ANALYSIS OF THIRD-ORDER METHODS
FOR SECULAR EQUATIONS

A. MELMAN

ABSTRACT. Third-order numerical methods are analyzed for secular equations. These equations arise in several matrix problems and numerical linear algebra applications. A closer look at an existing method shows that it can be considered as a classical method for an equivalent problem. This not only leads to other third-order methods, it also provides the means for a unifying convergence analysis of these methods and for their comparisons. Finally, we consider approximated versions of the aforementioned methods.

1. INTRODUCTION

Consider a real symmetric matrix $A \in \mathbb{R}^{n \times n}$ with known eigenvalues, to which a symmetric rank-one perturbation is added. The eigenvalues of the new matrix $A + \sigma z z^T$ ($z \in \mathbb{R}^n$) are then given by the solutions of a so called *secular equation* (see [11]). This and similar problems appear in several applications in numerical linear algebra, such as the singular value decomposition of a matrix and eigenvalue problems, to name but a few. The following list is a small sample of the many related references : [1, 4, 5, 6, 8, 9, 10, 11, 12, 13, 14]. The secular function $g(s)$ we will consider is a real function of the form

$$(1) \qquad g(s) = \mu + \nu s + \sum_{j=1}^{m} \frac{\zeta_j^2}{d_j - s} \ ,$$

where all parameters are real,

$$d_1 < d_2 < \cdots < d_m \ ,$$

$\zeta_j^2 \neq 0$ for all j and $\nu \geq 0$. This function has m simple poles and is monotonic increasing from $-\infty$ to $+\infty$ on the interval between two consecutive poles, which means that there is a root on each of those $m-1$ intervals. The problem then consists of solving $g(s) = 0$ for each of its roots. This function belongs to a class of functions called *Pick functions* (see [7]). Depending on the values of μ and ν, there may be one or two extra roots in addition to the $m-1$ roots on the intervals. For $\mu > 0$ and $\nu = 0$, there is one extra root to the right of d_m. For $\nu > 0$, there are two extra roots : one to the left of d_1 and one to the right of d_m.

Let us now consider the $m-1$ roots on the intervals (d_j, d_{j+1}) $(j = 1, ..., m-1)$. To compute the i-th root $(1 \leq i \leq m-1)$, i is fixed and, as in [3], the transformation

---

of variables $s = d_i + t$ is carried out, which, with $\delta_j = d_j - d_i$, yields the following root-finding problem on $(0, \delta_{i+1})$:

$$(2) \qquad\qquad f(t) = \alpha + \beta t + \sum_{j=1}^{m} \frac{\zeta_j^2}{\delta_j - t} \ ,$$

where

$$\delta_1 < \cdots < \delta_i = 0 < \delta_{i+1} < \cdots < \delta_m \ ,$$

$\alpha = \mu + \nu d_i$ and $\beta = \nu$. This transformation is not essential to our results. The basic problem then consists of solving $f(t) = 0$ on the interval $(0, \delta_{i+1})$. On this interval f is monotonic increasing and it has two singularities, one at $t = 0$ and one at $t = \delta_{i+1}$. It has exactly one real root on this interval. The two possible extra roots will not be treated separately as their computation is easier (the function is strictly convex or concave with only one singularity).

In the aforementioned applications, secular equations typically have to be solved to high accuracy many times as a subproblem of a larger one, which requires fast and stable methods. An important instance is the the Divide and Conquer method ([5]).

There exist several second-order methods, i.e., methods based on approximations to f that agree up to its first derivative at a given point. These approximations are usually rational functions, which attempt to take into account the singularities of f (see [3, 16, 17, 18]). An approximation based on a polynomial (such as Newton's method) is inappropriate because of the singular nature of f. Of course, it is always possible to use standard root-finding methods with proper safeguards. However, safeguards together with the failure to take into account the singularity of the equation lead to wasteful computations.

Third-order methods are less common. In fact only one method, Gragg's zero finder ([1, 13]), seems to be practically used. We start off with this method and show that hiding behind it, is the classical Euler[1] approximation (second-order Taylor polynomial) for a different function, but one having the same root. This interpretation leads to another third-order method, not less classical than Euler's, namely Halley's method, along with a modification of it. We prove their global convergence and present theoretical comparisons between them. Key to our analysis are a simplifying transformation of the problem and the geometry of the approximations that characterize the methods. Finally, we consider the aforementioned methods with finite difference approximations to the second derivative.

Part of the motivation for studying these methods is provided by the extensive numerical results in [16] which show the potential of Gragg's zero finder. Our intention is to present a theoretical convergence analysis of third-order methods and we do not treat implementation issues, such as numerical comparisons, stopping rules and the influence of inexact arithmetic.

We have used the term "approximation" where others might prefer "interpolation". An expression such as "approximation at $\bar{x}$ up to first derivative" means that the approximant (or interpolant) agrees with the function it approximates (interpolates) and with its first derivative, at a certain point $\bar{x}$.

---

[1]We have used Euler's name, which seems to be common practice. However, there is some doubt as to whom the method should be ascribed to. Cauchy emerges as a very likely candidate.

In Section 2 we present Gragg's zero finder and show its equivalence to Euler's method. In Section 3 a transformation of variables is introduced, Halley's method is presented and some lemmas are given. The convergence of our methods, as well as comparisons between them, is dealt with in Section 4. In Section 5, we present a modification of Halley's method and Section 6 concerns finite difference approximations to the second derivative.

2. Gragg's zero finder

Let us have another look at Gragg's zero finder ([1, 13]). We recall that we are solving the equation $f(t) = 0$ on the interval $(0, \delta)$, where $f(t)$ is given by (2) and where we have defined $\delta = \delta_{i+1}$. The method proposed in [1, 13] is an iterative method for the solution of this equation, based on a third-order approximation of the form

$$(3) \qquad a + \frac{b}{t} + \frac{c}{\delta - t} \; ,$$

with a, b and c chosen such that it agrees with f, f' and f'' at successive iterates, which are the roots of the successive approximations. This is a rational approximation which takes into consideration the two singularities of the function that are closest to the root and it can therefore be expected to be more appropriate than one based, e.g., on a Taylor expansion. In other words, we have

$$a + \frac{b}{t} + \frac{c}{\delta - t} \; \sim \; f(t) \; .$$

This is the same as saying that

$$-at^2 + (a\delta - b + c)t + b\delta \; \sim \; t(\delta - t)f(t) \; ,$$

where the approximating function is now quadratic. Since its coefficients are independent, this means that Gragg's zero finder can be interpreted as a method, based on the classical Euler approximation (Taylor expansion up to second-order terms) for a different function, but with the same root on the interval of interest.

Unfortunately, the function $t(\delta - t)f(t)$ is neither convex nor concave, which renders it rather useless for our purposes. The reason for this lies in the geometric nature of the analysis. We will therefore carry out a transformation of variables, which leads to another, equivalent, problem with a strictly concave function. This is the subject of the next section.

3. Transformation of variables and Halley's method

Consider the transformation of variables $t = 1/\gamma$ and denote $F(\gamma) \equiv f(1/\gamma)$. The problem is now to solve $F(\gamma) = 0$ on $(1/\delta, +\infty)$. The function F is monotonic decreasing from $+\infty$ to $-\infty$ and convex (see [17]). It therefore has a single root on the aforementioned interval, which we denote by γ^*. Gragg's approximation in the new variable γ can be written as :

$$p + q\gamma + \frac{r}{\gamma - \frac{1}{\delta}} \; \sim \; F(\gamma) \; ,$$

which is equivalent to

$$q\gamma^2 + \left(p - \frac{q}{\delta}\right)\gamma + r - \frac{p}{\delta} \; \sim \; \left(\gamma - \frac{1}{\delta}\right)F(\gamma) \; .$$

We observe that the approximating function is quadratic with independent coefficients, which means that we have once again obtained Euler's approximation, in this case for the function $(\gamma - 1/\delta)F(\gamma)$. As we shall see in Lemma 3.1, this function is strictly concave. In what follows, we denote

$$\phi(\gamma) = (\gamma - 1/\delta)F(\gamma) .$$

Its only root on the interval $[1/\delta, +\infty)$ is the root γ^* of $F(\gamma)$.

Using different approximations leads to different methods. One such method, not less classical than Euler's, is Halley's method, which is based on a rational approximation. This method dates back more than 300 years. In the words of J.F. Traub ([23]), Halley's iteration function (I.F.) *"must share with the secant I.F. the distinction of being the most frequently rediscovered I.F. in the literature"*. For a historical background to these methods, the reader is referred to [22] and references therein. In this case, the approximation to a function $G(\gamma)$ at some point $\bar{\gamma}$, if $G''(\bar{\gamma}) \neq 0$, takes the form

$$a + \frac{b}{\gamma + c} ,$$

where a, b and c (not the same as before) are again determined by the third-order approximation conditions. When $G''(\bar{\gamma}) = 0$, the approximation degenerates into a line. In fact, the standard way the approximation is presented (see [20, 21, 23]) is

$$\frac{\gamma + a}{b\gamma + c} \sim G(\gamma) ,$$

i.e., as a linear fractional (or Möbius) transformation. Here we prefer to use the previous form, which is equivalent to the standard form for $G''(\bar{\gamma}) \neq 0$, as it is more convenient for our purposes.

In what follows, we will examine Euler's and Halley's methods for the problem $\phi(\gamma) = 0$.

We will need the following lemmas.

Lemma 3.1. *On the interval $[1/\delta, +\infty)$, the higher order derivatives ($n \geq 2$) of ϕ satisfy*

$$\operatorname{sgn}\left(\phi^{(n)}\right) = (-1)^{n+1} .$$

Proof. After some algebra we obtain

$$F(\gamma) = \alpha + \sum_{\substack{j=1 \\ j\neq i}}^{m} \frac{\zeta_j^2}{\delta_j} - \zeta_i^2 \gamma + \frac{\beta}{\gamma} + \sum_{\substack{j=1 \\ j\neq i}}^{m} \frac{\left(\frac{\zeta_j}{\delta_j}\right)^2}{\gamma - \frac{1}{\delta_j}} ,$$

$$F'(\gamma) = -\zeta_i^2 - \frac{\beta}{\gamma^2} - \sum_{\substack{j=1 \\ j\neq i}}^{m} \frac{\left(\frac{\zeta_j}{\delta_j}\right)^2}{\left(\gamma - \frac{1}{\delta_j}\right)^2}$$

and for $n \geq 2$

$$F^{(n)}(\gamma) = \frac{(-1)^n n! \beta}{\gamma^{n+1}} + (-1)^n n! \sum_{\substack{j=1 \\ j\neq i}}^{m} \frac{\left(\frac{\zeta_j}{\delta_j}\right)^2}{\left(\gamma - \frac{1}{\delta_j}\right)^{n+1}} .$$

From Leibniz's rule, we have

$$\phi^{(n)}(\gamma) = \left(\gamma - \frac{1}{\delta}\right) F^{(n)}(\gamma) + n F^{(n-1)}(\gamma) \, .$$

For $n \geq 2$ we therefore obtain

$$\phi^{(n)}(\gamma) = \frac{(-1)^n n! \beta}{\gamma^n} \left(\frac{\gamma - \frac{1}{\delta}}{\gamma} - 1\right) + (-1)^n n! \sum_{\substack{j=1 \\ j \neq i}}^{m} \frac{\left(\frac{\zeta_j}{\delta_j}\right)^2}{\left(\gamma - \frac{1}{\delta_j}\right)^n} \left(\frac{\gamma - \frac{1}{\delta}}{\gamma - \frac{1}{\delta_j}} - 1\right).$$

Recalling that $\gamma > 1/\delta_j$ for all j, one has

(4) $$\frac{\gamma - \frac{1}{\delta}}{\gamma - \frac{1}{\delta_j}} - 1 = \frac{\frac{1}{\delta_j} - \frac{1}{\delta}}{\gamma - \frac{1}{\delta_j}} < 0 \qquad \text{and} \qquad \frac{\gamma - \frac{1}{\delta}}{\gamma} - 1 = -\frac{1}{\delta\gamma} < 0 \, .$$

Therefore, the sign of $\phi^{(n)}$ is given by $-(-1)^n$. $\qquad\square$

Lemma 3.2. *The function*

$$h(t) = \sum_{j=1}^{n} \alpha_j (t + \beta_j)^p \, ,$$

with p a nonzero integer and the α_j's nonnegative, satisfies

$$\frac{1-p}{p} h'^2(t) + h(t) h''(t) \geq 0 \, ,$$

for all t such that $t + \min_j\{\beta_j\} > 0$.

Proof. Let us first compute h' and h'' :

$$h'(t) = p \sum_{j=1}^{n} \alpha_j (t + \beta_j)^{p-1} \quad ; \quad h''(t) = p(p-1) \sum_{j=1}^{n} \alpha_j (t + \beta_j)^{p-2} \, .$$

We then have

$$\frac{p-1}{p} h'^2(t) = p(p-1) \left(\sum_{j=1}^{n} \alpha_j (t + \beta_j)^{p-1}\right)^2$$

$$= p(p-1) \left(\sum_{j=1}^{n} \sqrt{\alpha_j}(t + \beta_j)^{\frac{p}{2}} \sqrt{\alpha_j}(t + \beta_j)^{\frac{p}{2}-1}\right)^2 \, .$$

We note that $\frac{1}{p}(p-1)$ and $p(p-1)$ are nonnegative when p is a nonzero integer. Applying the Cauchy-Schwarz inequality yields

$$\frac{p-1}{p} h'^2(t) \leq p(p-1) \sum_{j=1}^{n} \left(\sqrt{\alpha_j}(t + \beta_j)^{\frac{p}{2}}\right)^2 \sum_{j=1}^{n} \left(\sqrt{\alpha_j}(t + \beta_j)^{\frac{p}{2}-1}\right)^2 = h(t) h''(t) \, .$$
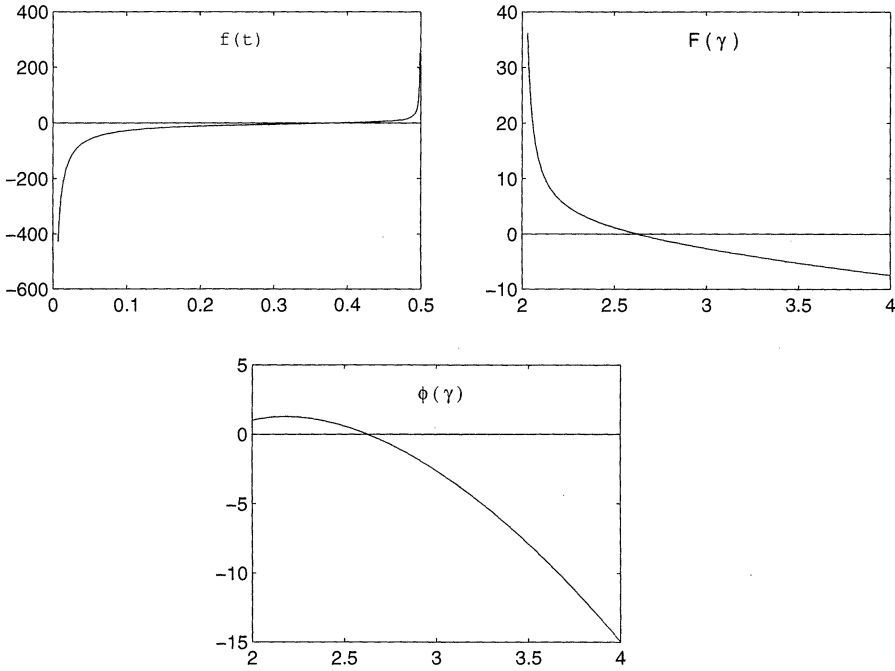
This completes the proof. $\qquad\square$

FIGURE 1. The functions $f(t)$, $F(\gamma)$ and $\phi(\gamma)$.

As an example, Figure 1 shows the functions $f(t)$, $F(\gamma)$ and $\phi(\gamma)$, with

$$f(t) = -8.5 + t + \frac{1}{-1-t} + \frac{3}{-t} + \frac{0.25}{0.5-t} + \frac{6}{1-t} + \frac{8}{2-t},$$

$$F(\gamma) = f\left(\frac{1}{\gamma}\right) = 1 - 3\gamma + \frac{1}{\gamma} + \frac{1}{\gamma+1} + \frac{6}{\gamma-1} + \frac{1}{\gamma-2} + \frac{2}{\gamma-0.5},$$

$$\phi(\gamma) = (\gamma-2)F(\gamma) = -3\gamma^2 + 7\gamma + 9 - \frac{2}{\gamma} - \frac{3}{\gamma+1} - \frac{6}{\gamma-1} - \frac{3}{\gamma-0.5}.$$

4. CONVERGENCE

We now turn to the convergence analysis of the two methods previously mentioned. The following theorem compares Newton's, Euler's and Halley's approximations to $\phi(\gamma)$. We denote these approximations, respectively, by $N(\gamma)$, $E(\gamma)$ and $H(\gamma)$.

Theorem 4.1. *At a point $\bar{\gamma}$ for which $\phi'(\bar{\gamma}) < 0$, the approximations $N(\gamma)$, $E(\gamma)$ and $H(\gamma)$ satisfy the following inequalities for $1/\delta \leq \gamma \leq \bar{\gamma}$:*

$$(5) \qquad\qquad \phi(\gamma) \leq E(\gamma) \leq H(\gamma) \leq N(\gamma) \, ;$$

For $\gamma \geq \bar{\gamma}$ they satisfy

$$(6) \qquad\qquad N(\gamma) \geq \phi(\gamma) \geq E(\gamma) \geq H(\gamma) \, .$$

Proof. For Eulers's approximation at the point $\bar{\gamma}$, we have

$$p\gamma^2 + q\gamma + r \sim \phi(\gamma),$$

which yields the following equations for the parameters p, q and r :

(7) $$p\bar{\gamma}^2 + q\bar{\gamma} + r = \phi(\bar{\gamma}),$$
(8) $$2p\bar{\gamma} + q = \phi'(\bar{\gamma}),$$
(9) $$2p = \phi''(\bar{\gamma}).$$

In other words, ϕ' is approximated up to first derivative at the point $\bar{\gamma}$ by the line $2p\gamma + q$. From the previous lemma, we know that ϕ''' is positive on $[1/\delta, +\infty)$. Therefore, on this interval, ϕ' is convex, which implies that $2p\gamma + q \leq \phi'(\gamma)$. This means that for $\gamma \geq \bar{\gamma}$

$$\int_{\bar{\gamma}}^{\gamma} (2pt + q)\, dt \leq \int_{\bar{\gamma}}^{\gamma} \phi'(t)\, dt ,$$

which yields

$$p\gamma^2 + q\gamma - p\bar{\gamma}^2 - q\bar{\gamma} \leq \phi(\gamma) - \phi(\bar{\gamma}) .$$

Adding and subtracting r in the LHS and using (7) gives

(10) $$p\gamma^2 + q\gamma + r \leq \phi(\gamma)$$

for $\gamma \geq \bar{\gamma}$. The inverse inequalities are obtained for $1/\delta \leq \gamma \leq \bar{\gamma}$.

On the other hand, since $\phi'' \neq 0$ on $[1/\delta, +\infty)$, Halley's approximation at a point $\bar{\gamma}$ takes the form

$$a + \frac{b}{\gamma + c} \sim \phi(\gamma) .$$

This yields the following equations for the parameters a, b and c :

(11) $$a + \frac{b}{\bar{\gamma} + c} = \phi(\bar{\gamma}),$$
(12) $$\frac{-b}{(\bar{\gamma} + c)^2} = \phi'(\bar{\gamma}),$$
(13) $$\frac{2b}{(\bar{\gamma} + c)^3} = \phi''(\bar{\gamma}) .$$

In other words, $\phi'(\gamma)$ is approximated up to first derivative at the point $\bar{\gamma}$ by the rational function $-b(\gamma + c)^{-2}$. Since $\phi'(\bar{\gamma}) < 0$, b must be positive. This means that $(-\phi'(\gamma))^{-1/2}$ is approximated up to first derivative by the line $-(1/\sqrt{b})(\gamma + c)$. Let us now have a look at $((-\phi')^{-1/2})''$:

(14) $$((-\phi')^{-1/2})'' = \left(\frac{1}{2}(-\phi')^{-\frac{3}{2}}\phi''\right)' = \frac{3}{4}(-\phi')^{-\frac{5}{2}}\phi''^2 + \frac{1}{2}(-\phi')^{-\frac{3}{2}}\phi''' .$$

Since ϕ''' is positive, we have obtained that $(-\phi')^{-1/2}$ is a convex function. Therefore

(15) $$0 < -(1/\sqrt{b})(\gamma + c) \leq (-\phi'(\gamma))^{-1/2} .$$

We are only interested in points where the line is strictly positive, as it has to be positive at $\bar{\gamma}$ and its zero is a singularity for the original approximation to ϕ'.

Once this inequality established, we immediately obtain that $-b(\gamma + c)^{-2} \le \phi'(\gamma)$. Proceeding as before, we now have for $\gamma \ge \bar{\gamma}$:

$$\int_{\bar{\gamma}}^{\gamma} \frac{-b}{(t+c)^2}\, dt \le \int_{\bar{\gamma}}^{\gamma} \phi'(t)\, dt \ ,$$

which yields

$$\frac{b}{\gamma + c} - \frac{b}{\bar{\gamma} + c} \le \phi(\gamma) - \phi(\bar{\gamma}) \ .$$

Adding and subtracting a in the LHS and using the first approximation condition gives

$$a + \frac{b}{\gamma + c} \le \phi(\gamma) \ .$$

The opposite inequality is obtained for $1/\delta \le \gamma \le \bar{\gamma}$. We now compare the different approximations : $N(\gamma)$, $E(\gamma)$ and $H(\gamma)$. Since $E(\gamma)$ and $H(\gamma)$ both approximate $\phi(\gamma)$ up to first order at $\bar{\gamma}$, we have

$$p\bar{\gamma}^2 + q\bar{\gamma} + r = a + \frac{b}{\bar{\gamma} + c} = \phi(\bar{\gamma}),$$

$$2p\bar{\gamma} + q = \frac{-b}{(\bar{\gamma} + c)^2} = \phi'(\bar{\gamma}),$$

$$2p = \frac{2b}{(\bar{\gamma} + c)^3} = \phi''(\bar{\gamma}) \ .$$

The meaning of this is that $-b(\gamma + c)^{-2}$ is approximated up to first derivative by the line $2p\gamma + q$. However, since $b > 0$, $-b(\gamma + c)^{-2}$ is a concave function and therefore $2p\gamma + q \ge -b(\gamma + c)^{-2}$. For $\gamma \ge \bar{\gamma}$ we obtain

$$\int_{\bar{\gamma}}^{\gamma} (2pt + q)\, dt \ge \int_{\bar{\gamma}}^{\gamma} \frac{-b}{(t+c)^2}\, dt \ ,$$

which yields

$$p\gamma^2 + q\gamma - p\bar{\gamma}^2 - q\bar{\gamma} \ge \frac{b}{\gamma + c} - \frac{b}{\bar{\gamma} + c} \ .$$

Adding and subtracting r in the LHS and a in the RHS and using (11) yields $E(\gamma) \ge H(\gamma)$. The opposite inequality holds for $1/\delta \le \gamma \le \bar{\gamma}$. Since both $E(\gamma)$ and $H(\gamma)$ are concave and since Newton's approximation is simply the tangent line to $\phi(\gamma)$ (and therefore to $E(\gamma)$ and $H(\gamma)$ as well), we obtain that $N(\gamma) \ge E(\gamma)$ and $N(\gamma) \ge H(\gamma)$. Because of the concavity of ϕ, $N(\gamma) \ge \phi(\gamma)$ on the entire interval. □

Here, Newton's method, when starting from a point $\bar{\gamma}$ such that $\phi(\bar{\gamma}) < 0$, yields a better approximation to the root the closer the starting point lies to that root. The reason for this is the concavity of ϕ.

Because of the inequalities in the previous theorem, we know that those methods converge at least as fast as Newton's method for the same problem with a starting point $\bar{\gamma}$ such that $\phi(\bar{\gamma}) < 0$. The convergence of Newton's method is immediate in this case (see, e.g., [15, 19, 23]). This means that Euler's and Halley's methods
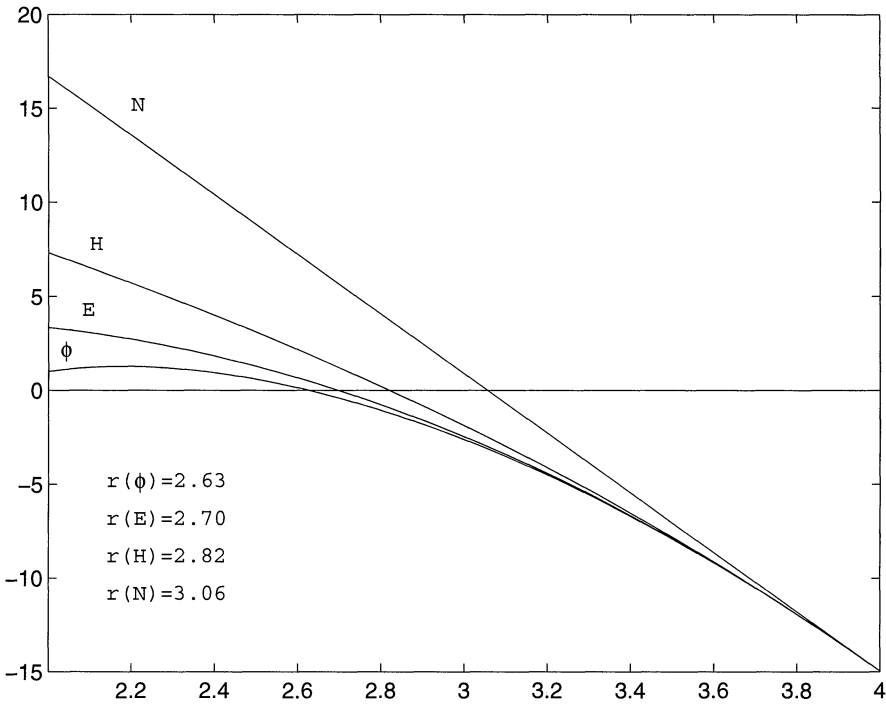
FIGURE 2. Comparison between Euler's, Halley's and Newton's approximations.

also converge and their cubic order of convergence folllows from, e.g., [23]. We also note that the inequalities imply that, for our problem, the step in Euler's method is larger than for Halley's. All this is summarized in the following theorem.

Theorem 4.2. *Euler's and Halley's methods converge from any point $\bar{\gamma}$ for which $\phi(\bar{\gamma}) < 0$ with cubic order of convergence. At each iteration, the Euler step is larger than Halley's.* □

Euler's method also converges from points where $\phi' \geq 0$, as in this case the proof of the inequalities does not depend on the sign of ϕ'. Newton's or Halley's method do not converge from such a point. Figure 2 compares Euler's, Halley's and Newton's approximations for the example in Section 3 at the point $\bar{\gamma} = 4$. The roots of the different approximations are denoted by "$r(\cdot)$".

The convergence of Euler's method for this particular case, or Gragg's zero finder, was also proven in [1]. Our proof was presented from a more general viewpoint, with the additional advantage of avoiding the relatively messy computation of the coefficients in the approximation.

A shorter way for proving the convergence of Halley's method in this case is provided by its equivalence to Newton's method (see [2]), applied to an equivalent problem which in our case is given by $(-\phi)/(-\phi')^{1/2} = 0$. Convergence will be monotonic from the right for any starting point to the right of the root of ϕ. The reason for this is the fact that $(-\phi)/(-\phi')^{1/2}$ is an increasing convex function to

the right of the root of ϕ. This easily follows from (14) and from

$$
\begin{aligned}
((-\phi)(-\phi')^{-1/2})'' &= \Big((-\phi)((-\phi')^{-1/2})' + (-\phi')((-\phi')^{-1/2})\Big)' \\
&= \Big((-\phi)((-\phi')^{-1/2})' + ((-\phi')^{1/2})\Big)' \\
&= (-\phi')((-\phi')^{-1/2})' + (-\phi)((-\phi')^{-1/2})'' - \frac{1}{2}(-\phi')^{-1/2}\phi'' \\
&= (-\phi)((-\phi')^{-1/2})'' \, .
\end{aligned}
$$

The convergence is then immediate from a classical result for Newton's method (see, e.g., [15, 19, 23]).

However, unlike the aforementioned equivalence approach, our techniques allow us to consider modifications in which only part of a function is approximated, which is precisely what we are about to do.

5. MODIFICATION OF HALLEY'METHOD

We recall that $\phi(\gamma) = (\gamma - 1/\delta)F(\gamma)$. From Lemma 3.1 we have

$$
F(\gamma) = \alpha + \sum_{\substack{j=1 \\ j \neq i}}^{m} \frac{\zeta_j^2}{\delta_j} - \zeta_i^2 \gamma + \frac{\beta}{\gamma} + \sum_{\substack{j=1 \\ j \neq i}}^{m} \frac{\left(\frac{\zeta_j}{\delta_j}\right)^2}{\gamma - \frac{1}{\delta_j}} \, .
$$

Multiplying this expression by $\gamma - 1/\delta$ and using (4), one obtains $\phi(\gamma) = Q(\gamma) + \psi(\gamma)$, where $Q(\gamma)$ is a quadratic function and $\psi(\gamma)$ is of the form

$$
\text{(16)} \qquad\qquad \psi(\gamma) = -\sum_i \frac{q_i^2}{\gamma - r_i} \, ,
$$

where for all i : $q_i, r_i \in \mathbb{R}$ and $r_i < 1/\delta$. Therefore, $\psi(\gamma) < 0$ on $[1/\delta, +\infty)$. Instead of approximating ϕ by either Eulers's or Halley's approximation, we could simply leave $Q(\gamma)$ as it is and approximate only ψ. Of course, this makes a difference only for Halley's method, as Euler's approximation is already quadratic. As we will see in the following theorem, this simple modification rehabilitates Halley's method by making it convergent on the whole interval $[1/\delta, +\infty)$ and, in addition, superior to Euler's method.

We denote by H_ψ Halley's approximation to ψ when there can be no misunderstanding about the point at which the approximation is carried out.

Theorem 5.1. *The method obtained by approximating $\phi(\gamma)$ by $Q(\gamma) + H_\psi(\gamma)$ is globally convergent on the interval $[1/\delta, +\infty)$ and, at each iteration, its step is at least as large as the step in Euler's method.*

Proof. Again we use the geometric interpretation of Halley's method, similarly to the proof of Theorem 4.1. The same conditions (11)–(13) hold with ϕ replaced by ψ. On $[1/\delta, +\infty)$, ψ satisfies : $\psi' > 0$, $\psi'' < 0$ and $\psi''' > 0$. In addition, one obtains from Lemma 3.2, that

$$
\left(\psi'^{-\frac{1}{2}}\right)'' = \frac{1}{2}\psi'^{-\frac{5}{2}}\left(\frac{3}{2}\psi''^2 - \psi'\psi'''\right) \leq 0 \, .
$$

A similar computation also appears in [24]. This means that $\psi'^{-1/2}$ is a concave
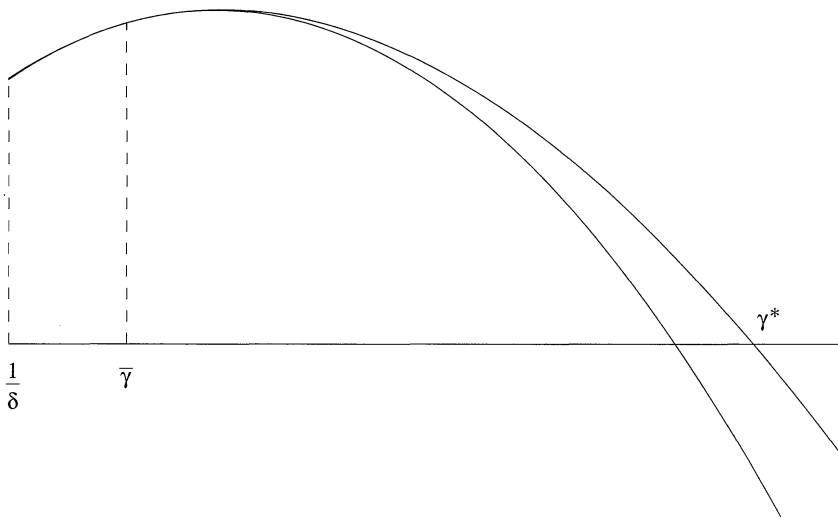
FIGURE 3. One iteration of the modified Halley method for $\bar{\gamma} < \gamma^*$.
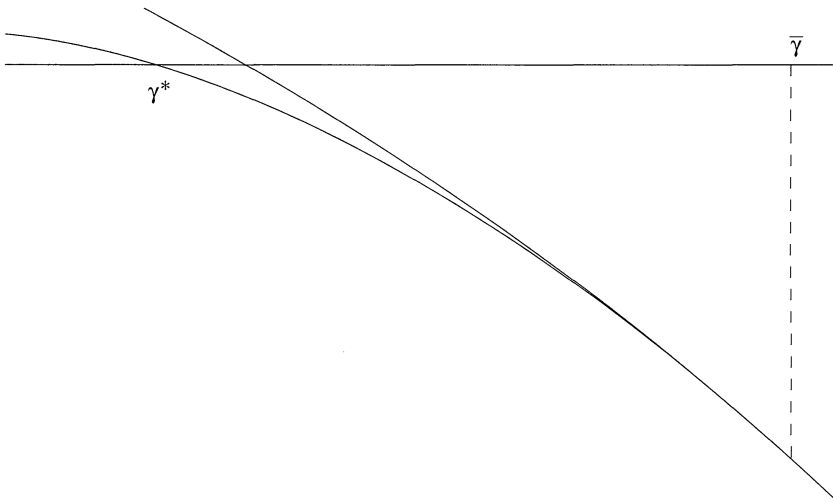

FIGURE 4. One iteration of the modified Halley method for $\bar{\gamma} > \gamma^*$.

function. Proceeding analogously as in the proof of Theorem 4.1, we find that the same inequalities that were derived for ϕ also hold for ψ. The difference is that whereas those inequalities hold for ϕ only at points where $\phi' < 0$, for ψ they hold at any point in the interval $[1/\delta, +\infty)$. For $\gamma \geq \bar{\gamma}$ this yields : $\phi(\gamma) \leq Q(\gamma) + H_\psi(\gamma)$, whereas for $1/\delta \leq \gamma \leq \bar{\gamma}$, the opposite inequality is obtained. This means that for $\bar{\gamma} > \gamma^*$ (where $\phi < 0$) the approximation is concave and lies above ϕ for points between $\bar{\gamma}$ and the root, while for $\bar{\gamma} < \gamma^*$ (where $\phi > 0$), it is also concave but lies below it for points between $\bar{\gamma}$ and the root. This is illustrated in Figures 3 and 4, where one iteration of the method is shown for each case.

Therefore, convergence will be monotonic from either side of the root. We obtain a monotonic increasing or decreasing bounded sequence, depending on the side from which the process converges. Those sequences must therefore converge to an accumulation point ζ, at which ϕ must be zero. To prove this, suppose that $\phi(\zeta) \neq 0$ and that convergence is from the left. Then, since $\phi'' \neq 0$ and $|\phi'| < +\infty$, the approximation can never be a vertical line and therefore one can always find $\eta_0 < \zeta$ with its next iterate η_1 such that $\eta_1 > \zeta$, contradicting the fact that ζ is an accumulation point for a monotonically converging sequence.

It remains to be shown that this modification is better than Euler's approximation. Since $p\gamma^2 + q\gamma + r$ approximates ϕ, it also approximates $Q(\gamma) + H_\psi(\gamma)$ up to second derivatives. Therefore, $p\gamma^2 + q\gamma + r - Q(\gamma)$ must approximate $H_\psi(\gamma)$. For any $\bar{\gamma} \in [1/\delta, +\infty)$, we obtain

$$p\bar{\gamma}^2 + q\bar{\gamma} + r - Q(\bar{\gamma}) = H_\phi(\bar{\gamma}) = \psi(\bar{\gamma}),$$
$$2p\bar{\gamma} + q - Q'(\bar{\gamma}) = H_\phi'(\bar{\gamma}) = \psi'(\bar{\gamma}),$$
$$2p - Q''(\bar{\gamma}) = H_\phi''(\bar{\gamma}) = \psi''(\bar{\gamma})$$

However, from the conditions on ψ, we have that H_ψ' is convex, implying $2p\gamma + q - Q'(\gamma) \leq H_\phi'(\gamma)$. Arguments analogous to those in the proof of Theorem 4.1 then lead to $E(\gamma) \leq Q(\gamma) + H_\psi$ for $\gamma \geq \bar{\gamma}$ and the inverse inequality for $1/\delta \leq \gamma \leq \bar{\gamma}$. Therefore Euler's method yields a smaller step in both cases.   □

In order to use a particular approximation for a function in an iterative method, it should be easy to compute its root, as this will be the next iterate. For Euler's and Halley's approximations, this means solving, respectively, a quadratic and a linear equation. For the modification of Halley's method however, we need to solve a cubic equation. Even though this is still possible analytically, it is preferable to use Newton's method. The approximation is concave, and Newton's method will converge from any point to the right of its root. In a typical application, the number of operations needed to carry this out is negligable compared to the number of operations involved in the computation of one function value of the original function f.

An iterative method also needs a starting point. We first note that, as a starting point, $\gamma_0 = 1/\delta$ is a perfectly good choice for Euler's method and the modified Halley method. We can obtain other starting points by considering (16), which leads to the following inequalities :

$$Q(\gamma) + \psi\left(\frac{1}{\delta}\right) \ \leq \ Q(\gamma) + \psi(\gamma) \ \leq \ Q(\gamma) \, .$$

Possible starting points are then provided by the roots of the quadratics that bound ϕ.

6. Second derivative approximations

To reduce the number of arithmetic operations per iteration, one can replace the second derivative by a suitable finite difference approximation. However, this approximation should be carried out so as to maintain convergence. The appropriate finite difference approximation turns out to be different for different methods. Of course, cubic convergence is lost, but the resulting algorithm will be at least of second order. The methods thus obtained could be called quasi-Euler and quasi-Halley.

We recall that, generally speaking, the convergence of our methods derives from inequalities such as (5), (6) and similar ones in the proof of Theorem 5.1. These inequalities, in turn, are based on the convexity of certain functions and these functions are the key to the correct finite difference approximation. This forms the subject of the following theorem.

Theorem 6.1. *The following finite difference approximations of the second deriv-ative, based on the derivatives at γ_1 and γ_2, maintain monotonic convergence for the resulting methods. For Euler's method :*

$$\phi''(\gamma_2) \approx \frac{\phi'(\gamma_2) - \phi'(\gamma_1)}{\gamma_2 - \gamma_1} \ . \tag{17}$$

For Halley's method :

$$\phi''(\gamma_2) \approx 2\left((-\phi'(\gamma_2))^{\frac{3}{2}}\right) \frac{(-\phi'(\gamma_2))^{-\frac{1}{2}} - (-\phi'(\gamma_1))^{-\frac{1}{2}}}{\gamma_2 - \gamma_1} \ . \tag{18}$$

For the modified Halley method :

$$\psi''(\gamma_2) \approx -2\left((\psi'(\gamma_2))^{\frac{3}{2}}\right) \frac{(\psi'(\gamma_2))^{-\frac{1}{2}} - (\psi'(\gamma_1))^{-\frac{1}{2}}}{\gamma_2 - \gamma_1} \ . \tag{19}$$

Proof. The proof relies heavily on the proofs of Theorem 4.1 and Theorem 5.1. In the proof of Theorem 4.1 we saw that inequality (10), which is responsible for the monotonic convergence of Euler's method, was based on the convexity of ϕ'. This convexity caused a linear approximation to ϕ' to lie below ϕ', but it also causes any secant approximation to ϕ', based on two points γ_1 and γ_2, to lie below ϕ' for points outside the interval determined by γ_1 and γ_2. Since this is precisely what we need, the finite difference approximation of the second derivative for Euler's method is given by (17).

For Halley's method, the function $(-\phi')^{-\frac{1}{2}}$ plays the role, which is played by ϕ' for Euler's method. Since

$$\left((-\phi')^{-\frac{1}{2}}\right)' = \frac{1}{2}(-\phi')^{-\frac{3}{2}}\phi'' \ ,$$

we obtain (18) immediately. The finite difference approximation in (19) is obtained analogously. $\qquad\square$

As we mentioned, the approximated methods, while still being at least of second order, are no longer of third order. For Euler's and Halley's methods, e.g., it can be proven that the order becomes $1 + \sqrt{2}$. The proofs are rather technical and we shall not consider them here. They are similar to results of the same type in [23].

Figures 5 and 6 compare the different methods for the example in Section 3. We have denoted the modified Halley method by "MH" and have distinguished the methods obtained by finite difference approximations to the second derivative by placing the letter "A" in front of their abbreviations. The approximations were carried out at the point $\bar{\gamma} = 4$. The second point used in the approximation of the second derivative is $\tilde{\gamma} = 8$. The potential of the approximated, or quasi-, modified Halley method, is clear from this example as it produces better results than the exact Euler method.

There are many possible combinations of methods, such as Newton-Secant, which leads to a third-order method if both steps are taken together, etc. (see [23]). The special form of our problem, obtained by transforming both the variable, $t = 1/\gamma$,
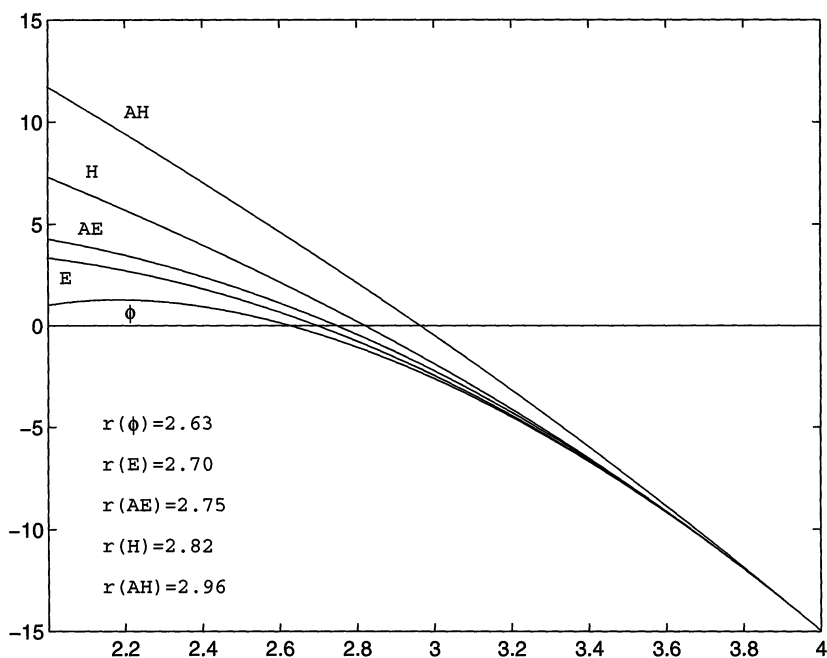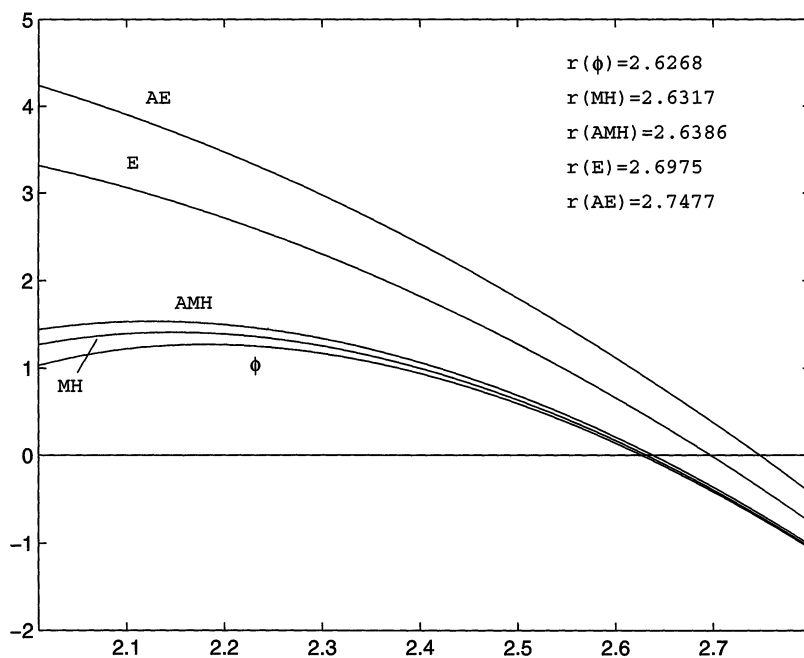
FIGURE 5. Comparison of the methods.

FIGURE 6. Further comparison of the methods.

and the function, $\phi(\gamma) = (\gamma - 1/\delta)F(\gamma)$, makes the convergence analysis of such methods relatively easy. If we take the Newton-Secant method as an example, then monotonic convergence from the right of the root is immediate because of the concavity of ϕ. We stress that without these transformations, the analysis of these methods would be much more complicated.

Other possible modifications include leaving more terms of ψ in the approximation. This does not affect the monotonicity of the convergence.

References

1. Borges, C.F., Gragg, W.B. (1993): A parallel divide and conquer algorithm for the generalized real symmetric definite tridiagonal eigenproblem. In Numerical Linear Algebra and Scientific Computing, L. Reichel, A. Ruttan and R.S. Varga, eds., pp. 10–28. de Gruyter, Berlin. MR **94k:**65051

2. Brown G.H., Jr. (1978): On Halley's variation of Newton's method. Amer. Math. Monthly **84**, pp. 726–728. MR **57:**1866

3. Bunch, J.R., Nielsen, C.P., Sorensen, D.C. (1978): Rank-one modification of the symmetric eigenproblem. Numer. Math. **31**, pp. 31–48. MR **80g:**65038

4. Bunch, J.R., Nielsen, C.P. (1978): Updating the singular value decomposition. Numer. Math. **31**, pp. 111–129. MR **80m:**65025

5. Cuppen, J.J.M. (1981): A divide and conquer method for the symmetric tridiagonal eigenvalue problem. Numer. Math. **36**, pp. 177–195. MR **82d:**65038

6. Dongarra, J.J., Sorensen, D.C. (1987): A fully parallel algorithm for the symmetric eigenvalue problem. SIAM J. Sci. Stat. Comput. **8**, No.2, pp. s139–s154. MR **88f:**65054

7. Donoghue W.F., Jr. (1974): Monotone matrix functions and analytic continuation. Springer-Verlag, Berlin Heidelberg. MR **58:**6279

8. Fuhrmann, D.R. (1988): An algorithm for subspace computation with applications in signal processing. SIAM J. Matrix Anal. Appl. **9**, No.2, pp. 213–220. MR **89f:**65040

9. Gander, W., Golub, G.H., von Matt, U. (1989): A constrained eigenvalue problem. Linear Algebra Appl. **114–115**, pp. 815-839. MR **90e:**15008

10. Gill, D., Tadmor, E. (1990): An $O(N^2)$ method for computing the eigensystem of $N \times N$ symmetric tridiagonal matrices by the divide and conquer approach. SIAM J. Sci. Stat. Comput. **11**, No.1, pp. 161–173. MR **91d:**65058

11. Golub, G.H. (1973): Some modified matrix eigenvalue problems. SIAM Rev. **15**, No.2, pp. 318–334. MR **48:**7569

12. Gragg, W.B., Reichel, L. (1990): A divide and conquer method for unitary and orthogonal eigenproblems. Numer. Math. **57**, pp. 695–718. MR **91h:**65052

13. Gragg, W.B., Thornton, J.R., Warner, D.D. (1992): Parallel divide and conquer algorithms for the symmetric tridiagonal eigenproblem and the bidiagonal singular value problem. In Modeling and Simulation, W.G. Vogt and M.H. Mickle, eds., vol. 3, part 1, pp. 49–56. University of Pittsburgh School of Engineering, Pittsburgh, PA.

14. Gu M., Eisenstat S.C. (1994): A stable and efficient algorithm for the rank-one modification of the symmetric eigenproblem. SIAM J. Matrix Anal. Appl. **15**, No.4, pp. 1266–1276. MR **96c:**65057

15. Henrici, P. (1964): Elements of numerical analysis. John Wiley & Sons, Inc., New York. MR **29:**4173

16. Li, R.C. (1994): Solving secular equations stably and efficiently. Technical Report UCB//CSD-94-851, Computer Science Division, University of California, Berkeley, CA. Also : LAPACK Working Notes 89.

17. Melman, A. (1995): Numerical solution of a secular equation. Numer. Math. **69**, pp. 483–493. MR **95j:**65050

18. Melman, A. (1997): A unifying convergence analysis of second-order methods for secular equations. Math. Comp. **66** (1997), 333–344. CMP 97:02

19. Ostrowski A.M. (1973): Solution of equations in Euclidean and Banach Spaces. Academic Press, New York. MR **50:**11760

20. Salehov G.S. (1952): On the convergence of the process of tangent hyperbolas. Dokl. Akad. Nauk SSSR **82**, pp. 525–528 (in Russian). MR **14:**91f

21. Šafiev R.A. (1963): The method of tangent hyperbolas. Dokl. Akad. Nauk SSSR **149**, pp. 788–791 (in Russian). Translation in Soviet Math. Dokl. **4**, pp. 482–485. MR **28:**734

22. Scavo T.R. and Thoo J.B. (1995): On the geometry of Halley's method. Amer. Math. Monthly **102**, pp. 417-426. MR **96f:**01019

23. Traub J.F. (1964): Iterative methods for the solution of equations. Prentice-Hall, Inc., Engle-wood Cliffs, New Jersey. MR **29:**6607

24. von Matt, U. (1993): Large constrained quadratic problems. Verlag der Fachvereine, Zürich.

DEPARTMENT OF INDUSTRIAL ENGINEERING, BEN-GURION UNIVERSITY, BEER-SHEVA 84105, ISRAEL

E-mail address: `melman@bgumail.bgu.ac.il`