JOURNAL OF RESEARCH of the National Bureau of Standards—B. Mathematics and Mathematical Physics
Vol. 64B, No. 4, October-December 1960

Sequence Transformations Based on Tchebycheff Approximations

John R. Rice [1]

(May 17, 1960)

Let $P=\{p_n\}$ be a real sequence and define the transformed sequence $\mathcal{C}(P)$ as follows: Consider the segment, $p_m, p_{m+1}, p_{m+2}, p_{m+3}$ of P and determine a_m, b_m, and c_m so that $\max_n|a_mb_m^n+c_n-p_n|$, $n=m, m+1, m+2, m+3$ is minimized. The mth term of $\mathcal{C}(P)$ is taken to be c_m. The effect of the transformation $\mathcal{C}$ on sequences of the type $p_0+\sum_{i=1}^k \alpha_i\lambda_i^n$, $p_0+(-1)^n\sum_{i=1}^k a_in^i/\sum_{i=1}^l b_in^i$, and $p_0+\frac{a\lambda^n}{n}$ is considered. In each case $\mathcal{C}$ is shown to be very effective in accelerating convergence or decelerating divergence. For example, if the second sequence behaves as an^k as $n\rightarrow\infty$, then the transformed sequence behaves as an^{k-3}. A similar transformation $\mathcal{D}$ is defined by approximating in the Tchebycheff sense a segment of P by $ab_n \cos(\theta+n\phi)+c$. The effect of $\mathcal{D}$ is studied for sequences of the above type and also for $p_0+\Sigma_{i=1}^k\alpha_i\lambda_i^n \cos(\theta_i+n\phi_i)$. These sequence transformations are similar in nature to Aitken's δ^2-process and its generalization. A comparison of the two types of transformations is made. Several examples are given to illustrate the effect of $\mathcal{C}$ and $\mathcal{D}$ on various sequences.

1. Introduction

Let $P=\{p_n|n=1,2, \ldots\}$ be a real sequence and let $\mathcal{T}$ be a sequence transformation with

$$Q=\mathcal{T}(P)=\{t_n(p_1,p_2, \ldots, p_n)|n=1, 2, \ldots\}.$$

A large variety of such transformations have been studied, many in the modified form of summability methods for infinite series, which are designed to accelerate or induce the convergence of the sequence P. Such transformations may be based on approximating a segment of P by a function of a given form. Assume that p_n behaves approximately as $f(A,n)$ where A represents some parameters to be determined. Consider a segment $\{p_i|i=1,2, \ldots, m\}$ of P. The parameters A_m may be determined so as to minimize the error of the approximation in some sense. Since $f(A_m,n)$ is a known function of n we may set

$$t_m(p_1,p_2, \ldots, p_m)=\lim_{n\rightarrow\infty} f(A_m,n).$$

The transformations of this paper result from assuming p_n to behave as ab^n+c or $ab^n\cos(\theta+n\phi)+c$. The transformations are denoted by $\mathcal{C}$ and $\mathcal{D}$, respectively. The approximation to give the mth term of $\mathcal{C}(P)$ and $\mathcal{D}(P)$ is based on four or six, respectively, terms from P. The parameters are determined so as to minimize

$$\max_n|a_mb_m^n+c_m-p_n|, \qquad n=m, m+1, m+2, m+3$$

or

$$\max_n|a_mb_m^n\cos(\theta_m+n\varphi_m)+c_m-p_n|, \qquad n=m, \ldots, m+5$$

If $b_m<1$, then $\lim_{n\rightarrow\infty} f(A_m,n)$ is just c_m; if $b_m>1$ then this limit does not exist, but the value c_m is still taken, c_m being the *antilimit.*

Explicit formulas for c_m in terms of the p_i have been obtained in [5].[2] These formulas are not linear in the p_i and the transformations exhibit some marked differences from the classical linear transformations. In many cases $\mathcal{C}$ and $\mathcal{D}$ are much more powerful and in other cases P may converge quite rapidly and the transformed sequence may be identically infinite. Although $\mathcal{C}$ and $\mathcal{D}$ are not regular they do have a property of *joint convergence,* namely that if both P and the transformed sequences converge then they have the same limit.

It is not surprising that $\mathcal{C}$ and $\mathcal{D}$ are effective for sequences of exponential type

$$P=\left\{p_0+\sum_{i=1}^k \alpha_i\lambda_i^n|n=1, 2, \ldots\right\}. \tag{1}$$

If $|\lambda_i|>|\lambda_{i+1}|$ then $\mathcal{C}(P)$ converges or diverges as $(\lambda_2)^n$ instead of $(\lambda_1)^n$. $\mathcal{D}(P)$ behaves as $(\lambda_3)^n$ or $(\lambda_2^2/\lambda_1)^n$. $\mathcal{D}$ is effective for sequences of the more general type

$$P=\left\{p_0+\sum_{i=1}^k\alpha_i\lambda_i^n \cos(\theta_i+n\phi_i)|n=1, 2, \ldots\right\}. \tag{2}$$

and C is useful for such sequences in some cases.

These transformations are also very powerful for alternating rational sequences:

$$P=\left\{p_0+(-1)^n\sum_{i=1}^{k_1}a_in^i/\sum_{i=1}^{k_2}b_in^i|n=1, 2, \ldots\right\}. \tag{3}$$

[1] Present address, General Motors Corp., Warren, Mich. This work was done under a National Research Council-National Bureau of Standards Postdoctoral Research Associateship.

[2] Figures in brackets indicate the literature references at the end of this paper.

It is shown that $\mathcal{C}$ accelerates the convergence by a factor of $1/n^3$. That $\mathcal{D}$ accelerates the convergence even more can be seen from the examples although the actual factor of acceleration has not been determined. An example shows that $\mathcal{C}$ and $\mathcal{D}$ are very effective for a sequence of the form

$$P=\left\{\frac{(-1)^n\lambda^n}{n}\,|n=1, 2, \ldots\right\}. \tag{4}$$

The basic principle of these transformations is the same as that of Aitken's δ^2-process and its generalizations. $\mathcal{C}$ is the direct analog of the δ^2-process. Shanks [8] has given the most complete account of this theory. $\mathcal{C}$ and $\mathcal{D}$ have many properties in common with these transformations. The analogs of $\mathcal{C}$ and $\mathcal{D}$ in Shanks' work have the same power as $\mathcal{C}$ and $\mathcal{D}$ for sequences of type (1) and (2). For sequences of type (3), $\mathcal{C}$ and $\mathcal{D}$ are definitely more powerful. Aitken's δ^2-process accelerates the convergence of (3) by a factor of only $1/n^2$. $\mathcal{C}$ and $\mathcal{D}$ are much more effective for the sequence (4). There are sequences for which the δ^2-process and its generalizations are more effective, but as a general guide $\mathcal{C}$ and $\mathcal{D}$ are as effective or more effective than their analogs in Shanks' paper.

A comparison of this paper with Shanks' will indicate that $\mathcal{C}$ and $\mathcal{D}$ have many properties which have not been developed. The main purpose here has been to establish these transformations as tools of numerical analysis and not of analysis. When reading Shanks' paper one has the feeling that much remains to be done in the study of these transformations and that a complete understanding of their behavior for sequences of real and complex functions will require a penetrating analysis.

It is typical but somewhat disappointing that these transformations may be most effective for easy problems and least effective for difficult problems. If one is solving three simultaneous linear equations by Gauss-Seidel iteration and the characteristic values determining the rate of convergence are .95, .5, and .1, then these transformations increase the rate of convergence dramatically. But if one has 100 equations with 10 characteristic values between .99 and .995, then the elimination of even the five largest ones is not very significant. On the other hand if the largest characteristic value is .995 and the others are, say, less than .9 the transformations are very effective in accelerating the convergence.

2. Derivation of the Transformations

2.1. Preliminaries

Sequences are denoted by P, Q, . . . and their elements by p_n, q_n, ... A sequence transformation is denoted by a script letter as $\mathcal{C}$ and $\mathcal{D}$. Braces, { }, denote a sequence or set and $\{x|\ \ldots\}$ is read "the set of x such that . . ." We denote by $O(x)$ and $o(x)$ two functions such that

$$\lim_{x\to\infty}\frac{o(x)}{x}=0$$

and, for x sufficiently large,

$$\left|\frac{O(x)}{x}\right|\leq\text{constant}.$$

Let P be a sequence of real numbers. A transformed sequence $Q=\{q_n|n=1, 2, \ldots\}$ of P, is obtained as follows: Take a segment, $\{p_i|i=m, \ldots, m+k\}$ of P, and approximate the values p_n as a function of n by ab_n+c or $ab_n\cos(\theta+n\phi)+c$. The corresponding c value is then assigned to q_m.

For these particular transformations to be effective, the sequence P must, in some sense, behave exponentially. As an example of such sequences consider the real k-vectors defined by

$$v_{n+1}=Av_n,$$

where A is a real $k\times k$ matrix. Let v^* be the solution of

$$(A-I)v^*=0.$$

Then $v_{n+1}-v^*=A_n(v_0-v^*)$. The sequence of vectors defined will converge to v^* if $A^n(v_0-v^*)$ tends to zero as n tends to infinity. If A is a normal matrix with characteristic vectors u_i and distinct real characteristic values λ_i, then

$$A^n(v_0-v^*)=\sum_{i=1}^{k}\alpha_i\lambda_i^n u_i.$$

If the λ_ir are complex then

$$A^n(v_0-v^*)=\sum\alpha_i|\lambda_i|^n\cos\,(\theta_i+n\phi_i)u_i.$$

2.2. Tchebycheff Approximations

We wish to approximate a segment of P by ab^n+c. The use of least squares approximation appears to be impossible due to the difficulty of the nonlinear equations involved. In [5] the theory of approximation by ab^x+c in the Tchebycheff sense is developed in detail. The following result is given. Let p_m, p_{m+1}, p_{m+2}, and p_{m+3} be four consecutive values of P to be approximated. Then c_m of the best approximation, $a_m(b_m)^n+c_m$, is given by

$$c_m=\frac{(p_m+p_{m+1})(p_{m+2}+p_{m+3})-(p_{m+1}+p_{m+2})^2}{2(p_m-p_{m+1}-p_{m+2}+p_{m+3})}. \tag{5}$$

Let p_m, . . . , p_{m+5} be consecutive values from P to be approximated by $ab^n\cos(\theta+n\phi)+d$. The theory of Tchebycheff approximation by $ab^x\cos(\theta+x\phi)+d$ is not complete and an explicit formula for d_m of the best approximation has not been rigorously established. However it is conjectured that the following formula is valid. Set $s_i=\frac{1}{2}(p_{m+i}+p_{m+i+1})$, $i=0, 1, \ldots, 4$. Then

$$d_m=\frac{2s_1s_2s_3+s_4(s_0s_2-s_1^2)-s_0s_3^2-s_2^3}{2s_2(s_0+2s_1-3s_2+2s_3+s_4)-2(s_1-s_3)^2-4(s_1s_4+s_0s_3)+2s_0s_4}. \tag{6}$$

This formula is based on the assumption that the best Tchebycheff approximation is characterized by the alternation five times of the error function. See [6] for a discussion of the characterization of best Tchebycheff approximations.

Note that (6) may be written in a much simpler form. Let $s_i'=(s_i-s_2)$, $i=0,1,3,4$; then

$$d_m=s_2-\frac{s_4'(s_1')^2+s_0'(s_3')^2}{s_0's_4'-(s_1'-s_3')^2-2(s_1's_4'+s_3's_0')}.$$

2.3. The Transformations $\mathcal{C}$ and $\mathcal{D}$

Let P be given, then a new sequence Q is found as follows: For each m such that the denominator of (5) is nonzero $q_m=c_m$; if the denominator of (5) is zero and the numerator does not vanish then we assign $q_m=\infty$; if both the numerator and denominator are zero $q_m=p_m$. This transformation of P into Q may be written in operator form as

$$Q=\mathcal{C}(P).$$

A transformation $\mathcal{D}$ is defined in a similar manner by (6).

The repeated application k times of $\mathcal{C}$ and $\mathcal{D}$ is denoted by $\mathcal{C}^k(P)$ and $\mathcal{D}^k(P)$.

3. Analysis of the Transformations

3.1. Algebraic Properties of $\mathcal{C}$

The study of the properties of these transformations begins with a simple algebraic property of $\mathcal{C}$.

If $P=\{p_n\}$ then $aP+b$ is defined as $\{ap_n+b|n=1,2,\ldots\}$. The following result may be established by direct computation from (5).

THEOREM 1: $\mathcal{C}(aP+b)=a\mathcal{C}(P)+b.$

It is not true that $\mathcal{C}(P_1+P_2)=\mathcal{C}(P_1)+\mathcal{C}(P_2)$.

3.2. Transformation of Exponential Sequences by $\mathcal{C}$

Let P be a real sequence of the form

$$\left\{p_n=p_0+\sum_{i=1}^{k}\alpha_i\lambda_i^n|n=1,2,\ldots\right\}$$

with $|\lambda_i|>|\lambda_{i+1}|$. If $|\lambda_1|>1$, P does not converge but *diverges from* p_0. In such a case p_0 is the *antilimit* of P. This sequence is of an exponential type and $\mathcal{C}$ should be effective in increasing the rate of convergence or decreasing the rate of divergence.

Let $P'=P-p_0$; then

$$\mathcal{C}(P)=\mathcal{C}(P')+p_0.$$

The general term of $\mathcal{C}(P')$ may be explicitly computed to be

$$\frac{\sum_{i<j}\alpha_i\alpha_j\lambda_i^n(\lambda_j/\lambda_1)^n(1+\lambda_i)(1+\lambda_j)(\lambda_i-\lambda_j)^2}{\alpha_1(1-\lambda_1-\lambda_1^2+\lambda_1^3)+\sum_{i=2}^{k}\alpha_1(\lambda_i/\lambda_1)^n(1-\lambda_i-\lambda_i^2+\lambda_1^3)}.$$

The denominator of this expression may have only a finite number of zeros as a function of n. We may write the denominator as

$$\alpha_1(1-\lambda_1-\lambda_1^2+\lambda_1^3)+o(1)$$

since $|\lambda_j/\lambda_1|<1$ for $j>1$.

The largest term in the numerator is λ_2^n. Others among the larger terms are

$$\lambda_3^n,\quad(\lambda_2\lambda_3/\lambda_1)^n.$$

Thus $\mathcal{C}$ is seen to eliminate the largest exponential term from P'.

The repeated application of $\mathcal{C}$ will eliminate the largest remaining term. The largest term in $\mathcal{C}^2(P')$ is λ_3^n.

The above analysis has established:

THEOREM 2: *Let* $P=\{p_0+\sum_{i=1}^{k}\alpha_i\lambda_i^n\}$, $|\lambda_i|>|\lambda_{i+1}|$; *then*

$$\mathcal{C}(P)=\left\{p_0+\frac{\lambda_2^n\sum_{i=0}^{k'}d_i(\mu_i)^n}{1+o(1)}\right\}\qquad \begin{matrix}\mu_0=1\\ |\mu_i|<1,\quad i\geq 1.\end{matrix}$$

Thus $\mathcal{C}$ is seen to have a desirable effect on sequences of this type, which was to be expected. In the next section it is seen that $\mathcal{C}$ is effective for some sequences of a completely different nature.

3.3. Transformations of Alternating Rational Sequences by $\mathcal{C}$

Let P be a sequence of the form

$$\left\{p_n=p_0+(-1)^n\sum_{i=0}^{k_1}a_in^i\Big/\sum_{i=0}^{k_2}b_in^i\right\}.$$

We shall consider $\mathcal{C}(P)$.

First take the special case

$$P=\{p_0+(-1)^nan^k\}.$$

The nth term of $\mathcal{C}(P)$ is

$$p_0+a\frac{(n^k-(n+1)^k)((n+2)^k-(n+3)^k)-((n+1)^k-(n+2)^k)^2}{2(-1)^n(n^k+(n+1)^k-(n+2)^k-(n+3)^k)}.$$

After some manipulation this term is seen to be of the form

$$p_0+\frac{an^{2k-4}k^2(k-1)}{8(-1)^n kn^{k-1}}+O(n^{k-4})$$

or

$$p_0+a(-1)^n n^k\left[\frac{k(k-1)}{8n^3}+O\left(\frac{1}{n^4}\right)\right].$$

Now consider the general situation. p_n may be written as

$$p_n=p_o+\alpha_0' n^k+\sum_{i=1}^{\infty}\alpha_1' n^{k-i} \qquad k=k_1-k_2$$

for n sufficiently large. It is clear that the denominator of c_m is $8(-1)^n kn^{k-1}+O(n^{k-2})$ since it is linear in the terms of P. The denominator of c_m has only a finite number of zeros as a function of m. The terms in the numerator are of the form

$$\alpha_i'\alpha_j'[n^{k-i}-(n+1)^{k-i}][(n+2)^{k-j}-(n+3)^{k-j}]$$
$$+\alpha_i'\alpha_j'[n^{k-j}-(n+1)^{k-j}][(n+2)^{k-i}-(n+3)^{k-i}]$$
$$-2\alpha_i'\alpha_j'[(n+1)^{k-i}-(n+2)^{k-i}][(n+1)^{k-j}-(n+2)^{k-j}].$$

By direct calculation it is seen that this term is $O(n^{2k-i-j-4})$.

Thus we have established:

THEOREM 3: *Let* $P=\{p_0+(-1)^n\sum_{i=0}^{k_1}a_i n^i/\sum_{i=0}^{k_2}b_i n^i\}$; *then for m sufficiently large the mth term of* $\mathcal{C}(P)$ *is*

$$p_0+\sum_{i=1}^{\infty}\alpha_i'(-1)^m m^{k-i-3}$$

where $k=k_1-k_2$, $\alpha_0'=a_{k_1}k(k-1)/8b_{k_2}$.

Examples of sequences which are approximately of this type are given by the partial sums of alternating infinite series of rational functions of n. One can consider such sequences directly and establish an analog of theorem 3.

3.4. Regularity and Joint Convergence

A transformation $\mathcal{T}$ is said to be regular if the existence of $\lim_{n\to\infty} p_n$ implies $\lim_{n\to\infty} q_n=\lim_{n\to\infty} p_n$ where $Q=\mathcal{T}(P)$. Since $\mathcal{C}$ is a transformation designed to accelerate or induce convergence one would like for $\mathcal{C}$ to be regular. That $\mathcal{C}$ is not regular is seen by the following example:

$$P=\{p_n|p_{2k}=1+1/k, \quad p_{2k+1}=1+1/k\}.$$

Every other term of $\mathcal{C}(P)$ is infinite, yet P converges. Although $\mathcal{C}$ is not a regular transformation it does have the property of *joint convergence.* A transformation $\mathcal{T}$ is said to have the property of joint convergence if the convergence of P and $\mathcal{T}(P)$ imply that P and $\mathcal{T}(P)$ have the same limit. The proof of the following theorem will be given in section 5.1 (see theorem 10).

THEOREM 4: *If both P and* $\mathcal{C}(P)=Q$ *converge then*

$$\lim_{n\to\infty} p_n=\lim_{m\to\infty} q_m.$$

3.5. Properties of the Transformation $\mathcal{D}$

The following theorem gives algebraic properties of $\mathcal{D}$ similar to those of $\mathcal{C}$. The proof is by direct evaluation.

THEOREM 5: $\mathcal{D}(aP+b)=a\mathcal{D}(P)+b$.

Consider the real sequence

$$P=\left\{\sum_{i=1}^{k}a_i\lambda_1^n\cos(\theta_i+n\phi_i)\right\}$$

with $|\lambda_i|>|\lambda_{i+1}|$. Expression (6) may be directly evaluated as the ratio of the following two terms:

$$\sum_{|i-j|+|j-l|>0} a_i a_j a_l(\lambda_i\lambda_j\lambda_l/\lambda_1^2)^n$$
$$\times\cos[\theta_i+\theta_j+\theta_l+n(\phi_i+\phi_j+\phi_l)]H_{ijl}$$

and

$$\sum_{i=1,j=1}^{k}a_i a_j(\lambda_i\lambda_j/\lambda_1^2)^n G_{ij},$$

where H_{ijl} and G_{ij} are functions of the λ's, θ's, and ϕ's that are independent of n. H_{ijl} and G_{ij} are trigonometric functions with coefficients depending on powers of the λ's.

An examination of the above expression and a simplification of notation leads to the following theorem.

THEOREM 6. *Let* $P=\{p_0+\sum_{i=1}^{k}a_i\lambda_i^n\cos(\theta_i+n\phi_i)\}$, $|\lambda_i|>|\lambda_{i+1}|$. *Then the nth term of* $\mathcal{D}(P)$ *is*

$$p_0+\frac{\lambda_2^n\sum_{i=0}^{m}b_i\mu_i^n\cos(\alpha_i+n\beta_i)}{1+o(1)} \qquad \begin{array}{l}\mu_0=1\\ |\mu_i|<1, \quad i\geq 1\end{array}$$

except for possibly a finite number of terms.

One may also establish the following result:

THEOREM 7: *Let* $P=\{p_0+\sum_{i=1}^{k}a_i\lambda_i^n\}$, $|\lambda_i|>|\lambda_{i+1}|$. *Then the nth term of* $\mathcal{D}(P)$ *is*

$$\frac{p_0+(\lambda')^n\sum_{i=0}^{m}b_i\mu_i^n}{1+o(1)} \qquad \begin{array}{l}\mu_0=1\\ |\mu_i|<1,\ i\geq 1\\ \lambda'=\max(\lambda_3,\ \lambda_2^2/\lambda_1)\end{array}$$

except for possibly a finite number of terms.

No attempt has been made to analyze the effect of $\mathcal{D}$ on alternating rational sequences; however the examples indicate that $\mathcal{D}$ is quite effective for increasing the rate of convergence.

It is a simple matter to construct sequences for which $\mathcal{D}$ is not regular. It is probable that $\mathcal{D}$ has the joint convergence property, but this has not been established.

4. Comparison With Aitken's δ^2-process and its Generalization

The transformations $\mathcal{C}$ and $\mathcal{D}$ are similar to a family of sequence transformations first studied by Aitken [1,2]. These transformations have been studied since by Lubkin [4], Samuelson [7], and Shanks [8]. The most general and complete treatment is by Shanks. A brief outline of these transformations will be given.

Let P be given and set

$$\Delta p_n = p_{n+1} - p_n.$$

From P new sequences $Q_k = \{q_{k,n} | n=k, k+1, \ldots\}$ are formed by

$$q_{k,n} = \frac{\begin{vmatrix} p_{n-k} & \cdots & p_{n-1} & p_n \\ \Delta p_{n-k} & \cdots & \Delta p_{n-1} & \Delta p_n \\ \Delta p_{n-k-1} & \cdots & \Delta p_n & \Delta p_{n+1} \\ \vdots & & & \vdots \\ \Delta p_{n-1} & \cdots & & \Delta p_{n+k+1} \end{vmatrix}}{\begin{vmatrix} 1 & \cdots & 1 & 1 \\ \Delta p_{n-k} & \cdots & \Delta p_{n-1} & \Delta p_n \\ \Delta p_{n-k-1} & \cdots & \Delta p_n & \Delta p_{n+1} \\ \vdots & & & \vdots \\ \Delta p_{n-1} & \cdots & & \Delta p_{n+k-1} \end{vmatrix}}.$$

This defines a set of sequence transformations $\mathcal{E}_k$, by

$$Q_k = \mathcal{E}_k(P).$$

Shanks studies in some detail the transformations $\mathcal{E}_k, \mathcal{E}_1^k, \mathcal{E}_1' = \{p_1, q_{12}, q_{13}^{(2)}, \ldots, q_{1,k+1}^{(k)}, \ldots\}$, and $\mathcal{E}_d = \{q_{k,k} | k=0,1,2, \ldots\}$. $q_{1n}^{(k)}$ denotes the nth term of $\mathcal{E}_1^k(P)$. q_{1n} is given explicitly by

$$p_{1n} = \frac{p_{n+1}p_{n-1} - p_n^2}{p_{n+1} + p_{n-1} - 2p_n},$$

which defines Aitken's δ^2-process.

These transformations are derived heuristically as follows: Given a segment of P, $p_{n-k}, \ldots, p_{n+k}$, assume that this segment is exactly of the form $\sum_{i=1}^{k} a_i b_i^n + c$. Then define $q_{k,n}$ to be this c. The basic principle is the same as for $\mathcal{C}$ and $\mathcal{D}$, but the determination of the approximation is different.

If $S = \{s_n | s_n = \frac{1}{2}(p_n + p_{n+1})\}$ then it is seen that

$$\mathcal{C}(P) = \mathcal{E}_1(S).$$

Likewise, a straightforward calculation shows

$$\mathcal{D}(P) = \mathcal{E}_2(S).$$

The reason for this relationship is simple. $\mathcal{E}_k$ has been found by solving eq (8) for c:

$$\sum_{i=1}^{k} a_i b_i^m + c = p_m, \qquad m = n-k, \ldots, n+k. \tag{8}$$

The fact that best Tchebycheff approximations are characterized by the alternation $k+1$ times of the error function leads to the following equations:

$$\sum_{i=1}^{k} a_i b_i^m + c = p_m + (-1)^m e, \qquad m = n, \ldots, n+2k+2$$

where e is the error of the best approximation. If one averages the mth and $(m+1)$st equations, then

$$\sum_{i=1}^{k} \frac{a_i(1+b_i)}{2} b_i^m + c = \frac{1}{2}(p_m + p_{m+1}), \qquad m = n, \ldots, n+2k+1.$$

These equations are of the same form as (8) except P has been replaced by S.

Aitken considers the application of $\mathcal{E}_1$ to sequences of the form $P = \{p_0 + \sum_{i=1}^{k} \alpha_i \lambda_i^m\}$ and establishes:

THEOREM 8 (Aitken): *Let* $P = \{p_0 + \sum_{i=1}^{k} \alpha_i \lambda_i^m\}$, $|\lambda_i| > |\lambda_{i+1}|$; *then*

$$\mathcal{E}_1(P) = \left\{ p_0 + \frac{(\lambda_2)^n \sum_{i=0}^{m} d_i \mu_i^n}{1 + o(1)} \right\} \qquad \begin{matrix} \mu_0 = 1 \\ |\mu_i| < 1 \quad i \geq 1. \end{matrix}$$

Thus for sequences of this type $\mathcal{E}_1$ and $\mathcal{C}$ are equally effective in order of magnitude. The coefficients of $(\lambda_2)^n$ for $\mathcal{E}_1$ and $\mathcal{C}$ are, respectively

$$\frac{\alpha_1\alpha_2(\lambda_1 - \lambda_2)^2}{(1-\lambda_1)^2}, \qquad \frac{\alpha_1\alpha_2(1+\lambda_2)(\lambda_1-\lambda_2)^2}{(1-\lambda_1)^2}.$$

$\mathcal{E}_1$ appears to be slightly superior if $\lambda_2 > 0$. Examples indicate that either $\mathcal{E}_1$ or $\mathcal{C}$ may be better but in any case the difference in the rates of convergence is small.

For alternating rational sequences the two transformations are no longer similar in effect. Shanks has established:

THEOREM 9 (Shanks): *Let*

$$P = \left\{ p_0 + (-1)^n \sum_{i=0}^{k_1} a_i n^i \Big/ \sum_{i=0}^{k_2} b_i n^i \right\},$$

$a_{k_1} \neq 0$, $b_{k_2} \neq 0$ *and* $k = k_1 - k_2$. *Then the* nth *term oj* $\mathcal{E}_1(P)$ *is*

$$p_0 + \frac{a_{k_1} k (-1)^n n^{k-2}}{4 b_{k_2}} + O(n^{k-3})$$

with possibly a finite number of exceptions.

Thus $\mathcal{E}_1$ accelerates the convergence by a factor of $1/n^2$ whereas $\mathcal{C}$ accelerates it by a factor of $1/n^3$.

The comparative effectiveness of $\mathcal{E}_2$ and $\mathcal{D}$ may be judged from the examples given later.

Theorem 10 shows that $\mathcal{E}_1$ has the property of joint convergence [4].

THEOREM 10 (Lubkin): *If both* P *and* $\mathcal{E}_1(P)=Q$ *converge, then*

$$\lim_{n\to\infty} p_n=\lim_{n\to\infty} q_n$$

If P converges then so does $S=\{\tfrac{1}{2}(p_n+p_{n+1})\}$ and

$$\lim_{n\to\infty} p_n=\lim_{n\to\infty} s_n$$

Thus theorem 4 follows directly from theorem 10.

5. Examples

Some examples are given which illustrate various characteristics of $\mathcal{C}$ and $\mathcal{D}$. In some cases comparisons are made with $\mathcal{E}_1$ and $\mathcal{E}_2$. For more general comparisons see [4] where $\mathcal{E}_1$ is compared with several classical linear sequence transformations for a large number of examples.

The first two sets of examples illustrate the general behavior of $\mathcal{C}$ and $\mathcal{D}$ for various exponential and alternating rational sequences. Some comparisons are made with $\mathcal{E}_1$ and $\mathcal{E}_2$ here. The next set of examples shows the effect of increasing the largest λ in exponential sequences. The following two sets show the effects of lower order terms and the loss of significant figures. Detailed comments are given with each set of examples.

All of these examples were computed using double precision arithmetic with about 20 significant decimal digits.

5.1. Effect of $\mathcal{C}$ and $\mathcal{D}$ on Exponential Sequences

The following five examples are considered:

$$P_1=\{2^n+(1.2)^n+(.9)^n+(.8)^n\},$$

$$P_2=\{(.98)_n+(.95)^n+(.5)^n+(.2)^n\},$$

$$P_3=\{(.98)^n+(.95)^n\},$$

$$P_4=\{(.9)^n+(.8)^n\cos(.1n)\},$$

$$P_5=\{(.9)^n\cos(1.5n)+(.5)^n\cos(.25n)\}.$$

Table 1 gives results of $\mathcal{C}$, $\mathcal{C}^2$, $\mathcal{C}^3$, $\mathcal{C}^4$, and $\mathcal{D}$ for 33 terms of P_1. P_1 diverges quite rapidly and $\mathcal{C}(P_1)$ diverges at a much slower rate. The factors in $\mathcal{C}(P_1)$ larger than .5 are 1.2, .9, .8, .72, and .54. $\mathcal{C}^2(P_1)$ converges as can be seen from the list of factors in $\mathcal{C}^2(P_1)$ larger than .5. They are .9, .8, .72, .68, .6, and .54. The four largest factors of $\mathcal{C}_3(P_2)$ are .8, .72, .71, and .68 are the largest factor of $\mathcal{C}^4(P_1)$ is .72.

TABLE 1.

P_1	$\mathcal{C}(P_1)$	$\mathcal{C}^2(P_1)$	$\mathcal{C}^3(P_1)$	$\mathcal{C}^4(P_1)$	$\mathcal{D}(P_1)$
4.90000000000000	2.9151478	2.7270889	2.7433842	2.7029023	2.6785648
6.89000000000000	2.7741418	2.7426402	2.8808482	3.0727475	2.7348679
10.9690000000000	2.7169114	2.6757849	3.2921161	69.180737	2.7088105
19.1393000000000	2.7430500	2.5368629	4.8080981	72.818042	2.6047836
35.4064900000000	2.8532800	2.3466523	140.43599	−.54172894	2.4392636
67.7795690000000	3.0509855	2.1287500	−2.4381772	−.04037382	2.2345869
132.271192900000	3.3426610	1.9033466	−.76363656	.04995813	2.0125094
260.898056330000	3.7381380	1.6847333	−.28755651	.06710598	1.7902083
517.681418569000	4.2508558	1.4814607	−.09493555	.06620907	1.5791104
1030.64778904490	4.8982545	1.2976421	−.00731646	.06136882	1.3855556
2055.82979364889	5.7023160	1.1344196	.03331326	.05701478	1.2121566
4105.26724946147	6.6902694	.99116645	.05084351	.05513287	1.0591391
8203.00848270213	7.8954805	.86633340	.05650783	.05808964	.92537882
16397.1119330351	9.3585517	.75798632	.05602878	.07372670	.80910184
32783.6480970788	11.128667	.66412067	.05246752	.17953296	.70831457
65554.7018754061	13.265229	.58282968	.04750712	−.11327597	.62104494
131094.375400883	15.839833	.51238289	.04208180	−.03107621	.54546427
262170.791442315	18.938651	.45125306	.03670509	−.01478933	.47993879
524320.097496628	22.665309	.39811524	.03164886	−.00839710	.42304196
1048614.47070579	27.144335	.35183192	.02704444	−.00521395	.37354352
2097198.12376227	32.525314	.31143263	.02294119	−.00341676	.33037177
4194359.31199968	38.987872	.27609212	.01934122	−.00231783	.29269020
8388674.34190501	46.747669	.24510968	.01622020	−.00160176	.25998290
16777295.5813360	56.063584	.21789028	.01353993	−.00114340	.23010570
33554527.4717843	67.246349	.19392818	.01125615	−.00081105	.20587142
67108978.5430931	80.668906	.17279283	.00932342	−.00056122	.18756025
134217865.431120	96.778842	.15411687	.00769611	−.00044434	.14815635
268435620.898931	116.11330	.13758611	.00633423	−.00029985	.12460304
536871109.862244	----------	----------	----------	.00001497	----------
1073742061.41994	----------	----------	----------	----------	----------
2147483932.89072	----------	----------	----------	----------	----------
4294967637.85702	----------	----------	----------	----------	----------
8589935002.21781	----------	----------	----------	----------	----------

It is seen that the first values of $\mathcal{C}^k(P_1)$ behave erratically as k increases. $\mathcal{C}^4(P_1)$ is a very erratic function of n for small n. This is the typical behavior and indicates that care must be exercised in using repeated applications of $\mathcal{C}$.

The largest factor in $\mathcal{D}(P_1)$ is .9 and hence $\mathcal{D}(P_1)$ converges.

Only selected values from the other examples are given. In table 2 it is seen that $\mathcal{C}$ is not very effective in increasing the rate of convergence of P_2.

Table 2

$P_2=\{(.98)^n+(.95)^n+(.5)^n+(.2)^n\}$

n	P_2	$\mathcal{C}(P_2)$	$\mathcal{C}^2(P_2)$	$\mathcal{D}(P_2)$
1	2.6300	1.6876	3.0466	1.2528
10	1.4168	.32965	.20621	.21299
20	1.0261	.18278	--------	.02770

This is due to the closeness of the two largest factors. In the next table it is seen that even with the lower order term deleted the convergence is quite slow even for $\mathcal{C}^4(P_3)$.

Table 3

$P_3=\{(.98)^n+(.95)^n\}$

n	P_3	$\mathcal{C}(P_3)$	$\mathcal{C}^2(P_3)$	$\mathcal{C}^3(P_3)$	$\mathcal{C}^4(P_3)$	$\mathcal{D}(P_3)$
1	1.93	.29914	−.07157	−.08733	−.07864	.0
25	.88085	.15891	−.04624	−4.9028	.36843	.0
40	.57421	.10160	−.02795	.02408	.00322	.0
60	.34362	.05175	−.01084	.00256	.00004	.0

The next table shows that $\mathcal{D}$ is fairly effective on P_4. It is to be noted however that convergence is no longer monotonic even in the large range considered. The denominator has a zero between the 67th and 68th terms of $\mathcal{D}(P_4)$. This same phenomenon is present in $\mathcal{D}^2(P_4)$.

Table 4

$P_4=\{(.9)^n+(.8)^n\cos(.1n)\}$

n	P_4	$\mathcal{C}(P_4)$	$\mathcal{D}(P_4)$	$\mathcal{D}^2(P_4)$
1	1.6960	.02354	.08971	.05260
8	.54735	.07107	.11293	−.47672
10	.40669	.06085	−1.8972	−.00681
20	.11678	.01132	−.00282	−.00287
30	.04117	−.00036	−.00051	−.00048
35	.02465	−.00042	−.00151	.00005
50	.00516	−.00002	3×10^{-6}	2×10^{-7}
60	.00180	4×10^{-8}	5×10^{-7}	3×10^{-7}
67	.00086	3×10^{-8}	3×10^{-6}	6×10^{-7}

The final table shows that $\mathcal{D}$ is very effective for P_5, while $\mathcal{C}$ is not.

Table 5

$P_5=\{(.9)^n\cos(1.5n)+(.5)^n\cos(.25n)\}$

n	P_5	$\mathcal{C}(P_5)$	$\mathcal{D}(P_5)$	$\mathcal{D}^2(P_5)$
1	.54812	−.11751	.10538	−.00135
5	.21454	−.21059	.00017	-1×10^{-6}
10	−.26567	3.2919	−.00022	-6×10^{-8}
15	−.17983	.07227	-5×10^{-6}	5×10^{-11}
20	.01875	.06441	1×10^{-7}	6×10^{-14}
25	.07037	−.04624	7×10^{-9}	--------

5.2. Effect of $\mathcal{C}$ and $\mathcal{D}$ on Alternating Sequences

Consider the sequences of partial sums of the following series:

$$\pi=4\sum_{i=1}^{\infty}\frac{(-1)^{i+1}}{2i-1}=3.1415926,$$

$$ln\,2=\sum_{i=1}^{\infty}\frac{(-1)^{i}}{i}=.6931472,$$

$$ln\,3=\sum_{i=1}^{\infty}\frac{(-1)^{i}2^{i}}{i}=1.098612.$$

These sequences are denoted by A_1, A_2, and A_3 respectively.

The first two examples are alternating rational sequences and tables 6 and 7 show the results for $\mathcal{C}$ and $\mathcal{E}_1$ to be as predicted by theorems 3 and 9. $\mathcal{D}(A_1)$, $\mathcal{D}(A_2)$, $\mathcal{E}_2(A_1)$, and $\mathcal{E}_2(A_2)$ converge much more rapidly and apparently $\mathcal{D}$ accelerates the convergence better than $\mathcal{E}_2$ by a factor in $1/n$.

Table 6

n	A_1	$\mathcal{C}(A_1)$	$\mathcal{E}_1(A_1)$	$\mathcal{D}(A_1)$	$\mathcal{E}_2(A_1)$
1	4.0	3.1466666	3.1666667	3.1417367	3.1423423
2	2.6666667	3.140136	3.1333333	3.1415567	3.1413919
3	3.4666667	3.1421516	3.1452381	3.1416042	3.1416627
4	2.8952381	3.1413354	3.1396825	3.1415882	3.1415634
5	3.3396825	3.1417266	3.1427129	3.1415946	3.1416065
6	2.9760462	3.1415163	3.1408813	--------	3.1415854
7	3.2837385	3.1416392	3.1420718	--------	--------
8	3.0170718	--------	3.1412548	--------	--------
9	3.2523659	--------	--------	--------	--------
10	3.0418396	--------	--------	--------	--------
error of last term	.099753	.0000466	.0003378	.0000020	.0000072

Table 7

n	A_2	$\mathcal{C}(A_2)$	$\mathcal{E}_1(A_2)$	$\mathcal{D}(A_2)$	$\mathcal{E}_2(A_2)$
1	1.0	.69444444	.7	.69318182	.69333333
2	.5	.69270833	.69047619	.69313724	.69308943
3	.83333333	.69333333	.69444444	.69315068	.69316940
4	.58333333	.69305555	.69242424	.69314573	.69313724
5	.78333333	.69319726	.69358974	.69314783	.69315212
6	.61666666	.69311755	.69285712	.69314682	.69314448
7	.75952379	.69316576	.69334732	.69314734	.69314872
8	.63452379	.69313490	.69300331	.69314706	.69314621
9	.74563490	.69315556	.69325395	.69314723	.69314778
10	.64563490	.69314123	.69306573	.69314713	.69314677
11	.73654400	.69315149	.69321067	.69314719	.69314745
12	.65321067	.69314395	.69309670	.69314715	.69314698
13	.73013375	.69314961	.69318793	.69314719	.69314731
14	.65870517	.69314529	.69311377	--------	.69314707
15	.72537184	.69314865	.69317488	--------	--------
16	.66287184	--------	.69312394	--------	--------
17	.72169538	--------	--------	--------	--------
18	.66613982	--------	--------	--------	--------
error of last term	.0220074	.0000015	.0000232	.00000001	.00000010

It is remarkable that π can be found so accurately from a few terms of the simple sequence A_1. Shanks considers $\mathcal{E}_1^k(A_1)$ and π is given to 8 places by the first term of $\mathcal{E}_1^4(A_1)$. It would require approximately 40,000,000 terms of A_1 to give the same accuracy.

The third example is a "mixed" sequence. All of the transformed sequences diverge, but they are *semiconvergent* or *asymptotic* [3, pp. 520, 536] sequences for $ln\ 3$. In this example $\mathcal{C}$ and $\mathcal{D}$ show a very marked improvement over $\mathcal{E}_1$ and $\mathcal{E}_2$. It appears that round off has begun to affect $\mathcal{D}(A_2)$. Although $\mathcal{D}(A_2)$ was computed using double precision, A_2 itself was computed to only eight significant figures.

TABLE 8

n	A_3	$\mathcal{C}(A_3)$	$\mathcal{E}_1(A_3)$	$\mathcal{D}(A_3)$	$\mathcal{E}_2(A_3)$
1	0	1.0	1.0	1.0980392	1.0909091
2	*2.0	1.1111111	1.1428571	1.0980392	1.1014493
3	0	1.0952381	1.0666666	1.1047618	1.0970464
4	2.6666667	*1.0986667	*1.1282052	1.0991870	1.0997246
5	−1.3333333	1.1008547	1.0666666	1.0984127	1.0976744
6	5.0666667	1.0938776	1.1368423	1.0986935	1.0995076
7	−5.6	1.1067668	1.0493503	*1.0985975	*1.0976739
8	12.685714	1.0852653	1.1657143	----------	1.0996695
9	−19.314286	1.1202030	1.0031748	----------	----------
10	37.574603	----------	1.2391205	----------	----------
11	−64.825397	----------	----------	----------	----------
12	121.35642	----------	----------	----------	----------
error of last term	120.254	.0215908	.140508	.0000147	.0010573
*min error	.801388	.0000545	.029593	.0000147	.0008954

The sequence A_1 may be modified by repeating each term in the sequence twice. Call the resulting sequence A_1'. The application of $\mathcal{C}$, $\mathcal{D}$, $\mathcal{E}_1$ and $\mathcal{E}_2$ to A_1' give some unusual results. It is seen that

$$\mathcal{E}_1(A_1')=A_1'$$

$$\mathcal{C}(A_1')=\{a_n|a_n=\infty, \quad n=1, 2, \ldots\}.$$

$\mathcal{E}_2(A_1')$ is of the same type as A_1' with each term repeated twice. The odd terms of $\mathcal{D}(A_1')$ agree with those of $\mathcal{E}_2(A_1')$, but the even terms do not. The even terms of $\mathcal{E}(A_1')$ converge much more rapidly than the odd ones, by a factor of approximately $1/n$. Both $\mathcal{E}_2(A_1')$ and $\mathcal{D}(A_1')$ converge more slowly than $\mathcal{E}_2(A_1)$ and $\mathcal{D}(A_1)$.

5.3. Three Related Sequences

The following three sequences are considered:

$$P_3=\{(.98)^n+(.95)^n\},$$
$$P_6=\{(1.96)^n+(.95)^n\},$$
$$P_7=\{(1.96)^n+(1.9)^n\}.$$

The purpose is to investigate the following two questions. Is it advantageous to increase the largest term in an exponential sequence before applying $\mathcal{C}$? Is it advantageous to keep the larger terms away from 1 at the expense of increasing them? The answer to the first question appears to be sometimes and the answer to the second appears to be no.

The limit or antilimit of each of these sequences is zero. The first converges quite slowly and the other two diverge (see table 9).

TABLE 9. *Behavior of P_3, P_6, and P_7*

n	P_3	P_6	P_7
10	1.416	837.3	1450
25	.8808	2.025×10^7	2.955×10^7
50	.4411	4.1×10^{15}	4.967×10^{15}
75	.2411	----------	----------

P_3, $\mathcal{C}(P_3)$, $\mathcal{C}^2(P_3)$, and $\mathcal{C}^3(P_3)$ all converge slowly. There is an improvement by a factor of approximately 2 in the range considered for each further application of $\mathcal{C}$.

Theorem 2 states that $\mathcal{C}(P_3)$ and $\mathcal{C}(P_6)$ converge at the same rate while $\mathcal{C}(P_7)$ diverges. This is borne out by table 10. $\mathcal{C}^2$ and $\mathcal{C}^3$ produce entirely different effects for each sequence. $\mathcal{C}^2(P_6)$ converges very rapidly to zero and $\mathcal{C}^3(P_6)$ converges even more rapidly. Errors due to the loss of significant figures have affected the 24th term of $\mathcal{C}^2(P_6)$ and the 14th term of $\mathcal{C}^3(P_6)$. All calculations here were made with 20 decimal digits.

In P_7, $\mathcal{C}(P_7)$, $\mathcal{C}^2(P_7)$, and $\mathcal{C}^3(P_7)$ the first term is the closest to zero. However it is apparent that the sequence formed by the nth terms of $\mathcal{C}^k(P_7)$, $k=0, 1, \ldots$ converges quite rapidly to the antilimit of P_7.

5.4. Effect of Lower-Order Terms

To illustrate the effect of lower order terms we consider

$$P_8=\{2^n+(1.2)^n\},$$

which may be compared to P_1 in section 6.2.

TABLE 10. *Effect of C^k on the sequences $P_3=\{(.98)^n+(.95)^n\}$, $P_6=\{(1.96)^n+(.95)^n\}$, $P_7=\{(1.96)^n+(1.9)^n\}$.*

n	$\mathcal{C}(P_3)$	$\mathcal{C}^2(P_3)$	$\mathcal{C}^3(P_3)$	$\mathcal{C}(P_6)$	$\mathcal{C}^2(P_6)$	$\mathcal{C}^3(P_6)$	$\mathcal{C}(P_7)$	$\mathcal{C}^2(P_7)$	$\mathcal{C}^3(P_7)$
1	.2991	−.0716	−.0873	1.024	−.0671	$-.55\times10^{-3}$	.0059	$-.88\times10^{-5}$	$.40\times10^{-7}$
5	.2708	−.0684	−.0923	.8350	−.0028	$-.13\times10^{-5}$	.0807	$-.12\times10^{-3}$	$.54\times10^{-6}$
10	.2384	−.0638	−.1032	.6462	$-.57\times10^{-4}$	$-.75\times10^{-9}$	2.129	−.0031	$.13\times10^{-4}$
15	.2091	−.0584	−.1269	.5000	$-.12\times10^{-5}$	$-.12\times10^{-10}$	55.83	−.0769	$.32\times10^{-3}$
20	.1827	−.0525	−.2017	.3869	$-.25\times10^{-7}$	$-.21\times10^{-9}$	1456	−1.901	.0076
25	.1589	−.0462	−4.903	.2994	$.15\times10^{-8}$	$-.19\times10^{-8}$	$.38\times10^{5}$	−46.27	.1726
30	.1376	−.0399	.1319	.2316	$.15\times10^{-7}$	$-.38\times10^{-9}$	$.98\times10^{6}$	−1109	3.852
35	.1186	−.0337	.0490	.1792	$.65\times10^{-7}$	$-.98\times10^{-6}$	$.25\times10^{8}$	$-.26\times10^{5}$	84.02
40	.1016	−.0279	.0241	.1387	$.94\times10^{-5}$	$-.12\times10^{-4}$	$.64\times10^{9}$	$-.61\times10^{6}$	1795
45	.0866	−.0227	.0131	.1073	$.75\times10^{-4}$	----------	$.16\times10^{11}$	$-.14\times10^{8}$	----------

TABLE 11. *Effect of lower order terms*

n	P_1	P_8	$\mathcal{C}(P_1)$	$\mathcal{C}(P_8)$	$\mathcal{C}^2(P_1)$	$\mathcal{C}^2(P_8)$	$\mathcal{C}^3(P_1)$	$\mathcal{C}^3(P_8)$
1	4.8426	3.1625	2.8810	.8205	2.695	−.0719	2.711	2×10^{-4}
10	1018.569885	1018.119177	4.8408	4.3071	1.282	−.00378	−.0072	3×10^{-7}
20	10,363,260.20	10,363,258.88	26.826	26.673	.3477	−.00014	.0267	5×10^{-9}
27	1,326,449,998	1,326,449,997	95.645	95.576	----------	----------	----------	----------

In table 11 it is seen that the lower order terms disappear rather rapidly in P_1, but the 27th terms of $\mathcal{C}(P_1)$ and $\mathcal{C}(P_8)$ only agree to two digits. Further $\mathcal{C}^2(P_1)$ and $\mathcal{C}^3(P_1)$ are completely unrelated to $\mathcal{C}^2(P_8)$ and $\mathcal{C}^3(P_8)$. Thus it is seen that these lower order terms are significant even for $\mathcal{C}(P)$ and they affect the rate of convergence of $\mathcal{C}^k(P)$, $k>1$, drastically and adversely.

5.5. Loss of Significant Figures

It has been seen that terms with a small effect on P can have a large effect on $\mathcal{C}^k(P)$. A related phenomenon is the loss of significant figures in the computation of $\mathcal{C}(P)$ and $\mathcal{D}(P)$. If P is the constant sequence $\{p_n | p_n=1\}$, then $\mathcal{C}(P)$ and $\mathcal{D}(P)$ may be written symbolically as

$$\frac{1\cdot(1-1)}{(1-1)}$$

Hence slight inaccuracies in the computation of $\mathcal{C}(P)$ and $\mathcal{D}(P)$ may distort the results badly if P is slowly varying. Often it is necessary to use double precision for the computation of $\mathcal{D}(P)$ even though P is only given to six or eight digits.

The following table shows where the loss of significant figures can occur. The first column of table 12 defines the sequence P, the second column gives the approximate size of the terms of P in the range that the transformation is applied and the third column gives the transformation used. The next four columns give n, the size of the terms appearing in the numerator, the actual value of the numerator and the number of significant figures lost from P.

TABLE 12. *Loss of significant figures*

Sequence	Size of terms	Transformation	n	Size of numerical terms	Numerator	Significant figures lost
$(1.96)^n+(.95)^n$	$.1\times10^9$	$\mathcal{C}^2$	25	.01	$.43\times10^{-4}$	20
$(.9)^n+(.8)^n\cos(.1n)$	$.4\times10^{-1}$	$\mathcal{D}$	30	.0002	$.3\times10^{-10}$	9
$(.9)^n+(.8)^n\cos(.1n)$	$.5\times10^{-3}$	$\mathcal{D}$	75	$.2\times10^{-9}$	$.2\times10^{-21}$	18
$(.999)^n+(.99)^n$	2	$\mathcal{D}$	all	10	10^{-18}	19
$(.98)^n+(.6)^n+(.4)^n$	1	$\mathcal{D}$	15	1	10^{-11}	11

6. References

[1] A. C. Aitken, On Bernoulli's numerical solution of algebraic equations, Proc. Roy. Soc. Edinburgh **46,** 289 (1926).

[2] A. C. Aitken, Studies in practical mathematics, II, Proc. Roy. Soc. Edinburgh **57,** 269 (1937).

[3] K. Knopp, Theory and application of infinite series (Blackie & Son, London, England, 1928).

[4] S. Lubkin, A method of summing infinite series, J. Research NBS **48,** 228 (1952).

[5] J. R. Rice, Tchebycheff approximations by ab^x+c, J. Soc. Ind. Appl. Math. (to be published).

[6] J. R. Rice, The characterization of best nonlinear Tchebycheff approximations, Trans. Am. Math. Soc. (to be published).

[7] P. A. Samuelson, A convergent iterative process, J. Math. Phys. **24,** 131 (1945).

[8] D. Shanks, Nonlinear transformations of divergent and slowly convergent sequences, J. Math. Phys. **34,** 1 (1955).

(Paper 64B4—38)