

On Accelerating the Convergence of Infinite Double Series and Integrals

By David Levin

Abstract. The generalization of Shanks' ϵ -transformation to double series is discussed and a class of nonlinear transformations, the $[A/S]_R$ transformations, for accelerating the convergence of infinite double series is presented. It is constructed so as to sum exactly infinite double series whose terms satisfy certain finite linear double difference equations; in that sense it is a generalization of Shanks' ϵ -transformation or its equivalent Wynn's ϵ -algorithm. A generalization of the $[A/S]_R$ transformation to N -dimensional series is also presented and their application to power series is discussed and exemplified. Some transformations for accelerating the convergence of infinite double integrals are also obtained, generalizing the confluent ϵ -algorithm of Wynn and the G -transformation of Gray, Atchison, and McWilliams for infinite 1-D integrals.

1. Introduction. One of the important properties of Shanks' ϵ -transformation or its equivalent Wynn's ϵ -algorithm is that it sums exactly converging series whose terms satisfy linear difference equations with constant coefficients (Shanks [8], Wynn [9]), i.e., if

$$(1.1) \quad a_i = \sum_{k=1}^n \alpha_k \Delta^k a_i, \quad i = 0, 1, 2, \dots,$$

then

$$(1.2) \quad \epsilon_{2n}(A_i) = \sum_{j=0}^{\infty} a_j \equiv S, \quad i = 0, 1, 2, \dots,$$

where $A_i = \sum_{j=0}^i a_j$.

The above property can be proved as follows: From the relation (1.1) it can be shown that

$$(1.3) \quad \sum_{j=i}^{\infty} a_j = - \sum_{k=1}^n \alpha_k \Delta^{k-1} a_i, \quad i = 0, 1, 2, \dots,$$

and, therefore,

$$(1.4) \quad A_i = S + \sum_{k=1}^n \alpha_k \Delta^{k-1} a_{i+1}, \quad i = 0, 1, 2, \dots$$

Regarding S and the constants α_k as unknowns, we can find the sum S of a series whose terms satisfy any difference equation of the form (1.1) by solving the $(n + 1)$ th order linear system of equations

$$(1.5) \quad A_i = S + \sum_{k=1}^n \alpha_k \Delta^{k-1} a_{i+1}, \quad i = m, m + 1, \dots, m + n.$$

Received November 19, 1979.

1980 *Mathematics Subject Classification.* Primary 65B10; Secondary 65D15.

© 1980 American Mathematical Society
 0025-5718/80/0000-0175/\$04.75

The solution of this system for S is

$$(1.6) \quad S = \frac{\begin{vmatrix} A_m & A_{m+1} & \cdots & A_{m+n} \\ a_{m+1} & a_{m+2} & \cdots & a_{m+n+1} \\ \Delta a_{m+1} & \Delta a_{m+2} & \cdots & \Delta a_{m+n+1} \\ \vdots & \vdots & & \vdots \\ \Delta^{n-1} a_{m+1} & \Delta^{n-1} a_{m+2} & \cdots & \Delta^{n-1} a_{m+n+1} \end{vmatrix}}{\begin{vmatrix} 1 & 1 & \cdots & 1 \\ a_{m+1} & a_{m+2} & \cdots & a_{m+n+1} \\ \Delta a_{m+1} & \Delta a_{m+2} & \cdots & \Delta a_{m+n+1} \\ \vdots & \vdots & & \vdots \\ \Delta^{n-1} a_{m+1} & \Delta^{n-1} a_{m+2} & \cdots & \Delta^{n-1} a_{m+n+1} \end{vmatrix}}$$

which is identical to $\epsilon_{2n}(A_m)$ (Shanks [8]).

The relations (1.4) can also be written as

$$(1.7) \quad A_i = S - \sum_{k=1}^n \beta_k a_{i+k}, \quad i = 0, 1, 2, \dots,$$

and these relations can be viewed as obtained by truncating the simple identity

$$(1.8) \quad A_i = S - \sum_{k=1}^{\infty} a_{i+k}$$

and compensating for it by inserting some constant factors $\beta_k, k = 1, 2, \dots, n$. A generalization of this last idea for double series is presented in the following section.

It is shown that the general order Padé-type approximants for double power series $[A/S]_M$ of Levin [6] can be obtained by this method. Using some structural analysis of the method, it is then suggested that certain choices of the set A are preferable. In Section 3 it is shown that the same particular choices of A also follow from a direct generalization of the procedure (1.3)–(1.5) to double series. With these choices of A , an $[A/S]$ approximant is obtained which sums exactly double series $\sum_{i,j=0}^{\infty} \alpha_{ij}$ whose coefficients satisfy any linear double difference relation of the form

$$(1.9) \quad \sum_{(k,l) \in S} \alpha_{kl} \Delta_1^k \Delta_2^l a_{ij} = 0, \quad \alpha_{00} \neq 0,$$

where $\Delta_1 a_{ij} = a_{i+1,j} - a_{ij}$ and $\Delta_2 a_{i,j} = a_{i,j+1} - a_{ij}$.

The generalization of $[A/S]$ to N -dimensional (N -D) series and its application to power series is discussed in Section 4, and a numerical example with comparison to Chisholm approximants [2] is presented in Section 5. In the last section we study the generalization of Wynn's confluent ϵ -algorithm and the G -transformation of Gray,

Atchison, and McWilliams [4] to infinite double integrals. For both 2-D series and 2-D integrals we have found it necessary to assume that the 1-D case is satisfactorily solved.

2. General Order Approximations to 2-D Sums. We consider approximations to the sum of a converging double series $\sum_{i,j=0}^{\infty} a_{ij}$ by using a finite number of its terms. Let Ω denote the set of all ordered pairs of integers (i, j) , referred to below as the lattice plane. Given a subset M of Ω , we define the complement \bar{M} of M as

$$\bar{M} = \Omega - M,$$

the “nonnegative part of M ” as

$$M^+ = \{(i, j) \mid (i, j) \in M, i \geq 0, j \geq 0\},$$

and the “ (i, j) translation of M ” as

$$M_{ij} = \{(k, m) \mid (k - i, m - j) \in M\}.$$

We also use M to denote the partial sum of order M of the series,

$$M = \sum_{(i,j) \in M^+} a_{ij},$$

hence Ω also denotes the sum of the double series

$$(2.1) \quad \Omega = \sum_{i,j=0}^{\infty} a_{ij}.$$

For a chosen subset A of Ω we have, in analogy to (1.8), the relations

$$(2.2) \quad A_{-m,-n} = \Omega - \sum_{(i,j) \in \Omega - A} a_{i-m,j-n}, \quad (m, n) \in \Omega.$$

To obtain finite order relations of the form (1.7), we assume that the remainder sums in (2.2) can be approximated by a linear combination of a finite number of their terms

$$(2.3) \quad A_{-m,-n} \approx \Omega - \sum_{(i,j) \in R} \beta_{ij} a_{i-m,j-n}, \quad (m, n) \in \Omega,$$

where R is a finite set of r elements, $R \subset \Omega - A$. We now choose another finite set $S \subset \Omega^+$ having $s = r + 1$ elements to obtain a system of s linear equations

$$(2.4) \quad A_{-m,-n} = \Omega' - \sum_{(i,j) \in R} \beta_{ij} a_{i-m,j-n}, \quad (m, n) \in S,$$

for the unknowns $\{\beta_{ij}\}_{(i,j) \in R}$ and the approximation Ω' to Ω . The sets $A, R,$ and S must be chosen so that the system (2.4) has a unique solution. For instance, at most one of the sets $R_{-m,-n}^+, (m, n) \in S$ can be empty; otherwise the matrix of equations (2.4) will have two dependent rows of the form $(0, 0, \dots, 0, 1)$.

For the case of a double power series, $a_{ij} = c_{ij}x^i y^j$, systems of the type (2.4) generate the general order Padé-type approximants $[A/S]_{M=A \cup R}$ presented in [6]. We note that the notation used here is identical to that used in [6] apart from a slight change in the definition of the “translation” of a set in the lattice plane. An approximant $[A/S]_M$ is shown there to be a rational function with a numerator of rank A and a denominator of rank S such that its power series expansion agrees with the original power series on M , i.e.,

$$(2.5) \quad \sum_{i,j=0}^{\infty} c_{ij}x^i y^j - [A/S]_M = O(x^i y^j, (i, j) \in \bar{M}^+).$$

It is also shown there that the Chisholm approximants [2] can be obtained from relations of the form (2.4) where identical factors, $\beta_{ij} = \beta_{i'j'}$ are taken for some selected pairs $(i, j), (i', j') \in R$.

Let us examine the structure of the system (2.4) in terms of geometrical patterns in the lattice plane Ω . $\Omega - A_{-m,-n}$ is the remainder of the sum $A_{-m,-n}$, and the sum $\sum_{(i,j) \in R} \beta_{ij} a_{i-m,j-n}$ consists of series terms with indices in $R_{-m,-n} \subset \Omega - A_{-m,-n}$. The geometrical structures of $\Omega - A_{-m,-n}$ and of $R_{-m,-n}$ are unchanged for different $(m, n) \in S$ apart from a translation. This however is not always true for $\Omega^+ - A_{-m,-n}$ and $R_{-m,-n}^+$, the sets which are actually active in (2.4) since $a_{ij} \equiv 0$ for $(i, j) \notin \Omega^+$. These sets might have different geometrical patterns in the lattice plane for different $(m, n) \in S$. For example, Chisholm $f_{N,N}$ approximants can be obtained from (2.4) with $A = \{(i, j) \mid i, j \leq N\}$, $S = \{(i, j) \mid 0 \leq i, j \leq N\}$ and $R = M - A$ where $M = \{(i, j) \mid i + j \leq 2N + 1\}$ and with $\beta_{2N+1-j,j} = \beta_{j,2N+1-j}$, for $j = 1, 2, \dots, N$ (see [6]). The set $\Omega^+ - A_{-m,-n}$ is thus the quarter plane Ω^+ taking away an $(N - m) \times (N - n)$ rectangle from its corner, and $R_{-m,-n}^+$ is mainly composed of two isosceles right triangles with legs of length $N - m$ and $N - n$. Obviously, the structures of these sets are geometrically different for different $(m, n) \in S$. In such a case the functional form assumed in (2.3) for the remainder sums $\Omega - A_{-m,-n}$ differs for different $(m, n) \in S$, and the system (2.4) is hence termed as a system with an irregular structure.

We claim that the irregularity in the structure pattern of the equations in the system (2.4) affects the accuracy and the efficiency of the approximation. To give an example we consider the 1-D Padé approximants $[m/n]$ to $\sum_{i=0}^{\infty} c_i x^i$. These can be obtained from the system of equations

$$(2.6) \quad A_i(x) = [m/n] + \sum_{k=1}^n \beta_k c_{i+k} x^{i+k}, \quad i = m - n, m - n + 1, \dots, m.$$

Here it is assumed that the remainder $\sum_{j=i+1}^{\infty} c_j x^j$ can be approximated by a linear combination of the first n terms in it, $c_{i+1} x^{i+1}, c_{i+2} x^{i+2}, \dots, c_{i+n} x^{i+n}$. The structure of the system (2.6) is irregular whenever $m - n < 0$, and the worst case is $m = 0$ where in each equation in (2.6) the same remainder is represented by a different algebraic form. The resulting approximation for $m = 0$ is the $[0/n]$ Padé approximant which is usually not better than the $[n/0]$ approximant which is simply the partial sum $A_n(x)$. On the other hand, the $[n/n]$ approximant which is usually the most efficient one is derived from a system of equations with a regular structure. This system, i.e. (2.6) with $m = n$, is actually the higher order system with a regular structure pattern which can be obtained using only $2n + 1$ terms of the power series.

In the 2-D case it is obvious that any choice of A such that A^+ is finite yields a system of equations in which each equation is of a different geometrical structure. Yet Chisholm approximants are known to be quite efficient and they possess some important properties; see [2]. In terms of the above discussion we can argue that in the systems of equations (2.4), which generate Chisholm approximants, there are still subsets of equations in which the structure is fully or partially preserved in certain directions in the lattice plane. However, more efficient approximants are expected to be obtained by choosing the sets A, R , and S so that the structure of the sets $\Omega^+ - A_{-m,-n}$ and

of $R_{-m,-n}$ in the lattice plane is unchanged for all $(m, n) \in S$. This is equivalent to demanding that $\Omega - A_{-m,-n} \subset \Omega^+ \quad \forall (m, n) \in S$ since $R_{-m,-n} \subset \bar{A}_{-m,-n}$. For a given $S \subset \Omega^+$, the smaller set A satisfying the above property is the three-quarter lattice plane

$$(2.7) \quad A = \{(i, j) \mid i < I \text{ or } j < J\},$$

where $I = \max_{(i,j) \in S} i$ and $J = \max_{(i,j) \in S} j$.

In the following section we present another motivation for choosing A of the form (2.7) which also clarifies the significance of the classes R and S .

3. Series With Coefficients Satisfying Linear Double Difference Equations. In analogy with the 1-D series with terms satisfying (1.1), let us consider 2-D series whose terms satisfy linear relations of the form

$$(3.1) \quad a_{ij} = \sum_{(k,l) \in T} \alpha_{kl} \Delta_1^k \Delta_2^l a_{i-k,j-l}, \quad i \geq I, j \geq J,$$

where T is a finite set, $T \subset \Omega^+$, $(0, 0) \notin T$, $\Delta_1 a_{ij} = a_{i+1,j} - a_{ij}$, and $\Delta_2 a_{ij} = a_{i,j+1} - a_{ij}$. To avoid the occurrence of negative indices in (3.1) we must take $I \geq \max_{(i,j) \in T} i$ and $J \geq \max_{(i,j) \in T} j$. Yet all the following results can be carried out for any $I, J \geq 0$ provided that terms with a negative index are defined as zeros. This additional assumption is, however, “unnatural” and usually unconstructive, and the $[0/n]$ Padé approximant may serve as an example for the outcome of such an assumption.

If $\Omega = \sum_{i,j=0}^\infty a_{ij} < \infty$ then, in analogy to (1.3), it can be shown that for $M \geq I$ and $N \geq J$,

$$(3.2) \quad \begin{aligned} \Omega = & \sum_{(i,j) \in A^+} a_{ij} + \sum_{\substack{(k,l) \in T \\ k \cdot l \neq 0}} \alpha_{kl} \Delta_1^{k-1} \Delta_2^{l-1} a_{M-k,N-l} \\ & - \sum_{(k,0) \in T} \alpha_{k,0} \sum_{j=N}^\infty \Delta_1^{k-1} a_{M-k,j} - \sum_{(0,l) \in T} \alpha_{0,l} \sum_{i=M}^\infty \Delta_2^{l-1} a_{i,N-l}, \end{aligned}$$

where

$$(3.3) \quad A = \{(i, j) \mid i < M \text{ or } j < N\}.$$

Equation (3.2) is an explicit expression for the double infinite sum Ω of a series whose terms satisfy the given linear relation (3.1). In order to use this expression one has to find the coefficients α_{kl} in (3.1) and that can, in general, be done by matching a relation of the form (3.1) to a given finite number of series’ terms. However, one also has to evaluate the infinite 1-D sums appearing in the expression (3.2), the same infinite sums as those resulting from the discussion in the previous section. In the following theorem we present the definition of the suggested approximations to double infinite sums and link together both motivations to it, the one just described and the one presented in Section 2.

THEOREM 3.1. *Let S be a subset of Ω^+ having s elements and being “full-ranked”, i.e., if $(k, l) \in S$, then $\{(i, j) \mid 0 \leq i \leq k, 0 \leq j \leq l\} \subset S$. Let $T = S - \{(0, 0)\}$ and let R be another subset of Ω^+ having $s - 1$ elements. Then the two following definitions of approximations to the double sum $\Omega = \sum_{i,j=0}^\infty a_{ij}$ are equivalent:*

Definition 3.1. For $M \geq I$ and $N \geq J$,

$$(3.4) \quad \begin{aligned} \Omega_{T,A,R} = A + \sum_{\substack{(k,l) \in T \\ k \cdot l \neq 0}} \alpha_{kl} \Delta_1^{k-1} \Delta_2^{l-1} a_{M-k,N-l} \\ - \sum_{(k,0) \in T} \alpha_{k,0} \sum_{j=N}^{\infty} \Delta_1^{k-1} a_{M-k,j} - \sum_{(0,l) \in T} \alpha_{0,l} \sum_{i=M}^{\infty} \Delta_2^{l-1} a_{i,N-l}, \end{aligned}$$

where A is the partial sum of order A of the series where A is given by (3.3), and the $s - 1$ α_{kl} 's satisfy the linear $(s - 1)$ th order system of equations

$$(3.5) \quad a_{ij} = \sum_{(k,l) \in T} \alpha_{kl} \Delta_1^k \Delta_2^l a_{i-k,j-l}, \quad (i,j) \in R.$$

Definition 3.2. The approximation $[A/S]_R$ is defined as the value Ω' in the solution vector $(\Omega', \{\beta_{ij}\}_{(i,j) \in R})$ of the linear system of equations

$$(3.6) \quad \Omega' - \sum_{(i,j) \in R} \beta_{ij} a_{i-k,j-l} = A_{-k,-l}, \quad (k,l) \in S,$$

where $A_{-k,-l}$ is the partial sum of order $A_{-k,-l}$, the $(-k, -l)$ translation of the above A .

Proof. Let $[A/S]_R$ and $\{\beta_{ij}\}_{(i,j) \in R}$ satisfy (3.6), i.e.,

$$(3.7) \quad A_{-k,-l} = [A/S]_R - \sum_{(i,j) \in R} \beta_{ij} a_{i-k,j-l}, \quad (k,l) \in S.$$

Since S is "full-ranked", we can operate with $\Delta_1^k \Delta_2^l$ on (3.7) to get

$$(3.8) \quad \Delta_1^k \Delta_2^l A_{-k,-l} = - \sum_{(i,j) \in R} \beta_{ij} \Delta_1^k \Delta_2^l a_{i-k,j-l}, \quad (k,l) \in T.$$

Let $\{\alpha_{kl}\}_{(k,l) \in T}$ satisfy (3.6), then, multiplying (3.8) by α_{kl} and summing over all $(k,l) \in T$, we obtain

$$(3.9) \quad \sum_{(k,l) \in T} \alpha_{kl} \Delta_1^k \Delta_2^l A_{-k,-l} = - \sum_{(k,l) \in T} \alpha_{kl} \sum_{(i,j) \in R} \beta_{ij} \Delta_1^k \Delta_2^l a_{i-k,j-l}.$$

Interchanging the order of summation, using (3.5) in the right-hand side of (3.9), and using the equalities

$$\Delta_1^k \Delta_2^l A_{-k,-l} = \begin{cases} -\Delta_1^{k-1} \Delta_2^{l-1} a_{M-k,N-l}, & k \cdot l \neq 0, \\ \sum_{j=N}^{\infty} \Delta_1^{k-1} a_{M-k,j}, & l = 0, \\ \sum_{i=M}^{\infty} \Delta_2^{l-1} a_{i,N-l}, & k = 0, \end{cases}$$

we obtain

$$(3.10) \quad \begin{aligned} \sum_{\substack{(k,l) \in T \\ k \cdot l \neq 0}} \alpha_{kl} \Delta_1^{k-1} \Delta_2^{l-1} a_{M-k,N-l} - \sum_{(k,0) \in T} \alpha_{k,0} \sum_{j=N}^{\infty} \Delta_1^{k-1} a_{M-k,j} \\ - \sum_{(0,l) \in T} \alpha_{0,l} \sum_{i=M}^{\infty} \Delta_2^{l-1} a_{i,N-l} = \sum_{(i,j) \in R} \beta_{ij} a_{ij}. \end{aligned}$$

Combining the (0, 0)th equation of (3.7),

$$[A/S]_R = A + \sum_{(i,j) \in R} \beta_{ij} a_{ij},$$

with (3.10), we thus obtain

$$[A/S]_R = \Omega_{T,A,R}. \quad \text{Q.E.D.}$$

From the motivation of $[A/S]_R$ via recursion relation, it is clear that the choice of the set R has no conceptual significance. Also the restriction $R \subset \Omega^+ - A$, stated in Section 2, can be reduced; R should only be chosen so that the matrix of the system (3.5) is full and nonsingular. It is of course preferred to use an R with which the number of series terms taking part in (3.5) is the minimal.

The two motivations, the structural considerations in the lattice plane, and the recursion relation assumption led us to the inevitable conclusion that, in order to obtain efficient approximations to the sum of 2-D series, the 1-D series which constitute the above partial sum A should be computed, or well approximated. In the generalization of these approximations to N -D series, as presented below, we assume that the $(N - 1)$ -D case is satisfactorily solved, i.e., that $(N - 1)$ -D series can be well approximated.

4. Generalization to N -D Series and Application to Power Series. The generalization of the above approximations to N -dimensional series is direct; let Ω be the set of all the ordered vectors of integers (i_1, i_2, \dots, i_N) and let also

$$\Omega = \sum_{i_1=0}^{\infty} \sum_{i_2=0}^{\infty} \cdots \sum_{i_N=0}^{\infty} a_{i_1, i_2, \dots, i_N}.$$

The previous lattice plane terminology will also be used for subsets in the N -D lattice space, its adaptation being evident. To simplify the notation, we introduce the vector notation $\mathbf{i} = (i_1, i_2, \dots, i_N)$, $a_{\mathbf{i}} = a_{i_1, i_2, \dots, i_N}$, and also $a_{\mathbf{i}-\mathbf{k}} = a_{i_1-k_1, i_2-k_2, \dots, i_N-k_N}$ and $\mathbf{x}^{\mathbf{i}} = x_1^{i_1} x_2^{i_2} \cdots x_N^{i_N}$.

Definition 4.1. Let S be a "full-ranked" subset of Ω^+ having s elements and $R \subset \Omega^+$ having $s - 1$ elements. The approximation $[A/S]_R$ to Ω is defined by the s order system of linear equations

$$(4.1) \quad A_{-\mathbf{k}} = [A/S]_R - \sum_{\mathbf{i} \in R} \beta_{\mathbf{i}} a_{\mathbf{i}-\mathbf{k}}, \quad \mathbf{k} \in S,$$

where

$$(4.2) \quad A = \{ \mathbf{m} \mid m_1 < M_1 \text{ or } m_2 < M_2 \dots \text{ or } m_N < M_N \}.$$

THEOREM 4.1. *The approximation $[A/S]_R$ is exact for all the convergent series whose terms satisfy linear recursion relations of the form*

$$(4.3) \quad a_{\mathbf{i}} = \sum_{\mathbf{k} \in S - \{(0,0)\}} \alpha_{\mathbf{k}} \Delta^{\mathbf{k}} a_{\mathbf{i}-\mathbf{k}} \quad \text{for } i_j \geq I_j, j = 1, 2, \dots, N,$$

where Δ is the vector of the difference operators $\Delta = (\Delta_1, \Delta_2, \dots, \Delta_N)$, provided that in (4.2) $M_j \geq I_j$ for $j = 1, 2, \dots, N$, and that the system (4.1) has a unique solution.

Proof. The result follows by a direct adaptation of the equivalence theorem 3.1 to the N -D case.

We now consider the application of the $[A/S]_R$ approximation to N -D power series

$$(4.4) \quad \Omega(\mathbf{x}) = \sum_{\mathbf{i} \in \Omega^+} c_{\mathbf{i}} \mathbf{x}^{\mathbf{i}}.$$

THEOREM 4.2. *The approximation $[A/S]_R$ to N -D power series can be written as the ratio of a “polynomial” of rank A and a polynomial of rank S (the quotation marks are used since A is not a finite set). Also*

$$(4.5) \quad \Omega(\mathbf{x}) - [A/S]_R = O(\mathbf{x}^{\mathbf{k}}, \mathbf{k} \in \overline{A \cup R}).$$

Proof. The $[A/S]_R$ approximation to power series is defined by (4.1) with $a_{\mathbf{i}} = c_{\mathbf{i}} \mathbf{x}^{\mathbf{i}}$. Multiplying (4.1) by $\mathbf{x}^{\mathbf{i}}$ and denoting $\beta_{\mathbf{i}} \mathbf{x}^{\mathbf{i}}$ by $\beta'_{\mathbf{i}}$, we obtain a more convenient defining system of equations for $[A/S]_R$,

$$(4.6) \quad \mathbf{x}^{\mathbf{k}} A_{-\mathbf{k}} = \mathbf{x}^{\mathbf{k}} [A/S]_R + \sum_{\mathbf{i} \in R} \beta'_{\mathbf{i}} c_{\mathbf{i}-\mathbf{k}}, \quad \mathbf{k} \in S.$$

The $s \times s$ matrix of coefficients of the system (4.6) has the $\mathbf{x}^{\mathbf{k}}, \mathbf{k} \in S$, in one column and the constant $c_{\mathbf{i}-\mathbf{k}}$ in all the other columns. Also the left-hand side of (4.6) is a “polynomial” of rank A . Therefore, in solving (4.6) for $[A/S]_R$ by Cramer’s rule, we obtain a “polynomial” of rank A over a polynomial of rank S . The determinant representation of $[A/S]_R$ which follows from solving (4.6) by Cramer’s rule is also needed for proving the second part of the theorem. Using this representation, the proof of the theorem in [6] can be easily adapted to prove (4.5).

THEOREM 4.3. *$[A/S]_R$ is exact for any function of the form*

$$(4.7) \quad f(x_1, x_2, \dots, x_N) = \frac{\sum_{i=1}^N \sum_{j=0}^{M_i-1} x_i^j g_{ij}(x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_N)}{\sum_{\mathbf{k} \in S} d_{k_1, k_2, \dots, k_N} x_1^{k_1} x_2^{k_2} \dots x_N^{k_N}}$$

in the domain of convergence D of its power series expansion at $\mathbf{x} = 0$, $f(\mathbf{x}) = \sum_{\mathbf{i} \in \Omega^+} c_{\mathbf{i}} \mathbf{x}^{\mathbf{i}}$, provided that the subset R is chosen such that the system (4.6) has a unique solution. In (4.7) we assume that the numerator and the denominator have no common factor and we denote them by $P(\mathbf{x})$ and $Q(\mathbf{x})$ respectively. The functions $\{g_{ij}\}_{i=1}^N \sum_{j=0}^{M_i-1}$ can be any functions of $N - 1$ variables analytic at the origin.

Proof. Multiplying (4.7) by $Q(\mathbf{x})$ and comparing the coefficients of the monomials $\mathbf{x}^{\mathbf{i}}$ in both sides, we get

$$(4.8) \quad \sum_{\mathbf{k} \in S} d_{\mathbf{k}} c_{\mathbf{i}-\mathbf{k}} = 0, \quad \mathbf{i} \in \bar{A}.$$

Multiplying (4.8) by $\mathbf{x}^{\mathbf{i}}$, we obtain for $\mathbf{i} \in \bar{A}$

$$(4.9) \quad \sum_{\mathbf{k} \in S} d'_{\mathbf{k}} a_{\mathbf{i}-\mathbf{k}} = 0,$$

where $d'_{\mathbf{k}} = d_{\mathbf{k}} \mathbf{x}^{\mathbf{k}}$ and $a_{\mathbf{i}} = c_{\mathbf{i}} \mathbf{x}^{\mathbf{i}}$. Equation (4.9) can be rewritten as

$$(4.10) \quad \sum_{k \in S} \alpha'_k \Delta^k a_{i-k} = 0,$$

where

$$\alpha'_0 = \sum_{k \in S} d_k x^k = Q(x).$$

For $x \in D$ it is clear that $Q(x) \neq 0$, and the terms a_i satisfy (4.3) with $\alpha_k = -\alpha'_k/Q(x)$ and $I_j = M_j$. All the conditions stated in Theorem 4.1 are thus satisfied for $x \in D$, therefore $[A/S]_R = f(x)$. Q.E.D.

So far we have assumed that the infinite $(N - 1)$ -D partial sums A_{-k} , $k \in S$, are known. In order to make the $[A/S]_R$ approximants practical, we must replace those $(N - 1)$ -D series by some appropriate approximations, e.g., by some $(N - 1)$ -D $[A/S]_R$ approximants. Of course, the above exactness theorem will no longer hold. However, it can still be shown that for a function of the form (4.7) the resulting approximation can be written as $\bar{P}(x)/Q(x)$, where the approximation $\bar{P}(x) \approx P(x)$ is controlled by those $(N - 1)$ -D approximations used to approximate the partial sums A_{-k} .

For the 2-D case we can use ordinary Padé approximants, $[m/n]$, to replace the infinite 1-D series appearing in $[A/S]$. The resulting approximation is denoted as $[[m/n]A/S]$. However, the assumption of a linear 2-D relation (3.2) (with constant coefficients) between the terms $\{a_{ij}\}$ of a 2-D series does not necessarily imply that a similar 1-D relation (1.1) holds between row or column elements of $\{a_{ij}\}$. Therefore, the ordinary Padé approximants might be unsatisfactory for our purpose, and some other summation methods, e.g., the recent d -transformation of Levin and Sidi [7] might be more suitable.

5. Numerical Example. We now apply the $[A/S]_R$ approximants to the Beta function

$$(5.1) \quad B(p, q) = B(1 + x, 1 + y) = \frac{\Gamma(1 + x)\Gamma(1 + y)}{\Gamma(2 + x + y)}.$$

This case has been considered by Graves-Morris, Hughes-Jones, and Makinson [3] using Canterbury approximants. Following their work, we consider approximations to

$$(5.2) \quad f(x, y) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} c_{ij} x^i y^j,$$

defined by

$$(5.3) \quad B(1 + x, 1 + y) = \frac{1 + xyf(x, y)}{(1 + x)(1 + y)}.$$

The coefficients $\{c_{ij}\}$, required for computing the $[A/S]_R$ approximants, were calculated using the first procedure suggested by Graves-Morris, Hughes-Jones, and Makinson. We considered symmetric subsets A , S , and R ,

$$A = \{(i, j) \mid i \leq m \text{ or } j \leq m\} \equiv A_m,$$

$$S = \{(i, j) \mid 0 \leq i, j \leq k\} \equiv S_k,$$

$$R = \{(i, j) \mid r \leq i, j \leq r + k, i + j \neq 2r + 2k\} \equiv R_{r,k}.$$

TABLE 5.1

Comparison between $f_{M/N}$ and $[A/S]_R$ approximants of various degrees to $B(p, q)$ at various points (p, q) .

p, q	1.75, -.25	-.5, -.5	-.75, -.75
$B(p, q)$	-6.77770	0.	9.88840
$f_{3/3}^{(31, 16)}$	-6.787	-.14	7.0
$f_{7/3}^{(79, 16)}$	-6.77774	-.0010	9.82
$[3; 3/1]_1^{(39, 4)}$	-6.794	-.026	9.94
$[5; 4/2]_3^{(85, 9)}$	-6.77775	-.00021	9.8877

p, q	-1.25, -1.25	-1.5, -1.5	-1.75, -1.75
$B(p, q)$	-16.266	0.	28.253
$f_{3/3}^{(31, 16)}$	-2.6	-3.7	-2.8
$f_{7/3}^{(79, 16)}$	-15.7	-3.8	4.0
$[3; 3/1]_1^{(39, 4)}$	-15.3	-.19	27.18
$[5; 4/2]_3^{(85, 9)}$	-16.271	-.023	28.259

p, q	1.75, -.25	1.75, -.5	1.75, -.75
$B(p, q)$	-5.08327851	-3.59442070	-4.44288294
$f_{7/2}^{(72, 9)}$	-5.083274	-3.59438	-4.4421
$[5; 2/2]_2^{(61, 9)}$	-5.083272	-3.59440	-4.44427

p, q	1.75, -1.25	1.75, -1.5	1.75, -1.75
$B(p, q)$	2.03331140	0.599070	0.
$f_{7/2}^{(72, 9)}$	2.05	0.62	0.10
$[5; 2/2]_2^{(61, 9)}$	2.0329	0.59901	0.010

For the infinite 1-D power series appearing in the definition of $[A/S]_R$, we have used diagonal ordinary Padé approximants $[n/n]$. The resulting approximation is thus denoted as $[[n/n]A_m/S_k]_{R_r}$, or shortly as $[n; m/k]_r$. $[[n/n]A_m/S_k]_{R_r}$ is an approximation of the form

$$(5.4) \quad \frac{\sum_{j=0}^m x^j \frac{\sum_{i=0}^n \alpha_{ij} y^i}{\sum_{i=0}^n \beta_{ij} y^i} + \sum_{j=0}^m y^j \frac{\sum_{i=0}^n p_{ij} x^i}{\sum_{i=0}^n q_{ij} x^i}}{\sum_{i=0}^k \sum_{j=0}^k \alpha_{ij} x^i y^j}.$$

Its computation involves the terms $\{c_{ij}\}$ with

$(i, j) \in \{(i, j) \mid 0 \leq i \leq 2n, 0 \leq j \leq m\} \cup \{(i, j) \mid 0 \leq i \leq m, 0 \leq j \leq 2n\} \cup R$,
 i.e., $2(2n + 1)(m + 1) - (m + 1)^2 + (k + r - m)^2 - 1$ terms, and the solution of a linear system of order $(k + 1)^2$.

We compare our results with those of Graves-Morris, Hughes-Jones, and Makinson [3] who have used symmetric-off-diagonal approximation (S.O.D.) of the form

$$(5.5) \quad f_{M/N}(x, y) = \frac{\sum_{i=0}^M \sum_{j=0}^M a_{ij} x^i y^j}{\sum_{i=0}^N \sum_{j=0}^N b_{ij} x^i y^j}.$$

The computation of $f_{M/N}$ involves the terms

$$\{c_{ij} \mid 0 \leq i + j \leq M + N + 1 \vee 0 \leq i, j \leq \max(M, N)\},$$

i.e., $(M + 1)^2 + (N + 1)^2 - 1$ terms, and the solution of a linear system of order $(N + 1)^2$. The special block structure of the system of equations can be used to reduce the computational effort; Hughes-Jones and Makinson [5].

Recently, a recursive procedure for computing 2-D Padé approximants was presented by Bose and Basu [1], and this procedure can be applied for our case as well. However, we have used direct computation by *LU* decomposition, which is accurate and fast enough for the low order approximations that we tested and of course is much simpler to program.

In Table 5.1, we compare the numerical values of the $[[n/n]A_m/S_k]_{R_r}$ approximants with those obtained by the $f_{M/N}$ approximants and with the exact values of the Beta function at various points. We add double upper subscripts to the notation of the different approximants, in which the first index stands for the number of series terms used and the second for the order of the system solved to obtain the approximant, e.g., $[4; 4/2]_2^{(54,9)}$ is the $[[4/4]A_4/S_2]_{R_2}$ approximant; it uses 54 series terms and is obtained by solving a 9th order system.

It is clear from the table above that the $[A/S]_R$ approximant can in some cases produce better approximations with less computational effort.

6. Infinite Double Integrals. In this section we consider the problem of accelerating the convergence of infinite double integrals. To the best of our knowledge, no direct method has been proposed for this problem yet. Of course it can be approached by considering a sequence of finite double integrals or by a consequential use of

methods for 1-D infinite integrals. In what follows we present a direct method which, unlike the indirect methods, makes use of 2-D properties of the integrand.

As in the case of 2-D series, we assume here that the 1-D case is satisfactorily solved. This assumption holds for a wide and important class of integrals for which the confluent ϵ -algorithm of Wynn [10], the G -transformation of Gray, Atchison, and McWilliams [4], or the D -transformation of Levin and Sidi [7] can be applied. The confluent ϵ -algorithm and the G -transformation serve us also as a model for building the 2-D approximations. Both these methods give the exact value of $\int_0^\infty f(t) dt$ for functions satisfying linear differential equations with constant coefficients,

$$(6.1) \quad f(t) = \sum_{k=1}^m p_k f^{(k)}(t), \quad t \in [0, \infty).$$

The confluent ϵ -algorithm of Wynn can be written as

$$(6.2) \quad \epsilon_m(t) = \int_0^t f(\tau) d\tau - \sum_{k=1}^m p_k f^{(k-1)}(t),$$

where the coefficients p_k satisfy

$$(6.3) \quad f^{(j)}(t) = \sum_{k=1}^m p_k f^{(k+j)}(t), \quad j = 0, 1, \dots, m-1.$$

The derivation of approximations to infinite double integrals is based upon the assumption that the integrand satisfies a linear partial differential equation with constant coefficients of the form

$$(6.4) \quad f(x, y) = \sum_{(k,l) \in T} \alpha_{kl} \partial_x^k \partial_y^l f(x, y), \quad T \subset \Omega^+, (0, 0) \notin T.$$

Here Ω stands for the lattice plane.

THEOREM 6.1. *Let f satisfy (6.4) for $x \geq a$ and $y \geq b$, and let*

$$\lim_{y \rightarrow \infty} \partial_x^k \partial_y^{l-1} f(x, y) = 0, \quad (k, l) \in T, l \neq 0,$$

and

$$\lim_{x \rightarrow \infty} \partial_x^{k-1} \partial_y^l f(x, y) = 0, \quad (k, l) \in T, k \neq 0.$$

Then

$$(6.5) \quad \begin{aligned} \int_b^\infty \int_a^\infty f(x, y) dx dy &= \sum_{\substack{(k,l) \in T \\ k \cdot l \neq 0}} \alpha_{kl} \partial_x^{k-1} \partial_y^{l-1} f(a, b) \\ &- \sum_{(k,0) \in T} \alpha_{k,0} \int_b^\infty \partial_x^{k-1} f(a, y) dy \\ &- \sum_{(0,l) \in T} \alpha_{0,l} \int_a^\infty \partial_y^{l-1} f(x, b) dx. \end{aligned}$$

Proof. The result follows by using representation (6.4) for the integrand and performing integration by parts.

Defining $A = \{(x, y) \mid x \leq a \text{ or } y \leq b, x, y \geq 0\}$, we can write the double integral in (6.5) as

$$(6.6) \quad \int_b^\infty \int_a^\infty f(x, y) dx dy = \int_0^\infty \int_0^\infty f(x, y) dx dy - \int \int_A f(x, y) dx dy.$$

The 2-D analogue of the confluent ϵ -algorithm is thus defined as follows:

Definition 6.1. Let T and R be two r -element subsets of Ω^+ , $(0, 0) \notin T$, then for $a, b \geq 0$

$$\begin{aligned} \epsilon_T^{(2)}(a, b) = & \iint_A f(x, y) dx dy + \sum_{\substack{(k,l) \in T \\ k \cdot l \neq 0}} \alpha_{kl} \partial_x^{k-1} \partial_y^{l-1} f(a, b) \\ (6.7) \quad & - \sum_{(k,0) \in T} \alpha_{k,0} \int_b^\infty \partial_x^{k-1} f(a, y) dy - \sum_{(0,l) \in T} \alpha_{0,l} \int_a^\infty \partial_y^{l-1} f(x, b) dx, \end{aligned}$$

where the coefficients α_{kl} are determined by the system of linear equations

$$(6.8) \quad \partial_x^i \partial_y^j f(a, b) = \sum_{(k,l) \in T} \alpha_{kl} \partial_x^{k+i} \partial_y^{l+j} f(a, b), \quad (i, j) \in R,$$

(assuming that all the necessary partial derivatives of f exist) provided that this system has a unique solution.

COROLLARY. Necessary and sufficient conditions that

$$\epsilon_T^{(2)}(a, b) = \int_0^\infty \int_0^\infty f(x, y) dx dy \quad \forall a, b \geq 0,$$

are that f satisfies the conditions of Theorem 6.1 and that the system (6.8) has a unique solution.

Proof. Sufficiency is proved by Theorem 6.1, and necessity follows by differentiating (6.5) with respect to a and b .

For $a = b = 0$ Eq. (6.5) provides an expression for $\int_0^\infty \int_0^\infty f(x, y) dx dy$ in terms of the derivatives of f at $(0, 0)$, integrals of the normal derivatives of f along the boundaries, and of course the α_{kl} of (6.4). As in the 1-D case, $\epsilon_T^{(2)}$ has the disadvantage of the need to provide high order derivatives of the integrand. However, in the 1-D case this difficulty can be avoided by using the G -transformation of Gray, Atchison, and McWilliams [4] which can be written as

$$(6.9) \quad G_m(a; \Delta t) = \int_0^a f(t) dt - \sum_{k=1}^m q_k \Delta^{k-1} f(a),$$

where the q_k 's satisfy

$$(6.10) \quad \int_{a+j\Delta t}^{a+(j+1)\Delta t} f(t) dt = \sum_{k=1}^m q_k \Delta^k f(a + j\Delta t), \quad j = 0, 1, \dots, m-1.$$

The relations (6.10) can be viewed as originating from the assumption that f satisfies

$$(6.11) \quad f(t) = \sum_{k=1}^m q_k \Delta^{k-1} f'(t), \quad t \geq 0.$$

The class of functions satisfying (6.11) is the same as the class of functions satisfying (6.1). Hence, although the G -transformation uses only function values and no derivatives, it still has the same class of exactness as that of the confluent ϵ -algorithm.

The 2-D analogue of the G -transformation is defined as follows:

Definition 6.2. Let T and R be two r -element subsets of Ω^+ , $(0, 0) \notin T$, and let $A = \{(x, y) | x \leq a \text{ or } y \leq b, x, y \geq 0\}$, then for $a, b \geq 0$

$$\begin{aligned}
 G_T^{(2)}(a, b; \Delta x, \Delta y) &= \iint_A f(x, y) dx dy + \sum_{\substack{(k,l) \in T \\ k \cdot l \neq 0}} \beta_{kl} \Delta_x^{k-1} \Delta_y^{l-1} f(a, b) \\
 (6.12) \quad &- \sum_{(k,0) \in T} \beta_{k,0} \int_b^\infty \Delta_x^{k-1} f(a, y) dy \\
 &- \sum_{(0,l) \in T} \beta_{0,l} \int_a^\infty \Delta_y^{l-1} f(x, b) dx,
 \end{aligned}$$

where the β_{kl} are determined by the linear system of r equations

$$\begin{aligned}
 \int_{b+j\Delta y}^{b+(j+1)\Delta y} \int_{a+i\Delta x}^{a+(i+1)\Delta x} f(x, y) dx dy \\
 (6.13) \quad &= \sum_{\substack{(k,l) \in T \\ k \cdot l \neq 0}} \beta_{k,l} \Delta_x^k \Delta_y^l f(a + i\Delta x, b + j\Delta y) \\
 &- \sum_{(k,0) \in T} \beta_{k,0} \int_{b+j\Delta y}^{b+(j+1)\Delta y} \Delta_x^k f(a + i\Delta x, y) dy \\
 &- \sum_{(0,l) \in T} \beta_{0,l} \int_{a+i\Delta x}^{a+(i+1)\Delta x} \Delta_y^l f(x, b + j\Delta y) dx, \quad (i, j) \in R.
 \end{aligned}$$

provided that this system has a unique solution. Here $\Delta_x f(x, y) = f(x + \Delta x, y) - f(x, y)$ and $\Delta_y f(x, y) = f(x, y + \Delta y) - f(x, y)$.

THEOREM 6.2. *Necessary and sufficient conditions that*

$$G_T^{(2)}(a, b; \Delta x, \Delta y) = \int_0^\infty \int_0^\infty f(x, y) dx dy \quad \forall a, b \geq 0,$$

are that f satisfies a relation of the form

$$\begin{aligned}
 f(x, y) &= \sum_{\substack{(k,l) \in T \\ k \cdot l \neq 0}} \beta_{kl} \Delta_x^{k-1} \Delta_y^{l-1} f_{xy}(x, y) + \sum_{(k,0) \in T} \beta_{k,0} \Delta_x^{k-1} f_x(x, y) \\
 (6.14) \quad &+ \sum_{(0,l) \in T} \beta_{0,l} \Delta_y^{l-1} f_y(x, y)
 \end{aligned}$$

$\forall x, y \geq 0$, that $\lim_{x \rightarrow \infty} f(x, y) = 0$ or $\lim_{y \rightarrow \infty} f(x, y) = 0$, and that R is chosen so that the system (6.13) has a unique solution.

Proof. Integrating (6.14) from a to infinity with respect to x and from b to infinity with respect to y and using integration by parts, we obtain

$$\begin{aligned}
 \int_b^\infty \int_a^\infty f(x, y) dx dy &= \sum_{\substack{(k,l) \in T \\ k \cdot l \neq 0}} \beta_{kl} \Delta_x^{k-1} \Delta_y^{l-1} f(a, b) \\
 (6.15) \quad &- \sum_{(k,0) \in T} \beta_{k,0} \int_b^\infty \Delta_x^{k-1} f(a, y) dy \\
 &- \sum_{(0,l) \in T} \beta_{0,l} \int_a^\infty \Delta_y^{l-1} f(x, b) dx.
 \end{aligned}$$

The equations (6.13) can be obtained for any $i, j \geq 0$ by integration of (6.14). Comparing (6.12) and (6.15) and using (6.6) we obtain

$$G_T^{(2)}(a, b; \Delta x, \Delta y) = \int_0^\infty \int_0^\infty f(x, y) dx dy$$

and thus sufficiency is proved. The necessity follows by differentiating (6.12) with respect to a and b .

Unlike the 1-D case, $\epsilon_T^{(2)}$ and $G_T^{(2)}$ do not, in general, share the same exactness class. That is due to the fact that a function satisfying a relation of the form (6.4) does not necessarily satisfy also a relation of the form (6.14) and vice versa. It is of course more natural to assume a relation of the form (6.4). However, the relations (6.14) are not less general than (6.4), and for many cases they can provide good approximations to the relations (6.4). Therefore, for cases in which the computation of the integrand's derivatives is expensive, $G_T^{(2)}$ is recommended.

Department of Mathematical Sciences
Tel-Aviv University
Ramat-Aviv, Israel

1. N. K. BOSE & S. BASU, "2-D matrix Padé approximants: existence, non-uniqueness and recursive computation," *IEEE Trans. Automat. Control*. (To appear.)
2. J. S. R. CHISHOLM, "Rational approximants defined from double power series," *Math. Comp.*, v. 27, 1973, pp. 841–848.
3. P. R. GRAVES-MORRIS, R. HUGHES-JONES & G. J. MAKINSON, "The calculation of some rational approximants in two variables," *J. Inst. Math. Appl.*, v. 13, 1974, pp. 311–320.
4. H. L. GRAY, T. A. ATCHISON & G. V. McWILLIAMS, "Higher order G -transformations," *SIAM J. Numer. Anal.*, v. 8, 1971, pp. 365–381.
5. R. HUGHES-JONES & G. J. MAKINSON, "The generation of Chisholm rational polynomial approximants to power series in two variables," *J. Inst. Math. Appl.*, v. 13, 1974, pp. 299–310.
6. D. LEVIN, "General order Padé-type rational approximants defined from double power series," *J. Inst. Math. Appl.*, v. 18, 1976, pp. 1–8.
7. D. LEVIN & A. SIDI, "Two new classes of non-linear transformations for accelerating the convergence of infinite integrals and series," *Appl. Math. and Comp.* (To appear.)
8. D. SHANKS, "Non-linear transformations of divergent and slowly convergent sequences," *J. Math. Phys.*, v. 34, 1955, pp. 1–42.
9. P. WYNN, "On a device for computing the $e_m(S_n)$ transformation," *MTAC*, v. 10, 1956, pp. 91–96.
10. P. WYNN, "Upon a second confluent form of the ϵ -algorithm," *Proc. Glasgow Math. Soc.*, v. 5, 1962, pp. 160–165.