REEL-C 3119 A

A.T.I 66568

		66568
An Analogy Between Transients and Mathematical Sequences and Some Nonlinear Sequence-to-Sequecne Transforms Suggested by It - Part I		(None)
	Shanks, Daniel	
	Naval Ordnance Lab., Silver Spring, Md.	NOLM 9994
	~~(Same)~~	(Same)

July' 49 Unclass. U.S. English 42 table

An analogy between sequences and transients has been introduced and developed. A uniform treatment for the evaluation of convergent and divergent sequences was developed on the basis of this analogy. Several nonlinear transforms were successfully applied to a large variety of sequences. The development of the theory is not completed, but some proofs are given and some connections with known algorithms are shown.

Copies of this report obtainable from CADO.

Sciences, General (33)
Mathematics (3)

Sequences, Mathematical

Naval Ordnance Laboratory
White Oak, Silver Spring 19, Maryland

NAVAL ORDNANCE LABORATORY MEMORANDUM 9994

26 July 1949

From: Daniel Shanks
To: NOL Files
Via: Chief, ~~Mechanics Division~~

Subj: An Analogy Between Transients and Mathematical Sequences and Some Nonlinear Sequence-to-Sequence Transforms Suggested by It. Part I. (Project NOL-4-Re9d-21-2)

Abst: In mathematics, and in applied mathematics especially, one wishes to obtain accurate answers rapidly. One obstacle often met with is that the simplest and most obvious analysis gives mathematical sequences which are slowly convergent or even divergent. The proper treatment of such sequences is therefore a general problem of real importance. This memorandum gives and discusses some methods of treating such sequences.

An analogy between mathematical sequences and the transients of linear systems is developed. Through each 2K + 1 consecutive values of the sequence A_n one passes a continuous function of the form $B + \sum_{i=1}^{k} a_i e^{\alpha_i n}$. The series of exponentials either converges to, diverges from, or is asymptotic to the constant, B. An explicit formula for B in terms of the A_n is given and this forms the basis of several nonlinear sequence-to-sequence transforms $A_n \to B_n$. The transforms are applied to a variety of convergent and divergent sequences. The complete theory is not given but some theorems are proven and some relations to the Padé Table, and to Thiele's Reciprocal Differences are discussed.

Fwrd: The data and conclusions presented here are the opinion of the author and do not necessarily represent the final judgment of the Laboratory.

Ref: (a) Whittaker, E. T. and Robinson, G. The Calculus of Observations, Blackie and Son, London, 369 (1944).
(b) Willers, F. A. Practical Analysis, Dover, New York, 355 (1948).

(c) Aitken, A. C. "On Bernoulli's Numerical Solution of Algebraic Equations," Proc. of the Roy. Soc. of Edinburgh, 46, 289-305 (1926)
(d) Shortley, G. H. and Weller, R. "Numerical Solution of Laplace's Equation," Jour. of Appl. Phys., 9, 343 (1938)
(e) Samuelson, P. A. "A Convergent Iterative Process," Jour. of Math. and Phys., 24, 131-134 (1945).
(f) Shanks, D. and Walton, T. S. The Use of Rational Functions as Approximate Solutions of Certain Trajectory Problems. NOLM 9524, 13 (1948).
(g) Milne-Thomson, L. M. The Calculus of Finite Differences, Macmillan, London, Chapter V, (1933).
(h) Wall, H. S. Continued Fractions, D. van Nostrand, New York, Chapter XX (1948).
~~(i) Dienes, P. The Taylor Series, Oxford, London, 321 (1931).~~
(j) Knopp, K. Theory and Application of Infinite Series. Blackie and Son, London. Chapter XIII (1928).
(k) Bromwich, T. J. An Introduction to the Theory of Infinite Series, Macmillan, London, 323 (1942).
(l) Moore, C. N. Summable Series and Convergence Factors, Am. Math. Soc., New York, 3 (1938).
(m) Wentworth, G. and Smith, D. E. Plane and Solid Geometry, Ginn, Boston, 249 (1913).
(n) Collatz, L. Eigenwertprobleme, Chelsea, New York, 137 (1948).
(o) Hardy, G. H. and Wright, E. M. An Introduction to the Theory of Numbers, Oxford, London, 272-282 (1938).
(p) Titchmarsh, E. C. Introduction to the Theory of Fourier Integrals, Oxford, London, 66 (1937).
(q) Lovitt, W. V. Linear Integral Equations, McGraw Hill, New York, 13 (1924).
(r) Goldstein, S. "Steady Flow of Viscous Fluid Past a Fixed Spherical Obstacle at Small Reynolds Number," Proc. Roy. Soc. A, 123, 225 (1929).
(s) Nörlund, N. E. Vorlesungen Über Differenzenrechnung, Springer, Berlin, 432 (1924).
(t) Kronecker, L. "Zur Theorie der Elimination einer Variabeln aus Zwei Algebraische Gleichungen," Berlin Berichte, 535-600 (1881).
(u) Encyklopädie der Mathematischen Wissenchaften II, 3, 1, p. 470 (1921).

TABLE OF CONTENTS

INTRODUCTION

1. In this paper we shall discuss an analogy between transients and mathematical sequences. By the term "physical transient" we mean a physical quantity, p, which, when expressed as a function of time, takes the form

$$p(t) = B + \sum_{i=1}^{K} a_i e^{\alpha_i t} \qquad (1)$$

It will appear below that it is useful to regard some mathematical sequences, A_n, as functions of n of the form

$$A_n = B + \sum_{i=1}^{K} a_i e^{\alpha_i n} \qquad (2)$$

and because of this we may call such sequences "mathematical transients."

2. We shall be concerned here with the analysis of mathematical transients. By the term "analysis" we mean the determination of the "amplitudes," a_i; the "frequencies," α_i; and the "base line constant," B, of these transients. If all the α_i have negative real parts, the transient converges to its limit, B. If one or more α_i has a nonnegative real part the transient is divergent and has no limit. In such cases we may call B the "antilimit" of the transient.

3. The analogy between transients and sequences is suggested by the graphs of some typical sequences in the (n, A_n) plane. Since, in general, the sequence is defined only at the integers, n, there is nothing to prevent us from drawing a smooth curve through these known discrete points.

4. If, then, the sequence converges and oscillates, the graph may look like

or perhaps

If it diverges it may look like

or

If it is asymptotic it may resemble

or

5. Generally, for sequences which arise naturally in analysis, the graph will <u>look like</u> a transient of a linear system of the form (1) and the idea naturally occurs to experiment with such forms, to treat the sequences <u>as if</u> they <u>were</u> transients, and to solve for the limit or antilimit B.

6. Suppose, for instance, we have 2K + 1 values of the sequence of partial sums of the slowly convergent series, $\ln 2 = 1 - 1/2 + 1/3 - 1/4 + \ldots.$ The graph of this sequence looks like the first graph above. It oscillates around and converges to ln 2. We can find a's, α's, and a B such that the resulting graph (2) would pass through these 2K + 1 pts. Intuitively, it would seem that the B should be a good approximation to ln 2.

7. Suppose we have 2K + 1 values of the sequence of partial sums of the divergent series, $\ln 3 = 2 - \frac{2^2}{2} + \frac{2^3}{3} - \frac{2^4}{4} + \ldots.$ The graph of this sequence looks like the fourth transient above. It does not <u>converge to</u> ln3 but it does <u>oscillate around</u> and <u>diverge from</u> ln3. And the corresponding B should be a good approximation to ln3. In this analogy the continuity between convergent and divergent sequences is similar to the continuity between stable and unstable transients. <u>This</u> continuity is a result of the continuity between α's with negative real parts (in (1)) and those with positive real parts. We will, therefore, take the same attitude toward divergent sequences as we take toward negative numbers. We accept them - at least tentatively. We will attempt to evaluate them by calculating the antilimit B. But to do this, we must have an algebra of transients.

8. This we will develop first. From this algebra we will obtain explicit (and relatively simple) formulae for the B's in terms of the A_n. On the basis of these formulae we will then develop some nonlinear sequence-to-sequence transforms which will convert the original A_n sequence into new B_n sequences. In fact we will have a variety of such transforms since (a) we may choose an arbitrary number, K, of exponentials and (b) we may then iterate the process or not.

9. If A_n is slowly convergent, we wish B_n to be more rapidly convergent to the same limit. If A_n is divergent, we wish B_n to be semi-convergent or, better still, convergent. In either case we are trying to filter out the exponential terms and to reduce the sequence to its static base, B.

10. We shall apply these transforms to a variety of mathematical sequences and discuss the results. We shall give some proofs of validity - but not a complete proof. We shall show the interrelations

between these transforms and some known algorithms.

I. The Algebra of Transients

11. The solution of a linear differential equation with constant coefficients, of order K, and with given initial conditions is well known and understood. The solution of the inverse problem, although it appears in the literature, is not as generally known. The problem is this: Given a tabulated function p(t), to find the unknown constants α_i, a_i, and B and the order K of a form (1) which will fit the given data. Related problems are the determination of the differential equation and the initial conditions and the extrapolation of the function p(t). Two examples of the inverse problem are:

(a) Analysis of a mixture of radioactively decaying substances. Given the radiation as a function of time, find the number of substances, the relative quantities, and the decay periods.

(b) Analysis of a portion of the "trajectory" of a device controlled by a linear servomechanism. Find the differential equation, extrapolate the trajectory, determine the stability.

12. Let us assume B = 0, K = 3 in (1) and that the data, p(t), is known exactly. (For least-square versions of the formulae which occur in the next two paragraphs see references a and b.) Given six values of the data:

$$p(t_n) \qquad (n = 0 \text{ to } 5)$$

we have six equations

$$p(t_n) = a_1 e^{\alpha_1 t_n} + a_2 e^{\alpha_2 t_n} + a_3 e^{\alpha_3 t_n} \qquad (n = 0 \text{ to } 5) \tag{3}$$

from which to evaluate the unknown a's and α's.

13. What appears to be a troublesome set of transcendental equations becomes quite simple if the t_n are equally spaced for if we take

$$t_n = 0, T, 2T, 3T, \text{etc.} \tag{4}$$

and

$$\left.\begin{aligned} e^{\alpha_i T} &= q_i \\ p(nT) &= p_n \end{aligned}\right\} \tag{5}$$

our equations:

$$p_n = \sum_{i=1}^{3} a_i \, q_i^n \qquad (n = 0 \text{ to } 5)$$

are seen to be algebraic. They are linear in the a's. If we consider the p's transferred to the other side, we obtain from these homogeneous equations, the conditions:

$$\begin{vmatrix} 1 & 1 & 1 & p_0 \\ q_1 & q_2 & q_3 & p_1 \\ q_1^2 & q_2^2 & q_3^2 & p_2 \\ q_1^3 & q_2^3 & q_3^3 & p_3 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & p_1 \\ q_1 & q_2 & q_3 & p_2 \\ q_1^2 & q_2^2 & q_3^2 & p_3 \\ q_1^3 & q_2^3 & q_3^3 & p_4 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & p_2 \\ q_1 & q_2 & q_3 & p_3 \\ q_1^2 & q_2^2 & q_3^2 & p_4 \\ q_1^3 & q_2^3 & q_3^3 & p_5 \end{vmatrix} = 0 \qquad (7)$$

For $q = q_1$, q_2, or q_3, we have the obvious:

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ q_1 & q_2 & q_3 & q_4 \\ q_1^2 & q_2^2 & q_3^2 & q_4^2 \\ q_1^3 & q_2^3 & q_3^3 & q_4^3 \end{vmatrix} = 0 \qquad (8)$$

Our four determinants form a set of four homogeneous equations in the common minors of the last columns. Therefore:

$$\begin{vmatrix} p_0 & p_1 & p_2 & 1 \\ p_1 & p_2 & p_3 & q \\ p_2 & p_3 & p_4 & q^2 \\ p_3 & p_4 & p_5 & q^3 \end{vmatrix} = 0 \qquad (9)$$

or

$$m_3 q^3 - m_2 q^2 + m_1 q - m_0 = 0 \qquad (10)$$

where m_i is the minor of q^i. Solving this cubic we obtain the three q's and the α's may be obtained by

$$\alpha_i = \frac{1}{T} \ln q_i \qquad (11)$$

Putting the q's back in (6) we may now find the a's.

14. If we had a seventh value, p_6, we would have:

$$\begin{vmatrix} p_0 & p_1 & p_2 & p_3 \\ p_1 & p_2 & p_3 & p_4 \\ p_2 & p_3 & p_4 & p_5 \\ p_3 & p_4 & p_5 & p_6 \end{vmatrix} = 0 \qquad (12)$$

This is a criterion to determine if K = 3. If the equation holds for all sets of seven consecutive p's then K = 3. If the determinant does <u>not</u> vanish, K $>$ 3 but if K does equal 3, (12) gives us the extrapolation formula:

$$p_6 m_3 = p_5 m_2 - p_4 m_1 + p_3 m_0 \qquad (13)$$

or more generally the recurrence formula

$$p_n m_3 = p_{n-1} m_2 - p_{n-2} m_1 + p_{n-3} m_0 \qquad (14)$$

On the other hand if all the m_i vanished, we would assume that K $<$ 3. The generalization of the formulae (3)-(14) for any K is obvious.

15. For <u>mathematical</u> transients, (2), the a's and α's are

of theoretical interest and they will be discussed in a later paper. But our present purpose is to evaluate the "limit" or "antilimit" (the base line constant) B.

16. Removing the restriction, B = 0, we have the seven equations for K = 3:

$$A_n = B + \sum_{i=1}^{3} a_i q_i^n \qquad (n = 0 \text{ to } 6) \tag{15}$$

and, therefore, from the first four:

$$B = \begin{vmatrix} A_0 & A_1 & A_2 & A_3 \\ 1 & q_1 & q_1^2 & q_1^3 \\ 1 & q_2 & q_2^2 & q_2^3 \\ 1 & q_3 & q_3^2 & q_3^3 \end{vmatrix} \Bigg/ \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & q_1 & q_1^2 & q_1^3 \\ 1 & q_2 & q_2^2 & q_2^3 \\ 1 & q_3 & q_3^2 & q_3^3 \end{vmatrix} \tag{16}$$

But since the q's are functions of the A's it should be possible to find B as a function of the A's only.

17. If we take

$$\Delta A_n = A_{n+1} - A_n \tag{17}$$

and

$$c_i = a_i (q_i - 1) \tag{18}$$

we have from the differences of equations (15), the six equations

$$\Delta A_n = \sum_{i=1}^{3} c_i q_i^n \qquad (n = 0 \text{ to } 5) \tag{19}$$

and, by comparison with (6), the two determinants:

$$\begin{vmatrix} 1 & q & q^2 & q^3 \\ 1 & q_1 & q_1^2 & q_1^3 \\ 1 & q_2 & q_2^2 & q_2^3 \\ 1 & q_3 & q_3^2 & q_3^3 \end{vmatrix} = \begin{vmatrix} 1 & q & q^2 & q^3 \\ \Delta A_0 & \Delta A_1 & \Delta A_2 & \Delta A_3 \\ \Delta A_1 & \Delta A_2 & \Delta A_3 & \Delta A_4 \\ \Delta A_2 & \Delta A_3 & \Delta A_4 & \Delta A_5 \end{vmatrix} = 0 \tag{20}$$

The ratios of the minors of <u>either</u> determinant are the symmetric functions on the roots, q_i, and, therefore, these ratios are equal to each other. That is

$$\frac{\begin{vmatrix} 1 & q_1 & q_1^3 \\ 1 & q_2 & q_2^3 \\ 1 & q_3 & q_3^3 \end{vmatrix}}{\begin{vmatrix} 1 & q_1 & q_1^2 \\ 1 & q_2 & q_2^2 \\ 1 & q_3 & q_3^2 \end{vmatrix}} = q_1 + q_2 + q_3 = \frac{\begin{vmatrix} \Delta A_0 & \Delta A_1 & \Delta A_3 \\ \Delta A_1 & \Delta A_2 & \Delta A_4 \\ \Delta A_2 & \Delta A_3 & \Delta A_5 \end{vmatrix}}{\begin{vmatrix} \Delta A_0 & \Delta A_1 & \Delta A_2 \\ \Delta A_1 & \Delta A_2 & \Delta A_3 \\ \Delta A_2 & \Delta A_3 & \Delta A_4 \end{vmatrix}} \tag{21}$$

and so forth. Therefore, (16) may be replaced by its equal:

$$B = \begin{vmatrix} A_0 & A_1 & A_2 & A_3 \\ \Delta A_0 & \Delta A_1 & \Delta A_2 & \Delta A_3 \\ \Delta A_1 & \Delta A_2 & \Delta A_3 & \Delta A_4 \\ \Delta A_2 & \Delta A_3 & \Delta A_4 & \Delta A_5 \end{vmatrix} \Bigg/ \begin{vmatrix} 1 & 1 & 1 & 1 \\ \Delta A_0 & \Delta A_1 & \Delta A_2 & \Delta A_3 \\ \Delta A_1 & \Delta A_2 & \Delta A_3 & \Delta A_4 \\ \Delta A_2 & \Delta A_3 & \Delta A_4 & \Delta A_5 \end{vmatrix} \tag{22}$$

Similar determinantal formulae are, of course, obtainable for $K \neq 3$. The general formula is

$$B = \frac{\begin{vmatrix} A_0 & A_1 & \cdots & A_K \\ \Delta A_0 & \cdots & \cdots & \Delta A_K \\ \vdots & & & \vdots \\ \Delta A_{K-1} & \cdots & \cdots & \Delta A_{2K-1} \end{vmatrix}}{\begin{vmatrix} 1 & 1 & \cdots & 1 \\ \Delta A_0 & \cdots & \cdots & \Delta A_K \\ \vdots & & & \vdots \\ \Delta A_{K-1} & \cdots & \cdots & \Delta A_{2K-1} \end{vmatrix}} \tag{23}$$

These determinants may be readily transformed into other interesting forms. Some of these are:

$$B = \frac{\begin{vmatrix} A_0 & A_1 & \cdots & A_K \\ \Delta A_0 & \Delta A_1 & \cdots & \Delta A_K \\ \Delta^2 A_0 & \Delta^2 A_1 & \cdots & \Delta^2 A_K \\ \vdots & \vdots & & \vdots \\ \Delta^K A_0 & \Delta^K A_1 & \cdots & \Delta^K A_K \end{vmatrix}}{\begin{vmatrix} 1 & 1 & \cdots & 1 \\ \multicolumn{4}{c}{\text{SAME AS (23)}} \end{vmatrix}} = A_K + \frac{\begin{vmatrix} (A_0 - A_K) & (A_1 - A_K) & \cdots & 0 \\ \Delta A_0 & & \cdots & \Delta A_K \\ \vdots & & & \vdots \\ \Delta A_{K-1} & & \cdots & \Delta A_{2K-1} \end{vmatrix}}{\begin{vmatrix} \text{SAME AS (23)} \end{vmatrix}} \qquad (23a,b,c)$$

$$B = \frac{\begin{vmatrix} A_0 & A_1 & \cdots & A_K \\ A_1 & A_2 & \cdots & A_{K+1} \\ \vdots & \vdots & & \vdots \\ A_K & A_{K+1} & \cdots & A_{2K} \end{vmatrix}}{\begin{vmatrix} \Delta^2 A_0 & \Delta^2 A_1 & \cdots & \Delta^2 A_{K-1} \\ \vdots & & & \vdots \\ \Delta^2 A_{K-1} & & \cdots & \Delta^2 A_{2K-2} \end{vmatrix}}$$

II. Two Nonlinear Sequence-to-Sequence Transforms

18. Let us return to the sequence of partial sums of the slowly convergent $\ln 2 = 1 - 1/2 + 1/3 - 1/4 + \ldots$ We have $A_0 = 1$, $A_1 = 1/2$, $A_2 = 5/6$, etc. Given the first seven partial sums, $A_0 \to A_6$,

we could assume this sequence to have the form $B + \sum_{i=1}^{3} a_i e^{\alpha_i n}$ with the seven unknowns, B, a_i and α_i , and solve for the constant B. We obtain from (22)

$$B = \left. \begin{vmatrix} 1 & \frac{1}{2} & \frac{5}{6} & \frac{7}{12} \\ -\frac{1}{2} & \frac{1}{3} & -\frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & -\frac{1}{4} & \frac{1}{5} & -\frac{1}{6} \\ -\frac{1}{4} & \frac{1}{5} & -\frac{1}{6} & \frac{1}{7} \end{vmatrix} \middle/ \begin{vmatrix} 1 & 1 & 1 & 1 \\ -\frac{1}{2} & \frac{1}{3} & -\frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & -\frac{1}{4} & \frac{1}{5} & -\frac{1}{6} \\ -\frac{1}{4} & \frac{1}{5} & -\frac{1}{6} & \frac{1}{7} \end{vmatrix} \right. \tag{24}$$

Therefore $B = \frac{1073}{1548} = 0.69315245$. But ln2 equals 0.69314718 and we have obtained a result accurate to five significant figures. This is encouraging.

19. To improve the approximation, <u>two</u> methods will be investigated. Assume two more values, A_7 and A_8 , to be known. The first method is to fit the <u>third</u> order form $B + \sum_1^3 a_i e^{\alpha_i n}$ to the <u>seven</u> quantities $A_1 \to A_7$ or to $A_2 \to A_8$. The second method is to fit a <u>fourth</u> order form $B + \sum_1^4 a_i e^{\alpha_i n}$ to the <u>nine</u> values $A_0 \to A_8$. The first method gives $B = \frac{9607}{13860} = 0.69314574$ for the set $A_1 \to A_7$ and $B = \frac{15679}{22620} = 0.69314766$ for the set $A_2 \to A_8$. The second method gives $B = \frac{14161}{20430} = 0.69314733$. Either method, therefore, improves the result.

20. Now we wish to give a general formulation of these two methods. Given 2K+1 members of the sequence A_n, <u>centered around $n = N$, $(N - K \leq n \leq N + K)$</u>, we may fit these 2K + 1 quantities by the K'th order form

$$A_n = B_{KN} + \sum_{i=1}^{K} a_{KNi}\, e^{\alpha_{KNi} n} \tag{25}$$

The quantity B_{KN} is obtained from formula (23) where all the suffixes are increased by N-K.
The transform

$$A_N \to B_{KN} \tag{26}$$

will be called a Kth order transform of A_N. If, in (26), N is varied from K to ∞ while K is held constant (the first method above), we say we have a Kth order sequence-to-sequence transform:

$$A_n \rightarrow B_{Kn} \qquad (n \geq K) \tag{27}$$

We designate the transform as e_K and tabulate it as follows:

$$\begin{array}{ll} A_0 & \\ A_1 & \\ \vdots & \\ A_K & B_{KK} \\ A_{K+1} & B_{K\,K+1} \\ A_{K+2} & B_{K\,K+2} \quad \text{etc.} \end{array}$$

21. We may call B_{Kn}, which in general will vary with n, the local base line constant of the sequence A_n. (See paragraphs 40, 53 for examples where all the B's are equal.) We expect the B's to lie closer together than the A's since they are all approximations to the base line constant of A_n. In the example above we saw that B_{33} (0.69315245), B_{34} (0.69314574), and B_{35} (0.69314766) do have this character. In fact, they are oscillating around and converging to ln2. We say the B sequence has the same base line constant as the A sequence. This suggests the iteration:

$$\left.\begin{aligned} B_n &= C_{KN} + \sum_{i=1}^{K} b_{KNi}\, e^{\beta_{KNi} n} \\ C_n &= D_{KN} + \sum_{i=1}^{K} c_{KNi}\, e^{\gamma_{KNi} n} \end{aligned}\right\} \tag{28}$$

and so forth. The Kth order iterated sequence-to-sequence transform:

$$A_n \rightarrow B_{Kn} \rightarrow C_{Kn} \rightarrow D_{Kn} \rightarrow \cdots\cdots \tag{29}$$

we shall designate as the $\tilde{e}_K$ transform of A_n.

22. The simplest of the $\tilde{e}_K$ transforms is $\tilde{e}_1$. It is symbolized as:

$$A_n \rightarrow B_{1n} \rightarrow C_{1n} \rightarrow D_{1n} \rightarrow \cdots \tag{30}$$

and tabulated as

$$\begin{array}{llll}
A_0 & & & \\
A_1 & B_{11} & & \\
A_2 & B_{12} & C_{12} & \\
A_3 & B_{13} & C_{13} & D_{13} \\
A_4 & B_{14} & C_{14} & . \\
A_5 & B_{15} & . & . \\
. & . & . & .
\end{array}$$

It is evaluated by the formulae:

$$B_{1,N} = \frac{A_N^2 - A_{N+1}A_{N-1}}{2A_N - (A_{N+1} + A_{N-1})}, \tag{31}$$

$$C_{1,N} = \frac{B_{1,N}^2 - B_{1,N+1}B_{1,N-1}}{2B_{1,N} - (B_{1,N+1} + B_{1,N-1})}, \tag{32}$$

etc. When using $\tilde{e}_1$, it is convenient to drop the subscript 1 from B, C, etc., and this done in the following pages.

23. The second type of transform is:

$$A_n \rightarrow B_{nn} \tag{33}$$

From A_0, A_1, and A_2 we derive B_{11} by a first-order transform. From A_0, A_1, A_2, A_3, and A_4 we derive B_{22} by a second-order transform, etc. This transform we will call the "diagonal" transform of A_n and we will designate it as e_d. It may also be iterated. The term "diagonal" should not be confused with the diagonal in the array of $\tilde{e}_1$ numbers above.

24. Formula (31) or its equivalent has been used by A. C. Aitken (reference c), G. Shortley and R. Weller (reference d), P. A. Samuelson (reference e), and D. Shanks and T. S. Walton (reference f). Shortley and Weller used it to extrapolate an iterative equation and the differences of which have a nearly constant ratio. Samuelson; and Shanks and Walton have used it to extrapolate an iterative sequence (the differences of which have a nearly constant ratio) which arises in the iterative solution of an equation of the form

$$x = f(x) \tag{34}$$

Aitken used what we call the $\tilde{e}_1$ process in its entitety, (30), to speed the convergence of a sequence which arises in Daniel Bernoulli's iterative solution of an algebraic equation (reference a, p. 98). In the examples which follow, the author applies $\tilde{e}_1$ and the other transforms to sequences arising from infinite series and products, continued fractions, integral and differential equations, eigenvalue convergents, etc. He believes this to be new.

25. The e_k and e_d transforms and their iterations are also believed to be new in a general sense. When applied to the partial sums of a power series, however, the e_k and e_d transforms are intimately related to Thiele's reciprocal differences (reference g), to the Padé Table (reference h), and to Kronecker's theorem on the power series of rational functions (reference i). This relationship will be discussed in Part VII. Samuelson (reference e) in reference to the solution of (34) suggests a form which is a special case of our (2) where $\alpha_2 = 2\alpha_1$, $\alpha_3 = 3\alpha_1$, etc., but he does not develop our (23), (27), or (29).

26. If we compare e_K or e_d with the Cesàro, Hölder, Abel, Euler, Riesz, Borel, LeRoy, and general Toeplitz summation processes (reference j) the most obvious difference is that all these processes are linear in the A_n whereas $e_{k,d}$ is nonlinear. If we consider the two determinants (23) to be expanded according to their first rows, we see that $e_{k,d}$ is a weighted average of the A_n. So are the above linear processes. In them, the weights are <u>preassigned</u> numbers or functions but in $e_{k,d}$ the weights are minors whose elements are differences of the A_n themselves. There is an obvious advantage of such a device. If we were summing a convergent series it would be desirable to weight the <u>later</u> A_n heavily - but if it were divergent it would be desirable to weight the <u>early</u> A_n heavily. No preassigned

numbers can do both. In the $e_{K,d}$ processes, we allow the sequence (so to speak) to choose its own appropriate weighting.

27. There is nothing to prevent us from combining e_K processes of various orders with each other, with e_d processes and with linear processes such as the Cesàro. Occasion for this will arise in some of the examples below. We now defer any further general discussion until some examples have been given.

III. Three Examples of $\tilde{e}_1$

28. $\ln 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \cdots$

Direct summation is imprudent since a billion terms would be required to obtain nine decimal places. The same accuracy however, is inherent in the first eleven (or even nine) terms and is obtainable if we concentrate not on the peaks of the graph but on the base line. Transforming the partial sums of this series by (30), (31), and (32) and keeping our results to 10 decimals places, we obtain:

A_0	1.00000 00000					
A_1	50000 00000	70000 00000				
A_2	83333 33333	69047 61905	69327 73109			
A_3	58333 33333	69444 44444	69310 57564	69314 88693		
A_4	78333 33333	69242 42424	69316 33407	69314 66820	69314 71961	
A_5	61666 66667	69358 97436	69313 99011	69314 73541	69314 71761	69314 71807
A_6	75952 38096	69285 71429	69315 08287	69314 71120	69314 71821	
A_7	63452 38096	69334 73390	69314 51963	69314 72107		
A_8	74563 49207	69300 33418	69314 83323			
A_9	64563 49207	69325 39683				
A_{10}	73654 40116					
	A_n	B_n	C_n	D_n	E_n	F_n

Each sequence is the local base line constant of the previous sequence, and oscillates around ln2 with a smaller amplitude than the previous sequences. $F_5 = 0.6931471807$ and $\ln 2 = 0.6931471806$. Since we are only keeping ten places we could not expect the last place of F_5 to be correct. It is possible to obtain the same result from the <u>nine</u> values $A_0 \rightarrow A_8$ if we treat the resulting <u>lower</u> diagonal A_8, B_7, C_6, D_5, etc. as a new sequence and subject it to $\widetilde{e}_1$. This, however, may be unreliable as a general procedure since the sequence terminates after 5 terms.

29. A similar calculation on the similar, but divergent series, $\ln 3 = 2 - \frac{2^2}{2}$ .. gives ln3 correct to 8 places from the first 11 terms. In greater detail, we take as our next example a more wildly divergent type

$$\int_0^\infty \frac{e^{-t}dt}{1+t} = \int_0^\infty e^{-t}(1-t+t^2-t^3+\cdots)dt = 1-1!+2!-3!+4!-\cdots$$

We wish to evaluate the integral. Proceeding formally, we obtain this alternating series of factorials which is of some fame in the theory of divergent series (see reference k, and j, p. 520). Knopp states that it cannot be summed by the Borel process, "the most powerful of the processes which are useful in practice". It does no good to tell us that the Stieltjes sum of the series is the integral $\int_0^\infty \frac{e^{-t}}{1+t}dt$ (reference j, p. 555) because that is what we started with. It is not our purpose to assign a "meaning" to the series. We already know its meaning. It is a series which arises from the integral and whose partial sums oscillate around the integral. Our sole purpose is a practical one. We wish to evaluate the integral numerically.

1					
0	.6666 6667				
2	.5000 0000	.6071 4286			
-4	.8000 0000	.5818 1818	.5977 8113		
20	.0000 0000	.6250 0000	.5943 5788	.5965 0983	
-100	2.8571 4286	.5194 8052	.6001 5250	.5961 0532	5963 6273
620	-10.0000 0000	.8620 6897	.5867 3183	.5968 1315	
-4420	60.0000 0000	-.5405 4054	.6272 4810		
35,900	-388.0000 0000	6.4347 8261			
-326,980	2910.9090 9091				
3,301,820					
A_n	B_n	C_n	D_n	E_n	F_n

Each sequence oscillates less vigorously than the previous sequences. They all diverge but the diagonal 1, .66666667, .60714286, etc. converges. Treating <u>this</u> sequence by $\tilde{e}_1$, we obtain

1.0000 0000		
6666 6667	5942 0290	
6071 4286	5960 3395	5963 5910
5977 8113	5963 1007	5963 4808
5965 0983	5963 4348	
5963 6273		

and from the <u>new</u> diagonal sequence we obtain 0.5963 4770.

30. But $\frac{1}{e}\int_0^\infty \frac{e^{-t}}{1+t}\,dt = \int_1^\infty \frac{e^{-x}}{x}\,dx = \int_0^\infty \left(e^{-x} - \frac{1}{1+x}\right)dx - \int_0^1 \frac{e^{-x}-1}{x}\,dx =$

$$-C + 1 - \frac{1}{2\cdot 2!} + \frac{1}{3\cdot 3!} - \frac{1}{4\cdot 4!} + \cdots$$

where C is Euler's constant. Thus we find that $\int_0^\infty \frac{e^{-t}}{1+t}\,dt$ is equal to 0.5963 4736. Our second example is thus also amenable to $\tilde{e}_1$. Perhaps the reader is surprised that the later A_n, 35,900; -326,980; 3,301,820, etc. do not spoil the calculation. The opposite is true. Each one improves the result. The reason is that each one gives further data on the base line constant around which the sequence is oscillating.

31. $C = \frac{1}{2} + \frac{B_2}{2} + \frac{B_4}{4} + \frac{B_6}{6} + \cdots$

This is a formula of Euler's and gives his constant (0.57721 56649) as a "sum" of Bernoulli numbers (reference j, p. 541). It is asymptotic and after the fourth term it diverges rapidly. Of these asymptotic series, Knopp says "we are not in position - not even in theory - to obtain <u>any degree of accuracy whatever</u> in the evaluation of f(x) ----. The degree of accuracy therefore cannot be lowered below the value of the least term of the series." Since the fourth partial sum of the series above is 0.57896 82540 and the fifth is 0.57480 15873 it would appear that the accuracy of Euler's formulae is not very high. Knopp, in fact, states that Euler's "sum" is "not valid, however, even from the general view-point of § 59", (summation of divergent series) "for the investigations of § 64" (asymptotic series) "have provided no process by which the sum in question may be obtained from the partial sums of the series by a <u>convergent</u> process, as was always supposed."

32. On the other hand Euler says (reference l), "Whenever an infinte series is obtained as the development of some closed expression, it may be used in mathematical operations as the equivalent of that expression, even for values of the variable for which the series diverges."

33. Bromwich (reference k, p. 325) says that Euler "regarded" his constant as the "sum" of the series but does not commit himself as to his own opinion. He does say "from this series we cannot obtain a closer approximation than S_4" (.5790).

34. Finally, the $\tilde{e}_1$ process gives:

		$\boxed{57721\ 56582}$			57721 56790
			57720 96151	57721 49167	57721 55832
A_3	.57896 82540				
A_4	.57480 15873	57748 97594			
A_5	.58237 73449	57680 35056	57725 42410		
•	.56128 45488	57811 68590	57715 13956	57722 05432	
•	.64461 78821	57447 20439	57736 24338	57720 65727	57721 62196
	.20135 80782	58843 62340	57676 63050	57723 77443	
	3.25531 24085	51740 63481	57902 02119		
	-23.20089 97127	98219 13367			
	258.25924 52148				

But C = 0.57721 56649 and we must agree with Euler. We now return to a more general discussion of $\tilde{e}_1$.

IV. General Discussion Resumed

35. Up to now our discussion has been based on qualitative and intuitive argument, on physical aralogy and on numerical evidence. A mathematician will naturally wish a more rigorous treatment. Some proofs are readily obtainable but the author does not have the complete theory at this time. A single e_1 process is not entirely regular. If we apply it to the rather artificial

convergent series

$$1+\frac{1}{2}+\frac{1}{2}+\frac{1}{4}+\frac{1}{4}+\frac{1}{8}+\frac{1}{8}+\cdots \tag{35}$$

we see from (31) that B_1, B_3, B_5, etc. are all ∞. (This does not mean that (35) is not summable by a second-order process, e_2.) Can we prove that $\tilde{e}_1$ is valid for the partial sums of the convergent ln 2 series (our first example)? Specifically, can we prove that each derived sequence (Bn, Cn, etc.) converges to the same limit as A_n (ln2) and further, that it converges more rapidly than the previous sequence?

36. Transforming (31) and using (17) we obtain:

$$\left.\begin{aligned} B_n &= A_N + \frac{\Delta A_N}{1-\dfrac{\Delta A_N}{\Delta A_{N-1}}} \\ B_{N+1} &= A_N + \frac{\Delta A_N}{1-\dfrac{\Delta A_{N+1}}{\Delta A_N}} \end{aligned}\right\} \tag{36}$$

If the A_n are the partial sums of a convergent infinite series with

$$\Delta A_n \rightarrow 0 \tag{37}$$

and

$$\frac{\Delta A_n}{\Delta A_{n-1}} \leq R < 1 \tag{38}$$

we see from (36) that B_n converges to the same limit as A_n. Therefore a single e_1 process is obviously valid if the series is either of the alternating type ($|\Delta A_n| < |\Delta A_{n-1}|$, $\Delta A_n/\Delta A_{n-1} < 0$) or of the ratio test type ($\Delta A_n \geq 0$, $\Delta A_n/\Delta A_{n-1} \leq R < 1$). But may it be iterated?

37. The difference of formulae (36) gives the term of the transformed series:

$$\Delta B_N = \Delta A_N \cdot \frac{\dfrac{\Delta A_{N+1}}{\Delta A_N} - \dfrac{\Delta A_N}{\Delta A_{N-1}}}{\left(1-\dfrac{\Delta A_{N+1}}{\Delta A_N}\right)\left(1-\dfrac{\Delta A_N}{\Delta A_{N-1}}\right)} \tag{39}$$

If we now take a sequence of the alternating type where

$$A_N = \sum_{i=0}^{N} t_i \tag{40}$$

$$\Delta A_N = t_{N+1} \tag{41}$$

and

$$t_n = (-1)^n \frac{f(n)}{g(n)} \tag{42}$$

where f(n) is a F'th degree polynomial in n and g(n) is a G'th degree polynomial in n, (39) becomes

$$\Delta B_N = \Delta A_N \frac{f^2(N+1)\, g(N)\, g(N+2) - g^2(N+1)\, f(N)\, f(N+2)}{\left[f(N) g(N+1) + f(N+1) g(N)\right]\left[f(N+1) g(N+2) + f(N+2) g(N+1)\right]} \tag{43}$$

Expanding the fraction in powers of N we obtain

$$\Delta B_N = \Delta A_N \left[\frac{F-G}{4N^2} + O\left(\frac{1}{N^3}\right)\right] \tag{44}$$

For $G > F$, and $g(n) \neq 0$, $\Delta A_n \rightarrow 0$ and the sequence A_n converges. From (36), (38) and (44), B_n converges to the same limit, but more rapidly. Likewise from (43) and (44) we may iterate the process since the B series is again of the alternating type with a term $(-1)^{n+1} \dfrac{f_1(n)}{g_1(n)}$

where

$$G_1 - F_1 = G - F + 2 \tag{45}$$

Since ln2 = 1 - 1/2 + 1/3 - 1/4 + ... is a special case of (42) we have formally proved that e_1 is valid for the partial sums of this series; (the numerical evidence was very convincing).

38. Likewise it would increase the rapidity of convergence of:-

$$\pi/4 = 1 - 1/3 + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \cdots \tag{46}$$

$$G = 1 - \frac{1}{3^2} + \frac{1}{5^2} - \frac{1}{7^2} + \cdots \qquad (47)$$

$$\frac{\pi^2}{12} = 1 - \frac{1}{2^2} + \frac{1}{3^2} - \frac{1}{4^2} + \cdots \qquad (48)$$

and so forth. Equation (45) shows further that although

$$1-2+3-4\ldots \qquad (49)$$

diverges, a single e_1 process converts it into a convergent sequence (series). This sequence converges to 1/4. Similarly two successive processes transform

$$1 - 2^2 + 3^2 - 4^2 + \ldots \qquad (50)$$

into a convergent sequence (series) which converges to 0. Two processes will also transform

$$1 - 2^3 + 3^3 - 4^3 + \ldots \qquad (51)$$

into a convergent sequence, and so forth. We will return to these divergent series later and show how they may be summed - not merely in a limiting sense - but exactly - to rational multiples of the Bernoulli Numbers.

39. We could undoubtedly show that $\tilde{e}_1$ is also valid for a more general series than (42), such as would include, for instance

$$1 - \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} - \frac{1}{\sqrt{4}} + \cdots \qquad (52)$$

and

$$\left(\tfrac{1}{2}\right) - \frac{\left(\frac{1}{2}\right)^2}{2} + \frac{\left(\frac{1}{2}\right)^3}{3} - \frac{\left(\frac{1}{2}\right)^4}{4} + \cdots \qquad (53)$$

but since we do not have an exhaustive theory of the range and speed of e_1 as applied to series (let alone sequences derived more simply from products, continued fractions, etc.), we will forego such proofs. We will show later that if a power series (convergent or divergent) is obtained by dividing out a rational function of x, we may sum the series to its sum - the rational function - not merely in a limiting sense - but exactly by a single e_K process (with sufficiently large K). Some other formal proofs will also be given.

40. First we wish to point out that if a single e_1 process is applied to any three consecutive partial sums of a geometric series

$$c + cr + cr^2 + cr^3 + \cdots \tag{54}$$

we obtain $\frac{c}{1-r}$. That is, our entire B_n sequence is a constant, $B = c/1-r$. Thus a single e_1 process sums a geometric series exactly. Divergence and convergence of the series are of course irrelevant. Since the partial sums are

$$A_n = \frac{c}{1-r} - \frac{c}{1-r} e^{\ln r \cdot n} \tag{55}$$

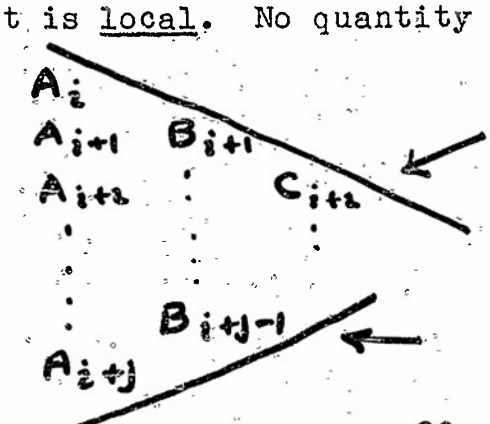
we see that we have a simple transient, (2), with $B = \frac{c}{1-r}$ and $a = -\frac{c}{1-r}$, $\alpha = \ln r$.

41. If we applied (32) to this constant B_n sequence we would obtain indeterminate expressions for C_n. Let us agree, then, to the consistent convention that if any three consecutive members of a sequence are equal, the e_1 transform of the center member is also equal to the same quantity. Let us also agree to the generalization of this convention for the case of the e_K transform of any 2K + 1 <u>equal</u>, <u>consecutive</u> members of a sequence. With this convention theory we may say that a geometric series is exactly summable by either e_1 or $\tilde{e}_1$.

42. The series for ln 2 is not geometric but it is "nearly geometric" in the sense that the ratio of the terms rapidly approach a constant, -1. Further, ln 2 is not rational. Thus, although we cannot sum ln2 exactly by e_1 we can sum it in a limiting sense. And since the series converges and becomes "more geometric" as n increases, it is clear that the $\tilde{e}_1$ process will have a more rapid convergence when applied to A_n with n large. As long as the series is making an effort to converge (n small) we may allow it to do so and defer the $\tilde{e}_1$ process until a nearly geometric character sets in. Thus from the <u>seven</u> partial sums of ln 2, A_{11} (.6532106783) to A_{17} (.6661398243), we obtain D_{14} (.6931471802) which is correct to 9 places. We do more addition to obtain the A_n with larger n but we more than make up for it by doing less $\tilde{e}_1$ calculation.

43. In connection with these calculations we wish to point out five characteristics of the $\tilde{e}_1$ transform.

I. It is <u>local</u>. No quantity above or below these two diagonals:



(56)

can affect any quantity between them. We used this idea in the calculation just performed and in the asymptotic series for Euler's C (see paragraph 34).

II. A_n may be multiplied by a constant. For e_1 applied to mA_{n-1}, mA_n, and mA_{n+1} gives mB_n

III. A constant may be added to or subtracted from A_n. For e_1 applied to $A_{n-1} + a$, $A_n + a$, and $A_{n+1} + a$ gives $B_n + a$. If we wish to transform $\begin{cases} 6931470376 \\ 6931472835 \\ 6931471053 \end{cases}$ we may transform $\begin{cases} 0376 \\ 2835 \\ 1053 \end{cases}$ and add 6931470000.

IV. Generally there is no loss of significant figures. From an A_n sequence accurate to 10 significant figures we will obtain a B_n sequence of the same accuracy. This may be seen from (36). Only when $\dfrac{\Delta A_n}{\Delta A_{n-1}}$ is nearly equal to +1 will be a less accurate B_n sequence result.

V. The $\tilde{e}_1$ process may be readily mechanized. With modern large-scale calculators the B_n, C_n, etc., sequences may be calculated almost as fast as the A_n sequence is fed in.

44. Finally we should give a rough indication of the types of sequences for which the $\tilde{e}_1$ process should be useful. If the sequence is nearly geometric, that is if $\dfrac{\Delta A_n}{\Delta A_{n-1}}$ approaches a constant ($\neq + 1$), we might expect $\tilde{e}_1$ to work. If the graph of A_n looks like a simple damped or growing oscillation or a simple growing or decaying exponential, we might expect $\tilde{e}_1$ to work. But if the sequence is derived from a structure like

$$1 - \frac{1}{2} - \frac{1}{3} + \frac{1}{4} + \frac{1}{5} - \frac{1}{6} - \frac{1}{7} + \cdots \quad \text{or} \quad \cfrac{1}{2 + \cfrac{1}{3 + \cfrac{1}{2 + \cfrac{1}{3 + \cfrac{1}{2 + \cdots}}}}}$$

it would seem more reasonable to use a second-order process We could, however, use a first-order process on the alternate member of such a sequence, that is, on A_0, A_2, A_4, etc., or A_1, A_3, A_5, etc. In no case where the sequence has been artificially and randomly put together (for example, out of numbers in a telephone book) would we attempt to use $\tilde{e}_1$ or any other of these transforms. Although certain classical mathematical operations frequently produce sequences of a transient character, this is no indication that all operations will do so. We now return to our examples and apply $\tilde{e}_1$ to some sequences not derived from infinite series.

V. Further Examples of $\tilde{e}_1$

45. $A_n = P_{6\cdot 2^n} \to 2\pi$

A limiting sequence of considerable antiquity is the sequence of the perimeters of 6.2^n-sided regular polygons inscribed in a unit circle. A table of such perimeters ($n = 0$ to 7) is given by Wentworth and Smith in reference (m). The sequence converges to 2π with moderate speed. The speed may be accelerated by $\tilde{e}_1$ but when this was first done the C_n sequence had a noticeably non-smooth character. This suggested an error in the Wentworth-Smith table. A new calculation was therefore made:

n	Wentworth-Smith	New Calculation
0	6.0000 0000	6.0000 00000
1	6.2116 5708	6.2116 57082
2	6.2652 5722	6.2652 57226
3	6.2787 0041	6.2787 00408
4	6.2820 6396	6.2820 63901
5	6.2829 0510	6.2829 04944

From the new data we obtain $C_3 = 6.2831\ 85307$ correct to the ten significant figures. Generally speaking, $\tilde{e}_1$ is sensitive to small errors and is a means of detecting them.

46. The reader may wish to try $\tilde{e}_1$ on other limiting sequences. Two such sequences are:

$$\left(1 + \frac{1}{2^n}\right)^{2^n} \to e \qquad (57)$$

and

$$\left(2^{\frac{1}{2^n}} - 1\right)\cdot 2^n \to \ln 2 \qquad (58)$$

He may also wish to experiment with linear transforms. If, in the perimeter sequence above, we take

$$\left.\begin{aligned} B_n &= \frac{4A_n - A_{n-1}}{3} \\ \text{and} \quad C_n &= \frac{16B_n - B_{n-1}}{15} \end{aligned}\right\} \qquad (59)$$

we again obtain $C_5 = 6.2831\ 85307$. This calculation is undoubtedly more rapid than the previous one but it assumes knowledge of the weighting factors in (59). These were obtained by:

$$\left.\begin{aligned} \frac{\Delta A_n}{\Delta A_{n-1}} &\to \frac{1}{4} \\ \frac{\Delta B_n}{\Delta B_{n-1}} &\to \frac{1}{16} \end{aligned}\right\} \qquad (60)$$

From these formulae the reader sees the connection between $\tilde{e}_1$ and this linear process. $\tilde{e}_1$ depends on local weighting factors but the linear process depends on extrapolated weights. The author calls this linear process "geometric extrapolation".

47. An Eigen-value Problem

Most approximate methods of finding eigenvalues have a rather low accuracy, (say 1/10 of 1%), unless a large number of approximations are calculated. A simple eigenvalue problem is:

$$\left.\begin{aligned} \ddot{y} + \lambda y &= 0 \\ y(0) = \dot{y}(1) &= 0 \end{aligned}\right\} \qquad (61)$$

Find the lowest eigenvalue. Collatz, (reference n) by using Courant's "Maximum-Minimum-Prinzip" finds the first three upper and lower bounds. Applying $\tilde{e}_1$ to these we obtain:

		A_n	B_n	C_n
3	=	3.00 00000		
$\frac{12}{5}$	=	2.40 00000	2.48 57143	
$\frac{5}{2}$	=	2.50 00000	2.47 09302	2.46 73381
$\frac{150}{61}$	=	2.45 90164	2.46 80403	2.46 74162
$\frac{42}{17}$	=	2.47 05882	2.46 75271	
$\frac{3416}{1385}$	=	2.46 64260		

Since the lowest eigenvalue is $\frac{\pi^2}{4}$ = 2.46 74011 we have obtained two extra decimal places. To the important question, "How can we know *a priori* what the accuracy is?" the author replies that if we have a long sequence, A_n, a study of the differences within and between the derived sequences will give a reasonable estimate of the accuracy; but if we have only a few terms of A_n, the question is more difficult.

48. In the Rayleigh-Ritz method the approximating functions are usually chosen in such a way that the resulting approximate-eigenvalue sequence converges as rapidly as possible. This choice, however, usually complicates the calculation. It seems possible that one could choose simple functions in a regular manner such that, while the resulting sequence converged slowly, a large number of terms could be readily calculated. We would then rely on $\tilde{e}_1$ or some other process to secure the necessary accuracy from the rough but abundant data.

49. A Continued Fraction Sequence

From $\sqrt{2} = 1 + \frac{1}{2+} \frac{1}{2+} \frac{1}{2+} \cdots$ we derive a sequence of convergents. We process these by $\tilde{e}_1$ and obtain convergents which are farther along the sequence.

$$
\begin{array}{llll}
\frac{1}{1} & & & \\
\frac{3}{2} & \frac{17}{12} & & \\
\frac{7}{5} & \frac{99}{70} & \frac{19601}{13860} & = 1.414213564 \\
\frac{17}{12} & \frac{577}{408} & & \\
\frac{41}{29} & & &
\end{array}
$$

C_3 is correct to 9 significant figures. The e_1 transform of any convergent here is the same as that obtained by:

$$
\left.
\begin{aligned}
B_N &= \tfrac{1}{2}\left(A_N + \frac{2}{A_N}\right) \\
C_N &= \tfrac{1}{2}\left(B_N + \frac{2}{B_N}\right)
\end{aligned}
\right\} \qquad (62)
$$

But this is Newton's iterative method of taking square roots. Each iteration doubles the number of correct decimal places. (See reference a, page 79.)

50. Euler's Partition Function

This function (reference o) may be expressed as:

$$f(x) = 1 + x + 2x^2 + 3x^3 + 5x^4 + 7x^5 + 11x^6 + \cdots \qquad (63)$$

where the coefficient of x^n is $p(n)$, the number of possible partitions of n. It may also be expressed as the infinte product:

$$f(x) = \frac{1}{(1-x)(1-x^2)(1-x^3)(1-x^4)\cdots} \qquad (64)$$

or as the reciprocal of a power series with gaps:

$$f(x) = \frac{1}{1 - x - x^2 + x^5 + x^7 - x^{12} - x^{15} + x^{22} + \cdots} \qquad (65)$$

Because of the increasingly long gaps in (65), the convergence of this series for $x = 1/2$ is rapid and from the first nine non-zero terms we obtain ten figure accuracy:

$$f(\tfrac{1}{2}) = 3.462746619 \qquad (66)$$

On the other hand from the first nine partial products of (64) we obtain only three figure accuracy:

$$A_8 = 3.459987843 \qquad (67)$$

But if we use $\tilde{e}_1$ on these nine partial products, we again find:

$$E_4 = 3.462746619 \qquad (68)$$

51. A Nonlinear Differential Equation

Given

$$\ddot{y} + 32.17 + 0.11534\,\dot{y}\,e^{-\frac{y}{27,800}} = 0 \qquad (69)$$

with the initial conditions:

$$y(0) = 0 \qquad \dot{y}(0) = 6400 \qquad (70)$$

evaluate y for $t = 9$. (See reference f, page 2.) We introduce

$$S = 0.11534t \quad \text{and} \quad z = \frac{y}{27,800} \qquad (71)$$

and solve the resulting equation by the power series:

$$z = 1.99598\,S - 1.04147\,S^2 + 1.01114\,S^3 - 1.10380\,S^4 + 1.28005\,S^5 - \cdots \tag{72}$$

The nearly-geometric character of this series is apparent. For $t = 9$ ($S = 1.03806$) the series oscillates. We apply e_1 to the first five partial sums and obtain:

$$y(9) = 41,456 \tag{73}$$

A solution of (69) by numerical integration gives

$$y(9) = 41,440 \tag{74}$$

while series (72) directly gives the false value:

$$y(9) = 65,106 \tag{75}$$

52. A Divergent Series for Catalan's Constant

We may derive a divergent series for Catalan's constant, (47), from a formula of Titchmarsh (reference p).

$$\left.\begin{aligned} &(\pi/2)^s\,\Gamma(1-s)\cos\left(\frac{s\pi}{2}\right) L(1-s) = L(s) \\ &\text{where } L(s) = 1 - \frac{1}{3^s} + \frac{1}{5^s} - \frac{1}{7^s} + \cdots \end{aligned}\right\} \tag{76}$$

Since $G = L(2)$, we find from (76) with $S = -1$

$$\frac{2}{\pi} G = \frac{1 - 3 + 5 - 7 + \cdots}{\cos(-\pi/2)} \tag{77}$$

But the denominator vanishes and the numerator may be summed to zero. Proceeding in the eulerian manner, we write:

$$\frac{2}{\pi} G = \left.\frac{\frac{d}{ds} L(s)}{\frac{d}{ds}\cos(s\pi/2)}\right|_{s=-1} \tag{78}$$

or $$\left(\frac{2}{\pi}\right)^2 G = 3\ln 3 - 5\ln 5 + 7\ln 7 - \cdots \tag{79}$$

The first seven partial sums of this divergent series may be summed to

$$\left(\frac{2}{\pi}\right)^2 G = 0.5831 \tag{80}$$

Since G = 0.9159655942 we calculate

$$\left(\frac{2}{\pi}\right)^2 G = 0.37122\,68726 \tag{81}$$

The agreement is not perfect. As it happens, though

$$\left(\frac{2}{\pi}\right) G = 0.58312\,18080 \tag{82}$$

and this suggests that formula (76) has a typographical error of a factor of $2/\pi$. That this is the case may be seen by taking S = 1/2 in (76). Then

$$(\pi/2)^{\frac{1}{2}}\Gamma\left(\tfrac{1}{2}\right)\frac{\sqrt{2}}{2} = 1 \quad \text{But actually} \left(\frac{\pi}{2}\right)^{\frac{1}{2}-1}\Gamma\left(\tfrac{1}{2}\right)\frac{\sqrt{2}}{2} = 1 \tag{83}$$

53. Other applications of $\tilde{e}_1$ could be given but we must now turn to the study of the higher order e_k transforms. A compelling reason for this study is the fact that $\tilde{e}_1$ does not work well for some simple series. For instance, take the divergent series

$$1 + 2\cdot 2 + 3\cdot 2^2 + 4\cdot 2^3 + 5\cdot 2^4 + \cdots \tag{84}$$

which may be derived from the rational function $\frac{1}{(1-x)^2}$ with x = 2. The partial sums of this series diverge from a definite value, namely, +1. However, the $\tilde{e}_1$ process requires at least four iterations before any appreciable convergence takes place (after that it goes much more rapidly). In other words, it is slow. On the other hand, from the first five (or any five consecutive) partial sums, a second-order process, e_2, yields the exact sum

$$\frac{\begin{vmatrix} 1 & 5 & 17 \\ 4 & 12 & 32 \\ 12 & 32 & 80 \end{vmatrix}}{\begin{vmatrix} 1 & 1 & 1 \\ 4 & 12 & 32 \\ 12 & 32 & 80 \end{vmatrix}} = +1 \tag{85}$$

54. Another type of slow convergence of the $\tilde{e}_1$ process, and two more difficulties - local nonuniform convergence and local nonconvergence - will be discussed in a later paper. Some of the main ideas of the e_k transforms are illustrated in the next two examples.

VI. Two Examples of e_2

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

The second-order process, e_2, is useful in the solution of this simple integral equation by the method of Successive Substitution (reference q). This method, when applied to the above equation, gives the divergent series:

$$u(x) = \left(\frac{x}{2} - \frac{1}{3}\right) - \left(\frac{x}{12} + 0\right) - \left(\frac{x}{24} + \frac{1}{36}\right) - \left(\frac{7x}{144} + \frac{1}{36}\right) - \left(\frac{15x}{288} + \frac{13}{432}\right) \cdots \tag{86}$$

We apply e_2 to the first five partial sums and obtain

$$u(x) = x \tag{87}$$

Any other five consecutive partial sums of (86) will give the same result. The <u>vector</u> series, (86), is therefore exactly summable by e_2 to the exact solution of the integral equation.

56. This example is a particularly clear one since we can show exactly why e_2 is successful. The integration process whereby each vector term of (86) is obtained from the previous term is equivalent to a matrix multiplication. That is

$$\begin{pmatrix} \frac{1}{2} & 1 \\ \frac{1}{3} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} v_{n-1} \end{pmatrix} = \begin{pmatrix} v_n \end{pmatrix} \tag{88}$$

or

$$\begin{pmatrix} \frac{1}{2} & 1 \\ \frac{1}{3} & \frac{1}{2} \end{pmatrix}^n \begin{pmatrix} v_0 \end{pmatrix} = \begin{pmatrix} v_n \end{pmatrix} \tag{89}$$

where

$$\begin{pmatrix} v_0 \end{pmatrix} = \begin{pmatrix} \frac{x}{2} \\ -\frac{1}{3} \end{pmatrix} \tag{90}$$

But the matrix can be expanded into a pair of unit orthogonal matrices (for instance, by the algebraic apparatus of Part I) and therefore we have

$$\begin{pmatrix} \frac{1}{2} & 1 \\ \frac{1}{3} & \frac{1}{2} \end{pmatrix}^n = \left(\frac{3+2\sqrt{3}}{6}\right)^n \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{6} & \frac{1}{2} \end{pmatrix} + \left(\frac{3-2\sqrt{3}}{6}\right)^n \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{6} & \frac{1}{2} \end{pmatrix} \tag{91}$$

Finally

$$A_N = \sum_{n=0}^{N} (v_n) = X - \left(\frac{3+2\sqrt{3}}{6}\right)^N \begin{pmatrix} \frac{3+2\sqrt{3}}{12} X \\ \frac{2+\sqrt{3}}{12} \end{pmatrix} - \left(\frac{3-2\sqrt{3}}{6}\right)^N \begin{pmatrix} \frac{3-2\sqrt{3}}{12} X \\ \frac{2-\sqrt{3}}{12} \end{pmatrix} \quad (92)$$

and we see that the sequence of partial sums is a second-order transient. The second term, $\left(\frac{3+2\sqrt{3}}{6}\right)^N \left(\frac{3+2\sqrt{3}}{12} X + \frac{2+\sqrt{3}}{12}\right)$, is monotonic and divergent; the third term is a damped oscillation; and the first term, x, is the base line constant.

57. Goldstein's Formula for Drag

Goldstein, (reference r), has investigated the drag of a sphere in a viscous fluid as a function of the Reynolds number. This investigation is based on the linearized (Oseen) theory and results in the following series formula

$$K_D = \frac{12}{R}\left[1 + \frac{3R}{16} - \frac{19R^2}{1,280} + \frac{71R^3}{20,480} - \frac{30,179R^4}{34,406,400} + \frac{122,519R^5}{560,742,400} - \right. \quad (93)$$

The series converges for R = 2, oscillates for R = 4, and diverges badly for R = 6. In view of this, Goldstein abandons the series (which had been obtained with much labor) and resorts to an approximate, numerical solution for R > 2.

58. Series (93) is nearly geometric and we could sum it by $\tilde{e}_1$. For any fixed value of R, this would be the simplest thing to do. However, we would like a formula for K_D in terms of R and for this purpose it is simpler to use a second-order process. We apply e_2 to the first five terms of (93) and obtain the rational approximation

$$K_D \approx \frac{12}{R} \cdot \frac{295,680 + 133,200R + 10,880R^2}{295,680 + 77,760R + 689R^2} \quad (94)$$

In the table which follows we compare (93) and (94) with Goldstein's numerical solution.

<u>Table 1</u>

R	(93) Series Solution	Numerical Solution	(94) Rational Approx.
1	14.106	14.11	14.1051
2	8.018	8.00	8.0043
3	6.018	5.93	5.9287
4	5.201	4.87	4.8688
5	5.122	4.22	4.2195
6	5.803	3.78	3.7776
8	6.331	3.21	3.2088
10	21.523	2.85	2.8532
12	43.282	2.60	2.6060
16	139.518	2.28	2.2777
20	351.232	2.08	2.0630

If we apply e_2 to the last five partial sums of (93) the agreement is not quite as good. The last term of (93) is wrong or at least it does not follow from the λ's of Goldstein's paper (reference s). When we correct it:

$$\frac{122{,}519\,R^5}{550{,}502{,}400} \quad \text{instead of} \quad \frac{122{,}519\,R^5}{560{,}742{,}400} \tag{95}$$

we find that the last five terms now give

$$K_D \simeq \frac{12}{R}\cdot\frac{18{,}520{,}320 + 8{,}332{,}800R + 677{,}120R^2 - 300R^3}{18{,}520{,}320 + 4{,}860{,}240R + 40{,}736R^2} \tag{96}$$

which agrees with (94) to three or four significant figures up to $R = 20$.

59. Series (93) may be written

$$\frac{12}{R}\left[1 + \frac{3}{4}\left(\frac{R}{4}\right) - \frac{19}{80}\left(\frac{R}{4}\right)^2 + \frac{71}{320}\left(\frac{R}{4}\right)^3 - \frac{30{,}179}{134{,}400}\left(\frac{R}{4}\right)^4 + \frac{122{,}519}{537{,}600}\left(\frac{R}{4}\right)^5\right. \tag{97}$$

while its reciprocal is

$$\frac{R}{12}\left[1 - \frac{3}{4}\left(\frac{R}{4}\right) + \frac{4}{5}\left(\frac{R}{4}\right)^2 - \left(\frac{R}{4}\right)^3 + \frac{559}{420}\left(\frac{R}{4}\right)^4 - \frac{38}{21}\left(\frac{R}{4}\right)^5\right. \tag{98}$$

The relative simplicity of (98) suggests that it might be simpler, in any future work on drag in this range of Reynolds number, to work with $\frac{1}{K_D}$ instead of K_D directly.

60. In the last two examples we have illustrated:

(a) the utility of reciprocal series
(b) the accuracy of rational approximations
(c) the numerical simplicity of e_1 versus the analytic simplicity of e_K or e_d, and
(d) the exact summation of series whose sum is rational.

These mathematical phenomena are tightly interwoven, together with the Padé Table, the algorithms of continued fractions and rational approximations, a theorem of Kronecker's, and Thiele's reciprocal differences in the theory of the e_K and e_d transforms.

VII. General Discussion of the e_K and e_d Transforms

61. We now return to the application of higher order processes to the series for ln(1+x), as introduced in paragraph 18. A zero order process applied to the first term of

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots \tag{99}$$

gives

$$\ln(1+x) \cong x = B_{00} \tag{100}$$

An e_1 process applied to the first three terms gives

$$\ln(1+x) \cong x \cdot \frac{6+x}{6+4x} = B_{11} \tag{101}$$

An e_2 process applied to the first five terms gives

$$\ln(1+x) \cong x \cdot \frac{30+21x+x^2}{30+36x+9x^2} = B_{22} \tag{102}$$

A continuation of this process (the e_d transform) gives:

$$\ln(1+x) \cong x \cdot \frac{420+510x+140x^2+3x^3}{420+720x+360x^2+48x^3} = B_{33} \tag{103}$$

and so forth. This sequence of rational approximations for $\ln(1+x)$ converges rapidly for all x (except the real cut -1 to $-\infty$) to the value ln(1+x). For example, from (103) we have

$$\left.\begin{aligned} \ln 2 &\cong \frac{1073}{1548} = 0.693152 \\ \text{and} \quad \ln 3 &\cong \frac{1012}{921} = 1.0988 \end{aligned}\right\} \tag{104}$$

correct to five and four significant figures. Rational approximations for any analytic function may be obtained by a similar calculation.

62. If we now take a known continued fraction for ln(1+x) (reference h, page 342)

$$\ln(1+x) = \frac{x}{1+}\ \frac{1^2x}{2+}\ \frac{1^2x}{3+}\ \frac{2^2x}{4+}\ \frac{2^2x}{5+}\ \frac{3^2x}{6+}\cdots \qquad (105)$$

we find that the first, third, fifth, etc., convergents of this fraction are exactly the rational approximations we have just obtained.

63. Similarly, if we apply e_d to the series

$$1 - 1! + 2! - 3! + 4! - \cdots$$

(see paragraph 29) we obtain the following rational approximations:

$$\int_0^\infty \frac{e^{-t}}{1+t}\,dt \approx \frac{1}{1},\ \frac{3}{3},\ \frac{8}{13},\ \frac{44}{73},\ \frac{300}{501},\ \frac{2420}{4051},\ \text{etc.} \qquad (106)$$

These approximations are likewise obtainable from the odd convergents of a known continued fraction (reference h, page 356).

$$\int_0^\infty \frac{e^{-t}}{1+t}\,dt = \frac{1}{1+}\ \frac{1}{1+}\ \frac{1}{1+}\ \frac{2}{1+}\ \frac{2}{1+}\ \frac{3}{1+}\ \frac{3}{1+}\cdots \qquad (107)$$

Euler, in fact, knew of this fraction and used it to sum the series to four decimal places (reference k). The fraction converges rather slowly. The 17'th convergent, A_{17} = 0.5964599995, is only correct to three decimal places. We can accelerate its convergence by applying $\tilde{e}_1$ to it. But the fraction, as it stands, (107), has a double structure such as was discussed in paragraph 44. We prefer, therefore, to apply $\tilde{e}_1$ to the alternate members of the convergents sequence - namely, to our rational approximations, (106). From the first nine approximations, that is, up to the A_{17} mentioned above, we obtain nine-place accuracy (0.596347362). Here, then, we have a combination process. First, a diagonal transform to produce a slowly convergent sequence, and then an $\tilde{e}_1$ process to extrapolate it.

64. We now shift to a somewhat different topic but we will soon show its relation to our nonlinear transforms. Here is an algorithm for the calculation of e. Consider the table of fractions:

$$\begin{array}{cccc}
\frac{N_{00}}{D_{00}} & \frac{N_{01}}{D_{01}} & \frac{N_{02}}{D_{02}} & \cdot \\[2ex]
\frac{N_{10}}{D_{10}} & \frac{N_{11}}{D_{11}} & \frac{N_{12}}{D_{12}} & \frac{N_{13}}{D_{13}} \quad \cdot \\[2ex]
\cdot & \frac{N_{21}}{D_{21}} & \cdot & \cdot
\end{array}$$

and the rules:

$$\frac{N_{Rc}}{D_{Rc}} = \frac{N_{R-1,c} + N_{R-1,c+1}}{D_{R-1,c} + D_{R-1,c+1}} \tag{108}$$

$$\frac{N_{0c}}{D_{0c}} = \frac{(c-1)\ N_{1,c-2}}{(c-1)\ D_{1,c-2}}\ ;\ \frac{N_{00}}{D_{00}} = \frac{1}{1}\ ;\ \frac{N_{01}}{D_{01}} = \frac{1}{0}$$

We thus obtain

$$\begin{array}{cccccccc}
\frac{1}{1} & \frac{1}{0} & \frac{2}{1} & \frac{6}{2} & \frac{24}{9} & \frac{120}{44} & \cdot & \\
\frac{2}{1} & \frac{3}{1} & \frac{8}{3} & \frac{30}{11} & \cdot & \cdot & \cdot & \\
\frac{5}{2} & \frac{11}{4} & \frac{38}{14} & \frac{174}{64} & \cdot & \cdot & & \\
\frac{16}{6} & \frac{49}{18} & \frac{212}{78} & \cdot & \cdot & \cdot & & \\
\cdot & \cdot & \cdot & & & & &
\end{array}$$

The first column; $\frac{1}{1}, \frac{2}{1}, \frac{5}{2}, \frac{16}{6}$, etc., is the sequence of partial sums of the series $1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \cdots$ and thus converges to e. The first row, $\frac{1}{1}, \frac{1}{0}, \frac{2}{1}, \frac{6}{2}$, etc., is the sequence of the <u>reciprocals</u> of the partial sums of the <u>reciprocal</u> series $1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \cdots$ and thus converges to e.

The step-like paths:

$$\left.\begin{array}{l}
\frac{1}{1},\ \frac{1}{0},\ \frac{3}{1},\ \frac{8}{3},\ \frac{38}{14},\ \frac{174}{64},\ \text{etc.} \\
\frac{1}{1},\ \frac{2}{1},\ \frac{3}{1},\ \frac{11}{4},\ \frac{38}{14},\ \frac{212}{78},\ \text{etc.}
\end{array}\right\} \tag{109}$$

$$\frac{2}{1},\ \frac{5}{2},\ \frac{11}{4},\ \frac{49}{18},\ \frac{212}{78},\ \frac{1370}{504},\ \text{etc.}$$

are sequences of convergents of the following continued fractions for e:

$$\left.\begin{array}{l}
e = \frac{1}{1-}\ \frac{1}{1+}\ \frac{1}{2-}\ \frac{1}{3+}\ \frac{2}{4-}\ \frac{2}{5+}\cdots \\
e = 1 + \frac{1}{1-}\ \frac{1}{2+}\ \frac{1}{3-}\ \frac{2}{4+}\ \frac{2}{5-}\cdots
\end{array}\right\} \tag{110}$$

$$e = 2 + \frac{1}{2-}\ \frac{2}{3+}\ \frac{1}{4-}\ \frac{3}{5+}\ \frac{2}{6-}\ \frac{4}{7+}\ \frac{3}{8-}\cdots$$

and the zig-zag path:

$$\frac{1}{1},\ \frac{1}{0},\ \frac{2}{1},\ \frac{3}{1},\ \frac{8}{3},\ \frac{11}{4},\ \frac{38}{14},\ \frac{174}{64},\ \text{etc.} \tag{111}$$

gives the regular continued fraction (numerators all +1) for e

$$e = 1 + \frac{1}{0+} \frac{1}{1+} \frac{1}{1+} \frac{1}{2+} \frac{1}{1+} \frac{1}{1+} \frac{1}{4+} \cdots \tag{112}$$

In general, one may proceed through the table in any south-easterly direction and obtain a convergent sequence for e. This is a Padé Table for e. (See reference h.)

65. What is the connection between the Padé Table and the e_K and e_d transforms? Simply this: If we apply e_1 to the first column we obtain the second column, N_{i1}/D_{i1}, from i = 1 to i = ∞. If we apply e_2 to the first column we obtain the third column, N_{i2}/D_{i2} from i = 2 to i = ∞, and so on. This accounts for the half of the Padé Table below the diagonal. (This array should not be confused with the similar-looking array of $\bar{e}_1$ sequences discussed in paragraph 22.) The other half of the table may be obtained by the reciprocal of the e_1 transform of the reciprocal of the first row, the reciprocal of the e_2 transform of the reciprocal of the first row and so forth. The diagonal may be obtained by applying the diagonal transform, e_d, to the first column or by the reciprocal of the diagonal transform of the reciprocal of the first row. Thus the name, "diagonal" transform.

66. The Padé Table for the more general e^x is

$$\begin{array}{cccc}
\frac{1}{1} & \frac{1}{1-x} & \frac{2}{2-2x+x^2} & \cdot \quad \cdot \\[2ex]
\frac{1+x}{1} & \frac{2+x}{2-x} & \frac{6+2x}{6-4x+x^2} & \cdot \quad \cdot \\[2ex]
\frac{2+2x+x^2}{2} & \cdot & \cdot & \\
\cdot & & &
\end{array}$$

By definition, in the normal Padé Table we have those rational approximations for e^x which have the following two properties: (a) The fraction N_{rc}/D_{rc} is made up of a numerator polynomial of the r'th degree and a denominator polynomial of the c'th degree; (b) When this fraction is divided through we obtain the power series for e^x correct to the term involving x^{r+c}

67. The relation between this more general Padé Table and the transforms is the same as that for the simpler table for e. This, in fact, is quite simple to prove. Consider the sequence of partial sums of a power series

$$A_N = \sum_{i=0}^{N} c_i x^i \tag{113}$$

Then the e_K transform of A_N is

$$B_{KN} = \frac{\begin{vmatrix} A_{N-K} & A_{N-K+1} & \cdots & A_N \\ C_{N-K+1}X^{N-K+1} & \cdots & \cdots & C_{N+1}X^{N+1} \\ \vdots & & & \vdots \\ C_N X^N & \cdots & \cdots & C_{N+K}X^{N+K} \end{vmatrix}}{\begin{vmatrix} 1 & 1 & \cdots & 1 \\ C_{N-K+1}X^{N-K+1} & \cdots & \cdots & C_{N+1}X^{N+1} \\ \vdots & & & \vdots \\ C_N X^N & \cdots & \cdots & C_{N+K}X^{N+K} \end{vmatrix}} \tag{114}$$

By multiplying and dividing the rows and columns by powers of x, this becomes

$$B_{KN} = \frac{\begin{vmatrix} x^K A_{N-K} & x^{K-1} A_{N-K+1} & \cdots & x^0 A_N \\ C_{N-K+1} & C_{N-K} & \cdots & C_{N+1} \\ \vdots & & & \vdots \\ C_N & \cdots & \cdots & C_{N+K} \end{vmatrix}}{\begin{vmatrix} x^K & x^{K-1} & \cdots & x^0 \\ C_{N-K+1} & \cdots & \cdots & C_{N+1} \\ \vdots & & & \vdots \\ C_N & \cdots & \cdots & C_{N+K} \end{vmatrix}} \tag{115}$$

In general, then, unless the lower right minor vanishes, the numerator is a polynomial of the N'th degree, and the denominator is a polynomial of the K'th degree. By adding all the rows of the upper determinant of (114) together, it becomes

$$B_{KN} = \frac{\begin{vmatrix} A_N & A_{N+1} & \cdots & A_{N+K} \\ & \text{SAME AS (114)} & & \end{vmatrix}}{\begin{vmatrix} \text{SAME AS (114)} \end{vmatrix}} \tag{116}$$

We may subtract the value A_{N+K} from the first row (see 23b) and therefore

$$B_{KN} = A_{N+K} + \frac{\begin{vmatrix} (A_N - A_{N+K}) & (A_{N+1} - A_{N+K}) & \cdots & 0 \\ & \text{SAME AS (114)} & & \end{vmatrix}}{\begin{vmatrix} \text{SAME AS (114)} \end{vmatrix}} \tag{117}$$

Finally, by the above transformation, (115), we find

$$B_{KN} = A_{N+K} + \frac{\begin{vmatrix} (A_N - A_{N+K})x^K & (A_{N+1} - A_{N+K})x^{K-1} & \cdots & 0 \\ & \text{SAME AS (115)} & & \end{vmatrix}}{\begin{vmatrix} \text{SAME AS (115)} \end{vmatrix}} \tag{118}$$

Since the smallest possible exponent of x in the numerator determinant is N + K + 1 and since the denominator has a constant term (unless the lower left minor vanishes) B_{KN} agrees with A_{N+K} up to, at least, the term involving x^{N+K}. If the minors vanish, see paragraph 72 below.

68. In their application to the partial sums of power series the $e_{K,d}$ transforms, then, give the same results as the Padé Table. But the transforms are a broader conception since (a) they may be iterated, applied to other types of sequences, and (b) they may be iterated, giving $\tilde{e}_K$ and $\tilde{e}_d$.

69. An interesting, and perhaps useful, conception is that of a Padé Surface. Through the discrete points of a Padé Table, for instance, that for e, paragraph 64, we may conceive a smooth surface. This would be a two-dimensional generalization of our transient-like graphs of paragraph 4. The Padé Surface for e is interesting. It has a pole at (1,0), an oscillating character along its rows, a monotonic character along its columns and a limiting plane towards the southeast. For e^x we have a family of Padé Surfaces. We cannot, at this time, develop this concept further.

70. We have shown above how we may start with a power series, apply a diagonal transform, obtain a sequence of rational approximations, and from these a continued fraction. The order of these operations may be altered. An interesting way of obtaining the continued fraction is by the method of interpolation known as Thiele's reciprocal differences (reference g). 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 had been done by Nörlund (reference s). His determinants are essentially the same as our (115), etc.

71. When applied to a power series, then, the diagonal transform is essentially equivalent to Thiele's Continued Fractions. The former algorithm, however, is broader in scope, and in addition, the author believes, it has a simpler intuitive basis.

72. A third mathematic related to our transforms is a theorem of Kronecker (references i, t, u) on the power series of rational functions. THEOREM: If

$$D_\lambda^{(\mu)} = \begin{vmatrix} a_\lambda & a_{\lambda+1} & \cdots & a_{\lambda+\mu} \\ a_{\lambda+1} & & & \vdots \\ \vdots & & & \vdots \\ a_{\lambda+\mu} & \cdots & \cdots & a_{\lambda+2\mu} \end{vmatrix} \qquad (119)$$

the necessary and sufficient condition that $\sum_0^\infty a_n z^n$ should represent a rational function is that $D_0^{(n)} = 0$ for $n \geq N$.

73. This resembles our criterion of the order of a transient, (12), paragraph 14. We have already shown examples where a power series is summed exactly to its rational sum. This suggests that in general a power series which represents a rational function can always be summed exactly by a single e_K process of a sufficiently large K. This is true. For from the rationality of the sum it

follows that $D_0^{(n)} = 0$ for $n \geq N$ and from this it follows that $D_\lambda^{(n)} = 0$ for $n \geq N$, $\lambda \geq 0$ (reference i). Hence it follows that

$$\begin{vmatrix} a_\lambda x^\lambda & a_{\lambda+1} x^{\lambda+1} & \cdots \\ a_{\lambda+1} x^{\lambda+1} & & \\ \vdots & & \\ & & a_{\lambda+2n} x^{\lambda+2n} \end{vmatrix} = 0 \qquad (120)$$

for all λ. Therefore, from our criterion, (12), it follows that the sequence $A_i = a_i x^i$ may be represented exactly by a transient of the form, (3). Finally, it follows, by summation of the geometric series, that the partial sums, $\sum_{i=0}^{n} a_i x^i$ may be represented by a form (2).

74. An example of such exact summation is that of the Riemann Zeta function. This function, while transcendental for positive integer arguments, is rational for negative integer arguments and may be expressed (in Knopp's terminology) as

$$\zeta(-s) = -\frac{B_{s+1}}{s+1} = \frac{1-2^s+3^s-4^s+\cdots}{1-2\cdot 2^s} \qquad (121)$$

where the B's are the Bernoulli numbers (reference j, page 533). These numbers are all rational and we should be able to obtain them by summing exactly the divergent series:

$$\left.\begin{array}{l} 1-1+1-1+1-\cdots \\ 1-2+3-4+5-\cdots \\ 1-4+9-16+25-\cdots \quad \text{etc.} \end{array}\right\} \qquad (122)$$

Compare paragraph 38. In fact, we sum the first series by e_1, the second by e_2, and so on, and thus obtain explicit formulae for the Bernoulli numbers as the ratio of two determinants. These may then be reduced and we find that the partial sums of (122) should be weighted by the binomial coefficients. Therefore

$$B_1 = \frac{1}{1}\cdot\frac{1\cdot 1 + 1\cdot 0}{1+1} = \frac{1}{2}$$

$$B_2 = \frac{2}{3}\cdot\frac{1\cdot 1 + 2(-1) + 1\cdot 2}{1+2+1} = \frac{1}{6} \qquad (123)$$

$$B_3 = \frac{3}{7}\cdot\frac{1\cdot 1 + 3(-3) + 3\cdot 6 + 1(-10)}{1+3+3+1} = 0$$

and in general

$$B_n = \frac{n}{2^n - 1} \cdot \frac{\binom{n}{0}\cdot 1 + \binom{n}{1}\left(1-2^{n-1}\right) + \cdots + \binom{n}{n}\left(1-2^{n-1} + \cdots (-1)^n (n+1)^{n-1}\right)}{2^n} \qquad (124)$$

It is not contended that (124) is a practical way to calculate the Bernoulli numbers. Its chief interest is that it was obtained by the exact summation of divergent series.

VIII. Summary

75. We have introduced and developed an analogy between sequences and transients. On the basis of this analogy we have developed a uniform treatment for the evaluation of convergent and divergent sequences. Several nonlinear transforms have been developed and applied successfully to a large variety of sequences. The complete theory has not been developed, but some proofs are given and some connections with known algorithms are shown.

76. In a forthcoming memorandum the author will discuss further aspects of these transforms. These are:

A. A generalization of $\tilde{e}_1$ which is especially adapted for the summation of monotonic sequences where $\Delta A_n / \Delta A_{n-1} \to +1$

B. Occasional nonuniform convergence of $\tilde{e}_1$ to the wrong answer.

C. Occasional nonconvergence of $\tilde{e}_1$ of sequences associated with multivalued functions and their branch points.

D. The analysis of mathematical transients into their spectra, and the relation between discrete and continuous spectra.

E. A relation between the e_α transform and Gauss' method of numerical integration.

F. The prejudice against divergent sequences.

Daniel Shanks

D. Shanks

DS:ked