THE CONVERGENCE OF MULTIPOINT ITERATIONS TO MULTIPLE ZEROS

G. W. Stewart

Department of Mathematics
Department of Computer Science
Carnegie-Mellon University
Pittsburgh, Pa. 15213

January, 1973

This research was supported in part by the Office of Naval Research under contract N00014-67-A-0314-0018.

ABSTRACT

This paper fills a gap in the theory of multipoint iteration function exemplified by the question: how does the secant method converge to a multiple zero. A general theory of the linear convergence of multipoint iterations is developed, and it is shown that two broad classes of iterative methods fit this theory. The results of numerical investigations based on the theory suggest that Muller's method applied to a multiple zero will inevitably produce complex iterates.

1. Introduction

Let θ be a complex valued function of a complex variable that is analytic in some region about the origin and has a zero of multiplicity p at the origin. The secant method for approximating a zero of θ starts with two initial approximations $\xi^{(0)}$ and $\xi^{(1)}$ and generates further approximations by means of the formula

$$\xi^{(i+1)} = \varphi(\xi^{(i)}, \xi^{(i-1)}),$$

where

$$\varphi(\xi_1, \xi_2) = \frac{\xi_1\theta(\xi_2) - \xi_2\theta(\xi_1)}{\theta(\xi_2) - \theta(\xi_1)}$$

It is known that if $p = 1$ and $\xi^{(0)}$ and $\xi^{(1)}$ are sufficiently near the origin, then the sequence $<\xi^{(i)}>$ converges to zero with order $\frac{1}{2}(1+\sqrt{5}) \cong 1.62$. When θ is a real function of a real variable and $p > 1$, it is obvious from geometric considerations that the secant method must converge to zero provided only that $\xi^{(0)}$ and $\xi^{(1)}$ are sufficiently small and lie on the same side of the origin. However, no analysis of the rate of convergence seems to have appeared in the literature.

Another popular iteration, Muller's method, generates approximate zeros by means of the formula

$$\xi^{(i+1)} = \varphi(\xi^{(i)}, \xi^{(i-1)}, \xi^{(i-2)}),$$

where $\varphi(\xi_1, \xi_2, \xi_3)$ is defined to be the zero lying nearest ξ_1 of the quadratic polynomial that interpolates θ at ξ_1, ξ_2, and ξ_3. When p is 1 or 2 and $\xi^{(0)}$, $\xi^{(1)}$, and $\xi^{(2)}$ are sufficiently small, the sequence $<\xi^{(i)}>$ is known to converge to zero with order about 1.8 or 1.2. When $p > 2$, geometric intuition is insufficient to guarantee the convergence of the method, although numerical experiments will readily convince one that the iteration converges linearly and that the convergence ratio is complex.

The object of this paper is to investigate the convergence of multipoint iterations of the form

$$(1.1) \qquad \xi^{(i+1)} = \varphi(\xi^{(i)}, \xi^{(i-1)}, \ldots, \xi^{(i-n+1)})$$

to a multiple zero of θ. The investigation is divided into two parts. In Section 2 we shall determine conditions under which the sequence of iterates generated by (1.1) can converge linearly to zero and exhibit an equation whose roots are potential convergence ratios. In Section 3 we shall apply this theory to two broad classes of iteration functions, those generated by linear interpolation, of which the secant method and Muller's method are examples, and those generated by inverse linear interpolation. Although an iteration such as Muller's method depends on both the function θ and the choice of an interpolating basis (in this case 1, ξ, and ξ^2), it turns out that the possible convergence ratios depend only on the multiplicity p of the zero of θ and the fact that linear interpolation on a nice set of functions is used to generate φ. This is analogous to, but slightly weaker than the well known fact that under rather general conditions a multipoint iteration of the form (1.1) converges to a simple zero of θ with an order that depends only on the number of points used by the iteration (e.g., see [4]).

The theory developed in Section 2 is closely related to techniques used by Kiho Lee Kim [2] and the author in a numerical investigation of the convergence of a variant of the Rayleigh quotient iteration.

Throughout this paper we shall use Householder's notational conventions [1]. In addition $\mathbb{C}^n$ will denote complex n space. We shall use the max vector norm defined by

$$\|x\| = \max\{|\xi_i| : i = 1,2,\ldots,n\}$$

and the subordinate matrix row sum norm defined by

$$\|A\| = \max\{\sum_{j=1}^{n} |\alpha_{ij}| : i = 1,2,\ldots,n\}$$

The symbol $\|.\|_1$ will denote the norm defined by

$$\|x\|_1 = \sum_{i=1}^{n} |\xi_i| .$$

Since we shall be concerned with sequences of numbers whose terms decrease in absolute value, it is convenient to introduce the notation $\mathcal{C}^n$ for the set

$$\mathcal{C}^n = \{x \in \mathbb{C}^n : 0 < |\xi_1| < |\xi_2| < \ldots < |\xi_n|\}.$$

2. The General Theory

Let $\varphi: \mathcal{C}^n \to \mathbb{C}$. In this section we shall be concerned with determining when a sequence $<\xi^{(i)}>$ generated by (1.1) converges linearly to zero. By linear convergence we mean

$$\lim_{i\to\infty} \frac{\xi^{(i+1)}}{\xi^{(i)}} = \varkappa, \tag{2.1}$$

where

$$|\varkappa| < 1.$$

Our approach will be to write φ in the form

$$\varphi = \varphi_1 + \varphi_2,$$

where φ_1 is a twice differentiable homogeneous function ($\varphi_1(\alpha x) = \alpha\varphi_1(x)$) and φ_2 is small compared with φ_1. It will then turn out that the behavior of the sequence $\langle\xi^{(i)}\rangle$ will be essentially determined by the behavior of the function φ_1.

In order to motivate the rather complicated Theorem 2.1, let us first consider the case where $\varphi_2 = 0$. With any vector $x \in \mathbb{C}^n$ associate the vector $y_x \in \mathbb{C}^{n-1}$ defined by

$$y_x = (\xi_1/\xi_n,\ \xi_2/\xi_n, \ldots, \xi_{n-1}/\xi_n).$$

If we set

$$x_i = (\xi^{(i)},\ \xi^{(i-1)}, \ldots, \xi^{(i-n+1)})^T,$$

then (2.1) will be satisfied if and only if

$$\text{(2.2)} \qquad \lim_{i\to\infty} y_{x_i} = (\varkappa^{n-1},\ \varkappa^{n-2}, \ldots, \varkappa)^T \equiv y_\varkappa .$$

Now it follows from the homogeneity of φ_1 that $y_{x_{i+1}} = g_1(y_{x_i})$, where

$$\text{(2.3)} \qquad g_1 = \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_{n-1} \end{pmatrix} \mapsto \eta_{n-1}^{-1} \quad \varphi_1 \begin{pmatrix} \eta_1,\ \eta_2, \ldots, \eta_{n-1},\ 1 \\ \eta_1 \\ \vdots \\ \eta_{n-2} \end{pmatrix}$$

Since g_1 is continuous, equation (2.2) implies that

$$g_1(y_\varkappa) = y_\varkappa,$$

which is satisfied only by vectors of the form $y_\varkappa = (\varkappa^{n-1}, \varkappa^{n-2},\ldots,\varkappa)^T$, where $\varkappa$ satisfies the equation

$$(2.4) \qquad \varkappa^{n-1} = \varkappa^{-1}\varphi(\varkappa^{n-1}, \varkappa^{n-2},\ldots,1).$$

Moreover if the spectral radius of the Jacobian of g_1 at $y_\varkappa$ is less than unity $(\rho[g_1'(y_\varkappa)] < 1)$, then the iteration $y_{x_{i+1}} = g(y_{x_i})$ will converge to $y_\varkappa$, provided only that $y_{x_{n-1}}$ is sufficiently close to $y_\varkappa$.

If φ_2 becomes small fast enough with x, then the behavior of the homogeneous function φ_1 determines the behavior of the general function $\varphi = \varphi_1 + \varphi_2$.

Theorem 2.1. Let $\varphi = \varphi_1 + \varphi_2$, where φ_1 is a homogeneous, twice differentiable function on $\mathbb{C}^n$ and φ_2: $\mathbb{C}^n \to \mathbb{C}$ satisfies

$$(2.5) \qquad |\varphi_2(x)| \le \|x\|^q \quad (q > 1).$$

Let $\xi^{(i)}$, x_i, y_i, and g_1 be defined as above. Let $\varkappa$ be a root of (2.4) satisfying

$$|\varkappa| < 1,$$

and let $y_\varkappa = (\varkappa^{n-1}, \varkappa^{n-2},\ldots,\varkappa)^T$. Then if

$$\rho[g_1'(y_\varkappa)] < 1,$$

there is a neighborhood $\mathcal{M} \subset \mathbb{C}^n$ at the origin and a neighborhood $\mathcal{N} \subset \mathbb{C}^{n-1}$ about $y_\varkappa$ such that whenever $x_{n-1} \in \mathcal{M}$ and $y_{x_{n-1}} \in \mathcal{N}$ the sequence $\langle\xi^{(i)}\rangle$ converges linearly to zero with ratio $\varkappa$.

<u>Proof</u>. Let $y_i = y_{x_i}$. Then

$$y_{i+1} = g_1(y_i) + g_2(y_i) \equiv g(y_i)$$

where

$$g_2(y_{x_i}) = \frac{\varphi_2(x_i)}{\xi^{(i-n+2)}} e_1 = \frac{\varphi_2(x_i)}{\xi^{(i-n+1)}\eta_{n-2}^{(i)}} e_1.$$

Moreover to say that $\langle \xi^{(i)} \rangle$ converges linearly to zero with ratio $\varkappa$ is equivalent to saying that $\langle \xi^{(i)} \rangle$ converges to zero and $\langle y_i \rangle$ converges to $y_\varkappa$. It is this latter proposition that we shall actually prove.

Because $\rho[g'(y_\varkappa)] < 1$, we can find a constant $\rho < 1$, a norm $\|\cdot\|_\rho$, and a neighborhood $\mathcal{N} = \{y \in \mathbb{C}^{n-1}: \|y_\varkappa - y\|_\rho < \pi\}$ such that if $y \in \mathcal{N}$ then

$$\|g(y) - y_\varkappa\|_\rho \leq \rho \|y - y_\varkappa\|_\rho.$$

Shrinking $\mathcal{N}$ if necessary, we can also find constants $\varkappa_-$ and $\varkappa_+$ with $0 < \varkappa_- \leq |\varkappa| \leq \varkappa_+ < 1$ such that for all $y \in \mathcal{N}$

$$\varkappa_-^i \leq |\eta_{n-i}| \leq \varkappa_+^i.$$

Now if $x \in \mathbb{C}^n$ and $y_x \in \mathcal{N}$, then $\|x\| = |\xi_n|$ and

$$\|g(y_x) - y_\varkappa\|_\rho \leq \rho \|y_x - y_\varkappa\|_\rho + \frac{|\xi_n|^{q-1}}{|\eta_{n-2}|} \|e_1\|_\rho$$

$$< \rho\pi + \frac{|\xi_\varkappa|^{q-1}}{\varkappa_-^2} \|e_1\|_\rho.$$

Hence if we take $\mathcal{M}$ to be the neighborhood in $\mathbb{C}^{n-1}$ of all x satisfying

$$\|x\|^{q-1} \le \frac{(1-\rho)\pi\, \varkappa_-^2}{\|e_1\|_\rho} ,$$

then it follows that whenever $x_i \in \mathcal{M}$ and $y_i \in \mathcal{N}$ we have $g(y_i) \in \mathcal{N}$. Moreover, $x_{i+1} \in \mathcal{M}$ and

$$|\xi^{(i-n+2)}| \le \varkappa_+ |\xi^{(i-n+1)}| .$$

Let $x_{n-1} \in \mathcal{M}$ and $y_{n-1} \in \mathcal{N}$. Then from the foregoing $x_i \in \mathcal{M}$ and $y_i \in \mathcal{N}$ for i = n, n+1,... . If we set $\xi_0 = |\xi^{(0)}|$ and $\epsilon_0 = \|y_{n-1} - y_\varkappa\|_\rho$ and define

(2.6) $$\xi_{k+1} = \varkappa_+ \xi_k$$

and

(2.7) $$\epsilon_{k+1} = \rho\epsilon_k + \sigma\xi_k^{q-1}$$

where $\sigma = \|e_1\|_\rho / \varkappa_-^2$, then

(2.8) $$|\xi^{(k)}| \le \xi_k$$

and

(2.9) $$\|y_{n+k-1} - y_\varkappa\|_\rho \le \epsilon_k .$$

The solution of the system (2.6) and (2.7) is easily seen to be

$$\xi_k = \varkappa_+^k \xi_0$$

and

(2.10) $$\epsilon_k = \rho^k \epsilon_0 + \sigma\xi_0^{q-1} \frac{\varkappa_+^{k(q-1)} - \rho^k}{\varkappa_+^{q-1} - \rho} .$$

Thus $<\xi_k>$ and $<\epsilon_k>$ converge to zero, which, in view of (2.8) and (2.9), establishes the theorem.

The nonhomogeneous perturbation has a curious effect on the behavior of the sequence $<\xi^{(i)}>$. As the theorem shows the $\xi^{(i)}$ may still converge linearly with ratio κ, and the rate at which the ratio κ is attained will be bounded by the rate at which the sequence $<\epsilon^{(k)}>$ approaches zero. Now in (2.10) ρ may be taken as near $\rho[g'(y_\kappa)]$ as we like. Moreover as y_i approaches y_κ, the constant κ_+ can be taken near κ. If $\kappa^{q-1} > \rho$, which will be true in most applications, the ϵ_k will approach zero as $\kappa^{k(q-1)}$. In the homogeneous case, of course, ϵ_k approaches zero as ρ^k. Otherwise put, in both the homogeneous and nonhomogeneous cases the $\xi^{(i)}$ converge linearly, but this behavior may be exhibited more slowly in the nonhomogeneous case.

3. Interpolatory Methods

In this section we shall consider the behavior of two broad classes of iterative methods for finding a zero of an analytic function θ and show that the behavior of these methods at multiple zeros is described by the theory of Section 2.

The first method is the method of direct linear interpolation, which includes the secant method and Muller's method. We shall assume that θ has a zero of multiplicity p at the origin. Let $\psi_1, \psi_2, \ldots, \psi_n$ be functions that are analytic in some neighborhood of the origin. Let the components of $x = (\xi_1, \xi_2, \ldots, \xi_n)^T$ be approximate zeros of θ. Determine constants $\gamma_{x1}, \gamma_{x2}, \ldots, \gamma_{xn}$ so that the function

$$\sigma_x = \gamma_{x1}\psi_1 + \gamma_{x2}\psi_2 + \ldots + \gamma_{xn}\psi_n$$

satisfies

$$\sigma_x(\xi_i) = \theta(\xi_i), \qquad i = 1,2,\ldots,n.$$

Thus σ_x is the linear combination of $\psi_1, \psi_2, \ldots, \psi_n$ that interpolates θ at $\xi_1, \xi_2, \ldots, \xi_n$. Let $\varphi(x)$ be the zero of σ_x that lies nearest ξ_1. Then $\varphi(x)$ is the new approximate zero of θ. The process is iterated as usual by repeating it with the vector $(\varphi(x), \xi_1, \xi_2, \ldots, \xi_{n-1})^T$.

Of course the interpolating function σ_x need not exist, nor need it have a zero. However, we shall show that, under mild restriction on the ψ_i, for sufficiently small $x \in \mathbb{C}^n$ the function σ_x is a small perturbation of the polynomial interpolant σ_{x1} of θ at $\xi_1, \xi_2, \ldots, \xi_n$ and consequently that $\varphi(x)$ is a slight perturbation of the zero $\varphi_1(x)$ of σ_{x1} lying nearest ξ_1. To establish this fact we shall need some additional notation.

For any $x \in \mathbb{C}^n$ let y_x be defined as in Section 2 and let

$$z_x = |\xi_n^{-1}|x.$$

Given any scalar α let

$$W_\alpha = \operatorname{diag}(1, \alpha, \ldots, \alpha^{n-1})$$

and

$$w_\alpha = W_\alpha e = (1, \alpha, \ldots, \alpha^{n-1})^T.$$

Given any vector x let

$$D_x = \operatorname{diag}(\xi_1, \xi_2, \ldots, \xi_n)$$

and let V_x be the Vandermonde matrix

$$V_x = \begin{pmatrix} 1 & \xi_1 & \cdots & \xi_1^{n-1} \\ 1 & \xi_2 & \cdots & \xi_2^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & \xi_n & \cdots & \xi_n^{n-1} \end{pmatrix}.$$

If θ is a function let

$$D_{\theta,x} = \mathrm{diag}(\theta(\xi_1),\ \theta(\xi_2),\dots,\theta(\xi_n))$$

and let

$$d_{\theta,x} = D_{\theta,x}e = (\theta(\xi_1),\ \theta(\xi_2),\dots,\theta(\xi_n))^T.$$

If $\psi_1,\ \psi_2,\dots,\psi_n$ are functions, let $V_{\psi,x}$ be the generalized Vandermonde matrix

$$V_{\psi,x} = (d_{\psi_1,x},\dots,d_{\psi_n,x}).$$

Note that the vector c_x of coefficients of σ_x, if it exists, must satisfy

$$V_{\psi,x}c_x = d_{\theta,x}.$$

Moreover if c_{x1} is the vector of coefficients of a polynomial σ_{x1}, then

$$\sigma_{x1}(\zeta) = w_\zeta^T\ c_{x1}.$$

We are now in a position to prove the results announced above.

Theorem 3.1. Let θ and $\psi_1,\ \psi_2,\dots,\psi_n$ be analytic at the origin and let $\theta(\xi) = \xi^p+\dots$ have a zero of multiplicity $p \geq n$ at the origin. Let $\psi_i(\xi) = \sum_{j=1}^{\infty} \alpha_{ij}\xi^{j-1}$ and let the matrix

$$(3.1)\qquad \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nn} \end{pmatrix}$$

be nonsingular. Let σ_x and $\varphi(x)$ be defined as above. Let σ_{x1} denote the polynomial of degree n-1 interpolating ζ^p at $\xi_1,\ \xi_2,\dots,\xi_n$ and let $\varphi_1(x)$ be the zero of σ_{x1} lying nearest ξ_1. Then φ_1 is a homogeneous function.

Further let $x_0 \in \mathcal{C}^n$ be such that $\varphi_1(x_0)$ is a simple zero of $\sigma_{x_0 1}$. Then there is a neighborhood $\mathcal{M} \subset \mathcal{C}^n$ at the origin and $\mathcal{N} \subset \mathcal{C}^{n-1}$ of $y_0 = y_{x_0}$ such that whenever $x \in \mathcal{M}$ and $y_x \in \mathcal{N}$ the interpolating function σ_x exists and has a unique simple zero $\varphi(x)$ that satisfies

$$|\varphi(x) - \varphi_1(x)| \leq \beta \|x\|^2,$$

for some constant β.

<u>Proof</u>. By the nonsingularity of the matrix (3.1) we may assume that

$$\psi_i(\xi) = \xi^{i-1} + \xi^n \pi_i(\xi),$$

where the functions π_i are analytic about the origin. We may also write θ in the form

$$\theta(\xi) = \xi^p + \xi^{p+1} \rho(\xi),$$

where ρ is analytic about the origin. Then if σ_x exists, its coefficients must satisfy

$$V_{\psi,x} c_x \equiv (V_x + D_x^n V_{\pi,x}) c_x = D_x^p e + D_x^{p+1} d_{\rho,x}$$

The first step in the proof is to determine conditions under which $V_{\psi,x} = V_x + D_x^n V_{\pi,x}$ is nonsingular.

Let $\mathcal{N}_1$ be a neighborhood of y_0 such that

$$\nu = \sup_{y_x \in \mathcal{N}_1} \|V_{z_x}^{-1}\|$$

is finite. Let $\mathcal{M}_1 = \{x \in \mathcal{C}^{n-1} : \|x\| \leq 1\}$ and let

$$\gamma = \sup_{x \in \mathcal{M}_1} \|V_{\pi,x}\| .$$

Now

$$V_x^{-1} D_x^n = V_{\|x\|z_x}^{-1} D_{\|x\|z_x}^n = \|x\|^n G_{\|x\|}^{-1} V_{z_x}^{-1} D_{z_x}$$

$$= \|x\| \operatorname{diag}(\|x\|^{n-1}, \|x\|^{n-2}, \dots, 1) V_{z_x}^{-1} D_{z_x}^n .$$

Hence if $x \in \mathcal{M}_1$ and $y_x \in \mathcal{N}_1$, then

$$\|V_x^{-1} D_x^n\| \le \nu \|x\| .$$

It follows from standard theorems on matrix inverses (e.g. see [4, Ch. 4]) that if

$$\mathcal{M}_2 = \mathcal{M}_1 \cap \{x \in \mathbb{C}^n : \nu\gamma\|x\| < 1/2\}$$

and $x \in \mathcal{M}_2$ with $y_x \in \mathcal{N}_1$ then $V_{\psi,x}$ is nonsingular and

$$V_{\psi,x}^{-1} = (I + E_x) V_x^{-1},$$

where

$$\|E_x\| < 1$$

and

$$E_x = -V_x^{-1} D_x^n V_{\pi,x}(I + E_x). \tag{3.2}$$

Thus if $x \in \mathcal{M}_2$ and $y_x \in \mathcal{N}_1$, then σ_x exists and its coefficient vector $c_x = V_{\psi,x}^{-1}(D_x^p e + D_x^{p+1} d_{\rho,x})$ can be written in the form $c_x = c_{x1} + c_{x2}$, where

$$c_{x1} = V_x^{-1} D_x^p e$$

and

$$c_{x2} = E_x V_x^{-1} D_x e + (I + E_x) V_x^{-1} D_x^{p+1} d_{\rho,x}$$

Moreover σ_x can be written in the form $\sigma_x = \sigma_{x1} + \sigma_{x2} + \sigma_{x3}$, where

$$\sigma_{x1}(\zeta) = w_\zeta^T c_{x1},$$
$$\sigma_{x2}(\zeta) = w_\zeta^T c_{x2},$$

and

$$\sigma_{x3}(\zeta) = \zeta^n \sum_{i=1}^{n} \gamma_{xi}\, \pi_i(\zeta)$$
$$= \zeta^n [d_\pi(\zeta)]^T c_x, \quad (d_\pi(\zeta) = (\pi_1(\zeta), \ldots, \pi_n(\zeta))^T).$$

Note that σ_{x1} is simply the polynomial that interpolates ξ^p at $\xi_1, \xi_2, \ldots, \xi_n$. For any $\alpha \neq 0$ we have

$$\sigma_{(\alpha x)1}(\alpha\zeta) = w_{\alpha\zeta}^T V_{\alpha x}^{-1} D_{\alpha x}^p e = (w_\zeta^T W_\alpha)(V_x W_\alpha)^{-1}(\alpha^p D_x^p) e$$
$$= \alpha^p w_\zeta V_x^{-1} D_x^p e = \alpha^p \sigma_{x1}(\zeta).$$

This implies that the zeros of σ_{x1} are homogeneous functions of x.

By hypothesis there is a simple zero $\varphi_1(x_0)$ of $\sigma_{x_0 1}$ lying nearest the first component of x_0. Since c_{x1} depends continuously on x and the zeros of σ_{x1} depend continuously on c_{x1}, we can find a neighborhood $\mathcal{N} \subset \mathcal{N}_1$ of y_0 such that whenever $y_x \in \mathcal{N}$ there is a unique simple zero $\varphi_1(x)$ of σ_{x1} lying nearest ξ_1. Moreover we can choose $\mathcal{N}$ so that there are positive constants ϵ and μ such that

$$y_x \in \mathcal{N} \text{ and } |\zeta - \varphi_1(z_x)| \le \epsilon \Rightarrow |\sigma_{z_x 1}(\zeta)| \ge \mu|\zeta - \varphi_1(\zeta_x)|.$$

It follows that if $y_x \in \mathcal{N}$, we have

(3.2)
$$|\zeta - \varphi_1(x)| \le \|x\|\,\epsilon \Rightarrow |\zeta/\|x\| - \varphi_1(z_x)| \le \epsilon$$
$$\Rightarrow |\sigma_{z_x 1}(\zeta/\|x\|)| \ge \mu|\zeta/\|x\| - \varphi_1(z_x)|$$
$$\Rightarrow \|x\|^{-p}|\sigma_{x1}(\zeta)| \ge \mu\|x\|^{-1}|\zeta - \varphi_1(x)|$$
$$\Rightarrow |\sigma_{x1}(\zeta)| \ge \mu\|x\|^{p-1}|\zeta - \varphi_1(x)|.$$

The rest of the proof consists in applying Rouché's theorem to show that there is a zero of σ_x near the zero $\varphi_1(x)$ of σ_{x1}. This requires a lower bound on the values of σ_{x1}, which has already been given in (3.3). We also require upper bounds on the values of σ_{x2} and σ_{x3}, which we now determine.

To obtain a bound on $\sigma_{x2}(\zeta)$, note that from (3.2)

$$w^T_{\omega\|x\|}\, E_x\, V_x^{-1}\, D_x^p = \|x\|^{p+1} w^T_{\omega}[V_{z_x}^{-1}\, D^n_{z_x}\, V_{\pi,x}(I + E_x)$$

$$\operatorname{diag}[(\|x\|^{n-1},\ldots,1)V_{z_x}^{-1}\, D^p_{z_x}].$$

Hence if $x \in \mathcal{M}_2$ and $y_x \in \mathcal{N}$

$$\|w^T_{\omega\|x\|}\, E_x\, V_x^{-1}\|_1 \le \|w_\omega\|_1 \|x\|^{p+1}[\nu\gamma(2)\nu]$$

$$= \quad 2\gamma\nu^2\|w_\omega\|_1\|x\|^{p+1}.$$

Let

$$\bar{\rho} = \sup_{x\in\mathcal{M}_2} \|d_{\rho,x}\|$$

Then if $x \in \mathcal{M}_2$, $y_x \in \mathcal{N}$ and $|\zeta| \le \omega\|x\|$,

$$|\sigma_{x2}(\zeta)| = |w^T_\zeta\, E_x\, V_x^{-1}(D^p_x\, e + D^{p+1}_x d_{\rho,x}) + w^T_\zeta\, V_x^{-1}\, D_x^{p+1}\, d_{\rho,x}|$$

(3.3)

$$\le \quad [2\gamma\nu^2(1 + \bar{\rho}) + \nu\bar{\rho}]\|w_\omega\|_1\ \|x\|^{p+1}.$$

The bound on $\sigma_{x3}(\zeta)$ depends more closely on the range of values of ζ. Let

$$\lambda = \sup_{y_x\in\mathcal{N}} |\varpi_1(z_x)|.$$

Then

$$(3.5) \qquad y_x \in \mathcal{N} \Rightarrow |\varphi_1(x)| \le \lambda\|x\| .$$

Let

$$\bar{\pi} = \sup_{\substack{|\xi| \le 2\lambda\|x\| \\ x \in \mathcal{M}_2}} \|d_\pi(\xi)\|_1 .$$

Then if $x \in \mathcal{M}_2$, $\omega \le 2\lambda$, and $|\zeta| \le \omega\|x\|$, we have

$$|\sigma_{x3}(\zeta)| = |\zeta^n [d_\pi(\zeta)]^T (I + E_x) V_x^{-1} (D_x^p e + D_x^{p+1} d_{\rho,x})|$$

$$(3.6) \qquad \le \omega^n \|x\|^n \bar{\pi} \quad (2) \ \|x\|^{p-n+1} \nu (1 + \|x\| \bar{\rho})$$

$$= 2 \bar{\pi}\nu\omega^n \|x\|^{p+1} (1 + \|x\| \bar{\rho}).$$

We are now ready to compare σ_{x1} with $\sigma_{x2} + \sigma_{x3}$. Specifically we shall seek a number $\tau_x < \varepsilon, \lambda$ such that for all sufficiently small $x \in \mathcal{M}_2$ with $y_x \in \mathcal{N}$ we have

$$(3.7) \qquad |\zeta - \varphi_1(x)| = \tau_x\|x\| \Rightarrow |\sigma_{x1}(\zeta)| > |\sigma_{x2}(\zeta)| + |\sigma_{x3}(\zeta)| .$$

Now if $|\zeta - \varphi_1(x)| = \tau_x\|x\|$, then from (3.5) $|\zeta| \le (\lambda+\tau_x)\|x\| \le 2\lambda\|x\|$. Hence

$$|\sigma_{x2}(\zeta)| + |\sigma_{x3}(\zeta)| < \delta\|x\|^{p+1},$$

where, from (3.4) and (3.6),

$$\delta = [2\gamma\nu^2(1+\bar{\rho}) + \nu\bar{\rho}]\|w_{2\lambda}\|_1 + 2 \bar{\pi}\nu(2\lambda)^n(1+\bar{\rho}).$$

But from (3.3)

$$\sigma_{x1}(\zeta) \ge \mu\tau_x\|x\|^p .$$

Hence if we set

$$\beta = \frac{\delta}{\mu}$$

and define

$$\tau_x = \beta\|x\|,$$

then for all x in

$$\mathcal{M} = \mathcal{M}_2 \cap \{x\colon \beta\|x\| < \epsilon, \lambda\}$$

such that $y_x \in \mathcal{N}$ the implication (3.7) holds. It follows from Rouché's theorem that σ_x has a single simple zero in the circle $|\zeta - \varphi_1(x)| \le \beta\|x\|^2$. Denoting this zero by $\varphi(x)$, we see that φ can be written in the form $\varphi_1 + \varphi_2$, where φ_1 is homogeneous and

$$|\varphi_2(x)| \le \beta\|x\|^2.$$

The importance of Theorem 3.1 when combined with Theorem 2.1 is that it allows the reduction of a large class of methods to methods based on polynomials. In particular the behavior of two point methods applied to a function with a zero of multiplicity p will be the same as the secant method applied to ξ^p. Similarly the behavior of Muller's method applied to ξ^p characterizes the behavior of three point methods.

However, even these simplified iterations are difficult to analyze. Consider, for example, the secant method applied to ξ^p. The iteration function is

$$\varphi(\xi_1, \xi_2) = \frac{\xi_1^p\xi_2 - \xi_1\xi_2^p}{\xi_1^p - \xi_2^p} .$$

The possible convergence ratios are the solutions of the equation (2.4), which takes the form

$$\varkappa = \frac{\varkappa^{p-1} - 1}{\varkappa^p - 1} \equiv \psi(\varkappa).$$

This equation can be written in the form

(3.8) $$(\varkappa-1)(\varkappa^p + \varkappa^{p-1} - 1) = 0.$$

The root $\varkappa=1$ of (3.8) is easily seen to be superfluous; hence the possible convergence ratios are the roots of the equation

(3.9) $$\gamma(\varkappa) \equiv \varkappa^p + \varkappa^{p-1} - 1 = 0.$$

Now $\gamma(0) = -1$ and $\gamma(1) = 1$. Since $\gamma(\zeta)$ is increasing for $\zeta \geq 0, \gamma$ has a simple zero $\varkappa$ in the interval $(0,1)$.

It remains to show that $\rho[g'(y_\varkappa)] < 1$. In this case

$$\rho[g'(y_\varkappa)] = |\psi'(\varkappa)|.$$

To show that $|\psi'(\varkappa)| < 1$ first note that

$$\psi(\zeta) = \frac{\zeta^{p-2} + \zeta^{p-3} +\ldots+ 1}{\zeta^{p-1} + \zeta^{p-2} +\ldots+ 1},$$

from which it is seen that ψ is nonincreasing on $(0,1)$. It follows that ψ' is nonpositive on $(0,1)$. From the identity

$$\ln(\psi)' = \frac{\psi'}{\psi}$$

and from the fact that $\psi(\varkappa) = \varkappa$, it follows that

$$-\psi'(\varkappa) = \frac{(p-1)\varkappa^{p-1}}{1 - \varkappa^{p-1}} - \frac{p\varkappa^p}{1 - \varkappa^p}.$$

Since $1 - \varkappa^{p-1}, 1 - \varkappa^p > 0$, the condition $-\psi'(1) < 1$ can be expressed in the form

$$(p-1)\varkappa^{p-1} - p\varkappa^p + \varkappa^{2p-1} < (1 - \varkappa^{p-1})(1 + \varkappa^p)$$

$$= 1 - \varkappa^{p-1} - \varkappa^p + \varkappa^{2p-1},$$

or

$$p\varkappa^{p-1}(1 - \varkappa) < 1-\varkappa^p = (1 - \varkappa)(1 + \varkappa + \varkappa^2 +\ldots+ \varkappa^{p-1}).$$

Thus the condition $-\psi'(\varkappa) < 1$ is equivalent to the condition

$$p\varkappa^{p-1} < 1 + \varkappa + \varkappa^2 +\ldots+ \varkappa^{p-1},$$

which is obviously true for $\varkappa \in (0,1)$.

This shows that the secant method can converge linearly to a multiple zero with a positive convergence ratio. However, this does not complete the picture, for equation (3.9) may have other roots whose moduli are less than unity, and these roots also represent possible convergence ratios. The complete analysis of the other roots of (3.9) is a difficult problem, and the author has had to content himself with a numerical investigation of the roots of (3.9) for values of p ranging from two to ten. The results show that equation (3.9) can indeed have other roots with moduli less than unity; however, for none of these roots is $|\psi'|$ less than unity, so that they can be effectively dismissed as possible convergence ratios. The remaining ratios, along with $|\psi'| = \rho(g')$ are given below.

n	$\varkappa$	$\rho[g'(\varkappa)]$
2	0.6180	0.3820
3	0.7549	0.3848
4	0.8192	0.3855
5	0.8567	0.3858
6	0.8813	0.3860
7	0.8987	0.3861
8	0.9116	0.3861
9	0.9216	0.3862
10	0.9296	0.3862

We are not able to perform for Muller's method even the modest amount of analysis that we were able to do for the secant method. Since $\omega(\xi_1, \xi_2, \xi_3)$ is a zero of the quatratic interpolating ξ^p at ξ_1, ξ_2, and ξ_3, it must satisfy the equation

$$\det\begin{pmatrix} 1 & \xi_1 & \xi_1^p \\ 1 & \xi_2 & \xi_2^p \\ 1 & \xi_3 & \xi_3^p \end{pmatrix} \omega^2(\xi_1, \xi_2, \xi_3) + \det\begin{pmatrix} 1 & \xi_1^p & \xi_1^2 \\ 1 & \xi_2^p & \xi_2^2 \\ 1 & \xi_3^p & \xi_3^2 \end{pmatrix} \omega(\xi_1, \xi_2, \xi_3)$$

(3.10)

$$+\det\begin{pmatrix} \xi_1^p & \xi_1 & \xi_1^2 \\ \xi_2^p & \xi_2 & \xi_2^2 \\ \xi_3^p & \xi_3 & \xi_3^2 \end{pmatrix} = 0.$$

Thus the roots of (2.4) will be among the roots of the equation

$$\det\begin{pmatrix} 1 & 1 & 1 \\ 1 & \varkappa & \varkappa^p \\ 1 & \varkappa^2 & \varkappa^{2p} \end{pmatrix} \varkappa^6 + \det\begin{pmatrix} 1 & 1 & 1 \\ 1 & \varkappa^p & \varkappa^2 \\ 1 & \varkappa^{2p} & \varkappa^4 \end{pmatrix} \varkappa^3$$

$$+\det\begin{pmatrix} 1 & 1 & 1 \\ \varkappa^p & \varkappa & \varkappa^2 \\ \varkappa^{2p} & \varkappa^2 & \varkappa^4 \end{pmatrix} = 0.$$

After some simplification, including the removal of some extraneous roots at 0 and 1, this equation becomes

$$\varkappa^{2p-1} + \varkappa^{2p-2} + \varkappa^{2p-3} - \varkappa^p - \varkappa^{p-1} - \varkappa^{p-2} + 1 = 0. \tag{3.11}$$

The derivatives of φ and hence the Jacobian of g can be obtained by differentiating (3.10) implicitly with respect to ξ_1, ξ_2, and ξ_3.

A numerical investigation of the solutions of (3.11) for values of p ranging from three to ten revealed the following facts. For each value of p, there is a pair of complex conjugat roots of (3.11) with moduli less than unity for which $\rho[g'] < 1$. For even p, there is also a negative root of modulus less than unity with $\rho[g'] < 1$. However, this root does not represent a convergence ratio, since, for this value of $\varkappa$, the number $\varkappa^3 = \varphi(\varkappa^2, \varkappa, 1)$ is not the root of the interpolating quadratic that lies nearest $\varkappa^2$. The convergence ratios are given below.

p	$\varkappa$	$\lvert\varkappa\rvert$	$\rho[g'(\varkappa)]$
3	$0.7132 \pm 0.2007i$	0.7409	0.4740
4	$0.8140 \pm 0.1464i$	0.8271	0.4717
5	$0.8617 \pm 0.1147i$	0.8693	0.4709
6	$0.8897 \pm 0.0942i$	0.8947	0.4705
7	$0.9083 \pm 0.0798i$	0.9118	0.4704
8	$0.9215 \pm 0.0693i$	0.9241	0.4702
9	$0.9313 \pm 0.0612i$	0.9333	0.4702
10	$0.9390 \pm 0.0548i$	0.9406	0.4701

These results suggest that when Muller's method is applied to finding a multiple zero it will produce complex values, a matter of practical importance when the function and its zero are real.

The second class of methods to be treated is the class of methods based on inverse linear interpolation. As with the direct interpolatory methods, one starts with a fixed set of basis functions $\{\psi_1, \psi_2, \ldots, \psi_n\}$. Given the vector x and the function values $\theta_i = \theta(\xi_i)$ $(i = 1, 2, \ldots, n)$ one determines coefficients $\gamma_{x1}, \gamma_{x2}, \ldots, \gamma_{xn}$ so that the function

$$\psi_x = \gamma_{x1}\psi_1 + \gamma_{x2}\psi_2 + \dots + \gamma_{xn}\psi_n$$

satisfies

$$\psi_x(\theta_i) = \xi_i \qquad (i = 1,2,\dots,n)$$

The new approximate zero $\varphi(x)$ is given by

$$\varphi(x) = \psi_x(0).$$

A result analogous to Theorem 3.1 holds for this class of methods.

Theorem 3.3. Let $\theta(\xi) = \xi^p + \xi^{p+1}\rho(\xi)$ be a function analytic in some neighborhood of the origin. Let $\psi_i(\xi) = \sum_{j=1}^{\infty} \alpha_{ij}\xi^{j-1}$ $(i = 1,2,\dots,n \le p)$ also be analytic at the origin, and suppose that the matrix

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} \end{pmatrix}$$

is nonsingular. Let $y_0 \in \mathcal{C}^{n-1}$. Then there is a neighborhood $\mathcal{M} \subset \mathcal{C}^n$ at the origin and a neighborhood $\mathcal{N} \subset \mathcal{C}^{n-1}$ of y_0 such that for all $x \in \mathcal{N}$ for which $y_x \in \mathcal{M}$, the function $\varphi(x)$ is well defined. Moreover if $\varphi_1(x)$ denotes the value at zero of the polynomial of degree n interpolating the values $\xi_1, \xi_2, \dots, \xi_n$ at the points $\xi_1^p, \xi_2^p, \dots, \xi_n^p$, then φ_1 is a homogeneous function of x that satisfies

$$\|\varphi(x) - \varphi_1(x)\| \le \beta\|x\|^{p+1}$$

for some constant β.

The proof of Theorem 3.2 is somewhat simpler than the proof of Theorem 3.1, since it does not involve an application of Rouché's theorem. Otherwise they are quite similar, and we shall not reproduce the proof of the latter theorem here.

The theorem, of course, permits the reduction of the general class of iterations applied to a general function to the polynomial case. When n = 2, the method of inverse polynomial interpolation is identical with the secant method, which has already been treated.

The author has investigated the case n = 3, again for values of p ranging from two to ten. In this case $\varphi(x)$ is given by the determinantal expression

$$(3.12)\qquad \varphi(\xi_1, \xi_2, \xi_3) = \frac{\det\begin{pmatrix} \xi_1 & \xi_1^p & \xi_1^{2p} \\ \xi_2 & \xi_2^p & \xi_2^{2p} \\ \xi_3 & \xi_3^p & \xi_3^{2p} \end{pmatrix}}{\det\begin{pmatrix} 1 & \xi_1^p & \xi_1^{2p} \\ 1 & \xi_2^p & \xi_2^{2p} \\ 1 & \xi_3^p & \xi_3^{2p} \end{pmatrix}}$$

Consequently the possible convergence ratios are roots of the equation

$$\det\begin{pmatrix} \varkappa^2 & \varkappa^{2p} & \varkappa^{4p} \\ \varkappa & \varkappa^p & \varkappa^{2p} \\ 1 & 1 & 1 \end{pmatrix} = \varkappa^3 \det\begin{pmatrix} 1 & \varkappa^{2p} & \varkappa^{4p} \\ 1 & \varkappa^p & \varkappa^{2p} \\ 1 & 1 & 1 \end{pmatrix},$$

which can be simplified to the equation

$$\varkappa(\varkappa^{p-1} + \varkappa^{p-2} + \ldots + 1)(\varkappa^{2p-1} + \varkappa^{2p-1} + \ldots + 1) = (\varkappa^{p-2} + \varkappa^{p-3} + \ldots + 1)({}^{2p-2} + {}^{2p-3} + \ldots + 1) \tag{3.13}$$

The derivatives of φ can be obtained explicitly from (3.12). For each p, there is only one real root of (3.13) for which $\rho(g') < 1$. These roots are given below.

p	$\varkappa$	$\rho[g'(\varkappa)]$
3	0.7200	0.3825
4	0.7906	0.3859
5	0.8327	0.3876
6	0.8607	0.3887
7	0.8806	0.3893
8	0.8956	0.3898
9	0.9072	0.3902
10	0.9165	0.3905

References

1. A. S. Householder, The Theory of Matrices in Numerical Analysis, Blaisdail, New York, 1964.

2. Kiko Lee Kim, A Numerical Study of a Variant of the Rayleigh Quotient Iteration, Masters Thesis, The University of Texas, Austin, Texas, 1970.

3. G. W. Stewart, Introduction to Matrix Computations, Academic Press, New York, 1972.

4. J. F. Traub, Iterative Methods for the Solution of Equations, Prentice-Hall, Englewood Cliffs, N. J., 1964.