1989

On Some Convergence Results of the k-Step Interactive Methods

S. Galanis

A. Hadjidimos

Report Number:

89-882

ON SOME CONVERGENCE RESULTS OF THE k-STEP ITERATIVE METHODS

S. Galanis[*,**]

and

A. Hadjidimos[*,**]

Computer Sciences Department
Purdue University

Technical Report CSD-TR-882
CAPO Report CER-89-15
April, 1989

1980 Mathematics Subject Classifications AMS(MOS): Primary 65F10. C.R. Categories: 5.14.

* Visiting, Department of Computer Science, Purdue University, West Lafayette, IN 47907.
The work of the second author was supported in part by NSF grant CCR-8619817.

** Department of Mathematics, University of Ioannina, GR-451 10, Ioannina, Greece.

On Some Convergence Results of the k-Step Iterative Methods

S. Galanis*,**
and
A. Hadjidimos*,**

Computer Sciences Department
Purdue University
Technical Report CSD-TR-882
CAPO Report CER-89-15
April, 1989

Abstract

For the iterative solution of the nonsingular linear system (1) $(I-T)x=c$ we consider the class of monoparametric k-step methods (2) $x^{(m)}=\omega Tx^{(m-1)}+(1-\omega)x^{(m-k)}+\omega c$ for $k=1, 2, 3, \ldots$, with ω being a real parameter. The main objectives of this paper are the following: i) To determine the value of $k=1, 2, 3, \ldots$ for which the method in (2) converges asymptotically as fast as possible under the assumption that $\sigma(T)\in[\alpha, \beta]$, $-\infty<\alpha\leq\beta<1$ and ii) For a given $\sigma(T)$, not necessarily on the real axis, and for a given $k\geq 3$ to make an attempt toward the determination of an "optimal" ω in the sense of (i) above. Finally based on a recent result by Eirmann, Neithammer and Ruttan for the k-cyclic SOR method we discuss and suggest possible ways of extending and improving the results in (i) and (ii) above.

1980 Mathematics Subject Classifications AMS (MOS): Primary 65F10. C.R. Categories: 5.14.

* Visiting, Department of Computer Science, Purdue University, West Lafayette, IN 47907.
The work of the second author was supported in part by NSF grant CCR-8619817.

** Department of Mathematics, University of Ioannina, GR-451 10, Ioannina, Greece.

1. INTRODUCTION AND PRELIMINARIES

Consider the nonsingular linear system

$$x = Tx + c \tag{1.1}$$

$T \in \mathbb{C}^{n,n}$, x, $c \in \mathbb{C}^n$, and assume that for its iterative solution the class of monoparametric k-step methods

$$x^{(m)} = \omega T x^{(m-1)} + (1-\omega)\, x^{(m-k)} + \omega c, \quad m = 1, 2, 3, \ldots, \tag{1.2}$$

is considered. In (1.2) $x^{(\ell)}$, $\ell = 0(-1) - k + 1$ are arbitrary and $\omega \in \mathbb{R} \setminus \{0\}$. As is known, except in very special cases, knowledge of the spectrum $\sigma(T)$ of T does not give any information as to the value of ω for which the scheme (1.2) converges asymptotically as fast as possible for a given $k(\geq 3)$. Even if we assume that for a given $\sigma(T)$ all the optimal schemes in (1.2) for all k's are known the question of which one is the best of them all is an open one.

In this paper we develop some background material which enlightens some of the difficulties one has to overcome before one attacks the two problems posed above. We then restrict ourselves to the following two main specific objectives: a) For a given $\sigma(T)$ satisfying

$$\sigma(T) \subset I := [\alpha, \beta], \; -\infty < \alpha \leq \beta < 1 \quad , \tag{1.3}$$

only those k's for which (1.2) converges for some ω's will be considered. In the case where convergence is guaranteed the corresponding optimal scheme, that is the one which converges asymptotically faster, will be selected and out of all optimal schemes the best one over all k's will be determined and b) It will be assumed that $\sigma(T)$ is known in some "sense" and an attempt will be made to determine the "optimal" ω for a given k by means of a simple algorithm. It is our belief that this effort will provide us with some of the background material needed in devising an algorithm for the determination of the optimal ω. It is hoped that this algorithm will be analogous to those by Hughes-Hallert [13] and Hadjidimos [11] (see also Hadjidimos [12] and Opfer and Schober [17] for complex ω's) for $k = 1$ and to that by Avdelas et al [2], which is nothing but the one by Young and Eidson [23] (see also Young [22]) devised for the 2-cyclic SOR method, for $k = 2$. For $k \geq 3$ a necessary condition for convergence is $\rho(T) < k / (k-2)$ and (optimal) convergence can be achieved whenever $\sigma(T)$ lies in the interior of a certain hypocycloid. In this paper the optima are determined by means of the "cusped" hypocycloids. The necessary background material can be found in the papers by Niethammer and Varga [16], Galanis et al [7] – [9] and in Wild and Niethammer [20], where an excellent account of the hypocycloids in connection with the (optimal) convergence of the k-cyclic SOR method is given.

The organization of this paper is as follows: In Section 2 the (optimal) 1-, 2- and k-step ($k \geq 3$) methods (1.2), under the assumption (1.3), are given and determined. Especially for $k \geq 3$ this is done by means of the "cusped" hypocycloids. A comparison among all optimal methods is made and the best one is given in Table 1.

In Section 3 we present, for a given k, an algorithm to determine the best optimal k-step method by means of the "cusped" hypocycloids again, under a more general assumption on the configuration of $\sigma(T)$, than that in (1.3). Finally in Section 4 we discuss an excellent recent result due to Eiermann, Niethammer and Ruttan [23] for the k-cyclic SOR method and "translate" it to its "equivalent" for the k-step methods (1.2), $k \geq 2$. This "translation" provided us with tools which could improve and extend the result of this paper. Suggestions for possible further investigations are also made.

2. OPTIMAL k-STEP METHODS

2.1. 1- and 2-Step Methods

As is known the 1-step method (1.2) converges for any I defined in (1.3) for infinitely many values of ω belonging to a certain interval (see e.g., Isaacson and Keller [14] Young [22], de Pillis and Neumann [3] and Hadjidimos [11]). The corresponding optimal values for $\hat{\omega}_1$ and the spectral radius $\hat{\rho}_1$ of the iteration matrix $\hat{\omega}_1 T + (1-\hat{\omega}_1)I$ (that is the optimal asymptotic convergence factor) are given by

$$\hat{\omega}_1 = \frac{2}{2-(\alpha+\beta)}, \quad \hat{\rho}_1 = \frac{\beta-\alpha}{2-(\alpha+\beta)} \quad , \tag{2.1}$$

which can also be obtained from the results of Young [21] (see also [22, pp. 364-365]) in connection with the cyclic Chebyshev acceleration by considering only one parameter (ω) in each cycle.

Remark: The optimal 1-step method (1.2) exists and is unique even for $I := [\alpha,\beta]$, $1 < \alpha \leq \beta < +\infty$, in which case all other k-step methods ($k \geq 2$) fail to converge for any real ω. The optimal parameters are again given by (2.1), with the only difference being that the denominator in the expression for $\hat{\rho}_1$ in (2.1) is $(\alpha+\beta)-2$ (see e.g., [11]).□

For the 2-step method (1.2) a necessary and sufficient condition for convergence for some ω's is that $I \subset (-1,1)$ and the corresponding optimal parameters are then given by

$$\hat{\omega}_2 = \frac{2}{1+(1-\gamma^2)^{1/2}}, \quad \hat{\rho}_2 = (\hat{\omega}_2 - 1)^{1/2}, \tag{2.2}$$

where $\gamma = \max\{|\alpha|, |\beta|\}$ (see e.g., Golub and Varga [10], Varga [19], Niethammer [15] and Young [22]).

Since the 2-step method (1.2) converges only for $-1 < \alpha$ we compare $\hat{\rho}_1$ and $\hat{\rho}_2$ under the assumption that $I \subset (-1,1)$. A straightforward comparison distinguishing the two cases $|\alpha| \le |\beta|$ (< 1) and $|\beta| \le -\alpha (< 1)$ reveals that:

$$\begin{aligned} &i) \ \ |\alpha| \le \beta \ \ : \ \ \hat{\rho}_1 \le \hat{\rho}_2 \ \text{ iff } \ 1-(1-\beta^2)^{1/2} \le \alpha \ , \\ &ii) \ |\beta| \le -\alpha \ : \ \ \hat{\rho}_1 \le \hat{\rho}_2 \ \text{ iff } \ -1 < \alpha \le 1-(1-\beta^2)^{1/2} \ . \end{aligned} \tag{2.3}$$

Note: Some of the results (2.3), in a special case, were obtained in Avdelas et al [1].□

2.2. The k-Step Method (k ≥ 3)

2.2.1. Background Material

We begin out analysis by giving a Lemma and a Theorem in connection with schemes (1.2) (see e.g., Galanis et al [7] – [9]) which are to be used in the sequel:

Lemma 1: a) The circle $\phi = \eta e^{i\theta}$, $\eta > 0$, $\theta \in [0,2\pi)$ in the complex plane is transformed through

$$z := p(\phi) := \frac{1-(1-\omega)\phi^k}{\omega\phi}, \ \omega \in (0,2)\setminus\{1\},\ k \ge 3 \ , \tag{2.4}$$

into a closed curve C_k (hypocycloid) consisting of k arcs symmetric w.r.t the lines through the origin with arguments $2\pi\ell / k$, $\ell = 0(1)k-1$. Rotations of the complex plane through angles of $2\pi / k$ but through no smaller angles carry C_k into itself. b) For $s := \operatorname{sign}(\omega - 1)$ and a given $\rho = |p(\hat{\eta}^{1/k} \exp((1-s)i\pi/(2k)))| \in (0, k/(k-s+1))$ (the case $\rho = 0$ is trivial) and for every $\eta \in (0,\hat{\eta}]$, with $\hat{\eta} = (s(k-1)(\hat{\omega}_k - 1))^{-1/k}$ and only for these values of η, C_k is a simple curve, where $\hat{\omega}_k$ is the unique positive real root in $(\min\{1, (k+s-1)/(k-1)\}, \max\{1,(k+s-1)/(k-1)\})$ of the equation

$$(\omega\rho)^k = sk^k (k-1)^{1-k} (\omega - 1) \ . \quad \square \tag{2.5}$$

Theorem 1: Let $\hat{R}_k := \overline{\operatorname{int}\hat{C}_k}$ of Lemma 1 ($\hat{C}_k$ is the C_k corresponding to $\hat{\omega}_k$). Then if $\sigma(T) \subset \hat{R}_k$ the method (1.2) converges and has an asymptotic convergence factor *(acf)* $\hat{\rho}_k \le 1/\hat{\eta}_k = (s(k-1)(\hat{\omega}_k - 1))^{1/k}$ with equality holding iff at least one element of $\sigma(T)$ lies on $\hat{C}_k$. □

Note 1: The graphs of $\hat{C}_k$ for $k = 3,4$ ($\hat{\omega}_k > 1$) and $k = 3,4$ ($\hat{\omega}_k < 1$) are given in Figures 1, 2, 3 and 4 respectively.□

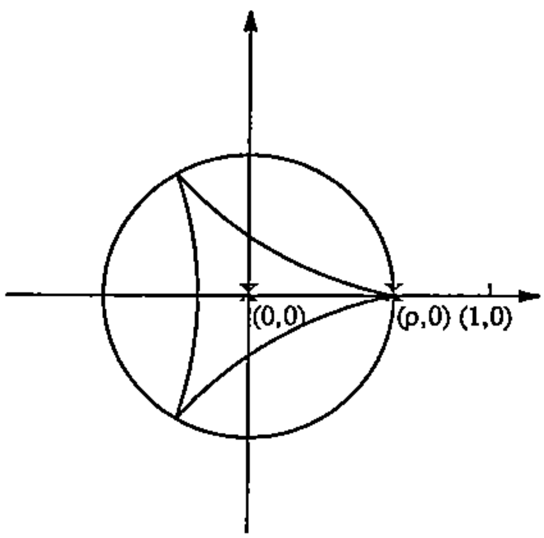

Figure 1.

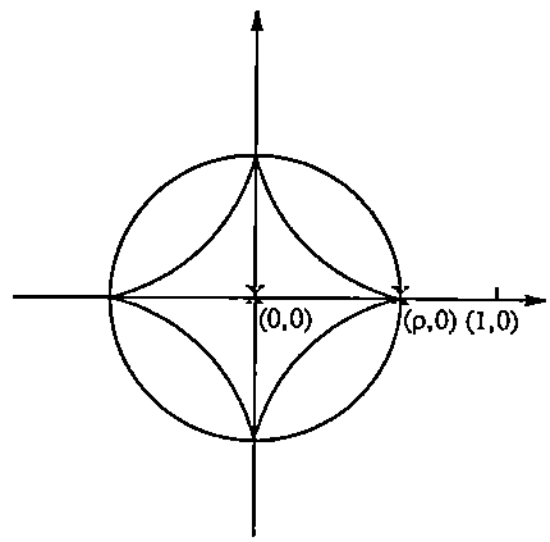

Figure 2.

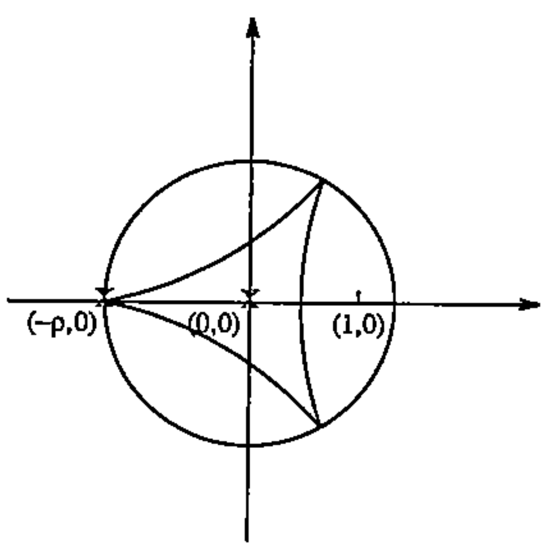

Figure 3.

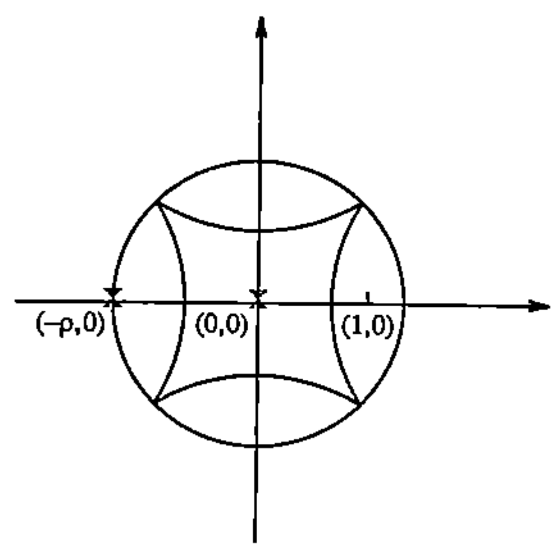

Figure 4.

Note 2: $\hat{C}_k$ is a cusped hypocycloid (see Wild and Niethammer [20]) with cusps on the rays, emanating from the origin, with arguments $2\pi\ell / k, \ell = 0(2)k - 1$, iff $\hat{\omega}_k > 1$ and $(2\ell + 1)\pi / k$, $\ell = 1(2)k - 1$, iff $\hat{\omega}_k < 1$. □

Note 3: From now on we shall refer to $\hat{C}_k$'s as being of type I or II iff $\hat{\omega}_k > 1$ or $\hat{\omega}_k < 1$ respectively. □

2.2.2. $[\alpha,\beta] \subset (-1,1)$

Before we go on with our analysis we make two points: i) The analysis can be restricted to intervals $I := [\alpha,\beta]$ such that $\alpha \le 0 \le \beta(\alpha+\beta \ne 0)$ since as we shall see all other possible cases are subcases of the ones which will be examined here. and ii) One should have in mind that in principle four categories of optimal k-step methods (1.2) can be considered depending on the way the cusped hypocycloid $\hat{C}_k$ associated with (1.2) is constructed. Thus we can have: A) $\hat{C}_k$ of type I with one cusp at the point β in the complex plane. B) $\hat{C}_k$ of type I with one cusp at $-\alpha\frac{k-2}{k}$. C) $\hat{C}_k$ of type II with one cusp at $-\alpha\exp(i\pi/k)$ and D) $\hat{C}_k$ of type II with one cusp at $\beta\frac{k}{k-2}\exp(i\pi/k)$. It is understood that in each case only those $\hat{C}_k$'s for which $[a,b] \subset \overline{\mathrm{int}\hat{C}_k}$ will give $\hat{\rho}_k$'s which should be compared against $\hat{\rho}_2$ based on $[\alpha,\beta]$. Two cases are then examined: I) $\alpha+\beta > 0$ and II) $\alpha+\beta < 0$.

I. $\alpha+\beta > 0$.

Consider all $\hat{C}_k$'s, $k \ge 2$, of the category A above. Because of Lemma 1 and Theorem 1 the optimal parameters for each $k \ge 2$, ignoring for the moment whether $\alpha \in \overline{\mathrm{int}\hat{C}_k}$ will satisfy

$$(\hat{\omega}_k\beta)^k = k^k(k-1)^{1-k}(\hat{\omega}_k - 1), \quad \hat{\rho}_k = \left[(k-1)(\hat{\omega}_k - 1) \right]^{1/k} . \tag{2.6}$$

Setting $y = y(k) = \hat{\rho}_k$ and considering it, by following the reasoning in the proof of Thm 4 of [7], which was also used successfully in [8], [9] and [18], as a (differentiable) function of the real $k \ge 2$, one obtains from (2.6) that

$$\beta = \frac{ky}{y^k + k - 1} . \tag{2.7}$$

Differentiating (2.7) w.r.t. k and rearranging terms we take

$$\frac{k}{y}(1-y^k)(1-k)y' = y^k - y^k \ln y^k - 1 \equiv \gamma , \tag{2.8}$$

that is (2.8) in [7]. Putting $y^k = \delta \in (0,1)$ and $\varepsilon = \gamma/\delta = 1 - \ln\delta - 1/\delta$ and $\partial\varepsilon/\partial\delta = (1-\delta)/\delta^2 > 0$. Hence ε increases with γ and because $\varepsilon = 0$ at $\delta = 1$ it is implied that ε is negative. So is γ and from (2.8) we have that y increases with k. Here we observe that for $k = 2$, $\hat{C}_2$:= the double line segment $[-\beta,\beta]$ contains $[\alpha,\beta]$. Also there exists a unique value $\ell = \min_{k=3,4,5,\ldots} \{k := -\beta(k-2)/k \le \alpha\}$. Therefore for $k = \ell, \ell+1, \ell+2, \ldots$ $[\alpha,\beta] \subset \overline{\mathrm{int}}\hat{C}_k$. Consequently $\hat{\rho}_2 < \hat{\rho}_k$, $k = \ell, \ell+1, \ell+2, \ldots$.

Consider all C_k's, $k \ge 3$, of the category B such that $-\alpha\dfrac{k}{k-2} < 1$ (if they exist) and let ℓ be the smallest value of k such that $\beta \le -\alpha\dfrac{k}{k-2}$. Following an analysis similar to the one before except that β is replaced by $-\alpha\dfrac{k}{k-2}$ and $k \ge \ell$, (2.7) becomes

$$-\alpha = \frac{(k-2)y}{y^k + k - 1} \quad . \tag{2.7'}$$

Therefore, instead of (2.8), we now have

$$\frac{(k-2)}{y}(1-y^k)(1-k)\,y' = y^k - y^k \ln y^k - 1 + 2y^k \ln y \tag{2.8'}$$

Because of the previous analysis and the fact that $2y^k \ln y < 0$, y increases with k, so $\hat{\rho}_\ell < \hat{\rho}_k$, $k = \ell+1, \ell+2, \ldots$. This value should be compared against $\hat{\rho}_2$ of type I whose cusp is at β. For this we consider the equation

$$f(y) := \alpha y^\ell + (\ell-2)y + \alpha(\ell-1) = 0 \quad , \tag{2.9}$$

which has a unique root $y \in (0,1)$. Since $f(0) = \alpha(\ell-1) < 0$ and $f(1) = \alpha\ell + (\ell-2) = (\ell-2)(1+\alpha\dfrac{\ell}{\ell-2}) > 0$ it is concluded that if $f(\hat{\rho}_2) < 0$, with $\hat{\rho}_2 = 2/(1+(1-\beta^2)^{1/2})$, then $\hat{\rho}_2 < \hat{\rho}_\ell$ otherwise the situation is reversed.

Consider now all $\hat{C}_k$, $k \ge 3$, of the category C. As is readily seen none of these curves can have $\beta \in \overline{\mathrm{int}}\hat{C}_k$ since their point of intersection with the positive real semiaxis is at $-\alpha\dfrac{k-2}{k}$ $(<\beta)$.

Finally consider all $\hat{C}_k$, $k \ge 3$, of the category D. It is apparent that for k even $\hat{C}_k$ contains $[\alpha,\beta]$ and therefore based on Thm 16 of [6] or Thm 6 of [4] it is concluded that the $\hat{\rho}_2$ found before is smaller than any $\hat{\rho}_k$, $k = 4, 6, 8, \ldots$. A further point should be made in case k is odd. In such a case $-\beta\dfrac{k}{k-2}$ is a cusp of $\hat{C}_k$ and by virtue of the previous Thms in [6] and [4] $\hat{\rho}_k$, $k = 3, 5, 7, \ldots$ can not be all good as $\hat{\rho}_2$ of type I with cusps at $-\beta$ and β.

II. $\alpha + \beta < 0$.

In this case it is obvous that the $\hat{C}_k$'s of the category A can not contain α.

Consider then all $\hat{C}_k$'s, $k \geq 3$, of the category B. For k even, $\hat{\rho}_k$ can not be better than $\hat{\rho}_2$ based on $\hat{C}_2$ with cusps at α and $-\alpha$. For k odd consider only those k's for which $-\alpha \frac{k}{k-2} < 1$ and let ℓ be the smallest k satisfying the inequality in question. Following an analysis identical to that in IB before we have that $\hat{\rho}_\ell < \hat{\rho}_k$, $k = \ell + 2, \ell + 4, \ldots$. But since $[\alpha, \beta] \subset [\alpha, -\alpha] \subset \overline{\mathrm{int}}\hat{C}_k$ it is concluded that $\hat{\rho}_2$ based on $[\alpha, -\alpha]$ is better than $\hat{\rho}_\ell$.

Consider now all C_k's, $k \geq 3$, of the category C. For k even again $\hat{\rho}_2$ based on $[\alpha, -\alpha]$ is better than any $\hat{\rho}_k$ because $[-\alpha, \alpha] \subset \overline{\mathrm{int}}\hat{C}_k$. For k odd let ℓ be the smallest k such that $[\alpha, \beta] \subset \overline{\mathrm{int}}\, \hat{C}_k$. This time $\hat{\rho}_2$ must be compared against $\hat{\rho}_\ell$ which is the unique root in $(0,1)$ of

$$(-\hat{\omega}_k \alpha)^k = k^k (k-1)^{1-k} (1 - \hat{\omega}_k),\ \hat{\rho}_k = ((k-1)(1-\hat{\omega}_k))^{1/k} \quad , \tag{2.10}$$

which give

$$-\alpha = \frac{ky}{-y^k + k - 1} \quad , \tag{2.11}$$

where $y = y(k) = \hat{\rho}_k$ is considered a (differentiable) function of the real variable $k \geq 3$. (2.11) is analogous to (2.7) and gives in turn

$$\frac{k}{y}(1 + y^k)(1 - k) y' = -y^k - y^k \ln y^k - 1 < 0 \quad . \tag{2.12}$$

Hence $y' > 0$ and y increases with k meaning that $\hat{\rho}_\ell < \hat{\rho}_k$, $k = \ell + 2, \ell + 4, \ldots$. Since

$$f(y) := \alpha y^\ell - \ell y - \alpha(\ell - 1) = 0 \tag{2.13}$$

and $f(0) = -\alpha(\ell - 1) > 0$ and $f(1) = -\alpha(\ell - 2) - \ell = -\ell(1 + \alpha\frac{(\ell-2)}{\ell}) < 0$, $f(\hat{\rho}_2) < 0$ implies $\hat{\rho}_2 > \hat{\rho}_\ell$ and $f(\hat{\rho}_2) > 0$ implied $\hat{\rho}_2 < \hat{\rho}_\ell$.

Finally consider all $\hat{C}_k$'s , $k \geq 3$, of the category D. For k even, $\alpha \notin \overline{\mathrm{int}}\hat{C}_k$. For k odd let ℓ be the largest $k = 3, 5, 7, \ldots$ such that $-\beta \frac{k}{k-2} \leq \alpha$, if such an ℓ exists. For $k = \ell, \ell - 2, \ldots, 3$ it will be

$$\left[\hat{\omega}_k \frac{k}{k-2} \beta \right]^k = k^k (k-1)^{1-k} (1 - \hat{\omega}_k) \quad ,$$

which gives

$$\beta = \frac{(k-2)y}{-y^k + k - 1} \quad , \tag{2.14}$$

where again $y = y(k) = \hat{\rho}_k$ is a function of the real $3 \leq k \leq \ell$. Following a similar analysis we get

$$\frac{(k-2)}{y}(1+y^k)(1-k)y' = -y^k + y^k \ell n y^k + 1 - 2y^k \ell n y > 0 \quad . \qquad (2.15)$$

This is because it can be proved that the minimum value of $-y^k + y^k \ell n y^k + 1$ is 0 and is attained for $y^k = 1$, and also $-2y^k \ell n y > 0$. So $y' < 0$, implying that $\hat{\rho}_\ell < \hat{\rho}_k$, $k = \ell - 2(-2)3$. Even in this case $\hat{\rho}_2$ corresponding to $[\alpha, -\alpha]$ has to be compared with $\hat{\rho}_\ell$.

This analysis concludes the case $[\alpha,\beta] \subset (-1,1)$.

2.2.3. $[\alpha,\beta] \subset (-3,1)$, $\alpha \leq -1$

III) In this case matters are a little simpler for two reasons. First (1.2), diverges for $k = 2$ and second the categories A and B of the $\hat{C}_k$ (cusped hypocycloids) do not exist any more.

Consider then all $\hat{C}_k$'s, $k \geq 3$ of the category C. Obviously k must be odd. Let ℓ be the smallest odd integer k such that $-\alpha\frac{k-2}{k} \geq \beta$. It will be

$$(-\hat{\omega}_k\alpha)^k = k^k(k-1)^{1-k}(1-\hat{\omega}_k), \quad \hat{\rho}_k = ((k-1)(1-\hat{\omega}_k))^{1/k} \quad , \qquad (2.10')$$

that is (2.10). Hence the best $\hat{\rho}_k$ is $\hat{\rho}_\ell$ as in 2.2.2 IIC.

Finally consider all $\hat{C}_k$'s, $k \geq 3$, of the category D. Again, as in the case 2.2.2.IID the results are the same except that $\hat{\rho}_\ell$ does not have to be compared with $\hat{\rho}_2$ any more. ℓ is given as the largest integer such that $-\beta\frac{k}{k-2} \leq \alpha$.

2.3. Comparison of the Optimal k-Step Methods

Having in mind the results of the various cases examined so far one may give Table 1 in which the best of all the optimal methods (1.2) is presented. To give an idea as to how the Table is easily constructed we have to make three points: i) The various subcases for $\alpha,\beta \in (1-1,1)$ are defined by the relationships (2.3) taken together with $\alpha+\beta > 0$ and $\alpha+\beta < 0$ in subsections 2.2.2.I and 2.2.2II respectively. ii) Whenever an $\ell(\geq 3)$ can not be defined in one of the subcases considered through the corresponding relationship this simplly means that there is no optimal ℓ-step method with $\ell \geq 3$ and therefore the best out of the 1- and 2-step methods is to be considered. and iii) Whenever an optimal ℓ-method exists and if to be compared with either the 1- or the 2-step method then the sign of the function $f(\hat{\rho}_\ell)$ of the left hand side of the equation from which $\hat{\rho}_\ell$ is determined (with the right hand side being zero as for example in (2.9) or (2.13)) gives the relative position of either $\hat{\rho}_1$ or $\hat{\rho}_2$ w.r.t. $\hat{\rho}_\ell$.

Table 1

<table>
<tr><th colspan="2">$\sigma(T) \in I := [\alpha, \beta]^*$</th><th>acf of the best optimal k-step method (1.2).</th><th>Section in the text where the index in $\hat{\rho}$ in the previous column is defined (if it exists).</th></tr>
<tr><td colspan="2">$1 < \alpha < \beta < \infty$</td><td>$\hat{\rho}_1$</td><td>2.1</td></tr>
<tr><td colspan="2">$-\infty < \alpha \leq -3 < \beta < 1$</td><td>$\hat{\rho}_1$</td><td>2.1</td></tr>
<tr><td colspan="2">$-3 < \alpha \leq -1 < \beta < 1$</td><td>$\hat{\rho}_1, \hat{\rho}_{\ell_1}, \hat{\rho}_{\ell_2}$</td><td>2.1, 2.2.3C, 2.2.3D respectively</td></tr>
<tr><td colspan="2">$-1 < \alpha < \beta \leq 0$</td><td>$\hat{\rho}_1, \hat{\rho}_{\ell_1}, \hat{\rho}_{\ell_2}$</td><td>2.1, 2.2.2IIB, 2.2.2.IIC >></td></tr>
<tr><td rowspan="3">$0 < \beta < 1$</td><td>$-1 < \alpha \leq -\beta$</td><td>$\hat{\rho}_1, \hat{\rho}_{\ell_1}, \hat{\rho}_{\ell_2}$</td><td>2.1, 2.2.2.IIC, 2.2.2IID >></td></tr>
<tr><td>$-\beta < \alpha \leq 1-(1-\beta^{2})^{1/2}$</td><td>$\hat{\rho}_2, \hat{\rho}_{\ell_1}, \hat{\rho}_{\ell_2}, \hat{\rho}_{\ell_3}$</td><td>2.1, 2.2.2.IA, 2.2.2.IB, 2.2.2.ID >></td></tr>
<tr><td>$1-(1-\beta^2)^{1/2} \leq \alpha < \beta$</td><td>$\hat{\rho}_1, \hat{\rho}_{\ell_1}, \hat{\rho}_{\ell_2}$</td><td>2.1, 2.2.2IA, 2.2.2ID >></td></tr>
<tr><td colspan="4">* It is always assumed that $\alpha < \beta$ because:
For $\alpha = \beta = 1$: None of the k-step methods converges.
For $\alpha = \beta \neq 0,1$: The optimal 1-step method is the best, $\hat{\rho}_1 = 0$.
For $\alpha = \beta = 0$: All optimal methods have $\hat{\omega}_k = 1$, because $\sigma(T) = 0$,
so the optimal method (1.2) is simply $x^{(m)} = Tx^{(m-1)} + c$.</td></tr>
</table>

3. THE OPTIMAL k-STEP METHOD FOR A GIVEN k

As has already been mentioned in principle for a given $k = 1, 2, 3, \ldots$ an optimal k-step method can be found whenever $\sigma(T)$ is given and lies in a certain region. Thus for $k = 1$ an optimal 1-step method can be found whenever $\sigma(T)$ lies strictly to the left (or strictly to the right) of the line $Rez = 1$ and for $k = 2$ whenever $\sigma(T)$ lies within the strip $|Rez| < 1$. However, for $k \geq 3$ an optimal k-step method can be found whenever $\sigma(T)$ lies strictly in the interior of the k-cusped hypocycloid $\hat{C}_{k,II}$ whole k cusps are at the points $\frac{k}{k-2}\exp((2\ell+1)\pi i / k)$, $\ell = 0(2)k-1$. This is an immediate consequence of the Lemma 1 and Theorem 1 (see also [20]). So, for some configurations of spectra $\sigma(T)$ only a 1- or a 2- or a k-step, with a specific $k \geq 3$, method can be applied to produce an optimal one. Since $\bigcap_{k=3}^{\infty} \overline{\mathrm{int}\hat{C}_{k,II}} = \overline{D}_1$ is the unit disc (centered at the origin) it is concluded that iff $\sigma(T) \subset D_1$ all k-step methods, $k = 1, 2, 3, \ldots$ can be applied to produce an optimal one.

Let us assume that a k is given and that $\sigma(T) \subset \mathrm{int}\, \hat{C}_{k,II}$. Assume further that we consider the images of all elements of $\sigma(T)$ in the following way: i) By rotating them about the origin by a multiple of $2\pi / k$ so that their images have argument in $(0, 2\pi / k)$ and, ii) By taking the mirror images of them all whose argument is in $(\pi / k, 2\pi / k)$ w.r.t. the line through the origin with argument in $[0, \pi / k]$. Obviously, due to the rotational symmetry, if one determines the optimal cusped hypocycloid taking into consideration only the images (of all points) with arguments in $[0, \pi / k]$ the problem of determining the optimal k-step method (1.2)) will have been solved. Assume that the images in question lie in a known line L consisting of all possible consecutive straight line segments. Let P_s, $s = 1(1)n$, be the vertices of this line in increasing order of their arguments. The solution will be given by using cusped hypocycloids of type I and type II only. For this consider the two limiting hypocycloids of type I and II $(\hat{C}_{k,I}, \hat{C}_{k,II})$ respectively which pass through the point 1. Obviously, if all vertices $P_s \in \mathrm{int}\, \hat{C}_{k,II}$, $s = 1(1)n$, there exists a unique optimal solution to our problem determined by a type II hypocycloid iff $P_s \notin \mathrm{int}\hat{C}_{k,I}$ for at least one s, while the unique optimal solution is determined by a hypocycloid of either type I or II iff $P_s \in \mathrm{int}\, \hat{C}_{k,I}$ for all $s = 1(1)n$.

Let $P_1(r_1, \psi_1)$, in polar coordinates, be the first vertex of L. The unique optimal hypocycloid of type I passing through P_1 is determined as follows: Find the unique $\theta_1 \in [-\pi / k, 0]$ by solving

$$\frac{-(k-1)\sin\theta_1 + \sin(k-1)\theta_1}{(k-1)\cos\theta_1 + \cos(k-1)\theta_1} = \tan\psi_1 \tag{3.1}$$

for θ_1 and then the unique $\hat{\rho}_{k,1} \in (0,1)$ through

$$r_1 \hat{\rho}_{k,1}^{k} - [(k-1)^2 + 1 + 2(k-1)\cos k\theta_1]^{1/2} \hat{\rho}_{k,1} + (k-1) r_1 = 0 \tag{3.2}$$

Equations (3.1) – (3.2) are obtained from

$$\begin{aligned} &i)\ r\cos\psi = x = \frac{1}{\rho^{k-1}}(\rho^k \sin\theta + (\omega - 1)\cos(k-1)\theta) \\ &ii)\ r\sin\psi = y = -\frac{1}{\rho^{k-1}}(\rho^k \sin\theta - (\omega - 1)\sin(k-1)\theta) \end{aligned} \tag{3.3}$$

(see equations (3.2) of [20], and take into consideration that

$\hat{\rho}_{k,1} = \dfrac{1}{\hat{\eta}_k} = ((k-1)(\hat{\omega}_k - 1))^{1/k}$ from Theorem 1 and that $\hat{\omega}_k$ satisfies (2.5)).

In the same way one applies (3.1) – (3.2) with 1 being replaced by $s = 2(1)n$ successively. Obviously from Lemma 1 and Theorem 1 one has $\hat{\omega}_{k,s} = \hat{\rho}^{\,k}_{k,s} / (k-1) + 1$ and therefore from (2.5)

$$(\hat{\omega}_{k,s}\rho_s)^k = k^k(k-1)^{1-k}(\hat{\omega}_{k,s} - 1) \quad , \tag{3.4}$$

where ρ_s is the unique real root in $(0,1)$ of equation (3.4). One then takes

$$\rho_I = \max_{s=1(1)\eta} \rho_s \quad . \tag{3.5}$$

Of course to the latter ρ_I there correspond a unique pair $\omega^{(I)}_{k,s}$ and a $\hat{\rho}^{(I)}_{k,s}$.

In a similar way one determines the optimal hypocycloid of type II passing through P_s, $s = 1(1)n$. The corresponding equations to (3.1) – (3.2) are now

$$\frac{(k-1)\sin\theta_1 + \sin(k-1)\theta_1}{-(k-1)\cos\theta_1 + \cos(k-1)\theta_1} = \tan\psi_1 \tag{3.1'}$$

and

$$r_1\hat{\rho}^{\,k}_{k,1} + [(k-1)^2 + 1 - 2(k-1)\cos k\theta_1]^{1/2}\hat{\rho}_{k,1} - (k-1)r_1 = 0 \quad , \tag{3.2'}$$

where of course $(\theta_1, \hat{\rho}_{k,1})$ of (3.1') – (3.2') are different from $(\theta_1, \hat{\rho}_{k,1})$ of (3.1) – (3.2). Equations (3.1') – (3.2') are obtained from (3.3) and the fact that $\hat{\rho}_{k,1} = \dfrac{1}{\hat{\eta}_k} = ((k-1)(1-\hat{\omega}_k))^{1/k}$ from Theorem 1 and that $\hat{\omega}_k$ satisfies (2.5)). Equations (3.1') – (3.2') are applied for all $s = 2(1)n$ and from $\hat{\omega}_{k,s} = 1 - \hat{\rho}^{\,k}_{k,s} / (k-1)$ and (2.5)

$$(\hat{\omega}_{k,s}\rho_s)^k = k^k(k-1)^{1-k}(1 - \hat{\omega}_{k,s}) \quad , \tag{3.4'}$$

where ρ_s is the unique real root in $(0, k/(k-2))$ of (3.4'). Then

$$\rho_{II} = \frac{k-2}{k} \max_{s=1(1)n} \rho_s \tag{3.5'}$$

is determined. To that specific ρ_{II} there corresponds a unique pair $\hat{\omega}^{(II)}_{k,s}$ and a $\hat{\rho}^{(II)}_{k,s}$, where the index s here is, in general, different from that corresponding to ρ_I previously.

The "best" of the two optimal hypocycloids corresponding to (3.5) and (3.5') is obviously the one for which

$$\hat{\rho}_k = \min\{\hat{\rho}^{(I)}_{k,s},\ \hat{\rho}^{(II)}_{k,s}\} \quad . \tag{3.6}$$

Depending on which of $\hat{\rho}^{(I)}_{k,s}$, $\hat{\rho}^{(II)}_{k,s}$ provides the minimum the optimal $\hat{\omega}_{k,s}$ is determined through $\hat{\rho}_k = ((k-1)(\hat{\omega}_k - 1))^{1/k}$ or $((k-1)(1-\hat{\omega}_{k}))^{1/k}$ respectively. Obviously in case at least one vertex $P_s \notin \operatorname{int} \hat{C}_{k,l}$ one has to work only with the formulas (3.1'), (3.2') and therefore (3.4') and (3.5'). Then $\hat{\rho}_k = \hat{\rho}_{k,s}^{(II)}$ and $\hat{\omega}_k = 1 - \hat{\rho}_k / (k-1)$.

4. EXTENSIONS AND IMPROVEMENTS

The analysis in the previous two sections was based on the optimal cusped hypocycloids. However, in a very recent paper Eiermann, Niethammer and Ruttan [5] gave among others Thm 5 (due to Wild and Niethammer [20]) concerned with cusped hypocycloids and Thm 6 (an excellent result indeed) concerned with shortened ones (see [20]) in connection with the optimal SOR method for k-cyclic matrices. However, if one has in mind how a k-ste method is related to a certain k-cyclic SOR one (see the theory develped in [7]) then one can give the equivalent forms of the aforementioned two Thms in connection with the k-step methods (1.2), $k \geq 3$. Thus we can have the following valid statement as an extension to our Theorem 1.

Theorem 2: Assume that for a given $k(\geq 2)$, $\sigma(T) \subset \bigcup_{\ell=0(2)k-1} [0, \beta\exp(2\pi i\ell / k)] \bigcup_{\ell=0(2)k-1} [0, \alpha\exp((2\ell+1)\pi i/k)]$, where $\alpha \in [0, k((k-2))$ and $\beta \in [0,1)$. Then for the optimal k-step method (1.2), $\hat{\omega}_k$ is the unique real positive root of

$$\left(\frac{\tilde{\beta}+\tilde{\alpha}}{2}\omega\right)^k - \frac{\tilde{\beta}+\tilde{\alpha}}{\tilde{\beta}-\tilde{\alpha}}(\omega - 1) = 0 \tag{4.1}$$

contained in

$$\left[\min\{1, 1+\frac{\tilde{\beta}-\tilde{\alpha}}{\tilde{\beta}+\tilde{\alpha}}\}, \max\left\{1, 1+\frac{\tilde{\beta}-\tilde{\alpha}}{\tilde{\beta}+\tilde{\alpha}}\right\}\right]$$

and there holds

$$\hat{\rho}_k = \left(\frac{\tilde{\beta}+\tilde{\alpha}}{\tilde{\beta}-\tilde{\alpha}}(\hat{\omega}_k - 1)\right)^{1/k} = \frac{\tilde{\beta}+\tilde{\alpha}}{2}\hat{\omega}_k \quad , \tag{4.2}$$

where $\tilde{\alpha}$ and $\tilde{\beta}$ are given by the expressions

$$\begin{array}{llll} i) & \tilde{\alpha} = \dfrac{k-2}{k}\beta, & \tilde{\beta} = \beta & \text{iff} \quad \alpha \leq \beta\dfrac{(k-2)}{k} \quad , \\ ii) & \tilde{\alpha} = \alpha, & \tilde{\beta} = \beta & \text{iff} \quad \alpha \in [\beta\dfrac{(k-2)}{k}, \beta\dfrac{k}{(k-2)}] \quad , \\ iii) & \tilde{\alpha} = \alpha, & \tilde{\beta} = \dfrac{k}{k-2}\alpha & \text{iff} \quad \alpha \geq \beta\dfrac{k}{(k-2)} \quad . \end{array} \tag{4.3}$$

Remark 1: The corresponding optimal curves $\hat{C}_k$ associated with $\hat{\omega}_k$ of Theorem 1 are cusped hypocycloids of type I or II in case $\alpha \le \beta \frac{(k-2)}{k}$ or $\beta \frac{k}{(k-2)} \le \alpha$ respectively and shortened hypocycloids in case $\alpha \in \left(\beta \frac{(k-2)}{k}, \beta \frac{k}{(k-2)}\right)$. □

Remark 2: In case $\alpha \notin (\beta \frac{(k-2)}{k}, \beta \frac{k}{(k-2)})$, Theorem 1 is recovered. This is because for $\alpha \le \beta \frac{(k-2)}{k}$ one obtains $\hat{\omega}_k \in (1, \frac{k}{k-1})$ and

$$\hat{\omega}_k^k \beta^k = k^k (k-1)^{1-k} (\hat{\omega}_k - 1) \quad , \tag{4.4}$$

where

$$\hat{\rho}_k = ((k-1)(\hat{\omega}_k - 1))^{1/k} \quad , \tag{4.5}$$

while for $\alpha \ge \beta \frac{k}{(k-2)}$, $\hat{\omega}_k \in (\frac{k-2}{k-1}, 1)$ and

$$\hat{\omega}_k^k \alpha^k = k^k (k-1)^{1-k} (1 - \hat{\omega}_k) \quad , \tag{4.6}$$

where

$$\hat{\rho}_k = ((k-1)(1 - \hat{\omega}_k))^{1/k} \quad . \qquad \square \tag{4.7}$$

One may explore further the result above. First in connection with our discussion in Section 3 where it is almost "obvious" that through a vertex P of the line L pass infinitely many "optimal" shortened (and one or at most two cusped) hypocycloids and the "best" of them all must be a unique one which "could be determined". If this, as is believed is the case, then through any two vertices there must pass a unique hypocycloid (which may lead to a convergent or a divergent k-step method) and so the development of an algorithm determining the optimal ω should be a complete analogue of the Young-Eidson one [23]. Secondly the previous ideas could be used to extend and improve the results of Section 2. In both these directions an investigation is being made.

References

1. Avdelas, G., J. de Pillis, A. Hadjidimos and M. Neumann, "A Guide to the Acceleration of Iterative Methods Whose Iterative Matrix is Nonnegative and Convergent", *SIAM J. Matrix Anal. Appl.*, 9 (1988), 329-342.

2. Avdelas, G., S. Galanis and A. Hadjidimos, "On the Optimization of a Class of Second Order Iterative Schemes", *BIT*, **23** (1983), 50-64.

3. de Pillis, J. and M. Neumann, "Iterative Methods with k-Part Splittings", *IMA J. Numer. Anal*, **1** (1981), 65-79.

4. Eiermann, M., X. Li and R.S. Varga, "On Hybrid Semi-iterative Methods", *SIAM J. Numer. Anal.* (to appear).

5. Eiermann, M., W. Niethammer and A. Ruttan, "Optimal Successive Overrelaxation Iterative Methods for p-Cyclic Matrices", manuscript. (Paper presented at the *Conference on Approximation Theory and Numerical Linear Algebra*, Kent, Ohio, March-April, 1989).

6. Eiermann, M., W. Niethammer and R.S. Varga, "A Study of the Semi-iterative Methods for Nonsymmmetric Systems of Linear Equations", *Numer. Math*, **47** (1985), 503--533.

7. Galanis, G., A. Hadjidimos and D. Noutsos, "On Different Classes of Monoparametric Stationary Iterative Methods for the Solution of Linear Systems", *Math. Comput. Simulation*, **28** (1986), 115-128.

8. Galanis, S., A. Hadjidimos and D. Noutsos, "On the Convergence of Monoparametric k-Step Iterative Euler Methods for the Solution of Linear Systems", *Intern. J. Computer Math.*, **26** (1988), 45-56.

9. Galanis, S., A. Hadjidimos and D. Noutsos, "On the Equivalence of the k-Step Iterative Euler Methods and Successive Overrelaxation (SOR) Methods for k-Cyclic Matrices", *Math. Comput. Simulation*, **30** (1988), 213-230.

10. Golub, G.H. and R.S. Varga, "Chebyshev Semi-Iterative Methods, Successive Over-relaxation Iterative Methods, and Second-Order Richardson Iterative Methods", *Numer. Math. Parts I and II*, **3** (1961), 147-168.

11. Hadjidimos, A., "The Optimal Solution of the Extrapolation Problem of a First Order Scheme", *Intern. J. Computer Math.*, **13** (1983), 153-168.

12. Hadjidimos, A., "The Optimal Solution to the Problem of Complex Extrapolation of a First Order Iterative Scheme", *Linear Algebra Appl.*, **63** (1984), 241-261.

13. Hughes-Hallett, A.J., "Some Extensions and Comparisons in the Theory of Gauss-Seidel Iterative Techniques for Solving Large Equation Systems", *Proceedings of the 1979 Econometric Society Meeting* (E.G. Charatsis, ed.), 279-318, North Holland, Amsterdam, 1981.

14. Isaacson, E. and H.B. Keller, *Analysis of Numerical Methods*, John Wiley, New York, 1966.

15. Niethammer, W., "Iterationsverfahren und Allgemeine Euler-Verfahren", *Math. Zeit.*, **102** (1967), 288-317.

16. Niethammer, W. and R.S. Varga, "The Analysis of k-Step Iterative Methods for Linear Systems from Summability Theory", *Numer. Math.*, **41** (1983), 177-206.

17. Opfer, G. and G. Schober, "Richardson's Iteration for Nonsymmetric Matrices", *Linear Algebra Appl.*, **58** (1984), 343-361.

18. Pierce, D., A. Hadjidimos and R.J. Plemmons, "Optimality Relationships for p-Cyclic SOR", manuscript. (Paper presented at the *Conference on Approximation Theory and Numerical Linear Algebra*, Kent, Ohio, March-April, 1989).

19. Varga, R.S., *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, New Jersey, 1962.

20. Wild, P. and W. Niethammer, "Over-and Underrelaxation for Linear Systems with Weakly Cyclic Jacobi Matrices of Index p", *Linear Algebra Appl.*, **91** (1987), 29-52.

21. Young, D.M., "On Richardson's Method for Solving Linear Systems with Positive Definite Matrices", *J. Math. Phys.*, **XXXII** (1954), 243-255.

22. Young, D.M., *Iterative Solution of Large Linear Systems*, Academic Press, New York, 1971.

23. Young, D.M. and H.E. Eidson, "On the Determination of the Optimum Relaxation Factor for the SOR Method When the Eigenvalues of the Jacobi Matrix are Complex", Report CNA-1, Center for Numerical Analysis, University of Texas, Austin, Texas, 1970.