On Calculating the Rate of Linear Convergence of Non-Linear Transformed Sequences

Johannes Grotendorst
Institute for Advanced Simulation
Forschungszentrum Jülich
52425 Jülich, Germany
and
Aachen University of Applied Sciences
j.grotendorst@fz-juelich.de

ABSTRACT
In this paper we calculate the rate of linear convergence of non-linear transformed sequences. Using symbolic computation new results are derived which quantify the convergence acceleration effects of the Aitken transformation \( S \) and the Shanks transformation \( S_2 \). The results are applied to linearly convergent fixed-point iteration methods, i.e. sequences \( \{x_n\}_{n \geq 0} \) generated by \( x_{n+1} = \phi(x_n) \), where the iteration function \( \phi(x) \) possesses a certain number of continuous derivatives. We verify the findings by numerical examples.

Categories and Subject Descriptors
G.1.2 [Numerical Analysis]: Approximation—rational approximation; G.4 [Mathematics of Computing]: Mathematical Software—algorithm analysis; I.1.4 [Symbolic and Algebraic Manipulation]: Applications

General Terms
Algorithms, Performance, Theory

Keywords
Aitken transformation, iterated Aitken transformation, Shanks transformation, rate of linear convergence, convergence acceleration, fixed-point iteration methods, symbolic computation

1. INTRODUCTION
The speed at which a convergent sequence \( \{x_n\}_{n \geq 0} \) approaches its limit is of great importance for practical computations. A basic concept to quantify the convergence speed is introduced in numerical analysis by the following definition.

**Definition.** Consider a sequence \( \{x_n\}_{n \geq 0} \) in \( \mathbb{R} \) with \( \lim_{n \to \infty} x_n = \xi \) and \( x_n \neq \xi \) for all \( n \in \mathbb{N} \). The sequence is said to have convergence order \( p \geq 1 \) if there exists a constant \( 0 < \rho < \infty \) such that

\[
\lim_{n \to \infty} \left| \frac{x_{n+1}}{x_n} - \rho \right| = 0.
\]

Here \( \{e_n\} \) is the sequence of errors \( e_n = x_n - \xi \), \( \rho \) is called asymptotic error constant. If \( p = 1 \), then \( 0 < \rho < 1 \) is required and \( \{x_n\}_{n \geq 0} \) is said to converge linearly and \( \rho \) is the rate of linear convergence. For \( p = 2 \) the convergence is called quadratic. In the numerical literature the order \( p \) is called more precisely the \( Q \)-order of convergence. The \( Q \) stands for quotient, because the definition uses the quotient between two successive error terms [1]. Many numerical methods may be regarded as special cases of fixed-point iterations \( x_{n+1} = \phi(x_n), n \geq 0 \). The convergence order of fixed-point iterations can be determined if \( \phi(x) \) is sufficiently many times continuously differentiable in a neighborhood of the fixed point.

**Theorem 1.** Let \( \{x_n\}_{n \geq 0} \) be a convergent iteration method, i.e. \( x_{n+1} = \phi(x_n) \) and \( \lim_{n \to \infty} x_n = \xi \). Assume that the real iteration function \( \phi(x) \) is \( p \) times continuously differentiable in a neighborhood of \( \xi \) with \( D^k(\phi)(\xi) = 0 \), \( k = 1 \ldots p-1 \), \( D^p(\phi)(\xi) \neq 0 \) \( (|D(\phi)(\xi)| < 1 \text{ for } p = 1) \) and \( \phi(\xi) = \xi \). Then the iteration method is of order \( p \) and

\[
\rho = \frac{1}{|D^p(\phi)(\xi)|}.
\]

This Theorem follows directly from the Taylor series expansion of \( \phi(x) \) in the neighborhood of \( \xi \) and the assumptions (see ref. [1], p. 623, Theorem 6.1.8). In 1926 Aitken introduced the famous \( \Delta^2 \) process [2], a simple but effective convergence acceleration method. The \( \Delta^2 \) process is a sequence transformation, i.e. a mapping \( S : \{x_n\} \to \{x'_n\} \), where

\[
x'_n = x_n - \frac{\Delta x_n^2}{\Delta^2 x_n} = x_n - \frac{(x_n - x_{n+1})^2}{x_{n+2} - x_{n+1} + x_n}.
\]

Here, the \( \Delta \) operator generates the forward difference difference \( \Delta x_n = x_{n+1} - x_n \). Aitken’s \( \Delta^2 \) process (whose name comes from the \( \Delta^2 \) operator in the denominator) combines three successive elements of the sequence \( \{x_n\} \) in a non-linear way. A basic result due to Henrici [3] shows that the Aitken transformed sequence \( \{x'_n\} \) derived from a linearly convergent sequence \( \{x_n\} \) converges faster to the limit than the sequence \( \{x_n\} \).
THEOREM 2. Let \( \{x_n\}_{n \geq 0} \) be a linearly converging sequence in \( \mathbb{R} \) with \( \lim_{n \to \infty} x_n = \xi \), i.e. there exists a constant \( 0 < |\xi| < 1 \) and a sequence \( \{\varepsilon_n\} \) with \( \varepsilon_{n+1} = (\lambda + \varepsilon_n) \cdot e_n \) and \( e_n \to 0 \) for \( n \to \infty \). Then the Aitken transformed sequence \( \{x'_n\} \) converges faster to \( \xi \) than the sequence \( \{x_n\} \) in the sense that \( \frac{x_n - \xi}{x'_n - \xi} \to 0 \) for \( n \to \infty \).

For a proof see ref. [3] (p. 73, Theorem 4.5). Convergence acceleration methods and in particular the Aitken transformation as perhaps the most popular one are studied extensively in monographs from Wimp [4], Weniger [5], Brezinski and Zaglia [6], and Sidi [7]. Numerical comparisons of non-linear convergence accelerators were published by Smith and Ford [8, 9]. In addition, several acceleration methods have been made available in computer algebra systems [10, 11, 12].

In the following we use the Maple software [11] for symbolic and numerical calculations, but the results could also be reproduced with other modern computer algebra systems. This paper also aims to demonstrate the capabilities of computer algebra systems for studying numerical methods in the sense of the recent articles of Gander et al. [13, 14]. Quantitative results become possible which are hard to find with paper and pencil only. In Maple we implement infinite sequences \( \{x_n\} \) as functions on \( \mathbb{N} \). Aitken’s transformation \( S \) is the difference operator \( \Delta \) and the composition operator \( \Delta^2 \).

Firstly, we introduce two successive error terms \( \varepsilon_n \), with \( \varepsilon_{n+1} = (\lambda + \varepsilon_n) \cdot e_n \), and \( e_n \to 0 \) for \( n \to \infty \). Now, the assumption \( \varepsilon_n = \omega_n \cdot e_n \) (p\(^{-1}\)) is substituted recursively in expression (10), starting with \( \varepsilon_{n+2} \):

\[
\tau_{n+1} := \frac{\tau_n \cdot \varepsilon_{n+1}}{\varepsilon_{n+1}}
\]

Substituting \( x_{n+1} = e_{n+1} + \xi, \) \( i = 0 \ldots 3 \), we get for the quotient \( \frac{\tau_{n+1}}{\tau_n} \):

\[
\frac{\tau_{n+1}}{\tau_n} = \frac{e_{n+1} - (\lambda e_n + \xi)}{\lambda e_n - (\lambda e_n + \xi)}
\]

Next, the error terms \( e_{n+i}, \) \( i = 1 \ldots 3 \), are replaced by

\[
e_{n+i} := (\lambda + \varepsilon_n) \cdot e_n
\]

and we arrive at the following factored expression

\[
\text{factor}(\text{rhs}(8));
\]

\[
\frac{(\varepsilon_{n+1} - \varepsilon_{n+2})}{(\varepsilon_{n} - \varepsilon_{n+1})} \times \frac{(\varepsilon_{n+1} + \varepsilon_{n+2})}{(\varepsilon_{n} + \varepsilon_{n+1})} = \frac{(\lambda + \varepsilon_{n+1})(\lambda + \varepsilon_{n})}{(\lambda + \varepsilon_{n+1})(\lambda + \varepsilon_{n})}
\]

\[
\text{Theorem 3. Let } \{x_n\}_{n \geq 0} \text{ be a linearly convergent sequence in } \mathbb{R} \text{ with } \lim_{n \to \infty} x_n = \xi, \text{ i.e. there exists a constant } 0 < |\xi| < 1 \text{ and a sequence } \{\varepsilon_n\} \text{ with } \varepsilon_{n+1} = (\lambda + \varepsilon_n) \cdot e_n \text{ and } e_n \to 0 \text{ for } n \to \infty. \text{ If } e_n = \omega_n \cdot e_n \text{ (p\(^{-1}\)) with } p \in \mathbb{N}, p > 1, \text{ and } \omega_n \to 0 \text{ for } n \to \infty, \text{ then the rate of linear convergence of the transformed sequence } \{S(x_n)\} \text{ is } |\lambda|^{p-1}.
\]

PROOF. Firstly, we introduce two successive error terms of the transformed sequence:

\[
> \text{tau}[n+1] := S(x(n)) - xi;
\]

\[
> \text{tau}[n+1] := S(x(n+1)) - xi;
\]

\[
> \text{tau}[n] := S(x(n+1)) - xi;
\]

In this paper we also use the more familiar notation \( S(x) \) for the transformed sequence elements \( S(x_n) \).

2. RATE OF LINEAR CONVERGENCE OF AITKEN TRANSFORMED SEQUENCES

The following result quantifies the convergence acceleration effect of Aitken’s \( \Delta^2 \) process in terms of the rate of convergence of the given sequence \( \{x_n\} \). We make an assumption as to how fast \( e_n \) approaches to 0 in the error term of the linearly convergent sequence \( \{x_n\} \).

THEOREM 3. Let \( \{x_n\}_{n \geq 0} \) be a linearly convergent sequence in \( \mathbb{R} \) with \( \lim_{n \to \infty} x_n = \xi \), i.e. there exists a constant \( 0 < |\xi| < 1 \) and a sequence \( \{\varepsilon_n\} \) with \( \varepsilon_{n+1} = (\lambda + \varepsilon_n) \cdot e_n \) and \( e_n \to 0 \) for \( n \to \infty \). If \( e_n = \omega_n \cdot e_n \) (p\(^{-1}\)) with \( p \in \mathbb{N}, p > 1 \), and \( \omega_n \to 0 \) for \( n \to \infty \), then the rate of linear convergence of the transformed sequence \( \{S(x_n)\} \) is \( |\lambda|^{p-1} \).

PROOF. Firstly, we introduce two successive error terms of the transformed sequence:

\[
> \text{tau}[n] := S(x(n)) - xi;
\]

\[
> \text{tau}[n+1] := S(x(n+1)) - xi;
\]

In this paper we also use the more familiar notation \( S(x) \) for the transformed sequence elements \( S(x_n) \).

2. RATE OF LINEAR CONVERGENCE OF AITKEN TRANSFORMED SEQUENCES

The following result quantifies the convergence acceleration effect of Aitken’s \( \Delta^2 \) process in terms of the rate of convergence of the given sequence \( \{x_n\} \). We make an assumption as to how fast \( e_n \) approaches to 0 in the error term of the linearly convergent sequence \( \{x_n\} \).
to the following limit:

\[
\lim_{n \to \infty} \frac{\tau_{n+1}}{\tau_n} = \lambda^p
\]

\[
\lambda := \frac{\omega(\lambda - 1)^2 (\lambda^p - \lambda)}{\lambda}
\]

Next, we compare the error terms \( \tau_n \) of the Aitken transformed sequence with the error terms \( \epsilon_n \) of the given sequence. Under the assumption of Theorem 3, i.e. with \( \varepsilon_n = \omega_n \cdot \epsilon_n^{p-1} \), \( p \in \mathbb{N}, p > 1 \), the quotient \( \frac{\tau_n}{\epsilon_n} \) converges to

\[
\lim_{n \to \infty} \frac{\tau_{n+1}}{\tau_n} = \lambda^p
\]

We now apply Theorem 3 to linearly convergent fixed-point iteration methods.

**Corollary 1.** Let \( \{x_n\}_{n \geq 0} \) be a convergent iteration method, i.e. \( x_{n+1} = \phi(x_n) \) and \( \lim_{n \to \infty} x_n = \xi \). Assume that the real function \( \phi(x) \) is \( p \)-times continuously differentiable in a neighborhood of \( \xi \) with \( 0 < |D(\phi)(\xi)| < 1 \), \( D^{(p)}(\phi)(\xi) = 0 \), \( k = 2 \ldots p - 1 \), \( D^{(p)}(\phi)(\xi) \neq 0 \) and \( \phi'(\xi) = \xi \). Then the rate of linear convergence of the transformed sequence \( \{S(x_n)\} \) is \( |D(\phi)(\xi)|^p \).

**Proof.** Using the Taylor series of \( \phi(x) \) around \( \xi \) with expansion order \( p \) and the assumptions in Corollary 1 we get

\[
e_{n+1} = x_{n+1} - \xi = \phi(x_n) - \xi = \frac{D(\phi)(\xi)}{2} \cdot (x_n - \xi)^2 + \frac{D^{(p)}(\phi)(\xi)}{p!} \cdot (x_n - \xi)^p
\]

\[
\lim_{n \to \infty} \frac{\tau_{n+1}}{\tau_n} = \lambda^p
\]

Here \( \epsilon_n = x_n - \xi \) and \( \eta_n \) lies in the interval determined by \( x_n \) and \( \xi \). Setting \( \lambda = D(\phi)(\xi) \) and \( \omega_n = \frac{D^{(p)}(\phi)(\xi)}{p!} \) we have \( \omega_n \to \frac{D^{(p)}(\phi)(\xi)}{p!} \neq 0 \) for \( n \to \infty \). Then Corollary 1 follows from Theorem 3.
3. NUMERICAL EXAMPLES: AITKEN TRANSFORMED ITERATION SEQUENCES

As a first example we consider a fixed-point iteration method \( x_{n+1} = \phi(x_n) \), \( n \geq 0 \), with starting value \( x_0 = 5 \) and iteration function

\[
\phi := x \mapsto e^{-x}
\]

The equation for the fixed point \( \phi(x) = x \) is calculated symbolically.

\[
\xi := \text{solve}(\phi(x)=x,x)
\]

The limit of the iteration sequence \( \{x_n\}_{n \geq 0} \) is the value of Lambert’s \( W \) function at \( x = 1 \).

\[
\text{plot}([\text{LambertW}(x),[[1,\text{LambertW}(1)]]],x=-1..3,
\text{style}=[\text{line,point}], \text{symbol} = \text{circle},
\text{symbolsize}=20, \text{color} = \text{black});
\]

The Lambert \( W \) function satisfies the equation \( W(x) \cdot e^{W(x)} = x \). In the literature the value \( W(1) \) is also known as \( \Omega \) constant \[18\] with the property \( \Omega = e^{-\Omega} \):

\[
\text{alias}(\Omega = \text{LambertW}(1));
\]

\[
\text{simplify}(\Omega = e^{-\Omega});
\]

\[
\Omega := \text{evalf}(\Omega,10);
\]

\[
\Omega = 0.5671432904
\]

We compute the first and second derivative of \( \phi(x) \) at the fixed point \( x = \Omega \):

\[
\text{seq}((\text{D@@}k)(\phi)(\Omega),k=1..2);
\]

Theorem 1 implies that \( \{x_n\} \) converges linearly with the convergence rate \( \rho = |D(\phi)(\Omega)| = \Omega \). From Theorem 2 it follows that Aitken’s \( \Delta^2 \) process accelerates the convergence of this sequence. We study the convergence speed of \( \{x_n\} \) and \( \{S(x_n)\} \) numerically using functional programming in Maple:

\[
x := n \mapsto (\phi \circ \phi)(n)(5.0);
\]

\[
\tau_n = \frac{x(n+1) - \phi(n)}{\phi(n) - \phi(n-1)}
\]

\[
\lim_{n \to \infty} \frac{\tau_{n+1}}{\tau_n} = \frac{x(n+1) - \phi(n)}{\phi(n) - \phi(n-1)}
\]

\[
\tau_n = \frac{x(n+1) - \phi(n)}{\phi(n) - \phi(n-1)}
\]

\[
\lim_{n \to \infty} \frac{\tau_{n+1}}{\tau_n} = \frac{x(n+1) - \phi(n)}{\phi(n) - \phi(n-1)}
\]

\[
\xi = 0.5671432904
\]

whereas \( S(x)(20) \) already has 10 valid decimal places. The rate of convergence \( |\lambda| = \Omega \) of the fixed-point iteration \( x_{n+1} = \phi(x_n) \) is improved by Aitken’s \( \Delta^2 \) process to \( |\lambda|^2 \):

\[
\text{abs}(\lambda)^2 = \text{evalf}(\Omega^2,10);
\]

\[
|\lambda|^2 = 0.3216515118
\]

With \( \omega = \frac{D(\phi)(\Omega)}{2!} = \frac{\Omega}{2} \) and \( \lambda = D(\phi)(\Omega) = -\Omega \) we get for the limit \( (16) \):

\[
\text{simplify}(\text{subs}(\Omega = \text{evalf}(\Omega,10), (16)));
\]

\[
\lim_{n \to \infty} \frac{\tau_{n+1}}{\tau_n} = \frac{x(n+1) - \phi(n)}{\phi(n) - \phi(n-1)} = \frac{x(n+1) - \phi(n)}{\phi(n) - \phi(n-1)}
\]

In addition, Corollary 1 is verified by calculating the rate of convergence of the transformed sequence \( \{S(x_n)\} \) symbolically:

\[
\tau_{n+1}/\tau_n = (S(x)(n+1)-\Omega)/(S(x)(n)-\Omega);
\]

\[
\frac{\tau_{n+1}}{\tau_n} = \frac{(x(n+1) - \phi(n))}{\phi(n) - \phi(n-1)}
\]

\[
\lim_{n \to \infty} \frac{\tau_{n+1}}{\tau_n} = 0.1026235169
\]
Using the substitutions \( x_{n+i} = \phi^{(i)}(x_n), i = 1 \ldots 3, \) and \( x_n \to \Omega \) for \( n \to \infty \) we obtain
\[
\lim_{n\to\infty} \frac{\tau_{n+1}}{\tau_n} = \Omega^2
\]  
(28)

As second example we consider the fixed-point iteration method \( x_{n+1} = \Phi(x_n), n \geq 0, \) with iteration function
\[
\phi := x \mapsto \sin(x)/2 + x^3/12;
\]
\[
\xi := 0.
\]  
(30)

We compute the first eight derivatives at \( \xi = 0: \)
\[
> \text{seq}(\text{D}@@k)(\phi)(0), k=1 \ldots 8;
\]
\[
\begin{array}{l}
0.5, 0, 0, 0, \frac{1}{2}, 0, -\frac{1}{2}, 0
\end{array}
\]  
(31)

Thus, the iteration sequence \( x_{n+1} = \Phi(x_n) \) converges linearly with \( \lambda = \frac{1}{2}. \) Corollary 1 implies that the rate of linear convergence of the Aitken transformed sequence \( \{S(x_n)\} \) is improved to \( \lambda^p = \frac{1}{32} \) (\( p = 5 \)). Using \( \omega = \frac{D^{(3)}(\phi)(0)}{3!} = \frac{1}{2 \cdot 3!} = \frac{1}{240} \) and \( \lambda = \frac{1}{2} \) we obtain for the limit (15):
\[
\lim_{n\to\infty} \frac{\tau_n}{\epsilon_n} = \frac{1}{128}
\]  
(32)

In the following table the convergence behavior is reproduced numerically using \( x_{n+1} = \Phi(x_n) \) with starting value \( x_0 = 2.0: \)
\[
> \text{seq}(\text{printf}("\%5.8f \%4.2e \%11.8f \%11.8f\n", n,x(n),S1(x)(n),abs(\tau(n)/\tau(1))));
\]
\[
\begin{array}{l|l|l}
0 & 2.00000 & -3.74e-01 \\
1 & 1.12132 & -1.41e-02 \\
2 & 0.56783 & -4.59e-04 \\
3 & 0.28416 & -1.44e-05 \\
4 & 0.14209 & -4.52e-07 \\
5 & 0.07104 & -1.41e-08 \\
6 & 0.03552 & -4.42e-10 \\
7 & 0.01776 & -1.38e-11 \\
8 & 0.00888 & -4.31e-13 \\
9 & 0.00444 & -1.35e-14 \\
10 & 0.00222 & -4.21e-16 \\
11 & 0.00111 & -1.32e-17 \\
12 & 0.00056 & -4.11e-19 \\
13 & 0.00028 & -1.29e-20 \\
14 & 0.00014 & -4.02e-22 \\
15 & 0.00007 & -1.26e-23
\end{array}
\]  
(34)

and of limit (32):
\[
> \text{evalf}(10/128);
\]
\[
\begin{array}{l}
0.007812500000
\end{array}
\]  
(35)

4. RATE OF LINEAR CONVERGENCE OF S_2 TRANSFORMED SEQUENCES

In 1955 Shanks introduced a generalization of Aitken's transformation [19]. The Shanks transforms \( S_k(x_n), k \geq 1, \) can be represented as the ratio of Hankel determinants
\[
S_k(x_n) = \frac{|H_{k+1}(x_n)|}{|H_k(\Delta^2 x_n)|}.
\]

Here \( H_k(i,j) = V_{i+j-1}, 1 \leq i,j \leq k \) is a Hankel matrix of order \( k \) and the vector \( V \) is given by
\[
V = (\Delta^2 x_n, \Delta^2 x_{n+1}, \ldots, \Delta^2 x_{n+k-1})
\]
in the numerator and \( V = (\Delta^2 x_n, \Delta^2 x_{n+1}, \ldots, \Delta^2 x_{n+k-2}) \) in the denominator. The Shanks transforms \( S_k(x_n) \) may be constructed symbolically by the following procedure:
\[
> \text{Shanks} := \text{proc}(k) \rightarrow \text{point}(n,x) \end{proc};
\]
\[
\begin{array}{l}
\text{uses LinearAlgebra;} \\
\text{Delta:=x->(m->x(m+1)-x(m));} \\
\text{H:=HankelMatrix(Vector(2*k+1,1->(x(n+i)-x(n+i-1))));} \\
\text{H:=HankelMatrix(Vector(2*k+1,1->(x(n+i)-x(n+i-1))));} \\
\text{Determinant(H1)/Determinant(H);} \\
\end{array}
\]  
(36)

For \( k = 1 \) we get Aitken's transformation, i.e. \( S_1(x_n) = S(x_n): \)
\[
> S[1] := x \mapsto \text{Shanks}(1,n,x); \\
\]
\[
> 'S[1](x)(n)' = 'S[1](x)(n)';
\]
\[
S_1(x)(n) = \frac{x(n) x(n+2) - x(n+1)^2}{x(n+2) - 2 x(n+1) + x(n)}
\]  
(36)
Aitken's transformation $S_1$ depends on three successive sequence elements, whereas Shanks' transformation $S_2$ needs five successive sequence elements. In another context formulae (36) and (37) were already used by James Clerk Maxwell in his treatise on electricity and magnetism [20]. The following result quantifies the convergence acceleration effect of the $S_2$ transformation in terms of the rate of convergence of the given sequence $\{x_n\}$. The assumption for the subclass of linearly convergent sequences $\varepsilon_n = \omega_n \cdot e_n$, $\lim_{n \to \infty} \omega_n = \omega \neq 0$ is chosen more simple than in Theorem 3 because of the computational complexity.

**Theorem 4.** Let $\{x_n\}_{n \geq 0}$ be a linearly convergent sequence in $\mathbb{R}$ with $\lim_n x_n = \xi$, i.e. there exists a constant $0 < |\lambda| < 1$ and a sequence $\{\varepsilon_n\}$ with $\varepsilon_{n+1} = (\lambda + \varepsilon_n) \cdot e_n$ and $\varepsilon_n \to 0$ for $n \to \infty$. If $\varepsilon_n = \omega_n \cdot e_n$ with $\omega_n \to \omega \neq 0$ for $n \to \infty$, then the rate of linear convergence of the transformed sequence $\{S_2(x_n)\}$ is $|\lambda|$. 

**Proof.** We have to calculate the ratio of two successive error terms of the transformed sequence:

\[
\tau_{n+1} := S_2(x)(n+1) - \xi
\]

\[
\tau_n := S_2(x)(n) - \xi
\]

Substituting $x_{n+1} = \varepsilon_{n+1} + \xi$, $i = 0 \ldots 5$, we get for the quotient $\tau_{n+1}/\tau_n$:

\[
\tau_{n+1} = \left(\varepsilon_{n+1} + \varepsilon_{n+2} + 2 \varepsilon_{n+3} + \varepsilon_{n+4} + \varepsilon_{n+5} \right) \times
\]

\[
\left(2 \varepsilon_{n+1} + \varepsilon_{n+2} + 3 \varepsilon_{n+3} \right) \times
\]

\[
\left(2 \varepsilon_{n+1} + \varepsilon_{n+2} + \varepsilon_{n+3} \right)
\]

Next, the error terms $\varepsilon_{n+i}$, $i = 1 \ldots 5$, are replaced by

\[
\text{for } i \text{ from 1 to 5 do }
\]

\[
e(i) := \lambda + \varepsilon_n \cdot e_n
\]

\[
\text{end do;}
\]

\[
\varepsilon_{n+1} := (\lambda + \varepsilon_n) \cdot e_n
\]

\[
\varepsilon_{n+2} := (\lambda + \varepsilon_n) \cdot e_n
\]

\[
\text{for } i \text{ from 1 to 5 do }
\]

\[
\varepsilon_{n+3} := (\lambda + \varepsilon_n) \cdot e_n
\]

\[
\varepsilon_{n+4} := (\lambda + \varepsilon_n) \cdot e_n
\]

\[
\varepsilon_{n+5} := (\lambda + \varepsilon_n) \cdot e_n
\]

and coefficients depending on $\lambda$ and $\omega_{n+i}$, $i = 0 \ldots 4$. As in the proof of Theorem 3 we calculate the limit separately for the numerator and denominator. Therefore, we determine the first non-vanishing terms of the polynomials by Taylor expansion:

\[
\text{LD} := 2 \lambda^3 \omega^4 (\lambda + 1)^3 (\lambda - 1)^4 \cdot e_n^4
\]

\[
\text{LN} := 2 \lambda^3 \omega^4 (\lambda + 1)^3 (\lambda - 1)^4 \cdot e_n^4
\]
Then, with the assumptions $0 < |\lambda| < 1$ and $\omega \neq 0$ it follows for the limit
\[
\lim_{n \to \infty} \frac{\tau_{n+1}}{\tau_n} = \lambda^3
\]
(46)

Again, Theorem 4 is readily applied to linearly convergent fixed-point iteration methods.

**Corollary 2.** Let $\{x_n\}_{n \geq 0}$ be a convergent iteration method, i.e. $x_{n+1} = \phi(x_n)$ and $\lim x_n = \xi$. Assume that the real function $\phi(x)$ is two times continuously differentiable in a neighborhood of $\xi$ with $0 < |D(\phi)(\xi)| < 1$, $D^{(2)}(\phi)(\xi) \neq 0$ and $\phi(\xi) = \xi$. Then the rate of linear convergence of the transformed sequence $\{S_2(x_n)\}$ is $|D(\phi)(\xi)|^3$.

**Proof.** Using the Taylor series of $\phi(x)$ around $\xi$ with expansion order 2 and the assumptions in Corollary 2 we get
\[
\begin{align*}
\epsilon_{n+1} &= x_{n+1} - \xi \\
&= \phi(x_n) - \xi \\
&= D(\phi)(\xi) \cdot (x_n - \xi) + \frac{D^{(2)}(\phi)(\xi)}{2!} \cdot (x_n - \xi)^2 \\
&= D(\phi)(\xi) \cdot \epsilon_n + \frac{D^{(2)}(\phi)(\eta_n)}{2!} \cdot \epsilon_n^2 \\
&= \left(D(\phi)(\xi) + \frac{D^{(2)}(\phi)(\eta_n)}{2!} \right) \cdot \epsilon_n.
\end{align*}
\]

Here $\epsilon_n = x_n - \xi$ and $\eta_n$ lies in the interval determined by $x_n$ and $\xi$. Setting $\lambda = D(\phi)(\xi)$ and $\omega_n = \frac{D^{(2)}(\phi)(\eta_n)}{2!}$ we have $\omega_n \to D^{(2)}(\phi)(\xi) \neq 0$ for $n \to \infty$. Then Corollary 2 follows from Theorem 4. \(\square\)

Next, we calculate the limit of the quotient $\frac{\tau_n}{\epsilon_n}$ for $n \to \infty$, where $\tau_n$ is the error term of a $S_2$ transformed sequence and $\epsilon_n$ is the error term of the given convergent sequence. If the sequence is generated by a three times continuously differentiable iteration function $\phi(x)$, then $\epsilon_{n+1}$ can be expressed in terms of a cubic polynomial in $\epsilon_n$:
\[
\begin{align*}
\epsilon_{n+1} &= \left(D(\phi)(\xi) + \frac{D^{(2)}(\phi)(\xi)}{2!} \right) \cdot \epsilon_n + \frac{D^{(3)}(\phi)(\eta_n)}{3!} \cdot \epsilon_n^2 \\
&= \left(D(\phi)(\xi) + \frac{D^{(2)}(\phi)(\eta_n)}{2!} \right) \cdot \epsilon_n.
\end{align*}
\]

Setting $\lambda = D(\phi)(\xi)$, $\omega = \frac{D^{(3)}(\phi)(\xi)}{3!}$, $\eta_n = \frac{D^{(3)}(\phi)(\eta_n)}{2!}$ and assuming $0 < |\lambda| < 1$ and $\omega \neq 0$ we obtain a linearly convergent sequence with $\epsilon_{n+1} = (\lambda + \epsilon_n) \cdot \epsilon_n$ and $\epsilon_n = \omega \cdot \epsilon_n + \theta_n \cdot \epsilon_n^2$. Now, polynomial $\epsilon_n$ is substituted into the error expression $\tau_n$.

\[
\begin{align*}
\text{tau[n]} := \text{subs(seq(epsilon[n+1]=omega*epsilon[n]+}
\text{theta[n]*epsilon[n]^2, i=1..3, n=0), tau[n])};
\end{align*}
\]

Considering $\theta_n \to \theta = \frac{D^{(3)}(\phi)(\xi)}{3!}$ for $n \to \infty$ we calculate the first term of the Taylor series with respect to $\epsilon_n$:
\[
\begin{align*}
\text{factor(limit(convert(taylor(tau[n], epsilon[n]=0, 7), polynom), seq(theta[n]=theta, i=0..3)));}
\end{align*}
\]

Hence, the limit is evaluated to
\[
\begin{align*}
\lim_{n \to \infty} \frac{\tau_n}{\epsilon_n^3} &= \frac{\lambda^4 (\lambda^2 - \lambda \theta + 2 \omega^2)}{(\lambda + 1)(\lambda - 1)^2} \cdot \epsilon_n^3
\end{align*}
\]
(47)

5. **NUMERICAL EXAMPLE: S_2 TRANSFORMED ITERATION SEQUENCE**

For the $S_2$ transformation we reuse the fixed-point iteration method analysed in the first example, i.e. we consider the iteration function
\[
\begin{align*}
\phi := x \mapsto \exp(-x); \\
\text{with fixed point}
\end{align*}
\]
(50)

\[
\begin{align*}
\xi := \Omega;
\end{align*}
\]
(51)

We compute the first six derivatives of $\phi(x)$ at the fixed point $x = \Omega$:
\[
\begin{align*}
\text{seq(simplify(D(D(D(D(D(D(phi)(Omega)))))))(k=1..6));}
\end{align*}
\]

The iteration sequence $x_{n+1} = \phi(x_n)$ converges linearly with $|\lambda| = \Omega$. Corollary 2 implies that the rate of linear convergence of the transformed sequence $\{S_2(x_n)\}$ is improved to $|\lambda|^3 = \Omega^3$. Using $\omega = \frac{D^{(3)}(\phi)(\Omega)}{2!} = \frac{\Omega}{2}$ and $\theta = \frac{D^{(3)}(\phi)(\Omega)}{3!} = \frac{\Omega}{6}$ we obtain for the limit (48):
\[ \lim_{n \to \infty} \frac{\tau_n}{e_n^3} = \frac{\Delta^6 (\Omega - 2)}{6 (\Omega - 1) (\Omega + 1)^2} \] (53)

In the following table the convergence behavior is reproduced numerically using \( x_{n+1} = \phi(x_n) \) with starting value \( x_0 = 5.0 \):

| \( n \) | \( x(n) \) | \( \tau(n) \) | \( \varepsilon_n \) | \( |\varepsilon_n| \) |
|---|---|---|---|---|
| 0 | 5.0000000 | 0.58000254 | 0.1706525 | 0.0001476 |
| 1 | 0.0067379 | 0.56494883 | 0.1680219 | 0.0124687 |
| 2 | 0.9932847 | 0.56751201 | 0.1877789 | 0.0047647 |
| 3 | 0.3703582 | 0.56707405 | 0.1783486 | 0.0090858 |
| 4 | 0.6904870 | 0.56715564 | 0.1844362 | 0.0065805 |
| 5 | 0.5013319 | 0.56714101 | 0.1811800 | 0.0079901 |
| 6 | 0.6057234 | 0.56714370 | 0.1830959 | 0.0071858 |
| 7 | 0.5456796 | 0.56714321 | 0.1820303 | 0.0076406 |
| 8 | 0.5794479 | 0.56714330 | 0.1826417 | 0.0073822 |
| 9 | 0.5602076 | 0.56714329 | 0.1822972 | 0.0075286 |
| 10 | 0.5710905 | 0.56714329 | 0.1824932 | 0.0074456 |
| 11 | 0.5649091 | 0.56714329 | 0.1823823 | 0.0074927 |
| 12 | 0.5684118 | 0.56714329 | 0.1824453 | 0.0074659 |
| 13 | 0.5664243 | 0.56714329 | 0.1824096 | 0.0074811 |
| 14 | 0.5675512 | 0.56714329 | 0.1824298 | 0.0074725 |
| 15 | 0.5659120 | 0.56714329 | 0.1824183 | 0.0074774 |
| 16 | 0.5672745 | 0.56714329 | 0.1824249 | 0.0074746 |
| 17 | 0.5670689 | 0.56714329 | 0.1824212 | 0.0074762 |
| 18 | 0.5671855 | 0.56714329 | 0.1824233 | 0.0074753 |
| 19 | 0.5671194 | 0.56714329 | 0.1824221 | 0.0074758 |
| 20 | 0.5671569 | 0.56714329 | 0.1824227 | 0.0074755 |

For comparison we compute the numerical values of the convergence rate of the transformed sequence \( \{S_2(x_n)\} \)

\[ \lim_{n \to \infty} \frac{\tau_n}{e_n^3} = 0.007475611683 \] (55)

6. RATE OF LINEAR CONVERGENCE OF \( S^{(2)} \) TRANSFORMED SEQUENCES

The iterated Aitken transformation \( S^{(2)} \) combines five successive sequence elements in a non-linear way (see, for instance, ref. [5], p. 225, or ref. [21]) just as the Shanks transformation \( S_2 \):

\[ 'S(x(n+i))' = normal(S(S(x(n+i))) \]

\[ S^{(2)}(x)(n) = \frac{S(x)(n)S(x)(n+2) - S(x)(n+1)^2}{S(x)(n+2) - 2S(x)(n+1) + S(x)(n)} \] (56)

where

\[ \lim_{n \to \infty} \frac{\tau_n}{e_n^3} = \frac{\Delta^6 (\Omega - 2)}{6 (\Omega - 1) (\Omega + 1)^2} \] (53)

for \( i = 0, 2 \). The following result quantifies the convergence acceleration effect of the \( S^{(2)} \) transformation in terms of the rate of convergence of the given sequence \( \{x_n\} \).

Theorem 5. Let \( \{x_n\}_{n=0}^\infty \) be a linearly convergent sequence in \( \mathbb{R} \) with \( \lim_{n \to \infty} x_n = \xi \), i.e. there exists a constant \( 0 < \lambda < 1 \) and a sequence \( \{\varepsilon_n\} \) with \( \varepsilon_{n+1} = (\lambda + \varepsilon_n) \cdot \varepsilon_n \) and \( \varepsilon_n \to 0 \) for \( n \to \infty \). If \( \varepsilon_n = \omega_n \cdot e_n \) with \( \omega_n \to \omega \neq 0 \) for \( n \to \infty \), then the rate of linear convergence of the transformed sequence \( \{S^{(2)}(x_n)\} \) is \( |\lambda|^3 \).

Proof. For the error terms of the transformed sequence

\[ \tau_{n+1} := (S^{(2)}(x)(n+1) - \xi) \] (58)

\[ \tau_{n+1} := (S^{(2)}(x)(n+1) - \xi) \] (59)

similar calculations as in the proof of Theorem 4 lead to

\[ \tau_{n+1} := \text{subs(seq}(x(n+i) = e[n+i]*xi, i=0..4),\tau[n]) \]

\[ \tau_{n+1} := \text{subs(seq}(x(n+i) = e[n+i]*xi, i=0..4),\tau[n+1]) \]

\[ Q3 := \text{normal}(tau[n+1]/tau[n]) \]

for \( i \) from 1 to 5 do

\( e[n+i] := (\text{lambda}+\varepsilon[n+i-1])*e[n+i-1] \)

end do:

\[ Q3 := \text{factor}(Q3) \]

\[ Q4 := \text{subs(seq}(\varepsilon[n+i]=\omega[n]*e[n], i=0..4),Q3) \]

\[ LN := \text{factor(limit}(\text{convert}(\text{taylor}(\text{numerator}(Q2),e[n],5),\text{polynomial}),\text{seq}(\omega[n+i] = \omega[n], i=0..4))) \]

\[ LD := \text{factor(limit}(\text{convert}(\text{taylor}(\text{denom}(Q4),e[n],5),\text{polynomial}),\text{seq}(\omega[n+i] = \omega[n], i=0..4))) \]

\[ \text{limit}(tau[n+1]/tau[n],n=\infty) = \text{limit}(LN/LD,e[n]=0); \]

\[ \lim_{n \to \infty} \frac{\tau_{n+1}}{\tau_n} = \lambda^3 \] (62)

Theorem 4 and 5 show that the transformations \( S_2 \) and \( S^{(2)} \) produce identical convergence improvements for the class of linearly convergent sequences \( \{x_n\} \) with \( \varepsilon_n = \omega_n \cdot e_n \) and \( \omega_n \to \omega \neq 0 \). For fixed-point iteration methods we have
Corollary 3. Let \( \{x_n\}_{n \geq 0} \) be a convergent iteration method, i.e. \( x_{n+1} = \phi(x_n) \) and \( \lim_{n \to \infty} x_n = \xi \). Assume that the real function \( \phi(x) \) is two times continuously differentiable in a neighborhood of \( \xi \) with \( 0 < |D(\phi)(\xi)| < 1 \), \( D^{(2)}(\phi)(\xi) \neq 0 \) and \( \phi(\xi) = \xi \). Then the rate of linear convergence of the transformed sequence \( \{S^{(2)}(x_n)\} \) is \( |D(\phi)(\xi)|^3 \).

Proof. Identical to the proof of Corollary 2.

Finally, we calculate the limit of the quotient \( \frac{\tau_n}{e_n^3} \) for \( n \to \infty \), where \( \tau_n \) is the error term of a \( S^{(2)} \) transformed sequence and \( e_n \) is the error term of the given convergent sequence. We consider a linearly convergent sequence \( \{x_n\} \) with \( e_{n+1} = (\lambda + e_n)^2 \cdot e_n \), \( 0 < |\lambda| < 1 \), \( e_n = \omega \cdot e_n + \theta_n \cdot e_n^{-2} \), \( \omega \neq 0 \) and \( \theta_n \to \theta \) for \( n \to \infty \). Substituting the polynomial \( e_n \) into the error term \( \tau_n \) and calculating the first term of the Taylor series with respect to \( e_n \) we obtain

\[
\tau_n = \text{subs(seq(} \epsilon_n^i = \omega \cdot e_n^i + e_n^{-2} \cdot \epsilon_n^i \text{), 1..3)),}
\]

\[
\frac{\lambda^4 \left( \lambda \theta - \omega^2 \right)}{(\lambda + 1)(\lambda - 1)} e_n^3 \quad (63)
\]

Thus, the limit is evaluated to

\[
\lim_{n \to \infty} \frac{\tau_n}{e_n^3} = \frac{\lambda^4 \left( \lambda \theta - \omega^2 \right)}{(\lambda + 1)(\lambda - 1)} \quad (64)
\]

7. NUMERICAL EXAMPLE: \( S^{(2)} \) TRANSFORMED ITERATION SEQUENCE

Corollary 3 implies that the iterated Aitken transformation \( S^{(2)} \) improves the linear convergence rate \( |\lambda| = \Omega \) of \( \{x_n\} \), where \( x_{n+1} = \phi(x_n) \) and \( \phi(x) = e^{-x} \), to \( |\lambda|^3 = \Omega^3 \).

Using \( \omega = D^{(2)}(\phi)(\Omega) \) and \( e = \frac{\Omega^3}{\theta} \), where \( D^{(2)}(\phi)(\Omega) = \frac{\Omega^3}{\theta} \) and \( \lambda = -\Omega \) we obtain for the limit (64):

\[
\lim_{n \to \infty} \frac{\tau_n}{e_n^3} = \frac{\Omega^6}{12(1 - \Omega)} \quad (65)
\]

In the following table the convergence behavior is verified numerically:

<table>
<thead>
<tr>
<th>( n )</th>
<th>( \frac{\tau_n}{e_n^3} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.000000</td>
</tr>
<tr>
<td>1</td>
<td>0.0067379</td>
</tr>
<tr>
<td>2</td>
<td>0.9932847</td>
</tr>
<tr>
<td>3</td>
<td>0.370382</td>
</tr>
<tr>
<td>4</td>
<td>0.6904870</td>
</tr>
<tr>
<td>5</td>
<td>0.5013319</td>
</tr>
<tr>
<td>6</td>
<td>0.6073234</td>
</tr>
<tr>
<td>7</td>
<td>0.5710905</td>
</tr>
<tr>
<td>8</td>
<td>0.5649091</td>
</tr>
<tr>
<td>9</td>
<td>0.5684118</td>
</tr>
<tr>
<td>10</td>
<td>0.5675512</td>
</tr>
<tr>
<td>11</td>
<td>0.5669120</td>
</tr>
<tr>
<td>12</td>
<td>0.5672745</td>
</tr>
<tr>
<td>13</td>
<td>0.5670689</td>
</tr>
<tr>
<td>14</td>
<td>0.5649091</td>
</tr>
<tr>
<td>15</td>
<td>0.5684118</td>
</tr>
<tr>
<td>16</td>
<td>0.5675512</td>
</tr>
<tr>
<td>17</td>
<td>0.5669120</td>
</tr>
<tr>
<td>18</td>
<td>0.5672745</td>
</tr>
<tr>
<td>19</td>
<td>0.5670689</td>
</tr>
<tr>
<td>20</td>
<td>0.5649091</td>
</tr>
</tbody>
</table>

The numerical values of the convergence rate of the transformed sequence \( \{S^{(2)}(x_n)\} \) and of limit (65) are given by:

\[
\text{abs}(\text{lambda})^3 = \text{evalf}(\Omega^3,10);
\]

\[
|\lambda|^3 = 0.1824224968 \quad (66)
\]

\[
\text{evalf}(\text{abs}(\text{lambda})^3,10), \quad \lim_{n \to \infty} \frac{\tau_n}{e_n^3} = 0.00408811042 \quad (67)
\]

8. CONCLUSION AND OUTLOOK

It is well-known that Aitken’s \( \Delta^2 \) process and its generalization the Shanks transformation are powerful non-linear convergence acceleration methods. In this paper we have presented new results which quantify the convergence acceleration effects of the Aitken transformation \( S \), the iterated Aitken transformation \( S^{(2)} \) and the Shanks transformation \( S_2 \) for subclasses of linearly convergent sequences. The formal results are applied to fixed-point iteration methods, i.e. sequences \( \{x_n\} \) generated by \( x_{n+1} = \phi(x_n) \), where the iteration function \( \phi(x) \) possesses a certain number of continuous derivatives. Furthermore, we have used the computer algebra system Maple for symbolic and numerical calculations. It is shown how this powerful mathematical system can be used for analysing numerical methods and generating results which are hard to find with paper and pencil only. In addition, we have investigated how the three transformations \( S, S^{(2)} \) and \( S_2 \) affect quadratically convergent sequences. With minor variations of the Maple code we could rerun the symbolic calculations in the proofs of Theorems 3, 4 and 5 for second order sequences, i.e. sequences \( \{x_n\} \) with the convergence behavior \( e_{n+1} = \rho_n e_n^2 \), where \( \rho_n \to \rho \neq 0 \) for \( n \to \infty \). It turns out, that the transformed sequences are again quadratically convergent with the same asymptotic error constant \( \rho \) (see also ref. [15], p. 268). Future studies will include the extension of the subclasses of linearly convergent sequences, the higher order transformations \( S_k \) and \( S^{(k)} \), \( k \geq 3 \), as well as special transformations such as \( \{S_k(x_n)\}_{k \geq 2} \) and \( \{S^{(k)}(x_n)\}_{k \geq 2} \), however the computational demand increases rapidly.
9. REFERENCES