NON-LINEAR TRANSFORMATIONS OF DIVERGENT AND SLOWLY CONVERGENT SEQUENCES*

By Daniel Shanks

Abstract

This paper discusses a family of non-linear sequence-to-sequence transformations designated as e_k, e_k^m, $\tilde{e}_k$, and e_d. A brief history of the transforms is related and a simple motivation for the transforms is given. Examples are given of the application of these transformations to divergent and slowly convergent sequences. In particular the examples include numerical series, the power series of rational and meromorphic functions, and a wide variety of sequences drawn from continued fractions, integral equations, geometry, fluid mechanics, and number theory. Theorems are proven which show the effectiveness of the transformations both in accelerating the convergence of (some) slowly convergent sequences and in inducing convergence in (some) divergent sequences. The essential unity of these two motives is stressed. Theorems are proven which show that these transforms often duplicate the results of well-known, but specialized techniques. These special algorithms include Newton's iterative process, Gauss's numerical integration, an identity of Euler, the Padé Table, and Thiele's reciprocal differences. Difficulties which sometimes arise in the use of these transforms such as irregularity, non-uniform convergence to the wrong answer, and the ambiguity of multivalued functions are investigated. The concepts of antilimit and of the spectra of sequences are introduced and discussed. The contrast between discrete and continuous spectra and the consequent contrasting response of the corresponding sequences to the e_1 transformation is indicated. The characteristic behaviour of a semiconvergent (asymptotic) sequence is elucidated by an analysis of its spectrum into convergent components of large amplitude and divergent components of small amplitude.

Table of Contents

* Thesis submitted to the Faculty of the Graduate School of the University of Maryland in partial fulfillment of the requirements for the degree of Doctor of Philosophy, 1954. This paper was written under Task FR-30-(54) at the Naval Ordnance Laboratory.

Introduction. Some non-linear transforms. In this paper we will be concerned with a family of non-linear sequence-to-sequence transformations, with their origin and relation to known algorithms, with their theory and with their application to divergent and slowly convergent sequences. The transforms are defined as follows. Let $\{A_n\}(n = 0, 1, 2, \cdots)$ be a sequence of numbers or functions and let:

$$\Delta A_n = A_{n+1} - A_n . \tag{1}$$

Let k be a positive integer and let a new sequence $\{B_{k,n}\}$ $(n = k, k + 1, k + 2, \cdots)$, "the k'th order transform of $\{A_n\}$", be defined by

$$B_{k,n} = \frac{\begin{vmatrix} A_{n-k} & \cdots & A_{n-1} & A_n \\ \Delta A_{n-k} & \cdots & \Delta A_{n-1} & \Delta A_n \\ \Delta A_{n-k+1} & \cdots & \Delta A_n & \Delta A_{n+1} \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ \Delta A_{n-1} & \cdots & & \Delta A_{n+k-1} \end{vmatrix}}{\begin{vmatrix} 1 & \cdots & 1 & 1 \\ \Delta A_{n-k} & \cdots & \Delta A_{n-1} & \Delta A_n \\ \Delta A_{n-k+1} & \cdots & \Delta A_n & \Delta A_{n+1} \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ \Delta A_{n-1} & \cdots & & \Delta A_{n+k-1} \end{vmatrix}} \tag{2}$$

for each n for which the denominator does not vanish. It will sometimes be convenient to consider $\{A_n\}$ as its own 0'th order transform. We will then write:

$$B_{0,n} = A_n . \qquad (n = 0, 1, 2, \cdots) \quad \text{(2a)}$$

For each n for which the denominator vanishes and the numerator does not we will assign:

$$B_{k,n} = \infty . \quad \text{(2b)}$$

For each n for which both numerator and denominator vanish we will assign:

$$B_{k,n} = B_{k-1,n} \quad \text{(2c)}$$

and if necessary we will repeat this lowering in order until we obtain a well-defined quantity by (2), (2a), or (2b).

We hasten to add that the 1st order transform ($k = 1$) looks very much simpler:

$$B_{1,n} = \frac{\begin{vmatrix} A_{n-1} & A_n \\ \Delta A_{n-1} & \Delta A_n \end{vmatrix}}{\begin{vmatrix} 1 & 1 \\ \Delta A_{n-1} & \Delta A_n \end{vmatrix}} = \frac{A_{n+1}A_{n-1} - A_n^2}{A_{n+1} + A_{n-1} - 2A_n} . \quad (3)$$

The transforms may be written in operator form, thus:

$$B_{k,n} = e_k(A_n) \quad (4)$$

where e_k is the non-linear operator defined by the right side of (2). Again, expanding the determinants in (2) by their first rows we could write

$$e_k(A_n) = \frac{c_{n-k}A_{n-k} + \cdots + c_{n-1}A_{n-1} + c_nA_n}{c_{n-k} + \cdots + c_{n-1} + c_n} \quad (5)$$

where the c's are the appropriate cofactors. Then $B_{k,n}$ is seen to be a weighted combination of the A's, but *not* a linear one since the weights, c_n , are themselves functions of the A's. Both the power and the difficulties of these transforms stem from this non-linearity.

If e_k were linear we would have

$$e_k(CA_n) = Ce_k(A_n) \quad \text{and} \quad e_k(A_n + A'_n) = e_k(A_n) + e_k(A'_n)$$

where C is a constant with respect to n and $\{A'_n\}$ is a second sequence. Now the first rule is seen from (5) to be correct:

$$e_k(CA_n) = Ce_k(A_n) \quad \text{(5a)}$$

but the second is not generally true and we must content ourselves with the weaker, but still useful, rule:

$$e_k(A_n + C) = e_k(A_n) + e_k(C) = e_k(A_n) + C. \quad \text{(5b)}$$

Another useful rule valid for any $r = 0, 1, \cdots k$, is:

$$e_k(A_n) = \frac{c_{n-k}A_{n-k+r} + \cdots + c_{n-1}A_{n-1+r} + c_nA_{n+r}}{c_{n-k} + \cdots + c_{n-1} + c_n} . \qquad (0 \leqq r \leqq k) \quad \text{(5c)}$$

That this is true is seen by replacing the first row of the numerator of (2) by the sum of the 1st + 2nd + 3rd + $\cdots$ + $(r + 1)$'st rows of the numerator and again expanding by the (new) first row. This generalizes (5).

We are also concerned, at least for $k = 1$, with the iteration of these transforms. That is:

$$\begin{aligned} B_{k,n} &= e_k(A_n) && (n \geqq k) \\ C_{k,n} &= e_k(B_{k,n}) = e_k^2(A_n) && (n \geqq 2k) \\ D_{k,n} &= e_k(C_{k,n}) = e_k^3(A_n) && (n \geqq 3k) \\ &\text{etc.} \end{aligned} \tag{6}$$

For $k = 1$, the first index will usually be suppressed and these successive sequences will be tabulated as follows:

$$\begin{array}{llllll} A_0 \\ A_1 & B_1 \\ A_2 & B_2 & C_2 \\ A_3 & B_3 & C_3 & D_3 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{array} \tag{7}$$

The operator $\tilde{e}_k$, the "kth *order iterated transformation*", is defined by:

$$\tilde{e}_k(A_0\,; A_1\,; A_2\,; A_3\,; \cdots) = A_0\,; \quad B_{k,k}\,; \quad C_{k,2k}\,; \quad D_{k,3k}\,; \cdots \tag{8}$$

If $k = 1$, it is seen that $\tilde{e}_1\{A_n\}$ is the diagonal, $\{A_0, B_1, C_2, \cdots\}$, of (7).

We are also concerned with a second two dimensional array which should not be confused with (7), namely:

$$\begin{array}{lllll} B_{0,0} & B_{0,1} & B_{0,2} & B_{0,3} & \cdots \\ & B_{1,1} & B_{1,2} & B_{1,3} & \cdots \\ & & B_{2,2} & B_{2,3} & \cdots \\ & & & B_{3,3} & \cdots \\ & & & & \cdots \end{array} \tag{9}$$

Although the first two rows of (9) are merely the first two columns of (7) in a different notation, the remaining portions of the two arrays are in general independent. The operator e_d, the "*diagonal transformation*", is defined by:

$$e_d(A_n) = B_{n,n}\,. \tag{10}$$

It is seen that $e_d\{A_n\}$ is the diagonal of (9).

The family of sequence transforms thus defined includes e_k, e_k^m, $\tilde{e}_k$, and e_d. Clearly other derivates are possible, e_d iterated, combinations of $e_{k'}$, and $e_{k''}$, of e_k and e_d, etc. but these variants will not be discussed here. We will deal mostly with e_1^m, e_k, and the two arrays (7) and (9). These latter include the important $\tilde{e}_1$ and e_d.

A brief history. The e_1 transform (but often only in fragmentary form) has been found independently by many authors. It has been applied mostly to slowly convergent sequences arising either from iterative processes or as the partial sums of infinite series. These authors include Delaunay [3], Samuelson [18], Shanks and Walton [23], Hartree [10], and Isakson [11]. The e_1^m transform was applied to certain convergent iterative sequences by A. C. Aitken [1]. More general discussions of the e_1, e_1^m, and $\tilde{e}_1$ transforms have been given by Shanks [20] and by Lubkin [14]. The late Otto Szász also studied e_1^m but he did not publish his findings. A somewhat fuller history of these transforms is given in the appendix.

The e_k and e_d transforms and their applications have been discussed by Shanks [20]. If $\{A_n\}$ is the sequence of partial sums of a power series then the array (9) and its diagonal (10) are intimately related to the Padé Table and related algorithms. This relationship is shown in our Theorem VI below.

This paper includes an improved version of the author's memorandum, [20]; and a number of new topics such as non-uniform convergence, spectra of sequences, meromorphic functions, and Gauss Numerical Integration, which have not been previously published.

An example (from Leibnitz). It is now desirable to illustrate these transforms, and in particular e_1, e_1^m, and the array (7) by an example. Let $A_0 = 4$, $A_1 = 4(1 - \frac{1}{3})$, and in general let $\{A_n\}$ be the sequence of partial sums of the very slowly convergent Leibnitz series:

$$\pi = 4 - \tfrac{4}{3} + \tfrac{4}{5} - \tfrac{4}{7} + \cdots \tag{11}$$

Concerning this series Knopp remarks [12, p 214, p 252] "beautiful" but "for numerical purposes—practically valueless". Some simple arithmetic, namely (3) and its iterates, gives the following array (7):

n	A_n	B_n	C_n	D_n	E_n	
0	4.0000000					
1	2.6666667	3.1666667				
2	3.4666667	3.1333333	3.1421053			
3	2.8952381	3.1452381	3.1414502	3.1415993		
4	3.3396825	3.1396825	3.1416433	3.1415909	3.1415928	(12)
5	2.9760462	3.1427129	3.1415713	3.1415933	3.1415927	
6	3.2837385	3.1408814	3.1416029	3.1415925		
7	3.0170718	3.1420718	3.1415873			
8	3.2523659	3.1412548				
9	3.0418396					

The tenth partial sum, A_9, is correct to only one figure; $A_{40,000,000}$ would be correct to eight figures. This illustrates Knopp's remark. But E_5 is already correct

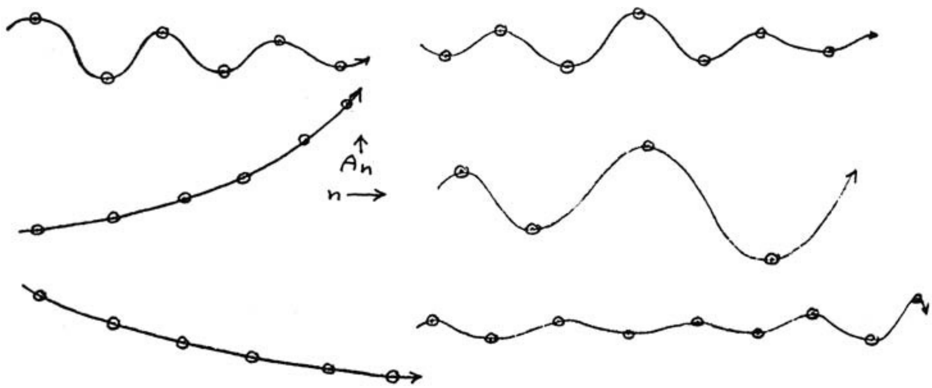

FIG. 1

to eight figures since $\pi = 3.1415926536$. And since E_5 is a function only of the nine arguments, A_1 to A_9, it is clear that the first ten terms of the Leibnitz series inherently possess this accuracy.

In Theorem I we shall give the theory corresponding to (12). But first it seems desirable not only to indicate why we might expect e_1^m to "work" so well on this series but more generally to motivate all of these transforms, (2), (3), (6), etc. To this motivation we now turn.

Heuristic motivation of the transforms. Consider a variety of typical numerical sequences $\{A_n\}$, convergent and divergent, monotonic and oscillatory, and plot A_n versus n. Draw a smooth curve through these discrete points. Typical graphs are as in Fig. 1.

One thing that all these graphs have in common is that they all look like the graphs of physical transients. By "*physical transient*" we mean any physical quantity, p, which is a function of time of the form:

$$p(t) = B + \sum_{i=1}^{k} a_i e^{\alpha_i t} \qquad (\alpha_i \neq 0) \quad (13)$$

where the α_i are arbitrary* complex numbers. Equivalently we write:

$$p(t) = B + \sum_{i=1}^{k} a_i q_i^t . \qquad (q_i \neq 1, 0) \quad (14)$$

In case of an m-fold degeneracy, that is, if

$$q_j = q_{j+1} = q_{j+2} = \cdots = q_{j+m-1} \quad (15)$$

we must modify the constants a_{j+1} to a_{j+m-1} in (14) according to the rule:

$$a_{j+r} \rightarrow a_{j+r} t^r . \qquad (r = 1, 2, \cdots m - 1) \quad (16)$$

* The author realizes that the term transient is often taken to mean something which is "dying out", i.e. all α's have negative real parts. However, the word is also used in the more general sense. Since we intend to apply the *same* transformations to *both* convergent and divergent sequences, it becomes desirable to emphasize the continuity between *stable* transients, (i.e. all α's have negative real parts), and *unstable* transients, (i.e. one or more α has a non-negative real part) by giving the term transient the more general meaning.

But for simplicity we ignore such degeneracy at present. Dependent upon its "*order*", k, its "*amplitudes*", a_i, and "*ratios*", q_i, the quantity p will oscillate or not, be stable or not, and in general manifest all behaviours indicated graphically above.

This close graphical resemblence suggests the possibility of regarding (some) sequences $\{A_n\}$ as if they were "*mathematical transients*", that is, as if they were functions of n of the form:

$$A_n = B + \sum_{i=1}^{k} a_i q_i^n . \qquad (q_i \neq 1, 0) \quad (17)$$

Analysis of such a mathematical transient would reveal its order and its "*spectrum*" of a's and q's. At present however we are primarily concerned with computing B, the "*base*" of the transient. For if $\{A_n\}$ is a mathematical transient, i.e. if it satisfies (17), and if each ratio satisfies $| q_i | < 1$ then clearly $B = \lim_{n\to\infty} A_n$. Again if $\{A_n\}$ is a transient and one or more $| q_i | \geqq 1$, A_n does not converge, but we will say that A_n "*diverges from B*". We will then call B the "*antilimit*" of $\{A_n\}$ and we intend to compute B as a summation method for $\{A_n\}$.

Now many sequences which arise naturally in analysis are indeed mathematical transients of some finite order k. Others are of infinite order ($k = \infty$ in (17)) and still others have a continuous spectrum:

$$A_n = B + \int_{q_0}^{q_1} a(q) q^n \, dq. \qquad (17a)$$

But in all of these cases it will often happen, at least for n greater than some fixed N, that a smaller number of components, say k, dominate the spectrum. In these cases we could say that $\{A_n\}$ is "*nearly of k'th order*".

Barring singular cases, *every* sequence $\{A_n\}$ is a k'th order transient *locally*. For consider the $2k + 1$ successive A's centered around $A_n : A_r$ with $n - k \leqq r \leqq n + k$ and $n \geqq k$. We can, in general, determine $2k + 1$ quantities, $B_{k,n}$; $a_{i,n}$; $q_{i,n}$ with $i = 1, 2, 3, \cdots k$ such that

$$\begin{aligned} A_r &= B_{k,n} + \sum_{i=1}^{k} a_{i,n}(q_{i,n})^r. && (q_{i,n} \neq 1, 0) \\ &(n - k \leqq r \leqq n + k) && (n \geqq k) \end{aligned} \qquad (18)$$

We call $B_{k,n}$ the "*local (k'th order) base*" of $\{A_n\}$. Holding k fixed, $B_{k,n}$ and the other $2k$ quantities are functions of n, but if $\{A_n\}$ is nearly of k'th order, $B_{k,n}$ should vary relatively little (with respect to the variations of A_n). We thus investigate the sequence of local bases, $\{B_{k,n}\}$ $(n = k, k + 1, \cdots)$ since its mode of generation suggests that if A_n converges to A, $B_{k,n}$ may converge to A more rapidly; while if A_n diverges, $B_{k,n}$ may either converge, or, more indirectly, serve to compute a suitable antilimit.

We shall see that if (18) is possible $B_{k,n} = e_k(A_n)$. The symbol e_k thus has reference to the k *exponentials* in (18). It is known [5] that if

$$p_s = \sum_{i=1}^{k} a_i q_i^s \qquad (s = 0, 1, \cdots 2k - 1) \quad (19)$$

then the k ratios q_i are the roots of the equation:

$$\begin{vmatrix} 1 & q & \cdot & q^{k-1} & q^k \\ p_0 & p_1 & \cdot & p_{k-1} & p_k \\ p_1 & p_2 & \cdot & \cdot & p_{k+1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ p_{k-1} & \cdot & \cdot & \cdot & p_{2k-1} \end{vmatrix} = 0. \tag{20}$$

Into the difference of successive equations in (18) we substitute p_s for ΔA_{n-k+s} and a_i for $a_{i,n}[(q_{i,n})^{n-k+1} - (q_{i,n})^{n-k}]$ and thus from (20) we see that the $q_{i,n}$'s of (18) are the roots of:

$$\begin{vmatrix} 1 & q & \cdot & q^k \\ \Delta A_{n-k} & \cdot & \cdot & \Delta A_n \\ \cdot & \cdot & \cdot & \cdot \\ \Delta A_{n-1} & \cdot & \cdot & \Delta A_{n+k-1} \end{vmatrix} = 0. \tag{21}$$

In the notation of (5) this becomes:

$$c_n q^k + c_{n-1} q^{k-1} + \cdots + c_{n-k} = 0. \tag{22}$$

Now inserting (18) into the right side of (5) and using (22) we find

$$e_k(A_n) = \frac{B_{k,n}(c_{n-k} + \cdots + c_{n-1} + c_n) + 0 + 0 + \cdots + 0}{c_{n-k} + \cdots + c_{n-1} + c_n} = B_{k,n} \tag{23}$$

provided the denominator does not vanish. If it does, then $q_{i,n} = 1$ is a root of (21) and therefore we have a singular case where (18) is not possible.

Further if (21) has an m-fold root, (15), and if (18) is modified as in the rule (16), it may be shown that(23) is still true. Thus the sequence of local bases is the k'th order transform and this motivates the sequence-to-sequence transform, $A_n \to B_{k,n}$, given by (2). One may similarly motivate the singular rules (2b) and (2c).

Now Leibnitz's series (11) is a "*nearly geometric*" series in the sense that $\lim_{n\to\infty} \Delta A_{n+1}/\Delta A_n$ exists (it equals -1). This implies, as is at once apparent, that the graph of A_n versus n may be approximated over a moderate range of n's by a first order transient: $A_n \doteq B + aq^n$ with $q \doteq -1$ and $B \doteq \pi$. Thus $B_n = e_1(A_n)$, the local base, should approximate π. But again B_n versus n resembles a first order transient and therefore we iterate $C_n = e_1(B_n)$, etc.

In short, given a sequence, we assume it to be a mathematical transient of order k or nearly of order k and attempt to filter out the exponential terms $\sum_{i=1}^k a_i q_i^n$ and thus reduce the sequence to its base (limit or antilimit) B. If necessary we repeat the transformation. With these remarks we conclude the heuristic motivation and the introduction.

Chapter I. e_1^m and some nearly geometric sequences.

Now returning to Leibnitz's series, $\pi = 4\sum_{m=0}^{\infty} (-1)^m/(2m+1)$, we wish to prove that each column of (12) converges to π more rapidly than the preceding column—in the sense that

$$\lim_{n\to\infty} \frac{\Delta e_1^k(A_n)}{\Delta e_1^{k-1}(A_n)} = 0$$

for $k = 1, 2, 3, \cdots$. At the same time we will generalize somewhat to the series $\sum_{m=0}^{\infty}(-1)^m f(m)/g(m)$ where $f(m)$ and $g(m)$ are polynomials in m of degree M_1 and M_2 respectively.

THEOREM I. Let $f(m) = \sum_{i=0}^{M_1} f_i m^i$ and $g(m) = \sum_{i=0}^{M_2} g_i m^i$ with $f_{M_1} \neq 0, g_{M_2} \neq 0$ and $g(m) \neq 0$ $(m = 0, 1, 2, \cdots)$, and let

$$A_n = \sum_{m=0}^{n}(-1)^m f(m)/g(m).$$

Then

(a) if $M_2 > M_1$, A_n converges and each derived sequence $B_n = e_1(A_n)$, $C_n = e_1(B_n)$, etc. converges more rapidly than its predecessor to the same limit.

(b) but if $M_2 \leqq M_1$, A_n diverges and each derived sequence diverges more slowly than its predecessor until $e_1^M(A_n)$ where $M = [\frac{1}{2}(M_1 - M_2)]$—the *integer* part of $\frac{1}{2}(M_1 - M_2)$. Then $e_1^{M+1}(A_n)$ converges and each further iterate converges faster to the same limit.

PROOF. First we note that (3),

$$B_n = \frac{A_{n+1}A_{n-1} - A_n^2}{A_{n+1} + A_{n-1} - 2A_n},$$

may be written

$$B_n = A_n + \frac{\Delta A_n}{1 - (\Delta A_n/\Delta A_{n-1})}$$

and also that

$$B_{n+1} = A_n + \frac{\Delta A_n}{1 - (\Delta A_{n+1}/\Delta A_n)}. \tag{24}$$

This implies

$$\Delta B_n = \Delta A_n \frac{\Delta A_{n+1}\Delta A_{n-1} - (\Delta A_n)^2}{(\Delta A_n - \Delta A_{n+1})(\Delta A_{n-1} - \Delta A_n)} \tag{25}$$

which in our case becomes

$$\Delta B_n = \Delta A_n \cdot \frac{f^2(n+1)g(n)g(n+2) - g^2(n+1)f(n)f(n+2)}{[f(n)g(n+1) - f(n+1)g(n)][f(n+1)g(n+2) - f(n+2)g(n+1)]}. \tag{26}$$

Expansion of (26) in powers of $1/n$ yields

$$\Delta B_n = \Delta A_n \left\{\frac{M_1 - M_2}{4n^2} + 0\left(\frac{1}{n^3}\right)\right\}. \tag{27}$$

Now since $f(m)$ and $g(m)$ do not vanish identically

$$\frac{\Delta A_n}{\Delta A_{n-1}} = -\frac{f(n+1)\,g(n)}{g(n+1)\,f(n)}$$

is negative for all sufficiently large n. Therefore if $M_2 > M_1$, the series $A_0 + \sum_{n=0}^{\infty}\Delta A_n$ is an alternating one (for sufficiently large n) with $|\Delta A_n| \to 0$

monotonically. Thus A_n converges, and by (24) B_n converges to the same limit. But by (27) $\Delta B_n/\Delta A_n \to 0$ and therefore B_n converges more rapidly.

The transformation may be repeated since $\{B_n\}$ is again a sequence of the same type (for sufficiently large n) with

$$\Delta B_n = (-1)^n \frac{f^{(1)}(n+1)}{g^{(1)}(n+1)}$$

where the degrees of $f^{(1)}$ and $g^{(1)}$, say $M_1^{(1)}$ and $M_2^{(1)}$, satisfy

$$M_2^{(1)} - M_1^{(1)} = M_2 - M_1 + 2 > 0 \tag{28}$$

by (26) and (27). This completes (a).

If $M_1 \geqq M_2$, A_n clearly diverges. If $M_1 - M_2 = 2N + 1$, $M = [\frac{1}{2}(M_1 - M_2)] = N$ and M iterations of e_1 give successively slower divergent series by (27) until $M_2^{(M)} - M_1^{(M)} = M_2 - M_1 + 2M = -1$. Then $M_2^{(M+1)} - M_1^{(M+1)} = +1$ and $e_1^{M+1}(A_n)$ converges. If $M_1 - M_2 = 2N$, M iterations of e_1 give $M_2^{(M)} - M_1^{(M)} = 0$. Therefore $M_2^{(M+1)} - M_1^{(M+1)} \geqq 3$ by (27) and $e_1^{M+1}(A_n)$ converges. By either alternative we now are in case (a) and this completes the proof.

In [20, p. 21] (1949) in which Theorem I was first proven, the author indicated that the theorem as stated covered the application of e_1^m to such famous convergent series as $\ln 2 = 1 - \frac{1}{2} + \frac{1}{3} - \cdots$ and $G = 1 - 3^{-2} + 5^{-2} - \cdots$ and to such divergent series as $\frac{1}{4} = 1 - 2 + 3 - \cdots$ and $0 = 1^2 - 2^2 + 3^2 - \cdots$ but did not cover some simple series such as $(1 - \sqrt{2})\zeta(\frac{1}{2}) = 1 - 2^{-\frac{1}{2}} + 3^{-\frac{1}{2}} - \cdots$ and $\ln 3/2 = \frac{1}{2} - \frac{1}{2}(\frac{1}{2})^2 + \frac{1}{3}(\frac{1}{2})^3 - \cdots$. Subsequently S. Lubkin, one of the independent discoverers of e_1^m, gave in [14] (written 1950, published 1952) a valuable generalization of Theorem I which does cover these alternating series together with a much larger class of nearly geometric series:

$$\{A_n\} \text{ convergent: } \frac{\Delta A_n}{\Delta A_{n-1}} = \alpha_0 + \frac{\alpha_1}{n} + \frac{\alpha_2}{n^2} + \cdots. \tag{28}$$

Further he shows e_1^m to be a suitable summation method for the (critically) divergent series:

$$\{A_n\} \text{ divergent: } \frac{\Delta A_n}{\Delta A_{n-1}} = -1 + \frac{\alpha_1}{n} + \frac{\alpha_2}{n^2} + \cdots. \tag{29}$$

The reader is referred to his paper and in particular to his Theorems 11 and 17.

Much earlier, (1926), A. C. Aitken [1] stated that if

$$A_n = \frac{\sum_{i=1}^{N} w_i z_i^{n+1}}{\sum_{i=1}^{N} w_i z_i^n} \tag{30}$$

with $z_1 > z_2 > \cdots z_N$, and the z's all real; them $\lim A_n = z_1$ and each derived sequence $e_1^m\{A_n\}$ converges to z_1 more rapidly than its predecessor. Here again the sequence $\{A_n\}$ of (30) is nearly geometric (with $\lim \Delta A_{n+1}/\Delta A_n = z_2/z_1$). The reader is referred to Aitken's papers [1], [2].

e_1^m and Newton's square root process. Still another type of nearly geometric sequence brings us to a new point. Let $A_1 = 1/1$, $A_2 = 3/2$, $A_3 = 7/5$, and in general let $\{A_n\}$ be the sequence of convergents of the continued fraction

$$\sqrt{2} = 1 + \frac{1}{2} \genfrac{}{}{0pt}{}{}{+ \frac{1}{2} \genfrac{}{}{0pt}{}{}{+ \frac{1}{2 + \cdots}}} \tag{31}$$

Applying e_1^m to this sequence, we obtain

n	A_n	B_n	C_n
1	1/1		
2	3/2	17/12	
3	7/5	99/70	19601/13860
4	17/12	577/408	
5	41/29		

Now $A_5 = 1.41379$ is correct to 3 figures but $C_3 = 1.414213564$ is correct to 9. To clarify this rapid convergence we shall prove

THEOREM II. If $\{A_n\}$ is the sequence of convergents of (31), then Newton's iterative square root process [28] and the e_1^m transform give identical derived sequences when applied to $\{A_n\}$. Specifically:

$$\begin{aligned} \tfrac{1}{2}(A_n + 2/A_n) &= e_1(A_n) = B_n \\ \tfrac{1}{2}(B_n + 2/B_n) &= e_1(B_n) = C_n\text{, etc.} \end{aligned} \tag{32}$$

PROOF. We note that if we write $A_n = N_n/D_n$ the recurrence law for continued fractions gives the equations

$$\begin{gathered} N_n = 2N_{n-1} + N_{n-2}, \qquad D_n = 2D_{n-1} + D_{n-2} \\ N_1 = D_1 = 1, \qquad N_2 = 3, \qquad D_2 = 2. \end{gathered} \tag{33}$$

The solution of these difference equations gives us

$$A_n = \sqrt{2}\left[\frac{1 + (2\sqrt{2} - 3)^n}{1 - (2\sqrt{2} - 3)^n}\right]. \tag{34}$$

More generally if we write

$$A_n = \sqrt{2}\left[\frac{1 + x^n}{1 - x^n}\right] \tag{34a}$$

we find from (3), after some algebra, that $\{B_n\}$ is a subsequence of $\{A_n\}$:

$$e_1(A_n) = B_n = A_{2n}. \tag{35}$$

And again some algebra yields:

$$\tfrac{1}{2}(A_n + 2/A_n) = A_{2n} = B_n. \tag{36}$$

Now writing

$$A_{2n} = \sqrt{2}\left[\frac{1 + (x^2)^n}{1 - (x^2)^n}\right] \tag{34b}$$

and repeating the computations, we find by induction further subsequences and

$$e_1^m(A_n) = A_{2^m \cdot n} = \frac{1}{2}\left[e_1^{m-1}(A_n) + \frac{2}{e_1^{m-1}(A_n)}\right]. \tag{37}$$

This proves the theorem.

Further it illustrates the rapid convergence which is often obtained by the e_1^m transform since it is known that each application of Newton's process doubles the number of correct decimal places.

This theorem will have analogues in later results. Of the many *special* algorithms which have been invented for the purpose of obtaining rapid convergence we call attention to Newton's method for square roots, Gauss numerical integration for Riemann sums, the Padé Table for rational approximations of analytic functions and the Euler and Gauss transformations of theta function infinite products. We shall find other instances, besides Theorem II, in which our non-linear transforms not only match the rates of convergence obtained by these *special* algorithms but also, as in Theorem II, yield identical results.

$e_1{}^m$, geometric series, and regularity. The previous section emphasized $e_1{}^m$ as an accelerator of convergence. We now turn to $e_1{}^m$ and e_1 as inducers of convergence, that is, as summation methods for divergent sequences. Consider first a trivial example. Let $\{A_n\}$ be the sequence of partial sums of the geometric series:

$$c + cz + cz^2 + cz^3 + \cdots . \tag{38}$$

Since

$$A_n = [c/(1 - z)] - [c/(1 - z)]z^n \qquad (z \neq 1) \tag{39}$$

it is clear that $\{A_n\}$ is an exact first order transient, ($k = 1$ in (17)), and that

$$B_n = e_1(A_n) = c/(1 - z) \tag{40}$$

for every n. We note that

$$B_n = \infty \qquad (z = 1) \tag{41}$$

and that

$$C_n = D_n = E_n = \cdots = B_n = c/(1 - z) \tag{42}$$

by rules (2b) and (2c). Thus we may say that e_1 (or e_1^m) sums (38) exactly. The divergence or convergence of (38) is clearly irrelevant. As special cases of (38) we obtain the well known sums:

$$\tfrac{1}{2} = 1 - 1 + 1 - 1 + \cdots \qquad -1 = 1 + 2 + 4 + 8 + \cdots$$

$$\tfrac{1}{3} = 1 - 2 + 4 - 8 + \cdots \qquad \infty = 1 + 1 + 1 + 1 + \cdots$$

and it is noted that e_1 (or e_1^m) sums a divergent series (38) to the "right" value, that is to the analytic continuation of the sums of convergent series (38). So far, so good, since this is desirable.

What is also desirable of a summation method is that it be "regular". If T is a transformation, T is regular means that if A_n converges, then $T(A_n)$ converges and $\lim T(A_n) = \lim A_n$. Now e_1 is (unfortunately) not regular [20, p 20].

THEOREM III. The transformation e_1 is not regular but if A_n and $e_1(A_n)$ both converge then $\lim e_1(A_n) = \lim A_n$.

PROOF (by counterexample). Let $\{A_n\}$ be the sequence of partial sums of

$$3 = 1 + \tfrac{1}{2} + \tfrac{1}{2} + \tfrac{1}{4} + \tfrac{1}{4} + \tfrac{1}{8} + \tfrac{1}{8} + \cdots \tag{43}$$

Then alternate B's, B_1, B_3, B_5, etc. are all ∞. But the even order B's $\to 3$ and thus $\lim B_n$ does not exist.

But if A_n and B_n both converge, their limits are equal. The reader is referred to [14] for Lubkin's Theorem I, the first published proof of this fact.

REMARK. The difficulty in (43) is in no way due to the infinities obtained for we could modify (43) to read

$$3 = 1 + (\tfrac{1}{2} + \epsilon_1) + (\tfrac{1}{2} - \epsilon_1) + (\tfrac{1}{4} + \epsilon_2) + (\tfrac{1}{4} - \epsilon_2) + \cdots$$

and thus alternate B's could be made arbitrarily large, although still finite. The difficulty is of a different origin; the series (43) is not nearly geometric and really calls for an e_2 transform. This point will be resumed later.

e_1^m, analytic continuation, and nonuniform convergence. The next example, though associated with the last two, (38) and (43), is more interesting and its theory is more intricate. Let $\{A_n\}$ be the sequence of partial sums of

$$f(z) = 1 + \tfrac{3}{2}z + \tfrac{7}{4}z^2 + \tfrac{15}{8}z^3 + \tfrac{31}{16}z^4 + \cdots, \tag{44}$$

the right side being the power series of the rational function

$$f(z) = 2/(1 - z)(2 - z).$$

We note, by breaking $f(z)$ into partial fractions, that (44) is the sum of two geometric series. From the partial sums of these two series we find

$$A_n = f(z) - \frac{2z}{1 - z}z^n + \frac{z}{2 - z}\left(\frac{z}{2}\right)^n. \qquad (z \neq 1, 2) \tag{45}$$

Since this is a second order transient, $k = 2$ in (17), it is clear that application of e_2 to *any* five successive members of $\{A_n\}$ would yield

$$e_2(A_n) = B_{2,n} = \frac{2}{(1 - z)(2 - z)} = f(z). \tag{46}$$

Again the value of z, and thus the convergence or divergence of (44) is irrelevant. We will return to the higher order transforms ($k \geqq 2$) later but now we are concerned with e_1^m and wish to investigate $e_1^m(A_n)$ and the step by step filtering process mentioned on p 8. Now A_n converges to $f(z)$ within the unit circle. We will prove that B_n converges to $f(z)$ within a circle of radius 2, C_n converges to $f(z)$ within a circle of radius 4, and in general $e_1^m(A_n)$ converges to $f(z)$ within a circle of radius 2^m with the following exceptional values of z. At the poles, $z = 1$ and 2, all sequences $\to \infty$. This is as it should be, but at the isolated point,

$z = 4$, which is in no way a singularity of $f(z)$, C_n, and all later sequences, converge *non-uniformly* to the *wrong* answer. That is $f(4) = \frac{1}{3}$ but as $n \to \infty$,

$$e_1{}^m(A_n) \to \tfrac{7}{27}. \qquad (z = 4)(m \geqq 2) \quad (47)$$

Similarly D_n and all later sequences converge non-uniformly to the wrong answer for $z = 8$ and the points $z = 2^m$ for $m > 3$ may also be irregular. The reason for the doubtful word "may" appears later.

We will need the following

LEMMA. Let $\{A_n\}$ be a sequence of the type:

$$A_n = B + x^n \frac{F(y^n)}{G(y^n)} \qquad \begin{matrix}(x, y \neq 0)\\(x, y \neq 1)\end{matrix} \quad (48)$$

with

$$F(y^n) = \textstyle\sum_{i=0}^{f} F_i(y^n)^i, \qquad (F_0, F_f \neq 0)$$

$$G(y^n) = \textstyle\sum_{i=0}^{g} G_i(y^n)^i. \qquad (G_0, G_g \neq 0)$$

Then $e_1\{A_n\} = \{B_n\}$ is either a sequence of the same type or $e_1(A_n) = \bar{B}$ (a constant) for every n.

PROOF. We compute

$$B_n = B + (\bar{x})^n \frac{\bar{F}(y^n)}{\bar{G}(y^n)} \quad (49)$$

with

$$\bar{G}(y^n) = \textstyle\sum_{i=0}^{\bar{g}} \bar{G}_i(y^n)^i, \qquad \bar{g} = f + 3g,$$

$$\bar{G}_0 = F_0G_0^3y(x-1)^2 \neq 0, \qquad \bar{G}_{\bar{g}} = F_fG_g^3y(x-1)^2 \neq 0,$$

$$\bar{F}(y^n) = \textstyle\sum_{i=0}^{\bar{f}} \bar{F}_i(y^n)^i, \qquad \bar{f} = 2(f + g - 1),$$

$$\bar{F}_0 = F_0G_0(F_1G_0 - F_0G_1)\,x(y-1)^2, \qquad \bar{x} = xy \neq 0,$$

$$\bar{F}_{\bar{f}} = F_fG_g(F_{f-1}G_g - F_fG_{g-1})x(y-1)^2.$$

Now (49) is of the same type as (48) except that we may have $\bar{F}_0 = 0$, $\bar{F}_f = 0$, $\bar{x} = 1$. But if every $\bar{F}_i = 0$, $e_1(A_n) = B$ for every n. If not, let $\bar{F}_{s-1}$ be the first non-vanishing coefficient. Shift the factor of $(y^n)^{s-1}$ in the numerator of (49) so that it reads

$$(\acute{\bar{x}})^n\acute{\bar{F}}(y^n) \quad \text{instead of} \quad (\bar{x})^n\bar{F}(y^n)$$

with $\acute{\bar{x}} = xy^s$ instead of $\bar{x} = xy$ and with a new $\acute{\bar{f}} < \bar{f}$ and new $F_0 \neq 0$, $\acute{\bar{F}}_{\acute{\bar{f}}} \neq 0$. Thus (49) (as revised if $s > 1$) is of the type (48) unless $\acute{\bar{x}} = 1$ (or if $s = 1$, unless $\bar{x} = 1$). But if $\acute{\bar{x}} = 1$ we have

$$B_n = B + \frac{\acute{\bar{F}}(y^n)}{\bar{G}(y^n)}$$

and $\acute{\bar{F}}(y^n)/\bar{G}y^n)$ is either a constant, ΔB, or it can be written as $\Delta B + (y^t)^n\ddot{\bar{F}}(y^n)/\bar{G}(y^n)$. Then (49) is of the type (48) with a new $\bar{B} = B + \Delta B$.

THEOREM IV. Let $\{A_n\}$ be the sequence of partial sums of the power series for

$$f(z) = \frac{c}{(z - z_0)(z - z_1)}, \qquad 0 < |z_0| < |z_1|, \quad (50)$$

and suppose $z \neq z_r (r = 0, 1, 2, \cdots)$ with

$$z_r = z_0(z_1/z_0)^r. \quad (51)$$

Then if $|z| < |z_m|$, $e_1^m(A_n) \to f(z)$. At the poles, $z = z_0$ and z_1, $e_1^m(A_n) \to \infty$ for all m. For $z = z_2 = z_1^2/z_0$ and $m \geqq 2$,

$$e_1^m(A_n) \to \left[1 - \frac{z_0 z_1}{(z_0 + z_1)^2}\right] f(z_2), \quad (52)$$

the *wrong* answer. In the neighborhood of z_2, $e_1^m(A_n)$ converges nonuniformly for $m > 2$.

PROOF. From (50) it follows for $z \neq z_0, z_1$ that

$$A_n = f(z) + f(z)\frac{z - z_1}{z_1 - z_0}\left(\frac{z}{z_0}\right)^{n+1}\left[1 - \frac{z - z_0}{z - z_1}\left(\frac{z_0}{z_1}\right)^{n+1}\right]. \quad (53)$$

From the lemma with $z/z_0 = x$, $z_0/z_1 = y$, $f = 1$, $g = 0$, and $B = f(z)$ we find $\bar{x} = xy = z/z_1$, $\bar{f} = 0$ and $\bar{g} = 1$:

$$B_n = f(z) - \left(\frac{z}{z_1}\right)^{n+2} \frac{f(z)(z_1 - z_0)}{(z - z_0) - (z - z_1)(z_0/z_1)^{n+2}}. \quad (54)$$

Therefore for $|z| < |z_1|$, $e_1(A_n) = B_n \to f(z)$. The series has been summed to the analytic continuation in a larger circle. (Remark: We note that the most divergent part of (53), the term involving $(z/z_0)^{n+1}$, has been removed—but what remains is *not* a first order transient. By dividing out the fraction in (54) we find that $\{B_n\}$ is an ∞ order transient:

$$B_n = f(z) - f(z)\frac{z_1 - z_0}{z - z_0}\sum_{i=0}^{\infty}\left(\frac{z - z_1}{z - z_0}\right)^i \left[\frac{z}{z_1}\left(\frac{z_0}{z_1}\right)^i\right]^{n+2}. \quad (55)$$

The nonlinear operator e_1 has *eliminated* one exponential but has also *introduced* infinitely many new ones.)

Now from (54) and (48) with $(z/z_1) = x$ and $(z_0/z_1) = y$, it follows that

$$C_n = f(z) - \frac{f(z)(z - z_1)(z_1 - z_0)^3 z_0^{-1} z_1^{-1}(zz_0/z_1^2)^{n+3}}{(z - z_0)^2(z - z_1)^2 z_0 z_1^{-3} + \sum_{i=1}^{3} a_i y^{in}}. \quad (56)$$

Since $|y| < 1$ we see that if $|z| < |z_2| = |z_1^2/z_0|$ then $|\bar{x}| = |zz_0/z_1^2| < 1$ and $C_n \to f(z)$. But for $z = z_2$, $\bar{x} = 1$, and

$$C_n \to f(z) - f(z)\frac{z_0 z_1}{(z_0 + z_1)^2}. \quad (57)$$

This anomaly may be attributed to the fact that the nonlinear e_1 operator, by combining two terms of B_n, the divergent term $x^n = (z_1/z_0)^n$, and the convergent term $y^n = (z_0/z_1)^n$ has produced a term in $(1)^n$ which is constant and thus becomes absorbed into the old base, $f(z)$ of $\{B_n\}$. The $\{C_n\}$ sequence for $z = z_2$ can again be put into the form (48) with the new B, $f(z)[1 - z_0z_1(z_0 + z_1)^{-2}]$ and a new $x = z_0/z_1 = y$. Further applications of e_1 will therefore make the derived sequences converge faster, but always to the same new B, since all positive powers of $| y |$ are *less* than one. Now let $z \neq z_r$, then $(z/z_0)(z_0/z_1)^r \neq 1$ and therefore no number of iterations of the transformation (48) $\to$ (49) can produce a term in $(1)^n$. Each iteration however multiplies x^n by a new factor of $y^n = (z_0/z_1)^n$ if the corresponding $\bar{F}_0 \neq 0$ or by y^{sn} if $F_0 = F_1 = \cdots = F_{s-2} = 0$, $F_{s-1} \neq 0$ or by 0 if all the F_i's vanish. Thus in any case if $| z | < | z_m |$ the dominant term in the transient part of $e_1{}^m(A_n)$ is $(z/z_0)^n(z_0/z_1)^{Mn}$ where $M \geqq m$ and therefore $e_1^m(A_n)$ and all later sequences converge to $f(z)$. At the poles $z = z_0$ or z_1 (53) must be replaced by a sequence of the type:

$$A_n = a + bn + cq^n. \qquad (b \neq 0) \quad (58)$$

It is easily seen that B_n is then of same type and further that $C_n = A_n$. Thus all sequences $e_1^m(A_n)$ for $z = z_0$ or $z = z_1$ tend to (the complex number) ∞. Finally it is clear that the convergence at z_2 is nonuniform since we do not obtain continuity there but instead the isolated little spike, $-f(z)z_0z_1(z_0 + z_1)^{-2}$ which is reminiscent of the Gibbs Phenomena. This completes the proof. Remark: The point $z = z_r$ for $r \geqq 3$ may also be irregular. The sequence $e_1^r\{A_n\}$, and later sequences, converge there with a possible term in $(1)^n$. However the above computation does not preclude the possibility that the corresponding F_i coefficient may vanish.

$\tilde{e}_1$ **and some numerical divergent series.** From $e_1{}^m$, "wrong" answers, and messy theory, we now turn to $\tilde{e}_1$, "right" answers, and (almost) no theory at all. Many interesting, useful, amazing, or amusing examples have been given of $\tilde{e}_1$ as a summation method of divergent sequences. We mention the following numerical examples:

$$\ln 3 = 0 + 2 - \tfrac{1}{2}2^2 + \tfrac{1}{3}2^3 - \tfrac{1}{4}2^4 + \cdots \qquad [20, \text{p } 17] \quad (59)$$

$$\int_0^\infty \frac{e^{-t}\,dt}{1+t} = 0.596347 + = 0! - 1! + 2! - 3! + 4! - \cdots \qquad [20, \text{p } 17] \quad (60)$$

$$C = \tfrac{1}{2} + \tfrac{1}{2}B_2 + \tfrac{1}{4}B_4 + \tfrac{1}{6}B_6 + \cdots \qquad [20, \text{p } 18] \quad (61)$$

$$G = \tfrac{1}{2}\pi(3 \ln 3 - 5 \ln 5 + 7 \ln 7 - \cdots) \qquad [20, \text{p } 29] \quad (62)$$

$$f(x) = \frac{x^2}{2!} - \frac{x^5}{5!} + \frac{11x^8}{8!} - \frac{375x^{11}}{11!} + \cdots \ (x > 3.12735) \qquad [21] \quad (63)$$

In (61) C is Euler's constant and the B's are the Bernoulli numbers. In (62) G is Catalan's constant and in (63) $f(x)$ is Blasius' boundary layer function.

Another numerical divergent series, which is much to the point in any discussion of "right" and "wrong" answers was given in a puzzle by G. H. Hardy [8]. Consider the series:

$$x = 1 - \tfrac{1}{2}(\tfrac{16}{9}) + \tfrac{1}{2}\cdot\tfrac{3}{4}(\tfrac{16}{9})^2 - \tfrac{1}{2}\cdot\tfrac{3}{4}\cdot\tfrac{5}{6}(\tfrac{16}{9})^3 + \cdots \qquad (64)$$

which was obtained from

$$f(z) = 1 + \tfrac{1}{2}\left(\frac{2z}{1 + z^2}\right)^2 + \tfrac{1}{2}\cdot\tfrac{3}{4}\cdot\left(\frac{2z}{1 + z^2}\right)^4 + \cdots \tag{64a}$$

by putting $z = 2i$. Consider the domains $D_1 : |z - i| \leqq \sqrt{2}$ and $D_2 : |z + i| \leqq \sqrt{2}$. In the intersection of D_1 and D_2 the right side of (64a) converges to $f_1(z) = (1 + z^2)/(1 - z^2)$ whereas in the complement of their union it converges to $f_2(z) = (z^2 + 1)/(z^2 - 1)$. In the remaining two lunes the series diverges. The point $z = 2i$ is in the upper lune and therefore (64) diverges. If we were to sum (64), say by $\tilde{e}_1$, what should the "right" answer be—should $x = f_2(2i) = +\frac{3}{5}$ or should $x = f_1(2i) = -\frac{3}{5}$?

We will answer this question later. Likewise in connection with the phenomenon of semi-convergence we will discuss (59) later. But in general the theory here is not developed and we must confine ourselves to a few remarks concerning such numerical divergent sequences.

1. In many cases, e.g. (59), (60), (61), no column of (7) converges. Nonetheless the diagonal of (7), $\tilde{e}_1\{A_n\}$, appears to converge quite rapidly to the right answer.

2. As already apparent (in the proof of Theorem IV) the algebra and analysis of (7) can become quite involved. But the arithmetic of (7), which is all that was used in examples (60) to (63) is very simple. It has been coded for the electronic Card Programmed Calculator by T. S. Walton of the Naval Ordnance Laboratory. Inserting a sequence $\{A_n\}$ into the machine, one rapidly obtains the array (7). The table in (12) was computed in this way.

3. It will soon be seen that the array (9) often has the opposite character—involved arithmetic *(because of the high order determinants) but relatively simple theory.

Chapter II. Introduction to e_2, an example (from Lovitt). Examples (43) and (44) showed mischief created by $e_1{}^m$ in cases where e_2 was really more appropriate. Other instances are easy to find. Modify (11) to read:

$$\pi = 4 + 0 - \tfrac{4}{3} + 0 + \tfrac{4}{5} + 0 - \tfrac{4}{7} + \cdots \tag{65}$$

that is, insert $x = 1$ in the power series:

$$4 \arctan x = 4x + 0x^2 - \tfrac{4}{3}x^3 + 0x^4 + \cdots.$$

One finds now not the pleasant table (12) but the discouraging one: $A_n = B_n = C_n = \cdots$. Again consider the divergent series:

$$1 = 1 + 2\cdot 2 + 3\cdot 2^2 + 4\cdot 2^3 + \cdots \tag{66}$$

which comes from putting $x = 2$ in the power series for $(1 - x)^{-2}$. With this second order pole, the situation is worse than in Theorem IV. Now no column

* Dr. A. S. Householder has pointed out to the author that the use of H. Rutishauser's "*Quotienten-Differenzen-Algorithmus*", ZAMP 5, 233 (1954) would simplify the arithmetic. (Note added in proof.)

of (7) converges and the diagonal converges only slowly. But e_2 yields the correct result at a blow. For

$$B_{2,2} = \begin{vmatrix} 1 & 5 & 17 \\ 4 & 12 & 32 \\ 12 & 32 & 80 \end{vmatrix} \Bigg/ \begin{vmatrix} 1 & 1 & 1 \\ 4 & 12 & 32 \\ 12 & 32 & 80 \end{vmatrix} = 1$$

and in fact $B_{2,n} = 1$ for every $n \geqq 2$. This is a special case of Theorem IX which is proven later. We now investigate these higher order transforms—beginning with an example.

Consider the simple integral equation [13, p. 13]

$$u(x) = \tfrac{1}{2}x - \tfrac{1}{3} + \int_0^1 (x + t)u(t)\, dt. \tag{67}$$

By the method of successive substitutions we obtain the sequence $u_0 = 0$, $u_1 = \frac{1}{2}x - \frac{1}{3}$, $u_2 = \frac{5}{12}x - \frac{1}{3}$, $u_3 = \frac{1}{3}x - \frac{13}{18}$, etc. But it is found that $u_n \to -\infty x - \infty$ and so the method "fails". Now it is true that u_n has no limit, but it does have an antilimit!

THEOREM V. The divergent sequence u_n may be summed to the solution of (67) by e_2.

PROOF. By the method of successive substitutions:

$$u_{n+1}(x) = u_1 + \int_0^1 (x + t)u_n(t)\, dt. \tag{68}$$

Writing $u_n = A_n x + B_n$ as a vector $u_n = \begin{pmatrix} A_n \\ B_n \end{pmatrix}$ we find from (68) that

$$\begin{pmatrix} A_{n+1} \\ B_{n+1} \end{pmatrix} = \begin{pmatrix} A_1 \\ B_1 \end{pmatrix} + \begin{pmatrix} \frac{1}{2} & 1 \\ \frac{1}{3} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} A_n \\ B_n \end{pmatrix}$$

and therefore if

$$M = \begin{pmatrix} \frac{1}{2} & 1 \\ \frac{1}{3} & \frac{1}{2} \end{pmatrix} \tag{69}$$

that

$$u_{n+1} = (M^n + M^{n-1} + \cdots + M + 1)u_1. \tag{70}$$

By expanding M into the form $M = aM_1 + bM_2$ with $M_1^2 = M_1$, $M_2^2 = M_2$ and $M_1M_2 = 0$ we obtain:

$$M^n = (\tfrac{1}{2} + \tfrac{1}{3}\sqrt{3})^n \begin{pmatrix} \frac{1}{2} & \frac{1}{2}\sqrt{3} \\ \frac{1}{6}\sqrt{3} & \frac{1}{2} \end{pmatrix} + (\tfrac{1}{2} - \tfrac{1}{3}\sqrt{3})^n \begin{pmatrix} \frac{1}{2} & -\frac{1}{2}\sqrt{3} \\ -\frac{1}{6}\sqrt{3} & \frac{1}{2} \end{pmatrix}. \tag{71}$$

Summing the geometric series of (70) thus yields

$$u_{n+1} = x - (\tfrac{1}{2} + \tfrac{1}{3}\sqrt{3})^{n+1}(\tfrac{1}{2}x + \tfrac{1}{6}\sqrt{3}) - (\tfrac{1}{2} - \tfrac{1}{3}\sqrt{3})^{n+1}(\tfrac{1}{2}x - \tfrac{1}{6}\sqrt{3}), \tag{72}$$

a divergent second order transient with a base $= x$. Thus $e_2(u_n) = x$ for all n.

That $u(x) = x$ is the solution of (67) may be readily verified but it is also clear a priori that it must be. For from (68) by a linear combination we have:

$$\frac{c_{n-2}u_{n-1}(x) + c_{n-1}u_n(x) + c_n u_{n+1}(x)}{c_{n-2} + c_{n-1} + c_n}$$
$$= u_1(x) + \int_0^1 (x + t) \frac{c_{n-2}u_{n-2}(t) + c_{n-1}u_{n-1}(t) + c_n u_n(t)}{c_{n-2} + c_{n-1} + c_n} dt$$

Now the right and left linear combinations appear to be different functions since the u indices are one greater on the left. But by (5c) they really add to the same function and thus this function is the solution of (67). Remark: From (72), with a proof similar to that of Theorem IV, one may show that e_1^m could also do the job—but not as simply as e_2.

Introduction to e_d, rational approximations. Returning to the partial sums of power series

$$A_n = \textstyle\sum_{i=0}^n a_i z^i, \tag{73}$$

we note that the k'th order transform

$$e_k(A_n) = \frac{\begin{vmatrix} A_{n-k} & A_{n-k+1} & \cdot & A_n \\ a_{n-k+1}z^{n-k+1} & \cdot & \cdot & a_{n+1}z^{n+1} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ a_n z^n & \cdot & \cdot & a_{n+k}z^{n+k} \end{vmatrix}}{\begin{vmatrix} 1 & 1 & \cdot & 1 \\ a_{n-k+1}z^{n-k+1} & \cdot & \cdot & a_{n+1}z^{n+1} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ a_n z^n & \cdot & \cdot & a_{n+k}z^{n+k} \end{vmatrix}} \tag{74}$$

may be reduced to

$$e_k(A_n) = \frac{\begin{vmatrix} z^k A_{n-k} & z^{k-1}A_{n-k+1} & \cdot & z^0 A_n \\ a_{n-k+1} & \cdot & \cdot & a_{n+1} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ a_n & a_{n+1} & \cdot & a_{n+k} \end{vmatrix}}{\begin{vmatrix} z^k & z^{k-1} & \cdot & z^0 \\ a_{n-k+1} & \cdot & \cdot & a_{n+1} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ a_n & a_{n+1} & \cdot & a_{n+k} \end{vmatrix}} \tag{75}$$

by multiplying the p'th column of each $p = 1, 2, \cdots k + 1$ by z^{k+1-p} and dividing the q'th row of each $q = 2, 3, \cdots k + 1$ by z^{n+q-1}. In this way a common factor of z^{kn} has been deleted from both determinants.

If $\{A_n\}$ is the sequence of partial sums of the logarithm series;

$$\ln (1 + z) = 0 + z - \tfrac{1}{2}z^2 + \tfrac{1}{3}z^3 - \cdots \tag{76}$$

one finds from (75) with $k = n$ that (10), the diagonal transform of $\{A_n\}$ is the sequence of rational approximations:

$$B_{0,0} = 0, \qquad B_{1,1} = \frac{2z}{2 + z}, \qquad B_{2,2} = \frac{6z + 3z^2}{6 + 6z + z^2}, \quad \text{etc.} \tag{77}$$

If A'_n is the same sequence with the zero term omitted, that is if $A'_0 = z$, $A'_1 = z - \frac{1}{2}z^2$, etc. we would find instead for $e_d\{A'_n\}$ the sequence:

$$B'_{0,0} = z, \qquad B'_{1,1} = z\,\frac{6 + z}{6 + 4z}, \qquad B'_{2,2} = z\,\frac{30 + 21z + z^2}{30 + 36z + 9z^2}, \quad \text{etc.} \tag{78}$$

Listing (77) and (78) alternately and computing the power series of these rational functions we obtain the sequence $\{K_n\}$:

$$\begin{aligned}
K_0 &= B_{0,0} = 0\\
K_1 &= B'_{0,0} = z\\
K_2 &= B_{1,1} = z - \tfrac{1}{2}z^2 + \tfrac{1}{4}z^3 - \tfrac{1}{8}z^4 + \cdots\\
K_3 &= B'_{1,1} = z - \tfrac{1}{2}z^2 + \tfrac{1}{3}z^3 - \frac{1}{4.5}\,z^4 + \frac{1}{6.75}\,z^5 - \cdots\\
K_4 &= B_{2,2} = z - \tfrac{1}{2}z^2 + \tfrac{1}{3}z^3 - \tfrac{1}{4}z^4 + \frac{1}{5.14}\,z^5 - \frac{1}{6.55}\,z^6 + \cdots\\
K_5 &= B'_{2,2} = z - \tfrac{1}{2}z^2 + \tfrac{1}{3}z^3 - \tfrac{1}{4}z^5 + \tfrac{1}{5}z^5 - \frac{1}{6.03}\,z^6 + \frac{1}{7.25}\,z^7 - \cdots
\end{aligned} \tag{79}$$

etc., and note the following:

1. K_n is a function only of $A_0, A_1, \cdots A_n$.
2. K_n agrees with A_n to the term in z^n.
3. The "predictions" made by K_n of the next few derivatives are increasingly good as n increases. Of course these future derivatives are in no way determined by the polynomial (73) and could be chosen arbitrarily. The good results of (79) must be attributed to the fact that the assumption of *smoothness* of the sequence $\{A_n\}$ implicit in the assumption (17) is indeed valid for the sequence of partial sums of the logarithm series (76).

It can be shown that $\{K_n\}$ is the sequence of convergents of the well known continued fraction [26, p 342]:

$$\ln (1 + z) = 0 + \cfrac{z}{1 + \cfrac{1^2 z}{2 + \cfrac{1^2 z}{3 + \cfrac{2^2 z}{4 + \cfrac{2^2 z}{5 + \cdots}}}}} \tag{80}$$

It will be shown later that the sequence (77) is obtainable by Gauss Numerical Integration and that it converges to the principal value of ln $(1 + z)$ for every z not on the real cut, $z \leqq -1$. It will be shown now that the sequence (78) divided by z is the diagonal of the Padé Table [26] of $(1/z) \log (1 + z)$.

e_k **and the Padé Table.** Given an analytic function:

$$f(z) = \sum_{i=0}^{\infty} a_i z^i \tag{81}$$

with $a_0 \neq 0$, the Padé Table for $f(z)$ is a two dimensional array of rational functions $R_{k,n}(k = 0, 1, 2, \cdots)(n = 0, 1, 2, \cdots)$

$$\begin{matrix} R_{0,0} & R_{0,1} & R_{0,2} & \cdot \\ R_{1,0} & R_{1,1} & R_{1,2} & \cdot \\ R_{2,0} & R_{2,1} & R_{2,2} & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{matrix} \tag{82}$$

with the two properties:

Property 1. $R_{k,n}$ may be written as the ratio of two polynomials:

$$R_{k,n} = N_{k,n}/D_{k,n} \tag{83}$$

with the degree of $N_{k,n} \leqq n$ and the degree of $D_{k,n} \leqq k$. But $D_{k,n}$ does not vanish identically.

Property 2. The power series of $R_{k,n}$ agrees with that of $f(z)$ to a higher power of z than any other rational function with degrees of numerator and denominator no greater than n and k respectively.

It is known that:

A. [26, p 378] The rational function $R_{k,n}$ is uniquely characterized by properties 1 and 2.

B. [26, p 378] Property 2 is equivalent to the condition

$$f(z)D_{k,n} - N_{k,n} = (z^{k+n+1}). \tag{84}$$

Here (z^r) means a power series $g(z) = \sum_{i=r}^{\infty} g_i z^i$ for which each term in z^s vanishes if $s < r$. Higher terms may also vanish; in particular we may have $g(z) \equiv 0$.

C. (Padé) [26, p 394] If $N_{k,n}$ and $D_{k,n}$ have no zero in common and if the degree of $D_{k,n}$ is exactly $p(p \leqq k)$ then

$$R_{k,n} = R_{k-1,n} = R_{k-2,n} = \cdots = R_{p,n}. \tag{85}$$

D. (Montessus de Balloire) [26, p 412] If $f(z)$ is regular for $|z| \leqq R$ except for p poles within this circle, then, as $n \to \infty$, $R_{p,n}$ converges to $f(z)$ uniformly in the domain obtained from $|z| \leqq R$ by removing the interiors of small circles with centers at these poles. Note: Higher order poles are counted according to their order.

We also need the following lemma similar to a part of theorem 98.2 of [26].

E. Lemma. If

$$\Delta_{k-1,n} = \begin{vmatrix} a_{n-k+1} & \cdots & a_n \\ \vdots & & \vdots \\ a_n & \cdots & a_{n+k-1} \end{vmatrix} = 0, \tag{86}$$

then

$$R_{k-1,n} = R_{k-1,n-1} = R_{k,n} .$$

After we have proved E we shall obtain the following theorem:

Theorem VI. Let $f(z)$ be given by (81) and $R_{k,n}$ by (82). If A_n is given by (73), then

$$B_{k,n} = R_{k,n} \tag{87}$$

and thus the array (9) is the upper half of the Padé Table (82)—above and including the diagonal.

From this result we obtain immediately:

Theorem VII. If $f(z)$ is regular for $|z| \leqq R$ except for p poles within this circle, and if $\{A_n\}$ is the sequence of partial sums of its power series, then as $n \longrightarrow \infty$,

$$B_{p,n} \xrightarrow[\text{uniformly}]{} f(z) \tag{88}$$

in the domain obtained from $|z| \leqq R$ by removing the interiors of small circles with centers at these poles.

Proof of VII. D and(87).

Theorem VIII. Let $f(z)$ be meromorphic and let z_0 be any regular point in the finite plane. Then if $\{A_n\}$ is the sequence of partial sums of its power series there is a p such that

$$\lim_{n\to\infty} e_p(A_n) = f(z_0). \tag{88a}$$

Proof of VIII. In the domain $|z| \leqq |z_0|$ there are only a finite number of poles. Let this number be p and apply Theorem VII.

Theorem IX. Let $f(z)$ be a rational function, which when reduced to its lowest terms has a q'th degree numerator and a p'th degree denominator, then if $k \geqq p$, $n \geqq q$,

$$B_{k,n} = f(z). \tag{89}$$

Proof of IX. If $k \geqq p$, $n \geqq q$, then $R_{k,n} = f(z)$ since $f(z)$ clearly has properties 1 and 2 and by A, $R_{k,n}$ is unique. Then (89) follows at once from (87).

We now return to the

Proof of E. To compute the Padé approximant $R_{k-1,n-1}$ we let

$$\left.\begin{aligned} R_{k-1,n-1} &= \frac{\sum_{i=0}^{n-1} s_i z^i}{\sum_{i=0}^{k-1} t_i z^i}, \qquad f(z) = \sum_{i=0}^{\infty} a_i z^i, \\ &\text{and let} \\ \sum_{i=0}^{\infty} a_i z^i &\sum_{i=0}^{k-1} t_i z^i = \sum_{i=0}^{\infty} b_i z^i. \end{aligned}\right\} \tag{90}$$

In order to satisfy (84) for the indices $k - 1$ and $n - 1$ we need

$$s_i = b_i \qquad (i = 0, 1, 2, \cdots n - 1) \tag{91}$$

and

$$b_i = 0. \qquad (i = n, n + 1, n + 2, \cdots n + k - 2)$$

The last equations, written out, are

$$\begin{aligned} b_n &= a_{n-k+1}t_{k-1} + a_{n-k+2}t_{k-2} + \cdots + a_n t_0 = 0 \\ b_{n+1} &= a_{n-k+2}t_{k-1} + a_{n-k+3}t_{k-2} + \cdots + a_{n+1}t_0 = 0 \\ \cdots\cdots &\quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ b_{n+k-2} &= a_{n-1}t_{k-1} + a_n t_{k-2} + \cdots + a_{n+k-2}t_0 = 0. \end{aligned} \tag{92}$$

These are $k - 1$ equations in the k unknowns t_0, t_1, $\cdots t_{k-1}$ and thus have a non-trivial solution. Choosing any such solution we now determine the s_i by (91). Therefore

$$f(z)D_{k-1,n-1} - N_{k-1,n-1} = (z^{k+n-1}) \tag{93}$$

and we have correctly computed $R_{k-1,n-1}$ since it is unique. However if (86) is true we may adjoin another equation,

$$b_{n+k-1} = a_n t_{k-1} + a_{n+1}t_{k-2} + \cdots + a_{n+k-1}t_0 = 0, \tag{94}$$

to (92) and solve the k homogeneous equations in k unknowns since the determinant of the system vanishes. Again we determine s_i by (91) and again we obtain $R_{k-1,n-1}$ since it is unique. However we now see from (94) that (93) can be improved to read:

$$f(z)D_{k-1,n-1} - N_{k-1,n-1} = (z^{k+n}). \tag{95}$$

This increase in degree means that we may set $N_{k-1,n} = N_{k-1,n-1}$ and $D_{k-1,n} = D_{k-1,n-1}$ and therefore

$$R_{k-1,n} = \frac{N_{k-1,n-1}}{D_{k-1,n-1}} = R_{k-1,n-1}. \tag{96}$$

For this $N_{k-1,n}$ and this $D_{k-1,n}$ clearly have degrees no greater than n and $k - 1$ respectively and thus property 1. Further, by (95), $f(z)D_{k-1,n} - N_{k-1,n} = (z^{k-1+n+1})$ so that B implies property 2. Again, we may set $N_{k,n} = zN_{k-1,n-1}$ and $D_{k,n} = zD_{k-1,n-1}$ and thus

$$R_{k,n} = \frac{zN_{k-1,n-1}}{zD_{k-1,n-1}} = R_{k-1,n-1} \tag{97}$$

since $N_{k,n}$ and $D_{k,n}$ have degrees no greater than n and k and since from (95)

$$f(z)D_{k,n} - N_{k,n} = z(f(z)D_{k-1,n-1} - N_{k-1,n-1}) = (z^{k+n+1}). \tag{98}$$

This proves E.

Proof of Theorem VI. Let the right side of (75) be written $\nu_{k,n}/\delta_{k,n}$ and consider the following four cases which may arise for any fixed $z = z_0$.

1. For $z = z_0$, $\delta_{k,n}$ does not vanish.
2. For $z = z_0$, $\delta_{k,n}$ does vanish but $\nu_{k,n}$ does not.
3. For $z = z_0$, $\delta_{k,n}$ and $\nu_{k,n}$ both vanish but $\delta_{k,n}$ does not vanish identically.
4. The determinant $\delta_{k,n}$ (and therefore also $\nu_{k,n}$) vanishes identically.

In cases 1 to 3, $\delta_{k,n}$ does not vanish identically—if it did in case 2, $\nu_{k,n}$ would also vanish. Therefore since $\nu_{k,n}$ is at most of degree n and $\delta_{k,n}$ is at most of degree k, $e_k(A_n)$ has property 1. In the numerator $\nu_{k,n}$ we add z^{n+1} times the second row, z^{n+2} times the third row, $\cdots$, and z^{n+k} times the last row to the first row. Then $\nu_{k,n}$ reads:

$$\nu_{k,n} = \begin{vmatrix} z^k A_n & z^{k-1} A_{n+1} & \cdot & zA_{n+k-1} & z^0 A_{n+k} \\ a_{n-k+1} & a_{n-k+2} & \cdot & a_n & a_{n+1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_n & \cdot & \cdot & a_{n+k-1} & a_{n+k} \end{vmatrix} \tag{99}$$

i.e. each index of A_i is merely increased by k. Therefore

$$A_{n+k}\,\delta_{k,n} - \nu_{k,n} = \begin{vmatrix} z^k(A_{n+k} - A_n) & z^{k-1}(A_{n+k} - A_{n+1}) & \cdot & 0 \\ a_{n-k+1} & a_{n-k+2} & \cdot & a_{n+1} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ a_n & a_{n+1} & \cdot & a_{n+k} \end{vmatrix} \tag{100}$$

Now the right side of (100) has only terms of degree $n + k + 1$ or higher. Therefore

$$f(z)\delta_{k,n} - \nu_{k,n} = A_{n+k}\delta_{k,n} - \nu_{k,n} + \delta_{k,n}\sum_{i=n+k+1}^{\infty} a_i z^i = (z^{n+k+1}) \tag{101}$$

and $e_k(A_n)$ has property 2. Thus in the first three cases $e_k(A_n) = R_{k,n}$. Now in case 1, $\delta_{k,n} \not\equiv 0$, and so by (2) $B_{k,n} = e_k(A_n) = R_{k,n}$. In case 2 we have assigned ∞ to $B_{k,n}$ by rule (2b). This agrees with the assignment of ∞ to a pole of $R_{k,n}$ and therefore $B_{k,n} = R_{k,n}$. In case 3 and 4 we have assigned to $B_{k,n}$ by rule (2c), the value

$$B_{k,n} = B_{k-1,n}$$

with repeated reduction of order as often as is necessary to obtain a well defined value. Now in case 3, $\nu_{k,n}$ and $\delta_{k,n}$ must have a common zero. If this and any other common factors are factored out of both polynomials, the resulting $\delta'_{k,n}$ has a degree which is less than k and therefore from (85)

$$R_{k,n} = R_{k-1,n}\,.$$

Repeating this as often as is necessary we find

$$B_{k,n} = B_{k-1,n} = \cdots = B_{p,n} = e_p(A_n) = R_{p,n} = \cdots = R_{k,n}\,.$$

Finally in case 4 if $\delta_{k,n}$ vanishes identically then in particular the lower left $k \times k$ minor of $\delta_{k,n} = \Delta_{k-1,n} = 0$ and therefore by Lemma E

$$B_{k,n} = B_{k-1,n} = R_{k-1,n} = R_{k,n} \; ;$$

with repetitions if necessary.

Thus for every z, (87) is true and this completes the proof of Theorems VI, VII, VIII, and IX.

Remarks on the foregoing. A number of remarks are in order.

1. From Theorem VI, and rule (5a), it is now easily seen that the sequence (78) is z times the diagonal of the Padé Table for $(1/z)\log(1 + z)$.

2. The lower half of the Padé Table—that is $R_{k,n}$ for $k \geqq n$—may similarly be obtained by transforming the reciprocal series. Specifically, let

$$g(z) = [f(z)]^{-1} = \textstyle\sum_{i=0}^{\infty} \acute{a}_i z^i; \qquad \acute{A}_n = \textstyle\sum_{i=0}^{n} \acute{a}_i z^i;$$

then for $k \geqq n$,

$$R_{k,n} = [e_n(\acute{A}_k)]^{-1}. \tag{102}$$

We will not use (102) and hence shall not prove it.

3. In Theorem VIII suppose that there are p_1 poles in the domain $|z| \leqq |z_1|$ with $p_1 > p$ and $|z_1| > |z_0|$. We will then have, in addition to (88a),

$$\lim_{n\to\infty} e_{p_1}(A_n) = f(z_0)$$

since the p_1 transform converges in the larger circle. But this should *not* be construed to read

$$\lim_{n\to\infty} e_{\acute{p}}(A_n) = f(z_0)$$

for every $\acute{p}$ which satisfies $p \leqq \acute{p} \leqq p_1$. This generalization would indeed be true if no two poles in the annulus $|z_0| < |z| \leqq |z_1|$ had the same absolute value but it is fallacious otherwise. The series:

$$(1 + \tfrac{1}{2}z)/(1 - \tfrac{1}{2}z^2) = 1 + \tfrac{1}{2}z + \tfrac{1}{2}z^2 + \tfrac{1}{4}z^3 + \tfrac{1}{4}z^4 + \cdots$$

serves as a counter example for it and its second order transform both converge to 3 for $z = 1$. But as already seen in Theorem II and (43), $e_1(A_n)$ does not converge.

4. From a naive point of view a divergent sequence is sometimes said to be "meaningless". We have attempted to give (some) divergent sequences greater dignity and *intuitive* as well as *quantitative* meaning by the concept of the anti-limit. In Theorem IX, a specific type of sequence, namely the sequence of partial sums of the power series of a rational function, is seen to have a particular simple "meaning". The power series is obtained by dividing out the rational function. The e_k transforms merely invert this operation and obtain the rational function from the power series. In both the direct and inverse processes, the convergence or divergence of the series is entirely irrelevant.

5. If λ is not an eigenvalue, the Fredholm equation [13]:

$$u(x) = f(x) + \lambda \int_a^b k(x, y)u(y)\, dy$$

has a unique solution,

$$u(x) = f(x) + \int_a^b R(x, y, \lambda) f(y)\, dy,$$

where the resolvent $R(x, y, \lambda)$ is a meromorphic function of λ whose power series is the Neumann series of iterated kernels, $\sum_{i=1}^{\infty} \lambda^i k^{(i)}(x, y)$. Since the resolvent is a meromorphic function—and sometimes a rational function—it is clear that for any fixed λ (not an eigenvalue) we can chose a p such that

$$\lim_{n\to\infty} e_p[\sum_{i=1}^{n} \lambda^i k^{(i)}(x, y)] = R(x, y, \lambda).$$

Further it can be seen that there are infinitely many valid choices for p.

6. For *the partial sums of power series*, the e_k transforms may be found in the rows of the Padé Table. But (2) is more general than (75) and the e_k transforms are more general than the Padé Table since they may be applied to other types of sequences. One such application was given in Theorem V and in [20d] an interesting example is given of the application of e_d to a sequence of partial products. Further the transforms have been iterated and combined with each other. For example in [20, p 35] e_d and then $\tilde{e}_1$ are successively and successfully brought to bear upon the wildly divergent sequence, (60).

7. Through Theorem VI and the known relations between the Padé Table and continued fractions, a power series such as (76) is related to a continued fraction such as (80) via the e_d transform. An alternative way of obtaining the continued fraction from the power series is by the method of interpolation known as Thiele's reciprocal differences [15]. The fraction thus obtained will have as its convergents a sequence of rational approximations. These may be expressed as the ratio of two determinants and this has been done by Nörlund [16]. His determinants [16, (35) and (36)] are special cases of our (75) for $k = n - 1$ and $k = n$.

8. Theorems VII through IX do not cover the sequence (77), nor its convergence to $\log(1 + z)$, for here we have a branch point and not a pole as a singularity. We will proceed to discuss (77) and the effect of branch points, but it is convenient to first discuss (briefly) the spectra of transients.

Chapter III. Introduction to the spectra of sequences. So far in our discussion of mathematical transients,

$$A_n = B + \sum_{i=1}^{k} a_i q_i^n, \qquad (q_i \neq 1, 0) \quad (17)$$

we have largely confined ourselves to the computation of the base B or of the $\{B_n\}$ sequence. But in equations (39), (45), and (72) we had occasion to exhibit transients explicity including the spectrum of amplitudes a_i and ratios q_i. These simple spectra have a finite order k.

A second type of spectrum is that of the sequence of equation (34). Since this sequence converges to $\sqrt{2}$ and since each A_n (see (31)) is a rational number, it is clear that $\{A_n\}$ cannot be of finite order. For if it were of order N, the limit of A_n, $\sqrt{2}$, could be obtained exactly by e_N. But e_N involves *only* rational

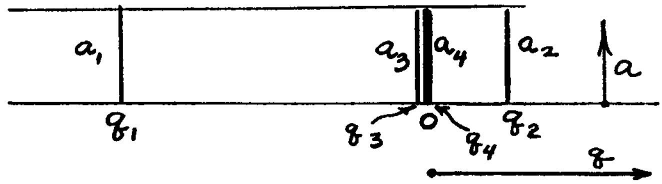

FIG. 2

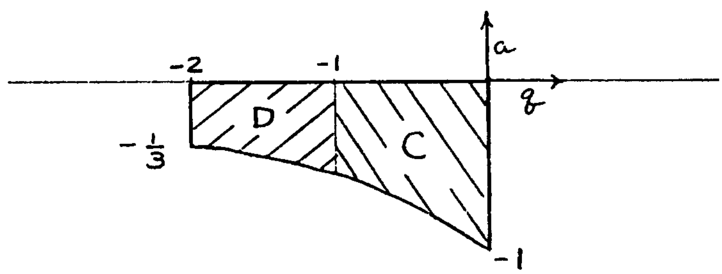

FIG. 3

operations and thus $\sqrt{2}$ would be rational. That the order of $\{A_n\}$ is infinite is also obvious by dividing out the fraction in (34). The explicit transient then reads:

$$A_n = \sqrt{2} + \sum_{i=1}^{\infty} 2\sqrt{2}([2\sqrt{2} - 3]^i)^n. \tag{103}$$

Here $B = \sqrt{2}$, $a_i = 2\sqrt{2}$ for all i, and $q_1 = -.1716$, $q_2 = +.0294$, $q_3 = -.005$, etc. We may sketch the spectrum (see Fig. 2).

Besides these line spectra, there exist continuous spectra. For example consider the sequence of equation (76). Here $A_0 = 0$ and

$$A_n = z - \tfrac{1}{2}z^2 + \tfrac{1}{3}z^3 - \cdots + (-1)^{n+1}z^n/n \text{ for } n > 0.$$

Therefore

$$A_n = \int_0^z \frac{1 - (-u)^n}{1 + u}\, du = \log(1 + z) + \int_{-z}^{0} \frac{dq}{q - 1}\, q^n. \tag{104}$$

Here $B = \log(1 + z)$ and the a's and q's range continuously from $-1/(1 + z)$ to -1 and from $-z$ to 0 respectively. If $z = 2$, we have $B = \log 3$ and the continuous spectrum (see Fig. 3), containing convergent components C and divergent components D.

An important problem concerning spectra is this. Given a sequence $\{A_n\}$, what is the spectrum of $e_k\{A_n\}$ with k arbitrary? In Theorem IV we have already noted the characteristic behaviour—some spectrum lines are annihilated—some change amplitude—and some new ones are introduced. We will return to this problem.

e_d and Gauss Numerical Integration. Equation (104) may also be derived by a longer, but more informative method. For each k there are $2k + 1$ corresponding polynomials, $A_0, A_1, \cdots A_{2k}$ of the sequence of (76). These can be fitted (locally) by a k'th order transient, (18). These transients, T_k, read:

$$(n = 0) \qquad T_0: A_n = 0$$

$$(n = 0, 1, 2) \qquad T_1: A_n = \frac{2z}{2+z} - \frac{2z}{2+z}\left(-\tfrac{1}{2}z\right)^n$$

$$(n = 0, 1, 2, 3, 4)\ T_2: A_n = \frac{6z + 3z^2}{6 + 6z + z^2} - \frac{3z}{6 + (3 - \sqrt{3})z}\left(\frac{\sqrt{3} - 3}{6}z\right)^n - \frac{3z}{6 + (3 + \sqrt{3})z}\left(\frac{-\sqrt{3} - 3}{6}z\right)^n \tag{105}$$

etc.

As $k \to \infty$ we should expect T_k to approach the right side of (104).

The three listed spectra were computed by solving equations of the type (19) and (20) but the computation of a high order transient is clearly a tedious job. Fortunately, the solutions are already known in terms of the roots of the Legendre Polynomials. The same equations are met with in the theory of Gaussian Integration [19].

In this theory one first reduces $\int_a^b F(x)\,dx$ to $\int_0^1 f(u)\,du$ by a change of variable $u = (x - a)/(b - a)$. One now wishes to approximate the integral by the sum:

$$G_k = \textstyle\sum_{i=1}^k R_i f(u_i) \tag{106}$$

and to determine the weights, R_i, and the ordinates, $f(u_i)$, such that the approximation is exact if $f(u)$ is any polynomial in u of degree $2k - 1$ or less. This leads one to the conditions:

$$1/(m + 1) = \textstyle\sum_{i=1}^k R_i u_i^m . \qquad (m = 0, 1, \cdots 2k - 1) \tag{107}$$

The solution of (107) yields the Gaussian weights and abscissae. This brings us to

THEOREM X. Let G_k be the k'th Gaussian approximant to $\log(1 + z) = \int_0^z dq/1 + q$ and let $B_{k,k}$ be given by (77), the diagonal transform of the power series of $\log(1 + z)$. Then

$$B_{k,k} = G_k . \tag{108}$$

PROOF. We multiply each equation in (107) of index m by $(-1)^m \cdot z^{m+1}$ and add the first n equations. This gives

$$z - \tfrac{1}{2}z^2 + \tfrac{1}{3}z^3 - \cdots + (-1)^{n+1}z^n/n = \textstyle\sum_{i=1}^k zR_i \sum_{m=0}^{n-1} (-zu_i)^m,$$

or summing both sides,

$$(n = 1, 2, \cdots 2k) \qquad A_n = \sum_{i=1}^k \frac{R_i z}{1 + zu_i} - \sum_{i=1}^k \frac{R_i z}{1 + zu_i}(-zu_i)^n. \tag{109}$$

Since both sides of (109) vanish for $n = 0$, (109) is also valid for $n = 0$ and therefore the right side is the explicit k'th order transient, T_k, of (105). Thus

$$B_{k,k} = \sum_{i=1}^{k} R_i z/(1 + zu_i). \tag{110}$$

But $\int_0^z dq/(1+q) = \int_0^1 z\,du/(1+zu)$ and therefore by (106)

$$G_k = \sum_{i=1}^{k} R_i z/(1 + zu_i) = B_{k,k}. \tag{111}$$

THEOREM XI. The sequence of (77) converges to the principal value of log $(1 + z)$ if z is not on the real cut $z \leqq -1$.

PROOF. It is known that max $|u_i - u_{i+1}| \to 0$ and hence that G_k will converge to the Riemann integral $R = \int_0^1 z\,du/(1 + zu)$ provided the integrand is continuous. Therefore if $1 + zu$ never vanishes as the real number u varies between 0 and 1, $G_k \to R$. Therefore if z is not on the real cut $z \leqq -1$, $B_{k,k} \to R$ by (111).

The shadows of branch points. Let z be taken as -2 in Theorem XI. Since it is on the real cut the theorem does not say whether or not the diagonal transform of

$$0 - 2 - \tfrac{1}{2}2^2 - \tfrac{1}{3}2^3 - \cdots$$

converges to log (-1). But it is obvious that it does not. For A_n is real and therefore $B_{n,n}$ is real and cannot converge to a pure imaginary. Alternatively, by symmetry and the Principle of Sufficient Reason, $B_{n,n}$ cannot converge to πi anymore than to $-\pi i$. Dr. Max Munk pointed this out to the author.

From the point of view of the Riemann surface, the branch point of log w is at $w = 0$ but the cut is arbitrary and could be made in any direction. Why should e_d discriminate against this particular (real) cut? The answer is that it does not. If we expand the logarithm function around $w = w_0$ instead of around $w = 1$ (i.e. $w = w_0 + z$ instead of $w = 1 + z$) we easily find that the diagonal transform of the *new* power series now converges to log $(w_0 + z)$ everywhere except along the infinite straight cut from the branch point *directly away* from w_0. As w_0 is moved, the cut moves.

We may describe this situation graphically by saying that $e_d(A_n)$ converges everywhere except in "*the shadow of the branch point*". This name for the appropriate cut was coined by a former colleague of the author, P. W. Zettler-Seidel. This result, convergence everywhere except in the shadow of the branch point, appears to be more general than is proven here—applicable to other multivalued functions and to $\tilde{e}_1$ as well as to e_d. Of $\tilde{e}_1$, nothing has been proven in this connection. But the reader is referred to Hardy's Puzzle in the appendix for an example.

The transformation of spectra and semi-convergence. Returning to (105) and its extension (109) we shall sketch a limiting process whereby

$$\lim_{k\to\infty} T_k = \log(1+z) + \int_{-z}^{0} \frac{dq}{q-1} q^n.$$

It is known [19, p 134] that $u_j = \frac{1}{2}(1 + t_j)$ where the t_j are the k roots of the Legendre Polynomial:

$$P_k(t) = \frac{1}{2^k k!} \frac{d^k}{dt^k} (t^2 - 1)^k.$$

For k large we have asymptotically:

$$t_j \sim -\cos\left(\frac{\pi j}{k+1}\right), \qquad R_j \sim \frac{\pi}{2(k+1)} \sin\left(\frac{\pi j}{k+1}\right)$$

from the inequalities of Markoff and Stieltjes [24]. Thus (109) becomes, for large k;

$$A_n \sim \sum_{j=1}^{k} \frac{\pi}{k+1} \frac{\sin[\pi j/k+1]}{(1+2/z) - \cos[\pi j/(k+1]} \cdot \left\{1 - z^n \left[\frac{\cos[\pi j/(k+1)] - 1}{2}\right]^n\right\}. \tag{112}$$

As $k \to \infty$ this becomes the Riemann integral

$$A_n = \int_0^{\pi} \frac{\sin v\, dv}{(1+2/z) - \cos v} \left[1 - z^n \left(\frac{\cos v - 1}{2}\right)^n\right] \tag{113}$$

and if $z \cos v = z - 2u$ this reduces to (104):

$$A_n = \int_0^{z} \frac{1 - (-u)^n}{1+u} du.$$

Upon returning to this continuous spectrum we return also to the previously posed problem—what is the spectrum of $e_k\{A_n\}$ with k an arbitrary integer and $\{A_n\}$ an arbitrary mathematical transient? We cannot give a complete treatment of the general problem—we will discuss $k = 1$ briefly and emphasize the different responses to e_1 of a continuous spectrum such as (104a) and a discrete spectrum such as (103a).

Considering the latter spectrum first we recall that in (103):

$$a_i = 2\sqrt{2}, \qquad q_i = (2\sqrt{2} - 3)^i. \qquad (i = 1, 2, 3, \cdots) \tag{114}$$

It follows from (103) and (37) that

$$e_1^m(A_n) = A_{2^m \cdot n} = \sqrt{2} + \textstyle\sum_{j=1}^{\infty} 2\sqrt{2}[(2\sqrt{2} - 3)^j]^{2^m \cdot n}.$$

Thus if $i = (2M + 1)2^m$ with M an integer, (and therefore i divisible by 2 exactly m times), the i'th spectrum line of $\{A_n\}$ remains unaltered in amplitude until e_1 has been applied $m + 1$ times. It then disappears abruptly. Of this interesting behaviour, (compare the Franck-Hertz experiment [17]), special attention may be directed to two features. 1.) At each transformation the most slowly convergent of the remaining components is annihilated. 2.) After the first transformation, $\{B_n\}$ has only even order q's left and therefore, from (114), only positive q's. Although $\{A_n\}$ is an oscillatory sequence, it follows that $\{B_n\}$, and all later sequences, are monotonic.

Somewhat more generally, let

with

$$\left.\begin{aligned} A_n &= B + \textstyle\sum_{i=1}^k a_i q_i^n && (q_i \neq 1, 0) \\ |q_1| &> |q_2| > |q_i|. && (i > 2) \end{aligned}\right\} \quad (115)$$

Then

$$B_n = B + \frac{\sum\sum_{i>j} a_i a_j (q_i - q_j)^2 (q_i q_j)^{n-1}}{\sum_{i=1}^k a_i (q_i - 1)^2 q_i^{n-1}}$$

or

$$B_n = B + \frac{\displaystyle\sum\sum_{i>j} \frac{a_i a_j q_1}{a_1 q_i q_j} \left(\frac{q_i - q_j}{q_1 - 1}\right)^2 \left(\frac{q_i q_j}{q_1}\right)^n}{\displaystyle 1 + \sum_{i=2}^k \frac{a_i q_1}{a_1 q_i} \left(\frac{q_i - 1}{q_1 - 1}\right)^2 \left(\frac{q_i}{q_1}\right)^n}. \quad (116)$$

It follows that

$$B_n = B + \textstyle\sum_{i=1}^\infty b_i r_i^n$$

with

and

$$\left.\begin{aligned} b_1 &= \frac{a_2}{q_2}\left(\frac{q_1 - q_2}{q_1 - 1}\right)^2 \\ q_2 &= r_1, \qquad |q_2| > |r_i|. && (i > 1) \end{aligned}\right\} \quad (117)$$

Therefore if A_n is given by (115), we see from (117) and (116) that

1. the spectrum of $\{B_n\}$ has a non-zero component in q_2 but the component in q_1 (which was the most rapidly divergent or most slowly convergent component of $\{A_n\}$) has vanished.

2. if $\{A_n\}$ has a "*purely positive spectrum*"—that is if all q's are real and positive—then $\{B_n\}$ has a purely positive spectrum.

3. if $\{A_n\}$ has a "*purely negative spectrum*" then $\{B_n\}$ has a purely negative spectrum.

These results suggest strongly that for the continuous and purely negative spectrum (104a) of the log 3 sequence, the spectrum of $\{B_n\}$ should again be purely negative and again have a continuous band of q's ranging from -2 to 0. For now "q_2" the "next largest" q to $q_1 = -2$, is infinitesimally close. This implies

$$\lim_{n\to\infty} \frac{\Delta B_{n+1}}{\Delta B_n} = -2$$

and thus that B_n (and similarly C_n, D_n, etc.) should (ultimately) diverge at the same rate that $(-2)^n$ does.

Now in fact a computation similar to (104) does show that while

$$A_n = \log 3 + \int_{-2}^{0} a_0(q)q^n \, dq, \qquad B_n = \log 3 + \int_{-2}^{0} a_1(q)q^n \, dq$$

with

$$a_0(q) = \frac{1}{q-1}, \qquad a_1(q) = \frac{1}{3}\left(\frac{q}{-2}\right)^{-2/3} + \frac{1}{q-1}. \tag{118}$$

A more extensive computation gives

$$C_n = \log 3 + \int_{-2}^{0} a_2(q)q^n \, dq$$

with

$$\left.\begin{aligned} a_2(q) &= \frac{1}{6}\left(\frac{q}{-2}\right)^{-2/3} + \frac{1}{6}\left(\frac{q}{-2}\right)^{-5/6} f(q) + \frac{1}{q-1} \\ &\text{where} \\ f(q) &= \cos\left(\frac{\sqrt{15}}{6}\ln\frac{q}{-2}\right) + \frac{1}{\sqrt{15}}\sin\left(\frac{\sqrt{15}}{6}\ln\frac{q}{-2}\right). \end{aligned}\right\} \tag{119}$$

Since $a_1(q)$ increases monotonically from $a_1(-2) = 0$ to $a_1(-1) = 0.0269$ to $a_1(-\frac{1}{2}) = 0.1733$ to $a_1(0) = \infty$ it is seen that the divergent components of $\{A_n\}$ have been greatly reduced in amplitude at the expense of the rapidly convergent components (see Fig. 4).

Similarly in $\{C_n\}$, since $a_2(-2) = 0$ and $a_2(q)$ decreases monotonically until $a_2(-1) = -0.000873$*, the divergent components are further reduced. If a_1 and a_2 are expanded around $q = -2$ it may be seen that

$$a_1(-2+\varepsilon) = \frac{\varepsilon^2}{108} + \cdots \quad \text{and} \quad a_2(-2+\varepsilon) = \frac{-\varepsilon^4}{7776} - \cdots \tag{120}$$

(More generally for the log $(1 + z)$ sequence, one finds:

$$a_1(-z+\varepsilon) = \frac{\varepsilon^2}{2z(z+1)^3} + \cdots \quad \text{and} \quad a_2(-z+\varepsilon) = \frac{-\varepsilon^4}{8z^2(z+1)^5} - \cdots.)$$

These very small amplitudes for the divergent components have the result that for n not too large the convergent components dominate the sequence. The

* For q nearly zero, $a_2(q)$ oscillates due to the trigonometric terms—see (119)—but this does not affect our conclusions. Further it may be shown that in the general case, log $(1 + z)$ instead of log 3, the trigonometric terms vanish for real z greater than $5 + 2\sqrt{6}$. In particular—for numerical reasons—$a_2(q)$ becomes particularly simple for $z = 10$ and 12. For the former we have:

$$a_2(q) = \frac{1}{22}\left[\left(\frac{q}{-10}\right)^{-10/11} + 5\left(\frac{q}{-10}\right)^{-14/11} - 4\left(\frac{q}{-10}\right)^{-15/11}\right] + \frac{1}{q-1}.$$

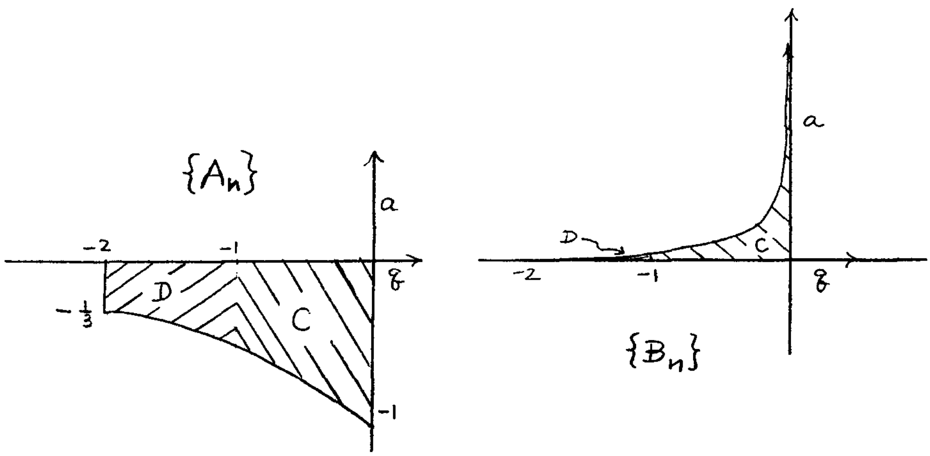

FIG. 4

old name "*semi-convergent*" [12, p 522, 541] is thus entirely appropriate for such a "temporarily converging" sequence with hidden (small amplitude) divergencies. A table (7) of $e_1^m(A_n)$ with A_n given by the log 3 sequence of (59) shows this semi-convergence very well. The asterisks indicate the points of closest approach to log 3. We may consider that the divergent components become dominant at these values of n.

n	A_n	B_n	C_n	D_n	E_n	F_n
0	0					
1	*2.0000000	1.0000000				
2	0.0000000	1.1428571	1.0931677			
3	2.6666667	1.0666667	1.1007092	1.0984266		
4	−1.3333333	*1.1282051	1.0974359	1.0986841	1.0986080	
5	5.0666667	1.0666667	1.0994536	1.0985761	1.0986141	1.0986122
6	−5.6000000	1.1368421	1.0989008	1.0986346	1.0986114	1.0986123
7	12.6857143	1.0493507	*1.0992921	1.0985862	1.0986128	
8	−19.3142857	1.1657143	1.0978997	1.0986254		
9	37.5746032	1.0031746	1.0994152			
10	−64.8253968	1.2391193				
11	121.3564214					

(121)

For comparison log 3 $= 1.098612289\cdots$. Therefore 12 terms of the semi-convergent sequence (59) yield 8 figure accuracy even though $A_{11} = 121.3564214$ has no accuracy whatsoever.

The graphs of transients and the Padé surface. A final representation of semi-convergence and of these successive sequences is obtained from our heuristic starting point—the graphs of sequences. Now we must admit that these graphs, although they served their purpose, are sometimes in need of improvement. For if the transient possesses a negative q, $q_1 = -v$ then the component

$$a_1(-v)^n$$

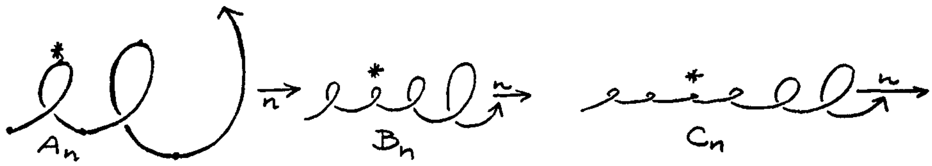

FIG. 5

is in general a complex number for non-integer n. Therefore we would do well to plot the real and imaginary parts of A_n versus n in a three dimensional Cartesian space.

Three dimensional graphs may thus be sketched for $\{A_n\}$, the log 3 sequence, and for its transforms (see Fig. 5). Finally the B coil should be inserted into the A coil, the C into the B, etc. in such positions that all the n axes coincide.

With this composite picture in mind one interprets e_1 geometrically as a transformation of one spiral into the next and may well ask whether there exists a continuous transformation T_x dependent upon a *continuous* parameter x which takes the m'th spiral into an $m + x$'th interpolating spiral for x not necessarily an integer. If so we would have an interpretation of e_1^x with x not necessarily an integer. Again, assume the existence of a "*Padé Surface*"—that is a continuous function $B(x, y)$ which interpolates the discrete points of an array (9)—and we would have an interpretation of the e_x transform with x not necessarily an integer. But to date these generalizations are merely speculative.

Appendix

Several miscellaneous aspects of the e_k transforms are discussed here. Each aspect is of interest but would have been digressive if given earlier.

The transforms and number theory. A.) An application of e_d to Euler's Partition Function [9] was given in [20d]. The coefficient of x^n in

$$f(x) = [\textstyle\prod_{i=1}^{\infty} (1 - x^i)]^{-1} \tag{122}$$

is P_n, the number of possible partitions of n. The partial product A_n :

$$A_0(x) = 1, \qquad A_n = [\textstyle\prod_{i=1}^{n} (1 - x^i)]^{-1}$$

has a power series which agrees with $f(x)$ to the term in x^n and thus yields the first n partition numbers correctly. But it is known [20*d*] that

$$e_d(A_n) = [1 + \textstyle\sum_{i=1}^{n} (-1)^i \{x^{i(3i-1)/2} + x^{i(3i+1)/2}\}^{-1}. \tag{123}$$

Since an identity of Euler reads:

$$\textstyle\prod_{i=1}^{\infty} (1 - x^i) = 1 + \sum_{i=1}^{\infty} (-1)^i \{x^{i(3i-1)/2} + x^{i(3i+1)/2}\} \tag{123a}$$

it follows that the power series of $e_d(A_n)$ agrees with that of $f(x)$ to the term in $x^{n(3n+5)/2}$. This, then, is another instance where our transforms give the identical result as some well-known special device.

We recall that $e_d(A_n)$ is a function only of the partial products A_0 to A_{2n}. If $n = 8$, the last partial product, A_{16}, yields the first 16 partition numbers correctly; but $e_d(A_8)$ yields the first 116 correctly. This result outdoes that of equations (79). There we "predicted" the next few coefficients of a power series approximately—here we predict the next 100 coefficients exactly.

B.) Similarly consider an application of e_1^m to a sequence akin to the Eratosthenes Sieve. Let the decimal number A_n $(n = 1, 2, \cdots)$:

$$A_1 = .11111\ 11111\ 11111\cdots, \qquad A_2 = .12121\ 21212\ 12121\cdots,$$

$$A_3 = .12221\ 31222\ 13122\cdots, \qquad A_4 = .12231\ 31322\ 14122\cdots,$$

$$A_5 = .12232\ 31323\ 14123\cdots, \qquad A_6 = .12232\ 41323\ 15123\cdots$$

represent the number of pebbles in contiguous boxes after n stages of the following operation. First a pebble is dropped into every box, then a pebble is dropped into every second box, then into every third box, etc. It is seen that the ultimate population in box n is $d(n)$, the number of divisors of n. The number A_n yields $d(n)$ correctly* up to but not beyond the first n integers and thus A_6 is correct only to $d(6) = 4$.

But consider the transforms B_n and C_n:

n	B_n	C_n
2	.12232 32323 23232··	
3	.12232 42333 24233··	.12232 42434 26161··
4	.12232 42433 25233··	.12232 42434 26244··
5	.12232 42434 25234··	

and note that C_4 is already correct in all 15 decimal places. Thus we read: $d(7) = d(11) = d(13) = 2$ and therefore 7, 11, and 13 are primes; $d(9) = 3$ and therefore 9 is the square of a prime, etc. This seems incredible—at first—but what we are really doing is summing the very smooth Lambert series [12, p 451].

$$\sum_{n=1}^{\infty} \frac{x^n}{1 - x^n} = \sum_{n=1}^{\infty} d(n)x^n \tag{124}$$

for the argument $x = \frac{1}{10}$.

The transforms and the detection of errors. Three types of errors have been detected by these nonlinear transforms.

A.) *Errors in Numerical Sequences.* In [27] Wentworth and Smith give the sequence of perimeters of regular $6 \cdot 2^n$-sided polygons inscribed in the unit circle. As is known, this sequence, $\{A_n\}$, converges to 2π. Upon transformation of $\{A_n\}$ by e_1^2 it was found that $\{C_n\}$ was noticeably not smooth. This suggests an error in $\{A_n\}$ and indeed this was the case [20, p 25]. Generally speaking

* For $n = 48$, $d(n) = 10$ and for $n = 72$, $d(n) = 12$, etc. Therefore the numbers A_n cannot be correctly written to the base 10 if we wish n to be carried beyond 47. However we could use the base 100 ($x = 1/100$ in (124)) and this would be satisfactory until $d(n) > 99$.

$\tilde{e}_1^m$ is sensitive to small errors and, like a difference table, is a means of detecting them.

B.) *Errors in Power Series.* In [7] Goldstein gives the series

$$K_d = \frac{12}{R}[1 + a_1 R - a_2 R^2 + a_3 R^3 - a_4 R^4 + a_5 R^5 -] \tag{125}$$

with

$$a_1 = 3/16, \qquad a_2 = 19/1{,}280, \qquad a_3 = 71/20{,}480$$

$$a_4 = 30{,}179/3{,}440{,}640, \qquad a_5 = 122{,}519/560{,}742{,}400,$$

for the drag on a sphere in a viscous fluid as a function of R, the Reynolds Number. But the series is not suitable for computation (at least without a summation method) for $R > 2$. It appears to have a radius of convergence of approximately $R = 4$ and so Goldstein also gives a table of K_d versus R obtained by a different method. If $\{A_n\}$ is the sequence of partial sums of the series in square brackets, the rational approximation:

$$\frac{12}{R}[B_{2,2}(R)] = \frac{12}{R}\left[\frac{295{,}680 + 133{,}200R + 10{,}880R^2}{295{,}680 + 77{,}760R + 689R^2}\right] \tag{126}$$

obtained from the first *five* terms of (125), is found to agree very well with Goldstein's table of K_d out to $R = 20$. But $(12/R)[B_{2,3}(R)]$ which utilizes the *last* term also does not agree as well and this suggests that this last term is in error. In fact it should read $a_5 = 122{,}519/550{,}502{,}400$ [20, p 33]. Errors in higher order terms of such power series can easily escape detection because where the series converges the term is negligible, and where the contribution of the term is appreciable the series diverges. But a serious attempt to use the divergent series may reveal the error.

C.) *A Typographical Error.* A typographical error detected with the aid of $\tilde{e}_1$ (in an apparently reckless manner) concerns a functional equation given by Titchmarsh in [25]:

$$\left(\frac{\pi}{2}\right)^s \Gamma(1 - s)\cos\left(\frac{s\pi}{2}\right) L(1 - s) = L(s) \tag{127}$$

where

$$L(s) = 1 - 3^{-s} + 5^{-s} - 7^{-s} + \cdots.$$

Now $L(2)$ is Catalan's constant, G. We wish to compute G from (127) and letting $s = -1$ we write tentatively:

$$\frac{2}{\pi}G \overset{?}{=} \frac{1 - 3 + 5 - 7 + \cdots}{\cos(-\pi/2)}.$$

But the denominator vanishes and the divergent series in the numerator may be easily summed to zero. Proceeding in this heuristic manner we write instead

$$\frac{2}{\pi} G \stackrel{?}{=} \left. \frac{(d/ds)L(s)}{(d/ds)\cos(s\pi/2)} \right|_{s=-1}$$

and therefore

$$G \stackrel{?}{=} 3 \ln 3 - 5 \ln 5 + 7 \ln 7 - \cdots . \tag{128}$$

(A convergent series is differentiated term by term only with caution—since the result may be divergent—but if the series is already divergent there seems to be no occasion to hesitate.) Summing the right side of (128) by $\tilde{e}_1$ we obtain from the first seven terms the approximation:

$$G \stackrel{?}{=} 0.5831 \cdots \tag{129}$$

But actually

$$G = 0.9159655942 \cdots$$

Where is the error?

Since from the true value of G we note:

$$(2/\pi)G = 0.5831218080 \cdots ,$$

equation (129) suggests that the exponent of $\pi/2$ in (127) should read $s - 1$ instead of s. This supposition is readily verified by putting $s = \frac{1}{2}$ in (127). Thus the error is not in $\tilde{e}_1$, nor in the heuristic manner, but is in equation (127).

However the real point of this example is not that this is a good technique for detecting typographical errors. It is instead the known fact—that judicious use of divergent series usually leads to correct results. We quote Abel "... Pour la plus grande partie, les résultats sont justes, il est vrai, mais c'est une chose bien étrange. Je m'occupe à en chercher la raison, problème très intéressant". [4, p 320]

Hardy's puzzle. In (64) we introduced Hardy's puzzle—should $x = 3/5$ or should $x = -3/5$? Let $\{A_n\}$ be the sequence of partial sums of the power series of $g(w) = (1 - w)^{-\frac{1}{2}}$. By similar arguments as were used previously for $\log(1 + z)$ —The Principle of Sufficient Reason or the reality of the sequence—it is clear that $\tilde{e}_1(A_n)$ cannot converge to either value of $g(w)$ in the shadow of the branch point—that is on the real cut $w > 1$. Therefore the right value for $g(w_0)$ with w_0 not in the shadow is determined by the *chosen* value of $g(0)$—that is, the first term of the power series—and by the corresponding sheet of the Riemann Surface cut by the shadow of the branch point.

Now (64a) is the binomial series for $[1 - (2z/1 + z^2)^2]^{-\frac{1}{2}}$. Under the mapping $(2z/1 + z^2)^2 = w$ both $z = 0$ and $z = \infty$ map into $w = 0$. The unit circle (Fig. 6) in the z plane maps into the real cut, $w \geqq 1$. Therefore we can find a path from $z = 2i$ to $z = \infty$, but *not* to $z = 0$, the map of which does not cross the real cut in the w plane. At ∞, (64a) sums to $f_2(z)$. Therefore by analytic continuation the right answer is $x = f_2(2i) = +\frac{3}{5}$. It may be noted that a short table (7) for the sequence (64):

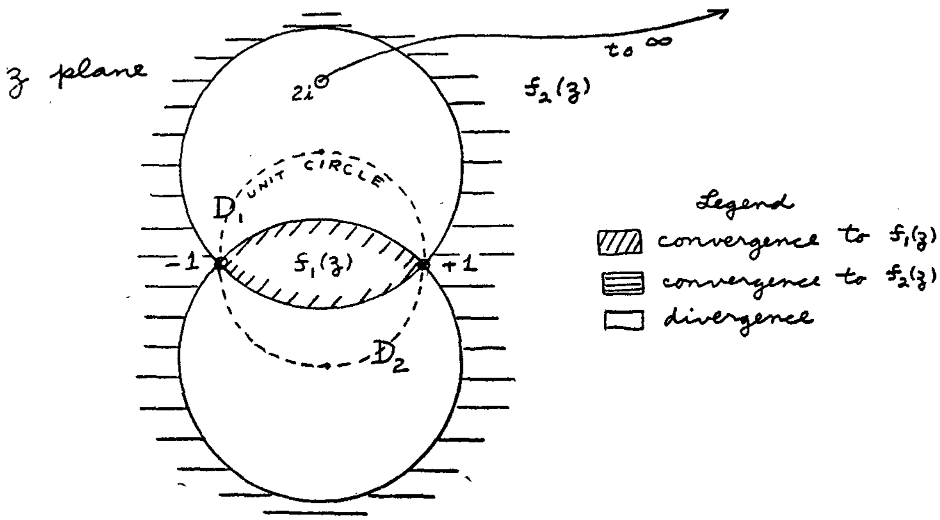

FIG. 6

n	A_n	B_n	C_n
0	1.000000		
1	0.111111	0.619047	
2	1.296296	0.588723	0.600957
3	−0.459534	0.609232	
4	2.271757		

makes it all but certain that $\tilde{e}_1(A_n)$ converges to the *right* answer.

Remarks on the history of e_k et al. No attempt has been made to give a complete history of these transforms but a few remarks may be useful.

In [1] A. C. Aitken used e_1 and e_1^m as sequence-to-sequence transformations for a very limited class of sequences—namely slowly convergent iterative sequences of the type (30). Both before and after [1] a number of authors independently found and used formula (3) or its equivalent. Usually these authors were concerned with a slowly convergent nearly-geometric series or with a slowly convergent iterative process.

Thus Delaunay [3], faced with complicated slowly convergent series for some lunar inequalities, notes that these series *appear* to be roughly geometric and he corrects for the uncomputed terms by assuming that they constitute a geometric series and summing. This is equivalent to applying (3) to the last three partial sums. In [23] Shanks and Walton do the same thing for other series in applied mathematics, some of which are convergent and some divergent.

Samuelson [18], Shanks and Walton [23], and others have noted that in the iterative solution of the equation

$$x = f(x)$$

the iterates often converge roughly geometrically and have thus deduced an approximation to A_∞,

$$\frac{A_{n+1}A_{n-1} - A_n^2}{A_{n+1} + A_{n-1} - 2A_n},$$

from the last three iterates. The following variation has been suggested more than once. Let ϕ be an iterative operator. Assume a value for A_0 and compute successively:

$$A_1 = \phi(A_0), \qquad A_2 = \phi(A_1), \qquad A_3 = e_1(A_1);$$

$$A_4 = \phi(A_3), \qquad A_5 = \phi(A_4), \qquad A_6 = e_1(A_4); \quad \text{etc.}$$

Even in this restricted use of e_1 all authors do not achieve the same generality. Thus Isakson [11] believes that (3) should be applied only to monotonic sequences whereas Samuelson allows even divergent sequences. Samuelson, still in connection with $x = f(x)$, suggests a generalization of (3) but he does not obtain our (2) since he assumes a special case of our (17) with $q_2 = q_1^2$, $q_3 = q_1^3$, etc.

In January 1949 Dr. M. Slawsky proposed a problem to the author: Given some radioactive decay data, $p(t)$, determine the constants k, a_i and q_i so that $p(t) \doteq \sum_{i=1}^{k} a_i q_i^t$ (in the sense of least squares). Consideration of this problem suggested the analogy between transients and sequences to the author and this in turn suggested e_k (and its relatives) as general sequence-to-sequence transformations [20], [20a, b, c, d]. In this connection it is of interest to contrast the *smooth* graphs of A_n versus n sketched on p. 6 with the jagged first six figures of Bromwich [4].

Two other transformations closely related to e_1 but not discussed in this paper are G, a semi-linear version of e_1:

$$G(A_n) = \frac{A_n - (\lim_{n\to\infty} \Delta A_n/\Delta A_{n-1})A_{n-1}}{1 - (\lim_{n\to\infty} \Delta A_n/\Delta A_{n-1})} \tag{130}$$

and $e_1^{(s)}$, a generalization which reduces to e_1 if s is taken to be ∞:

$$e_1^{(s)}(A_n) = \frac{sB_n - A_n}{s - 1} \tag{131}$$

with

$$s = \lim_{n\to\infty} \frac{\Delta A_n}{\Delta B_n}; \qquad B_n = e_1(A_n).$$

The transformation G was mentioned in passing in [20, p 25, 26] under the name "geometric extrapolation". The transformation $e_1^{(s)}$ was merely alluded to in [20, paragraph 76, A] but had been treated at some length in an earlier, longer version of [20] which was circulated privately but not published. This transformation, $e_1^{(s)}$ and its iteration, was devised primarily as a supplement to e_1 and e_1^m since the latter are not efficient in the summation of some monotonic sequences where $\Delta A_n/\Delta A_{n-1} \to 1$. Examples of this type of sequence where $e_1^{(s)}$ is effective are the sequences of partial sums of the famous:

$$\frac{\pi^2}{6} = 1 + \tfrac{1}{2}^2 + \tfrac{1}{3}^2 + \tfrac{1}{4}^2 + \cdots \tag{132}$$

and of the more elaborate:

$$\ln 2 \ln K = \ln \tfrac{2}{1} \ln \tfrac{3}{2} + \ln \tfrac{3}{2} \ln \tfrac{4}{3} + \ln \tfrac{4}{3} \ln \tfrac{5}{4} + \cdots . \tag{133}$$

In (133) K is Khintchine's constant [22].

In connection with Theorem I and (28), (29) we have already referred to Lubkin's paper [14]. There he proves theorems on the applicability, to certain classes of numerical series, of the e_1 transformation (which he calls T), of the G transformation, (130), (which he calls the Ratio transformation) and of a more complicated version of $e_1^{(s)}$, (131), (which he calls W). He also gives examples, but a number of these had already been discussed at length in [20a] and [20].

In [6], [6a] Forsythe refers to e_1 as the δ^2 process and to e_k as the *generalized* δ^2 process. Now the name "δ^2 process" was given by Aitken and thus has priority over the name "e_1 transformation" given by the author. Nonetheless, to refer to e_2, e_3, e_4, etc. all as the generalized δ^2 process is clearly ambiguous. Further we need names for e_1^m, $\tilde{e}_1$, and e_d. Still further, since we apply e_k et al. to sequences in general rather than to iterative processes only, the usual word transformation seems preferable to the word process. For these, and still other reasons, the author believes that the "e_k transform" terminology has much to recommend it.

The literature on non-linear transforms is not very large. In contrast, as is known, the literature on linear transforms, Cesàro, Hölder, Toeplitz, etc. is enormous [12, pp 464–477]. Yet if we wish to transform *both* convergent and divergent sequences in a *uniform* manner—to accelerate the convergence of the former and to sum the latter, the linear transforms have a serious handicap. These transforms may be expressed as $\sum_{i=0}^{n} c_{i,n} A_i / \sum_{i=0}^{n} c_{i,n}$ where the $c_{i,n}$'s are preassigned weights which are independent of the A's. Now if $\{A_n\}$ is convergent it is generally effective if the later A's are weighted more heavily whereas if $\{A_n\}$ diverges it is generally more effective if the early A's are weighted more heavily. Clearly no set of preassigned weights can do both. That e_k can and does often do both is seen in (75). For z large the early A's are weighted heavily and for z small, the later A's are weighted heavily.

Acknowledgement. I wish to express my sincere thanks to Professor G. S. S. Ludford for the many improvements which he suggested. I also wish to thank Mr. Ronald Smelt and many other colleagues at or formerly at the Naval Ordnance Laboratory for their friendly advice.

REFERENCES

1. Aitken, A. C. "On Bernoulli's Numerical Solution of Algebraic Equations." *Proc. Roy. Soc. Edinburgh* Vol. 46 (1926) p 289–305.
2. ——— "Studies in Practical Mathematics II" *Proc. Roy. Soc. Edinburgh* Vol. 57 (1937) p 269–304.
3. See Andoyer, H. *La Theorie de la Lune* Gauthier-Villars, Paris, (1926) p 59.
4. Bromwich, T. J. *An Introduction to the Theory of Infinite Series*, Macmillan, London, (1942).

5. See Cesàro, E. *Elementares Lehrbuch der Algebraischen Analysis und der Infinitesimalrechnung* B. G. Teubner, Leipzig, (1904) p 79.
6. Forsythe, G. E. "Solving Linear Equations Can Be Interesting" *Bull. Am. Math. Soc.* Vol. 59 no 4 (1953) p 309, 310.
6a. ——— "Tentative Classification of Methods and Bibliography on Solving Systems of Linear Equations" in *Simultaneous Linear Equations and the Determination of Eigenvalues* National Bur. Stand., Applied Math. Series 29 (1953) p 7, item 21b 2.
7. Goldstein, S. "Steady Flow of Viscous Fluid Past a Fixed Spherical Obstacle at Small Reynolds Number", *Proc. Roy. Soc.* A, Vol. 123 (1929) p 225.
8. Hardy, G. H. *Divergent Series* Oxford, London, (1949) p 16.
9. Hardy, G. H. and Wright, E. M. *An Introduction to the Theory of Numbers*, Oxford, London (1938) p 272–282.
10. Hartree, D. R. "Notes on Iterative Processes" *Cam. Phil. Soc.* Vol. 45 part 2 (Apr '49) p 230–236.
11. Isakson, Gabriel "A Method for Accelerating the Convergence of an Iteration Procedure" *Jour. Aero. Soc.* Vol 16, no 7 (Jul '49) p 443.
12. Knopp, K. *Theory and Application of Infinite Series* Blackie & Son, London (1928).
13. Lovitt, W. V. *Linear Integral Equations* McGraw Hill, New York (1924).
14. Lubkin, Samuel "A Method of Summing Infinite Series" *Jour. Research Nat. Bur. Stand.* Vol. 48 No. 3 March 1952 (dated 2 Nov '50).
15. Milne-Thomson, L. M. *The Calculus of Finite Differences*, Macmillan, London, Chapter V, (1933).
16. Nörlund, N. E. *Vorlesungen Uber Differenzenrechung* Springer, Berlin, (1924) p 432.
17. See Ruark, A. E. and Urey, H. C. *Atoms, Molecules and Quanta* McGraw Hill, New York, (1930) p 81.
18. Samuelson P. A. "A Convergent Iterative Process" *Jour. of Math. and Phys.* Vol. 24 (1945) p 131–134.
19. Scarborough, J. B. *Numerical Mathematical Analysis*, Johns Hopkins, Baltimore, (1930) p 133.
20. Shanks, Daniel *An Analogy Between Transients and Mathematical Sequences And Some Nonlinear Sequence-to-Sequence Transforms Suggested by It. Part I* Naval Ordnance Laboratory Memorandum 9994 (26 Jul '49) White Oak, Md.
20a. ——— "Mathematical Sequences treated as Transients" *Bull. Am. Math. Soc.* Abstract 55–7–366 (17 Mar '49).
20b. ——— "Non-Linear Transformations of Divergent and Slowly Convergent Sequences and an Example from Hydrodynamics" *Phys. Rev.* Vol. 76 No. 6 p 876 Abstract A3 (30 June '49).
20c. ——— "Non-Linear Transformations of Divergent and Slowly Convergent Sequences" *Am. Math. Mon.* Vol. 57, No. 6 p 439 Abstract 2 (10 Dec '49).
20d. ——— "Short Proof of an Identity of Euler" *Proc. Am. Math. Soc.* Vol. 2 No 5 pp 747–749 Oct 1951 (dated 6 Oct '50).
21. ——— "The Blasius and Weyl Constants in Boundary Layer Theory" *Phys. Rev.* Vol. 90 No. 2 (1953) p 377 Abstract U 11.
22. ——— "Note on an Absolute Constant of Khintchine" *Math. Tables and other Aids to Computers* (Jan 1950) p 28.
23. Shanks, D. and Walton, T. S. *The Use of Rational Functions as Approximate Solutions of Certain Trajectory Problems* Naval Ordnance Laboratory Memorandum 9524 (1948) White Oak, Md. p 2; p 13, equation (23).
24. Szegö, Gabriel "Inequalities for the Zeros of Legendre Polynomials and Related Functions" *Trans. Am. Math. Soc.* Vol. 39, No. 1, (Jan '36) p 1.
25. Titchmarsh, E. C. *Introduction to the Theory of Fourier Integrals*, Oxford, London, (1937) p 66.
26. Wall, H. S. *Continued Fractions*, D. von Nostrand, New York, Chapter XX (1948).

27. Wentworth, G. and Smith, D. E. *Plane and Solid Geometry*, Ginn, Boston, (1913) p 249.
28. Whittaker, E. T. and Robinson, G. *The Calculus of Observations*, Blackie & Son, London, (1944) p 79.

U. S. NAVAL ORDNANCE LABORATORY;
WHITE OAK, MARYLAND

(Received June 9, 1954)