2) the recursive implementation of the D-transformation by the $W^{(m)}$-algorithm. Of course, the simple nature of the $\phi_k(y)$ is a consequence of Theorem 1 of Levin and Sidi (1981) that produces (6.3).

7. GREP AND INFINITE OSCILLATORY INTEGRALS

When the integrand $f(x)$ in the integral $I[f] = \int_a^{\infty} f(t)dt$ is oscillatory at infinity the D-transformation may be made less expensive computationally by picking the x_l in a special way. Two such modifications of the D-transformation, namely, the $\bar{D}$- and $\tilde{D}$-transformations, have been proposed for such integrals in Sidi (1980). Later, a very efficient extrapolation method, designated the W-transformation, was proposed in Sidi (1982a) for very oscillatory infinite integrals (see also Davis and Rabinowitz (1983, pp. 233-235)). This work was later extended in Sidi (1987) to cover divergent ocillatory infinite integrals that are defined in the sense of Abel summability. We do not intend to go into any of the methods mentioned above. We shall only mention that all of these methods are GREP's of different forms.

Now the W-algorithm involves some asymptotic analysis of the integrand $f(x)$ as $x \to \infty$. This analysis covers both the phase of oscillations and the amplitude of $f(x)$ and is partly quantitative in nature. A modification of the W-transformation that involves the analysis of the phase of oscillations only was recently developed in Sidi (1988), and is as efficient as the W-transformation. We give a brief description of this method in the remainder of this section. As will become clear, the modified W-transformation too is a form of GREP.

7.1 The Modified W-transformation

We consider integrals of the form $I[f] = \int_a^\infty f(t)dt$, $a \ge 0$, with $f(x)$ being of the form

$$f(x) = \sum_{j=1}^{r} u_j(\theta_j(x))H_j(x), \tag{7.1}$$

where

(1) $u_j(z)$ is e^{iz} or e^{-iz} or any linear combination of those.

(2) roughly speaking, $\theta_j(x)$ are real functions of the form

$$\theta_j(x) \sim \sum_{i=0}^{m-1} \mu_i x^{m-i} + \alpha_{j0} + \alpha_{j1}x^{-1} + \alpha_{j2}x^{-2} + \dots \quad \text{as } x \to \infty, \ \mu_0 > 0\,. \tag{7.2}$$

i.e., the polynomial parts of $\theta_j(x)$, $j=1,\dots,r$ are identical. Denote them by $\overline{\theta}(x) = \sum_{i=0}^{m-1} \mu_i x^{m-i}$.

(3) $H_j(x)$ are arbitrary functions that do not oscillate as $x \to \infty$, or they oscillate very slowly compared with $u_j(\theta_j(x))$.

The integral $\int_a^\infty f(t)dt$ may be convergent or divergent, and in case it diverges, it exists in the sense of Abel summability, at least for a subset of functions $f(x)$ considered above.

The modified W-transformation involves the following steps:

(1) Pick x_0 to be largest zero of $\sin(\overline{\theta}(x)) = 0$ which is larger than a. x_0 is then the largest solution of $\overline{\theta}(x) = q\pi$ for some integer q. Next, determine x_l to be the largest solution of $\overline{\theta}(x) = (q+l)\pi$, $l=1,2,\dots$. Consequently, $a < x_0 < x_1 < \dots$, and $\lim_{l\to\infty} x_l = \infty$.

(2) Set $x_{-1} = a$, and compute the finite integrals

$$\psi(x_l) = \int_{x_l}^{x_{l+1}} f(t)dt, \quad l = -1,0,1,2,\dots, \tag{7.3}$$

and

$$F(x_l) = \int_a^{x_l} f(t)dt = \sum_{i=-1}^{l-1} \psi(x_i), \quad l = 0,1,\dots\,. \tag{7.4}$$

(3) The approximation $W_n^{(j)}$ to $I[f]$ is then defined as the solution of the linear system of equations

$$W_n^{(j)} = F(x_l) + \psi(x_l)\sum_{i=0}^{n}\beta_i/x_l^i, \quad j \le l \le j+n+1. \tag{7.5}$$

The W-algorithm of Sidi (1982b) can be used to compute $W_n^{(j)}$ very efficiently as follows: Set

$$M_{-1}^{(s)} = F(x_s)/\psi(x_s), \quad N_{-1}^{(s)} = 1/\psi(x_s), \quad s = 0,1,\ldots\,, \tag{7.6}$$

and compute for $s=0,1,\ldots,\ p=0,1,\ldots,$

$$\begin{aligned} M_p^{(s)} &= \left[M_{p-1}^{(s)} - M_{p-1}^{(s+1)}\right] / \left[x_s^{-1} - x_{s+p+1}^{-1}\right] \\ N_p^{(s)} &= \left[N_{p-1}^{(s)} - N_{p-1}^{(s+1)}\right] / \left[x_s^{-1} - x_{s+p+1}^{-1}\right] \\ W_p^{(s)} &= M_p^{(s)}/N_p^{(s)}. \end{aligned} \tag{7.7}$$

Again, $W_n^{(j)}$ has the best convergence properties for fixed j and increasing n.

Example 7.1. $I[f] = \int_0^\infty x^2 J_0(x)\,dx = -1.$

$J_0(x)$ is the Bessel function of order zero, and $I[f]$ exists only in the sense of Abel summability. The integrand $f(x) = x^2 J_0(x)$ is of the form described in the beginning of this subsection. Specifically, $f(x) = (\cos x)H_1(x) + (\sin x)H_2(x)$, since $J_0(x) = (\cos x)h_1(x) + (\sin x)h_2(x)$, where $h_1(x)$ and $h_2(x)$ do not oscillate at infinity. Thus $\theta_1(x) = \theta_2(x) = \bar{\theta}(x) = x$. Hence $x_l = (l+1)\pi$, $l=0,1\ldots$. Table 7.1 displays some of the results of the computation for this example .

n	$F(x_n)$	relative error in $W_n^{(0)}$
0	-0.776 D+01	0.865 D+00
2	-0.233 D+02	0.171 D-01
4	-0.437 D+02	0.443 D-05
6	-0.681 D+02	0.615 D-08
8	-0.959 D+02	0.744 D-11
10	-0.127 D+03	0.143 D-12

Table 7.1 - Modified W-transformation for Example 7.1

Sidi (1988) provides a detailed convergence analysis of $W_n^{(j)}$ both for $j \to \infty$, n fixed, and for j fixed, $n \to \infty$, when the integrand is in the class of functions denoted **B** there. The class **B** is a subset of that considered at the beginning of this subsection. For j fixed and $n \to \infty$, it can be shown that, when $f \in \mathbf{B}$

$$I[f] - W_n^{(j)} = o(n^{-\mu}) \quad \text{as } n \to \infty, \text{ any } \mu > 0. \tag{7.8}$$

The results for n fixed and $j \to \infty$ are more complicated to state, in general, without introducing more notation. For the case in which the functions $H_j(x)$ in (7.2) have the simple structure $H_j(x) \sim \sum_{s=0}^{\infty} h_{js} x^{\gamma_j - s}$ as $x \to \infty$, with γ_j arbitrary but $\gamma_j - \gamma_p$ = integer for all j and p when $r > 1$, we have

$$|I[f] - W_n^{(j)}| \sim C\, x_j^{\gamma - mn - m - n - 1} \quad \text{as } j \to \infty, \gamma = \max\{\gamma_1, \ldots, \gamma_r\}, \tag{7.9}$$

for some positive constant C. This should be compared with $|I[f] - F(x_{j+n+1})| \sim C_1 x_j^{\gamma - m + 1}$ as $j \to \infty$. For these and other additional results we refer the interested reader to the original paper.

REFERENCES

1. C. Brezinski (1980), A general extrapolation algorithm, *Numer. Math.*, **35**, pp. 175-187.
2. R. Bulirsch and J. Stoer (1964), Fehlerabschätzungen und Extrapolation mit rationalen Funktionen bei Verfahren vom Richardson-Typus, *Numer. Math.*, **6**, pp. 413-427.
3. P.J. Davis and P. Rabinowitz (1983), *Methods of Numerical Integration*, 2nd Edition, Academic Press, New York.
4. W.F. Ford and A. Sidi (1987), An algorithm for a generalization of the Richardson extrapolation process, *SIAM J. Numer. Anal.*, **24**, pp. 1212-1232.
5. J.F. Hart et al. (1968), *Computer Approximations*, SIAM Series in Applied Mathematics, John Wiley and Sons, Inc., New York.
6. T. Håvie (1979), Generalized Neville type extrapolation schemes, *BIT*, **19**, pp. 204-213.
7. D. Levin (1973), Development of non-linear transformations for improving convergence of sequences, *Intern. J. Comp. Math.*, **B3**, pp. 371-388.
8. D. Levin (1975), Numerical inversion of the Laplace transform by accelerating the convergence of Bromwich's integral, *J. Comp. Appl. Math.*, **1**, pp. 247-250.
9. D. Levin and A. Sidi (1981), Two new classes of non-linear transformations for accelerating the convergence of infinite integrals and series, *Appl. Math. Comp.*, **9**, pp. 175-215.
10. J.N. Lyness (1976), An error functional expansion for N-dimensional quadrature with an integrand function singular at a point, *Math. Comp.*, **30**, pp. 1-23.
11. J.N. Lyness and G. Monegato (1980), Quadrature error expansions for the simplex when the integrand function has singularities at vertices, *Math. Comp.*, **34**, pp. 213-225.
12. J.N. Lyness and B.W. Ninham (1967), Numerical quadrature and asymptotic expansions, *Math. Comp.*, **21**, pp. 162-178.

13. I. Navot (1961), An extension of the Euler-Maclaurin summation formula, *J. Math. and Phys.*, **40**, pp. 271-276.

14. I. Navot (1962), A further extension of the Euler-Maclaurin summation formula, *J. Math. and Phys.*, **41**, pp. 155-163.

15. C. Schneider (1975), Vereinfachte Rekursionen zur Richardson-Extrapolation in Spezialfällen, *Numer. Math.*, **24**, pp. 177-184.

16. A. Sidi (1979), Some properties of a generalization of the Richardson extrapolation process, *J. Inst. Maths. Applics.*, **24**, pp. 327-346.

17. A. Sidi (1980), Extrapolation methods for oscillatory infinite integrals, *J. Inst. Maths. Applics.*, **26**, pp. 1-20.

18. A. Sidi (1982a), The numerical evaluation of very oscillatory infinite integrals by extrapolation, *Math. Comp.*, **38**, pp. 517-529.

19. A. Sidi (1982b), An algorithm for a special case of a generalization of the Richardson extrapolation process, *Numer. Math.*, **38**, pp. 299-307.

20. A. Sidi (1983), Euler-Maclaurin expansions for integrals over triangles and squares of functions having algebraic/logarithmic singularities along an edge, *J. Approx. Th.*, **39**, pp. 39-53. See also TR#161, Computer Science Dept., Technion, Haifa, 1979.

21. A. Sidi (1987), Extrapolation methods for divergent oscillatory infinite integrals that are defined in the sense of summability, *J. Comp. Appl. Math.*, **17**, pp. 105-114.

22. A. Sidi (1988), A user-friendly extrapolation method for oscillatory infinite integrals, *Math. Comp.*, **51**, (in press).