Super-quadratic Convergence in Aitken Δ^2 Process

Kazufumi Ozawa
Department of Electronics and Information Systems,
Faculty of Systems Science and Technology,
Akita Prefectural University,
84-4 Tsuchiya-Ebinokuchi, Honjo City, Akita 015-0055, Japan
e-mail: ozawa@akita-pu.ac.jp
phone: +81-184-27-2079
fax: +81-184-27-2187

Running title: Super-quadratic convergence in Aitken Δ^2 process

Super-quadratic Convergence in Aitken Δ^2 Process

Abstract

In this short note, the existence of a super-quadratic convergence sequence in the Aitken Δ^2 process is shown, when the original sequence is generated by the fixed-point iteration with the iteration function of the form $T(x) = (r\,x + s)/(p\,x + q)$, where p, q, r and s are constants satisfying $p\,s - q\,r \neq 0$ and $p \neq 0$.

Key words: Aitken Δ^2 method, linear fractional transformation, super-quadratic convergence, fixed-point iteration, Steffensen's method

1 Introduction

The Aitken Δ^2 method is a powerful mean for accelerating slowly converging sequences. It is used, for example, to evaluate infinite range integrals with oscillatory integrands [1]. In this short note, we show that a super-quadratic convergence sequence is derived from the Aitken Δ^2 method, if the original sequence is generated by the fixed-point iteration with the iteration function of the form $T(x) = (r\,x + s)/(p\,x + q)$, where p, q, r and s are constants such that $p\,s - q\,r \neq 0$ and $p \neq 0$.

2 Linearly Convergent Sequence and Aitken Δ^2 Method

Let us consider the linearly convergent sequence $\{x_n^{(0)}\}_{n=0}^{\infty}$ given by

$$x_n^{(0)} = x^* + c_1\,\lambda^n + c_2\,\lambda^{2n} + c_3\,\lambda^{3n} + \cdots, \qquad 0 < |\lambda| < 1, \tag{1}$$

where c_i does not depend on λ and n. The convergence of the sequence $\{x_n^{(0)}\}$ to the limit x^* can be accelerated by using the Aitken Δ^2 method [3, 5]. The algorithm of the method is given by

$$x_{n+2}^{(k+1)} = x_{n+2}^{(k)} - \frac{(x_{n+2}^{(k)} - x_{n+1}^{(k)})^2}{x_{n+2}^{(k)} - 2\,x_{n+1}^{(k)} + x_n^{(k)}}, \quad n \geq 2k, \quad k = 0, 1, 2, \ldots. \tag{2}$$

This process evolves as in Table 1. Although we can start the acceleration process at any time, we assume, hereafter, without loss of generality, that we start the process at $n = 0$. The method, unlike the Richardson extrapolation [3, 5], can be used even when the convergence ratio λ is unknown, since λ does not appear in (2).

Table 1. Aitken Δ^2 process.

n	$x_n^{(0)}$	$x_n^{(1)}$	$x_n^{(2)}$	$x_n^{(3)}$	$\cdots$	$\cdots$
0	$\boldsymbol{x_0^{(0)}}$					
1	$x_1^{(0)}$					
2	$x_2^{(0)}$	$\boldsymbol{x_2^{(1)}}$				
3	$x_3^{(0)}$	$x_3^{(1)}$				
4	$x_4^{(0)}$	$x_4^{(1)}$	$\boldsymbol{x_4^{(2)}}$			
5	$x_5^{(0)}$	$x_5^{(1)}$	$x_5^{(2)}$			
6	$x_6^{(0)}$	$x_6^{(1)}$	$x_6^{(2)}$	$\boldsymbol{x_6^{(3)}}$		
$\vdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$		
$\vdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\ddots$

In general, faster convergence is expected for larger k. In fact, for the case given by (1), the unknown $\mathrm{O}(\lambda^n)$ term vanishes in $x_{n+2}^{(1)}$, which can be shown by an elementary calculation. Likewise the $\mathrm{O}(\lambda^{2n})$ term vanishes in $x_{n+4}^{(2)}$, the $\mathrm{O}(\lambda^{3n})$ term vanishes in $x_{n+6}^{(3)}$, and so on. Therefore, the convergence is improved as k increases, although the convergence of each of the accelerated sequence $\{x_n^{(k)}\}_n$ is still linear. More specifically, the convergence ratio of the accelerated sequence is given by

$$\lim_{n\to\infty} \frac{x_{n+1}^{(k)} - x^*}{x_n^{(k)} - x^*} = \lambda^{k+1}, \qquad k = 0, 1, \ldots$$

In spite of this observation, we show that the convergence becomes much faster for the particular case that the constants c_i's in (1) are given by

$$c_i = \alpha^{i-1}\,\beta, \qquad i = 1, 2, \ldots, \tag{3}$$

where α and β are independent of i and λ, and $\alpha\,\beta \neq 0$. In this case, assuming $|\alpha\,\lambda^n| < 1$, we have for $k = 0$

$$\begin{aligned} e_n^{(0)} &= x_n^{(0)} - x^* \\ &= \beta\,\lambda^n + \alpha\,\beta\,\lambda^{2n} + \alpha^2\,\beta\,\lambda^{3n} + \cdots \\ &= \frac{\beta\,\lambda^n}{1 - \alpha\,\lambda^n}, \end{aligned} \tag{4}$$

and for general k we have the following lemma:

LEMMA 1 *Let c_i's be given by* (3) *and satisfy* $|\alpha\,\lambda^n| < 1$, *then we have for the error* $e_{n+2k}^{(k)}$

$$e_{n+2k}^{(k)} = x_{n+2k}^{(k)} - x^* = \frac{\alpha^{2^k-1}\,\beta\,\lambda^{2^k k}\,(\lambda^{2^k})^n}{1 - \alpha^{2^k}\lambda^{2^k k}\,(\lambda^{2^k})^n}, \qquad k = 0, 1, 2, \ldots \tag{5}$$

Proof. We prove this by induction. When $k = 0$ this is clearly true. Here we suppose (5) is true for some $k > 0$. For simplicity, we put

$$u = \alpha^{2^k-1}\,\lambda^{2^k k}, \qquad v = \lambda^{2^k},$$

then from (5)

$$e_{n+2k}^{(k)} = \frac{u\,v^n\,\beta}{1 - \alpha\,u\,v^n}.$$

To see that (5) is valid for $k+1$, we use the relation

$$e_{n+2k+2}^{(k+1)} = \frac{e_{n+2k+2}^{(k)}\, e_{n+2k}^{(k)} - (e_{n+2k+1}^{(k)})^2}{e_{n+2k+2}^{(k)} - 2\, e_{n+2k+1}^{(k)} + e_{n+2k}^{(k)}},$$

which is easily derived from (2). We denote this by N/D, that is,

$$\begin{aligned} N &= e_{n+2k+2}^{(k)}\, e_{n+2k}^{(k)} - (e_{n+2k+1}^{(k)})^2 \\ &= \left\{ \frac{u\, v^{n+2}}{1-\alpha\, u\, v^{n+2}} \cdot \frac{u\, v^{n}}{1-\alpha\, u\, v^{n}} - \left(\frac{u\, v^{n+1}}{1-\alpha\, u\, v^{n+1}} \right)^2 \right\} \beta^2, \\ D &= e_{n+2k+2}^{(k)} - 2\, e_{n+2k+1}^{(k)} + e_{n+2k}^{(k)} \\ &= \left\{ \frac{u\, v^{n+2}}{1-\alpha\, u\, v^{n+2}} - \frac{2\, u\, v^{n+1}}{1-\alpha\, u\, v^{n+1}} + \frac{u\, v^{n}}{1-\alpha\, u\, v^{n}} \right\} \beta. \end{aligned}$$

We have for the numerator N,

$$\begin{aligned} N &= u^2\, v^{2n+2}\, \beta^2 \left\{ \frac{1}{(1-\alpha\, u\, v^{n+2})\,(1-\alpha\, u\, v^{n})} - \frac{1}{(1-\alpha\, u\, v^{n+1})^2} \right\} \\ &= \frac{\alpha\, u^3\, v^{3n+2}\, (v-1)^2\, \beta^2}{(1-\alpha\, u\, v^{n+2})\,(1-\alpha\, u\, v^{n+1})^2\,(1-\alpha\, u\, v^{n})}, \end{aligned}$$

and for the denominator D,

$$\begin{aligned} D &= u\, v^n\, \beta \left\{ \frac{v^2}{1-\alpha\, u\, v^{n+2}} - \frac{2\, v}{1-\alpha\, u\, v^{n+1}} + \frac{1}{1-\alpha\, u\, v^{n}} \right\} \\ &= \frac{u\, v^n\, \beta}{(1-\alpha\, u\, v^{n+2})\,(1-\alpha\, u\, v^{n+1})\,(1-\alpha\, u\, v^{n})}\, (v-1)^2\, (1+\alpha\, u\, v^{n+1}). \end{aligned}$$

Thus, we have

$$\begin{aligned} e_{n+2k+2}^{(k+1)} = \frac{N}{D} &= \frac{\alpha\, \beta\, u^2\, v^{2n+2}}{1-\alpha^2\, u^2\, v^{2n+2}} \\ &= \frac{\alpha^{2^{k+1}-1}\, \beta\, \lambda^{2^{k+1}\cdot(k+1)}\, (\lambda^{2^{k+1}})^n}{1-\alpha^{2^{k+1}}\, \lambda^{2^{k+1}\cdot(k+1)}\, (\lambda^{2^{k+1}})^n}, \end{aligned} \tag{6}$$

which completes the proof of this lemma. ∎

This lemma shows

$$x_{n+2k}^{(k)} = x^* + \mathrm{O}\,(\lambda^{2^k n}), \quad n \to \infty, \quad k = 0, 1, 2, \ldots,$$

which means a considerable improvement over the general case given by (1), since in that case

$$x_{n+2k}^{(k)} = x^* + \mathrm{O}\,(\lambda^{(k+1)\, n}), \quad n \to \infty, \quad k = 0, 1, 2, \ldots$$

Next we consider the convergence of the **bold-face sequence** in Table 1. Let us define the new sequence by

$$z_k = x_{2k}^{(k)}, \qquad k = 0, 1, 2, \ldots \tag{7}$$

For the convergence of $\{z_k\}$ we have the following lemma:

LEMMA 2 *The sequence $\{z_k\}_{k=0}^{\infty}$ defined by (7) converges to x^* super-quadratically, that is,*

$$\lim_{k\to\infty} \frac{z_{k+1} - x^*}{(z_k - x^*)^2} = 0.$$

Proof. Since $|\lambda| < 1$, it is easy to show that

$$\lim_{k\to\infty} \frac{e_{2k+2}^{(k+1)}}{\left(e_{2k}^{(k)}\right)^2} = \lim_{k\to\infty} \left(\alpha\,\lambda^{2^{k+1}}\right) \cdot \frac{\left(1 - \alpha^{2^k} \lambda^{2^k \cdot k}\right)^2}{\left(1 - \alpha^{2^{k+1}} \lambda^{2^{k+1}\cdot(k+1)}\right)\beta} = 0, \tag{8}$$

which leads to the conclusion. ∎

By the way, it is often the case that the error of an iterative method has an asymptotic expansion that

$$\begin{aligned} e_{n+1}^{(0)} &= b\,e_n^{(0)} + a\,b\,(e_n^{(0)})^2 + a^2 b\,(e_n^{(0)})^3 + \cdots \\ &= \frac{b\,e_n^{(0)}}{1 - a\,e_n^{(0)}}, \qquad n = 0, 1, \ldots, \end{aligned} \tag{9}$$

where we assume

$$a\,b \neq 0, \tag{10}$$

and $e_n^{(0)}$ is so small that $|a\,e_n^{(0)}| < 1$. In this case, $e_n^{(0)}$ has the closed form given by

$$e_n^{(0)} = \frac{b^n\,e_0^{(0)}}{1 - a\,\dfrac{b^n - 1}{b - 1}\,e_0^{(0)}}, \qquad n = 0, 1, \ldots \tag{11}$$

If we set in (11)

$$b = \lambda, \quad a = \alpha\,\beta^{-1}\,(\lambda - 1), \quad e_0^{(0)} = (1 - \alpha)^{-1}\,\beta, \tag{12}$$

then we can find the equivalence between (9) and (4). Thus we have:

Lemma 3 *The sequence* (9) *is equivalent to* (4), *if the inequalities*

$$|\alpha\,\lambda^n| < 1 \quad \textit{and} \quad |a\,e_n^{(0)}| = \left|\frac{(\lambda-1)\,\alpha\,\lambda^n}{1-\alpha\,\lambda^n}\right| < 1$$

hold for all n.

Proof. The first and the second inequalities in the condition imply the convergences of (4) and (9), respectively. ∎

Remark. In this lemma, the two inequalities are required to hold for all n. However, there may be the case that neither or only one of them holds for small n. In this case, let N be an integer such that all the inequalities hold for all $n \geq N$, and then we should replace n and α by $n' = n - N$ and $\alpha' = \alpha\,\lambda^N$, respectively.

As an application of this lemma, consider the fixed-point iteration

$$x_{n+1}^{(0)} = T(x_n^{(0)}), \qquad n = 0, 1, 2, \ldots \tag{13}$$

We assume that the iteration function $T(x)$ is defined on an interval $I \subset \mathbb{R}$ and satisfies the condition of contraction mapping [2] on this interval. Then the unique fixed point, say x^*, the point at which $x^* = T(x^*)$, exists in the interval, and the sequence $\{x_n^{(0)}\}_{n=0}^{\infty}$ with $x_0^{(0)} \in I$ is guaranteed to converges to the point. If $T(x)$ is analytic in the neighbourhood of x^*, then we have from the Taylor series expansion

$$\begin{aligned} e_{n+1}^{(0)} &= x_{n+1}^{(0)} - x^* \\ &= T'(x^*)\,e_n^{(0)} + \frac{1}{2}\,T''(x^*)\,(e_n^{(0)})^2 + \frac{1}{3\,!}\,T^{(3)}(x^*)\,(e_n^{(0)})^3 + \cdots \end{aligned} \tag{14}$$

For the particular case that $T(x)$ is a *linear fractional transformation*, we have:

Theorem 1 *Let* $T(x)$ *be*

$$T(x) = \frac{r\,x + s}{p\,x + q}, \qquad x \in I, \tag{15}$$

where p, q, r *and* s *are constants satisfying the conditions*

$$p \neq 0 \quad \textit{and} \quad p\,s - q\,r \neq 0.$$

Here we assume that $T(x)$ *satisfies the condition of contraction mapping in* I. *Then* $\{z_k\}_{k=0}^{\infty}$ *defined by* (7) *converges to the fixed point super-quadratically, if we start the iteration* (13) *with a suitably chosen* $x_0^{(0)} \in I$.

Proof. In this case, we have

$$\frac{1}{i\,!}\,T^{(i)}(x) = -\frac{(p\,s - q\,r)\,(-p)^{i-1}}{(p\,x + q)^{i+1}}, \qquad i = 1, 2, \ldots$$

Therefore, if we set

$$a = -\frac{p}{p\,x^* + q}, \qquad b = -\frac{p\,s - q\,r}{(p\,x^* + q)^2}, \tag{16}$$

then $T^{(i)}(x^*)/i\,! = a^{i-1}\,b$, and as a result, (14) reduces to (9). Note that the constants a and b defined above satisfy (10) because of the conditions of this theorem. Thus the super-quadratic convergence of $\{z_k\}$ is established. ∎

3 Numerical Example

We have proved that the Aitken Δ^2 process includes a super-quadratically converging sequence for the particular case given by (15). In this section we will show this by a numerical experiment, and compare the result with that of another iterative method, using multiple-precision arithmetic of Mathematica.

3.1 Aitken Δ^2 method

Let $T(x)$ be

$$T(x) = 1/(x + 1), \tag{17}$$

then the iteration function $T(x)$ satisfies the condition of contraction mapping on the interval $I = (0, \infty)$. The fixed point in the interval is $x^* = (-1 + \sqrt{5})/2$.

Here we investigate the conditions of Lemma 3 for the present case. First, we have from (16)

$$\begin{aligned} a &= -\frac{1}{x^* + 1} = \frac{1 - \sqrt{5}}{2} = -0.6180340\cdots \\ b &= \lambda = -\frac{1}{(x^* + 1)^2} = \frac{-3 + \sqrt{5}}{2} = -0.3819660\cdots \end{aligned} \tag{18}$$

Moreover, if we start the iteration with $x_0^{(0)} = 1$, then

$$e_0^{(0)} = x_0^{(0)} - x^* = \frac{3 - \sqrt{5}}{2} = 0.3819660\cdots, \tag{19}$$

so that $a\,e_0^{(0)} = -0.2360680\cdots$. Since $T(x)$ is a contraction mapping, we have

$$1 > |a\,e_0^{(0)}| > |a\,e_1^{(0)}| > |a\,e_2^{(0)}| > \cdots .$$

Thus the second inequality of Lemma 3 is shown to be valid. Solving (12) for α and β, we have

$$\alpha = \frac{7 - 3\sqrt{5}}{2} = 0.1458980\cdots , \qquad \beta = \frac{-15 + 7\sqrt{5}}{2} = 0.3262379\cdots , \tag{20}$$

from which the first inequality of Lemma 3 easily follows.

Now we calculate and accelerate $\{x_n^{(0)}\}$ using the multiple-precision arithmetic of Mathematica. In the experiment we set the working precision to 8000 decimal places, and start the fixed-point iteration with $x_0^{(0)} = 1$. The convergence behaviors of $\{x_n^{(k)}\}_n$ for $k = 0$ to 5 are shown in Figure 1, and that of $\{z_k\}$ is shown in Table 2.

In the table, we can observe the super-quadratic convergence of $\{z_k\}$, since d_k, a measure of the number of the correct digits in z_k, grows at the rate of greater than 2, although $\{x_n^{(k)}\}_n$ converges linearly as $n \to \infty$ for each of k, as shown in the figure.

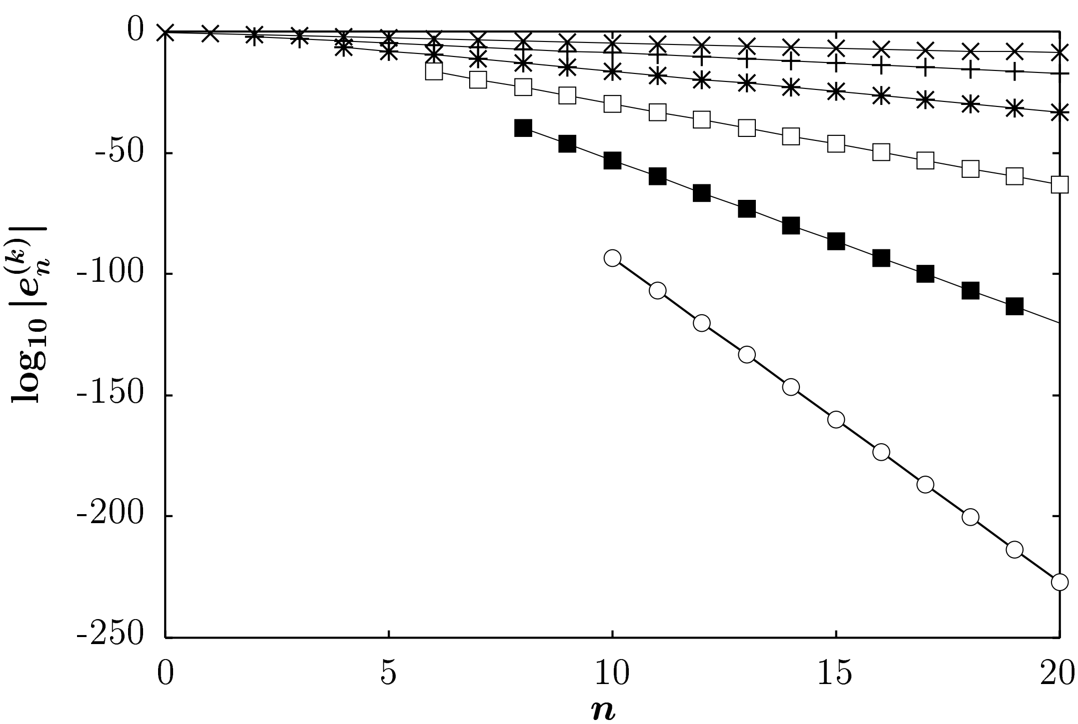

Fig. 1. Plots of $\log_{10} |e_n^{(k)}|$ for $k = 0$ (top) to 5 (bottom).

Table 2. Error $e_{2k}^{(k)}$ $(= x_{2k}^{(k)} - x^*)$ and the number of correct digits d_k in $x_{2k}^{(k)}$ for the Aitken Δ^2 method.

k	$d_k = -\log_{10}\lvert e_{2k}^{(k)}\rvert$	$\log_2 d_k$
0	0.41798	-1.2585
1	2.1570	1.1090
2	6.3381	2.6641
3	16.370	4.0329
4	39.776	5.3138
5	93.277	6.5434
6	213.65	7.7391
7	481.16	8.9103
8	1069.7	10.063
9	2353.7	11.201
10	5135.7	12.326

Next we show the Mathematica program used to get these values.

```
(*   Aitken Delta^2 method for x_{k+1}=T[x_k]   *)
   Aitken[a_, b_, c_] := a-(a-b)^2/(a - 2 b + c);
   T[x_] := 1/(x+1);

   prc  = 8000;  (*  working precision is 8000 decimal places  *)
   kmax = 10;    (*  k=0,1,,,kmax  *)
   nmax = 20;    (*  n=0,1,,,nmax  *)

   x[0, 0] = N[1, prc];  (*  starting value is x_0=1  *)

   (*   Generation of x_n^{(0)}   *)
   Do[
      x[n+1, 0] = T[x[n, 0]],
      {n, 0, nmax-1}
   ];

   (*   Aitken acceleration   *)
   Do[
      Do[
         x[n+2, k+1] = Aitken[x[n+2, k], x[n+1, k], x[n, k]],
         {n, 2 k, nmax-2}
      ],
      {k, 0, kmax-1}
   ];
```

3.2 Steffensen's method

For comparison we compute the problem by Steffensen's method. Steffensen [4] has proposed an Aitken like iterative method, which achieves a quadratic convergence locally without evaluating the derivative. The recurrence relation of Steffensen's method for $T(x)$ is given by

$$x_{k+1} = T(T(x_k)) - \frac{(\,T(T(x_k)) - T(x_k)\,)^2}{T(T(x_k)) - 2\,T(x_k) + x_k}, \qquad k = 0, 1, 2, \ldots \tag{21}$$

The sequence $\{x_k\}$ by the method converges to x^* quadratically, if $T'(x^*) \neq 1$ (see Sidi [3], pp. 90–92). As before we calculate the sequence by using Mathematica with the same working precision and the same initial value. The result is shown in Table 3.

From the table we can observe the quadratic convergence of the sequence, since $\log_2 d_k$ increases by 1 for every step.

Table 3. Error $e_k\ (= x_k - x^*)$ and the number of correct digits d_k in x_k for Steffensen's method (21).

k	$d_k = -\log_{10} \lvert e_k \rvert$	$\log_2 d_k$
0	0.41798	-1.2585
1	2.1570	1.1090
2	5.5022	2.4600
3	12.190	3.6076
4	25.565	4.6761
5	52.315	5.7092
6	105.82	6.7254
7	212.82	7.7334
8	426.82	8.7375
9	854.83	9.7395
10	1710.8	10.741
11	3422.8	11.741
12	6846.9	12.741

The Mathematica program for this calculation is the following:

```
(*    Steffensen's method    *)
   Aitken[a_, b_, c_] := a-(a-b)^2/(a - 2 b + c);
   T[x_] := 1/(x+1);

   prc  = 8000;  (*  working precision is 8000 decimal places  *)
   kmax = 12;    (*  k=0,1,,,kmax  *)

   x[0] = N[1, prc];    (*  starting value is x_0=1  *)

   (*   Steffensen iteration   *)
   Do[
      u=x[k];
      v=T[u];
      w=T[v];
      x[k+1] = Aitken[w, v, u],
      {k, 0, kmax-1}
   ];
```

Acknowledgment. The author is very grateful to Dr. C. Hirota, the research assistant of Akita Prefectural University, who has stimulated the research.

References

[1] G. Evans, Practical Numerical Integration, John Wiley & Sons, New York, 1993.

[2] J.M. Ortega and W.G. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York, 1970.

[3] A. Sidi, Practical Extrapolation Methods — Theory and Applications, Cambridge University Press, Cambridge, 2003.

[4] J.F. Steffensen, Remarks on iteration, Skand. Aktuar. Tidskr., **16** (1933), 64–72.

[5] J.F. Traub, Iterative Method for the Solution of Equations, Chelsea, New York, 1982.