AN ABSTRACT OF THE THESIS OF

Donald Allen Celarier for the degree of Doctor of Philosophy

in Mathematics presented on May 8, 1978

Title: A Study of the Global Convergence Properties of Newton's Method

Signature redacted for privacy.

Abstract approved: Joel Davis

A construct is developed which is useful in the investigation of the global convergence properties of Newton's method.

This construct is used to study the application of Newton's method to polynomials. A proof that Newton's method converges almost globally for polynomials with only real zeroes is extended to a larger class of polynomials. A conjecture is advanced concerning conditions which are necessary and sufficient for almost global convergence for both real and complex polynomials.

Another application of the construct involves the use of Bairstow's method on cubic polynomials having one real zero. The question of whether a certain fixed set is stable is resolved.

A Study of the Global Convergence
Properties of Newton's Method

by

Donald Allen Celarier

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Commencement June 1979

APPROVED:

Signature redacted for privacy.

Associate Professor of Mathematics
in charge of major

Signature redacted for privacy.

Chairman of Department of Mathematics

Signature redacted for privacy.

Dean of Graduate School

Date thesis is presented May 8, 1978

Typed by Lois F. Celarier for Donald Allen Celarier

ACKNOWLEDGEMENT

I would like to thank Dr. Joel Davis for his assistance and guidance. Without his help, understanding and patience, this study could not have been made.

I would also like to thank my family, particularly my wife, Lois, whose phenomenal adaptability to poverty and crisis has been invaluable in the culmination of this effort.

TABLE OF CONTENTS

A STUDY OF THE GLOBAL CONVERGENCE PROPERTIES OF NEWTON'S METHOD

I. INTRODUCTION

1.1. Preliminary Remarks

Let B be a Banach space [5]. Let P map B into B and be twice (Fréchet) differentiable at every point x in B. Then the Newton iteration

$$N(x) = x - (P'(x))^{-1}P(x)$$

is defined at every x such that $P'(x)$ has an inverse [1]. The function $N(x)$ maps a subset of B into B.

In this discussion of the global convergence properties of Newton's iteration (Newton's method), certain sets arise which contain their own images under N. This is the reason for mapping B into itself.

Some of the notation employed deserves explanation:

cl(S) is the closure of the set S;

int(S) is the interior of the set S;

fr(S) is the frontier (boundary) of S;

$||x||$ is the norm of x in B.

Occasionally there is a comment to be made which is incidental to the discussion. The existence of such a remark is indicated by the notation (#n). The comment will be found under 5.1.n in the final section.

1.2. Newton Sequences and Fixed Sets

A Newton sequence (generated by x_0) is a sequence $\{x_i\}$ of points in B such that N is defined at x_i and $x_{i+1} = N(x_i)$. Such a sequence may be said to terminate at x_i if $N(x_i)$ is undefined.

The definitions of Newton iteration and Newton sequence are extended to sets by

$$N(S) = \{N(x) \mid x \in S\} .$$

The inverse iteration is not in general a function but is a "set-valued function" defined by

$$N^{-1}(x) = \{y \mid N(y) = x\}$$

and the trail of x is

$$tr(x) = \{x_0 \mid x = x_i \text{ for some } i\} .$$

The extensions to sets are given by

$$N^{-1}(S) = \{y \mid N(y) \in S\} \quad \text{and}$$

$$tr(S) = \{y_0 \mid \text{for some } i,\ y_i \in S\} .$$

Notice that $tr(S) = S \cup N^{-1}(S) \cup N^{-1}(N^{-1}(S)) \cup \ldots$

A set S is fixed under N if $N(S) \subset S$. A fixed set S is attractive if every sufficiently small ε-neighborhood of S is fixed and repulsive if there is no fixed open ε-neighborhood of S. If B = R (the real line), a fixed singleton set $\{x\}$ is either attractive or repulsive.

A fixed set S is stable at one of its points x if $\|y_0 - x\| < \varepsilon$ implies $\text{dist}(y_i, S) < \varepsilon$ for all i. (Here dist(a,b) is the metric induced by the norm on B.)

1.3. Admissible Functions

Certain restrictions on the function P will be needed to discuss the global convergence of N. A function P is admissible if P is twice continuously differentiable at every point x in B and has the following property: each x_* in B such that $P(x_*) = 0$, there is an open ball $S(x_*,\rho)$ centered at x_* with radius ρ such that, for every x_0 in $S(x_*,\rho)$, $\{x_i\}$ converge to x_*.

Observe that if P is admissible and for some Newton sequence $\{x_i\}$ converges to x_*, then $P(x_*) = 0$. This is usually shown by writing $\|P'(x_i)(x_{i+1} - x_i)\| = \|P(x_i)\|$ and estimating the left-hand side using a mean-value theorem.

This behavior of P near a zero is exactly what is used in sequel. The way to ensure admissibility depends on the particular function under consideration and is not discussed in general here (#1).

II. THE G-CONSTRUCT

2.1. The Set $G_0(x_*)$

Let P be an admissible function. Suppose $P(x_*) = 0$ and let $T_0 = S(x_*,\rho)$ be the open ball required in the definition of admissible functions. Consider the set $U_1 = N^{-1}(T_0)$; there is an open connected component, T_1, of U_1 such that $T_0 \subset T_1$.

Recursively, let $U_{i+1} = N^{-1}(T_i)$ and let T_{i+1} be the open connected component of U_{i+1} which contains T_i.

Let $G_0(x_*) = \bigcup T_i$ where the union is over all positive integers. Then, $G_0(x_*)$ is open since it is the union of open sets and connected since $\bigcap T_i = T_0 \neq \emptyset$ (#2).

Clearly, $G_0(x_*)$ is fixed under N. It is the largest open connected set fixed under N for which every Newton sequence $\{x_i\}$ generated by a point of the set converges to x_*.

The boundary $fr(G_0(x_*))$ is also fixed under N (by a continuity argument).

2.2. The Sets $G_k(x_*)$ and $G(x_*)$

With $G_0(x_*)$ as described in 2.1, let $G_1(x_*) = N^{-1}(G_0(x_*))$. Then, each connected component of $G_1(x_*)$ is open. Also, $N(\mathrm{fr}(G_1(x_*))) \subset \mathrm{fr}(G_0(x_*))$. By construction, then, $G_1(x_*)$ is fixed and $\mathrm{fr}(G_1(x_*))$ is fixed.

Recursively define $G_{k+1}(x_*) = N^{-1}(G_k(x_*))$ and note that $G_k(x_*)$ has properties corresponding to those cited for $G_1(x_*)$.

Now let $G(x_*) = \cup(G_k(x_*))$ where the union extends over the non-negative integers k. The set $G(x_*)$ is open and contains all the points x_0 such that the Newton sequence $\{x_i\}$ generated by x_0 converges (eventually) to x_*.

Other characterizations of $G(x_*)$ which are useful are

$$G(x_*) = \mathrm{tr}(G_0(x_*)) \quad \text{and}$$
$$G(x_*) = \mathrm{tr}(S(x_*,\rho)).$$

2.3. The Sets G_k and G

The sets G_k are constructed from $G_k(x_i)$ where $P(x_i) = 0$. Specifically, let $X = \{x_i | P(x_i) = 0\}$ and define (for $k = 0,1,2,\ldots$) $G_k = \cup G_k(x_i)$ where the union is over all x_i in X. (X is countable since P is admissible.)

The set G has several useful characterizations. As a definition, put $G = \cup G_k$ where the union is over the non-negative integers k. Then,

$$\begin{aligned} G &= \cup_k(\cup_i G_k(x_i)) \\ &= \cup_i(\cup_k G_k(x_i)) \\ &= \mathrm{tr}(G_0) \\ &= \mathrm{tr}(\cup_i S(x_i, \rho_i)) \ . \end{aligned}$$

The set G contains all the points of B which generate Newton sequences which (eventually) converge.

Newton's method will be said to be <u>almost globally convergent</u> if $\mathrm{cl}(G) = B$; that is, if G is everywhere dense in the Banach space.

III. REAL POLYNOMIALS

3.1. The Theorem

The real line, R, (with, for example, the Euclidean norm) is a Banach space. For this space, every polynomial $f(x) \not\equiv 0$ is admissible and has a finite number of zeroes. These facts follow from Liouville's theorem, the fundamental theorem of algebra and the characterization of $f'(x)$ and $f''(x)$ as polynomials.

For proving the theorem which follows, it will be convenient to have described certain sets and inequalities. Let

$$X = \{x_i \mid f(x_i) = 0\} ,$$

$$Y = \{y_i \mid f(y_i) \neq 0,\ f'(y_i) = 0\} , \text{ and}$$

$$Z = \{z_i \mid f(z_i) \neq 0,\ f''(z_i) = 0\} .$$

Each of these sets is finite. Assume that all of the zeroes of $f(x)$ are real. Then the sets X and Y can be indexed so that

$$x_0 < y_1 < x_1 < \ldots < x_{n-1} < y_n < x_n \quad .$$

The set Z can be decomposed into two sets $Z_+ = \{z_{i+}\}$ and $Z_- = \{z_{i-}\}$ such that

$$x_0 < z_{0+} < y_1 < z_{1-} < x_1 < z_{1+} < y_2 < \ldots < y_n < z_{n-} < x_n$$

(where z_{i+} or z_{i-} fails to exist in exactly those cases where x_i is simple).

The following abuse of notation is also convenient: $G_k(x_0,x_n) = G_k(x_0) \cup G_k(x_n)$ and $G(x_0,x_n) = G(x_0) \cup G(x_n)$.

<u>Theorem</u>: For a real polynomial, if all of the zeroes of the polynomial are real, Newton's method is almost globally convergent.

<u>3.2. Proof of the Theorem</u>

The theorem is not new [2, 3] but the proof is new in some details. The proof allows a weaker hypothesis to be found (#3). A reasonable conjecture can be made concerning conditions which are necessary and sufficient for the almost global convergence of Newton's method for both real and complex polynomials.

The proof consists of establishing several lemmas.

<u>Lemma 1</u>: $X \cup Y \cup Z \subset \mathrm{cl}(G_1)$.

In fact, $X \subset G_0$, by definition.

$Y \subset \mathrm{cl}(G_1(x_0,x_n))$, since $G_0(x_0) = (-\infty, y_1)$ and $G_0(x_n) = (y_n, \infty)$ with the details supplied by noting that for an arbitrary sequence with y_i as a limit, the image under N has no cluster point and "diverges to infinity".

Finally, $z_{i+} \cup z_{i-} \subset G_0(x_i)$ is shown by noting that f(x) lies on one side of its tangent (the correct side) in the appropriate intervals. This results in monotone convergence of a Newton sequence generated by z_{i+} or z_{i-}. The convergence on the open subintervals defined by the

the terms of these sequences establishes the lemma. (#4)

<u>Lemma 2</u>: On each interval of points not in $cl(G_1)$, $N'(x)$ is continuous and monotone.

Let $u_i = \min(z_{i-}, x_i)$ and $v_i = \max(z_{i+}, x_i)$. Write $N'(x) = f(x)f''(x)/(f'(x))^2$. The continuity of $N'(x)$ on the intervals (y_i, u_i) and (v_i, y_{i+1}) follows from the previous lemma. Also, $N'(x) < 0$ on these intervals.

$N'(x)$ is unbounded at each y_i so the equation $N'(x) = a$ $(a<0)$ has at least 2n solutions. Comparison of the degrees of the numerator and denominator of $N'(x)$ shows that there are no more solutions. On each of these intervals $N'(x)$ takes each negative value exactly once and is monotone.

For $i = 1,2,\ldots,n-1$, let the interval $G_0(x_i)$ be (B_i, C_i). Define $J = \bigcup G_0(x_i)$ $(i = 1,2,\ldots,n-1)$.

Let H_* be the union of the two connected components of $G(x_0, x_n)$ which contain x_0 and x_n. Let $H = N^{-1}(H_*)$. There are n connected components of H, each containing a point of Y. With $D_0 = -\infty$ and $A_{n+1} = \infty$, designate the component of H containing y_i as (D_i, A_{i+1}).

Finally, let $E = \{A_1, D_n\}$. Then $A_i \in N^{-1}(E)$ and $D_i \in N^{-1}(E)$ for $i = 1,2,\ldots,n$.

<u>Lemma 3</u>: B_i is a cluster point of $tr(E)$ and C_i is a cluster point of $tr(E)$ for each $i = 1,2,\ldots,n-1$.

The closed intervals $[A_i, B_i]$ and $[C_i, D_i]$ occur in pairs such that each contains a component of the inverse image of the other.

Suppose that B_i is not a cluster point of tr(E). Then there is a point $E_i \in cl(G(x_o, x_n))$ such that the interval $[E_i, B_i]$ contains no other point of tr(E). On this interval, define the iteration $M(x) = N(N(x))$. Both E_i and B_i are repulsive fixed points under M.

By the previous lemma, $M'(x)$ is continuous and monotone so that either B_i or E_i must be attractive. This contradiction shows that $[E_i, B_i]$ must be a singleton set and that B_i is a cluster point of tr(E).

<u>Lemma 4:</u> $(A_i, B_i) \subset cl(G)$, $(C_i, D_i) \subset cl(G)$.

Set $J_0 = J$ and $J_{h+1} = N^{-1}(J_h)$. Between each pair of components of $tr(B_i, C_i)$ which are in J_h there is (at least) one additional component of $tr(B_i, C_i)$ in J_{h+1}. There is a component of J_{h+1} in each non-empty open interval in the complement of $J_j \cup G(x_0, x_n)$. This, coupled with the previous lemma, proves this lemma and the theorem.

3.3. Discussion of the Hypotheses

Consider, first, the polynomials

$$f(x;k) = x^4 - 2x^2 - k\ .$$

For all these polynomials, $Y = \{-1,0,1\}$ and $Z = \{-\sqrt{3}/3, \sqrt{3}/3\}$. Except for $x \in Y$,

$$N(x;k) = \frac{3x^4 - 2x^2 + k}{4x^3 - 4x}$$

For $k < -1$, there are no real zeroes of $f(x;k)$ and Newton's method converges nowhere.

For $-1 \leq k \leq 0$, the theorem holds and Newton's method converges almost globally.

For $0 < k < 11/9$, there are two intervals, each containing an inflection point of $f(x;k)$, which iterate into each other and Newton's method fails to converge in these intervals. (#5)

For $k = 11/9$, $Z \subset \mathrm{int}(\mathrm{cl}(G))$ but there is no i such that $Z \subset \mathrm{int}(\mathrm{cl}(G_i))$. Newton's method converges almost globally.

For $11/9 < k$, $Z \subset \mathrm{int}(\mathrm{cl}(G_i))$ for some i and Newton's method converges almost globally.

On reviewing the proof of the theorem, it is seen that the crucial fact is that (lemma 1) all of the critical values of $f(x)$ lie in $\mathrm{int}(\mathrm{cl}(G_1))$. It is clearly sufficient that all of the critical values lie in $\mathrm{int}(\mathrm{cl}(G_i))$ for some value of i (since the remainder of the proof would be essentially unchanged). As the example above shows, this is a weakening of the hypothesis of the theorem.

<u>Theorem</u>: Let $f(x)$ be a real polynomial. Let $X = \{x_i | f(x_i) = 0\}$, $Y = \{y_i | f'(y_i) = 0\}$, $Z = \{z_i | f''(z_i) = 0\}$. If, for some integer j, $X \cup Y \cup Z \subset \mathrm{int}(\mathrm{cl}(G_j))$, then Newton's method converges almost globally.

As indicated by the example, it is sufficient that $Y \subset \mathrm{int}(\mathrm{cl}(G_1))$ and $Z \subset \mathrm{int}(\mathrm{cl}(G))$. It is obviously necessary that $Y \cup Z \subset \mathrm{int}(\mathrm{cl}(G))$. These facts lead to the following conjecture.

<u>Conjecture</u>: For a real polynomial, Newton's method is almost globally convergent if and only if $Y \subset \mathrm{int}(\mathrm{cl}(G_1))$ and $Z \subset \mathrm{int}(\mathrm{cl}(G))$.

In addition, the same conjecture is offered for Newton's method applied to complex polynomials.

IV. A RESULT ON BAIRSTOW'S METHOD

4.1. Source of the Question

Bairstow's method [6, 7] is a scheme for finding real quadratic factors of a real polynomial. A common description of the method follows (#6).

Let $f(x)$ be a polynomial with real coefficients and let $p(x,u,v) = x^2 - ux - v$. The Euclidean algorithm gives

$$f(x) = p(x,u,v)q(x,u,v) + xF(u,v) + G(u,v)$$

where q, F and G are polynomials in the indicated variables. Newton's method is applied to the vector

$$P(u,v) = \begin{pmatrix} F(u,v) \\ G(u,v) \end{pmatrix}$$

resulting in

$$N(u,v) = \begin{pmatrix} u \\ v \end{pmatrix} - \left(\frac{\partial(F,G)}{\partial(u,v)}\right)^{-1} \begin{pmatrix} F \\ G \end{pmatrix} .$$

If this iteration converges, it converges to a point (u,v) for which $p(x,u,v)$ is a factor of $f(x)$.

The general question of when $P(u,v)$ is admissible deserves further study. It can be seen that $P(u,v)$ is admissible if $f(x)$ has no non-simple zeroes. (This obviates the possibility that $P'(u,v)$ is zero at some point in every punctured neighborhood of a zero of $P(u,v)$.)

By writing $p(x,u,v) = (x-z)(x-y)$ it becomes apparent that the convergence properties of Bairstow's method can be discussed in terms of the zeroes of p as well as in terms of the coefficients of p.

Boyd [4] obtains the following useful results:

Theorem: If, for real r, $f(r) = 0$, the line $v = r^2 - ru$ contains a dense subset which is fixed under N.

The proof consists in showing that, for such an r and all u_0, v_0,

$$0 = (ru_1 - ru_0 + v_1 - v_0)q(r,u_0,v_0) .$$

The exceptional points, at which $q(r,u_0,v_0) = 0$, are sparse enough to omit the trails of such points.

Theorem: If r is real, $f(r) = 0$, $v_0 = r^2 - ru_0$, $q(r,u_i,v_i) \neq 0$ and $p(x,u_i,v_i) = (x-z_i)(x-y_i)$, then $z_i = r$ and $y_{i+1} = y_i - g(y_i)/g'(y_i)$ where $g(x) = f(x)/(x-r)$.

The proof is a calculation.

Corollary: If $f(x)$ is a cubic with one real zero, r, no Newton sequence generated by (u_0,v_0) can converge if $v_0 = r^2 - ru_0$ (with a countable set of possible exceptions).

The fixed subset of the line $v = r^2 - ru$ can *a priori* be attractive, repulsive or neither. The question addressed here is: At what points and under what conditions is this fixed set stable?

4.2. Reduction of the Problem

If $f(x)$ is a cubic with one real zero and two non-real zeroes, $P(u,v)$ is admissible.

Without loss of generality,

$$f(x) = x(x^2 + 2ax + b) \quad \text{with} \quad 0 \leq a^2 < b .$$

Leaving $a = 0$ as a special case, assume $0 < a$. The line containing the fixed set is the line $v = 0$.

With the identifications

$$w = a(u + 2a) \quad \text{and} \quad c = a^2 ,$$

the Newton iteration can be written

$$N(u,v) = \begin{pmatrix} (w(w^2-cv-cb)-2a^2\Delta)/a\Delta \\ v(w^2-2cw+cb)/\Delta \end{pmatrix} = \begin{pmatrix} s_* \\ t \end{pmatrix}$$

where

$$\Delta = 2w^2 - 2cw - cv .$$

And, by performing the required transformation on s_*, namely

$$s = a(s_* + 2a) ,$$

the iteration is equivalent to the pair of homogeneous equations:

$$s = w(w^2-cv-cb)/\Delta$$

$$t = v(w^2-2cw+cb)/\Delta$$

with $0 < c < b$ and is defined except when $2w(w-c)-cv = 0$.

The line containing the fixed set is still the line $v = 0$. Also, $t = 0$ if and only if $v = 0$. For $v = 0$, $\Delta = 0$ only for $w = 0$ and $w = c$. Except at these two points, along $v = 0$,

$$s = (w^2 - cb)/2(w-c) = \psi(w) .$$

There are two points $(w,v) = (\alpha,0)$ and $(w,v) = (\beta,0)$, each of which iterates into the other. They are solutions of $\psi(\psi(w)) = w$ or $(w^2-2cw+cb)(3w^2-6cw+4c^2-cb) = 0$. Identify, then,

$$\alpha = c - \sqrt{(b-c)c/3} \qquad \text{and}$$

$$\beta = c + \sqrt{(b-c)c/3} .$$

The following lemma will be useful.

Lemma 1: For every $\gamma < \alpha$ and for every point $(w_0,0)$ not in $tr((0,0))$, there is an iterate $(w_i,0)$ in the (half-open) interval $\gamma < w \leq c$, $v = 0$.

The proof consists of showing that there is an iterate $(w_i,0)$ in the closed interval $\alpha \leq w \leq c$, $v = 0$. (Stipulate that $v = 0$ for the remainder of this proof.) The values α and β are repulsive fixed points of the iteration $\psi_2(w) = \psi(\psi(w))$.

Every point w_0 generates a w_i for which $\alpha \leq w_i \leq \beta$ so it is sufficient to establish the lemma for $c < w_0 < \beta$. This interval iterates, under ψ, into the interval $(-\infty,\alpha)$. By considering the iterates of subintervals of $(-\infty,\alpha)$, it is seen that there is a single point $\alpha_1 < \alpha$ such that $\psi(\alpha_1) = c$ and it is sufficient to establish the lemma for $\beta_1 < w_0 < \beta$ where $\psi(\beta_1) = \alpha_1 = c - \sqrt{c(b-c)}$.

Repetition of this process yields a sequence $\{\beta_i\}$ which converges to β and for each term of the sequence it is sufficient to establish the lemma for $\beta_i < w_0 < \beta$. That $\psi(\beta) = \alpha$ is sufficient to complete the proof.

For any w_0 such that $(w_0,0) \notin tr((0,0))$ and for each $\gamma < \alpha$, define the <u>representative</u> of w_0 with respect to γ as $\rho(w_0:\gamma)$ is the first iterate $(w_i,0)$ such that $\gamma < w_i \leq c$. If $w_0 \in tr((0,0))$, set $\rho(w_0:\gamma) = (0,0)$. It is apparent that the line $v = 0$ is stable at $(w_0,0)$ if and only if it is stable at $(\rho(w_0:\gamma),0)$.

4.3 The Trail of (0,0)

The investigation of the exceptional cases -- tr(0,0) -- proceeds by forming a "skeleton" of G_0. This construction uses the following information to establish the unique existence of certain arcs.

If $v < 0$, $cv < w^2-cb$, $w < 0$,

then $w < s < 0$, $v < t < 0$.

If $v < 0$, $w^2-cb < cv$, $w < 0$,

then $w < 0 < s$, $t < v < 0$.

If $v < 0$, $w^2-cb < cv$, $cv < 2w(w-c)$, $w < 0$,

then $s < 0 < w$, $t < v < 0$.

If $v < 0$, $cv < w^2-cb$, $w < 0$,

then $0 < s < w$, $v < t < 0$.

From the observation that for $v \neq 0$ and $w = 0$, $(s,t) = (0,-b)$ it can be seen that the arc $\Gamma_{0,0}$ defined next is in $cl(G_0)$. Let $\Gamma_{0,0}$ be the line segment described by $w = 0$, $0 \leq v \leq -b$. Call the endpoints of this arc B (the point $(0,-b)$) and $a_{0,0}$ (the point $(0,0)$). It is clear that if $N(w,v) = B$ then $w = 0$.

Let $\Gamma_{0,i+1}$ be the (unique) arc with the properties that $N(\Gamma_{0,i+1}) = \Gamma_{0,i}$ and $a_{0,i+1} < a_{0,i}$ where $N(a_{0,i+1}) = a_{0,i}$.

Let $\Gamma_{j+1,i}$ be the (unique) arc with the properties that $N(\Gamma_{j+1,i}) = \Gamma_{j,i}$ and $a_{j+1,i} < a_{j,i}$ where $N(a_{j+1,i}) = a_{j,i}$.

Since, for all i and j, $B \in \Gamma_{j,i}$, $\Gamma_{j,i} \subset cl(G_0)$. The sequence $\{a_{1,i}\}$ converges to $(c,0)$. All of the points $a_{j,i}$ are in $tr((0,0))$. If $b \leq 2c$, $\{a_{j,i}\} = tr((0,0)) \cap cl(G_0)$. (The last assertion is the observation that for $w < 0$ and $b \leq 2c$, $\psi(w) < c$.)

By way of summary:

<u>Lemma 2</u>: The line $v = 0$ is not stable at any point in $tr((0,0))$. Every point in $\{a_{j,i}\}$ is in $cl(G_0)$. The point $(w,v) = (c,0)$ is in $cl(G_0)$. If $b \leq 2c$, $\{a_{j,i}\} = tr((0,0)) \cap cl(G_0)$.

4.4 Intervals of Stability

There is no isolated point at which the line $v = 0$ is stable. If there is an interval in which every point is a point of stability, then there is such an interval contained in (γ, c) where $\gamma < \alpha$.

For investigation of these intervals, first observe that for v small

$$t = v\phi(w) + \mathcal{O}(v^2)$$

where

$$\phi(w) = (w^2-2cw+cb)/(2w^2-2cw) .$$

So $(w,0)$ cannot be a point of stability for the line $v = 0$ if $|\phi(w)| > 1$. An easy calculation shows that there are no points at which $v = 0$ is a stable set if $4c \le 3b$. At the same time it is seen that, if there is any interval of stability, one must lie in (γ, δ) where

$$\gamma = (2c - \sqrt{c(4c-3b)}\,)/3$$

and

$$\delta = (2c + \sqrt{c(4c-3b)}\,)/3 .$$

Observe that

$$0 < \alpha_1 < \gamma < \alpha < \delta < c .$$

For $w = a(u+2a)$ define $M(w,v) = N(N(u,v))$. There is some interval (μ_1, ν_1) containing α such that $M(\mu_1, 0) = (\alpha_1, 0)$ and $M(\nu_1, 0) = (c, 0)$. The interval (μ_1, ν_1) cannot be a stable interval. Applying this argument to

(M(M(w,v)) and succeeding iterations, a nest of such intervals containing α is obtained. These intervals converge to $[\alpha]$ and none of them can be a stable interval. Thus, $(\alpha,0)$ is not a point of stability. If there is an interval of stability, there must be one contained in (α,δ) .

In the same manner that lemma 1 is proved, the following lemma can be established.

Lemma 3: For every point $(w_0,0)$ such that $\alpha < w_0 < \delta$, there is an iterate $(w_i,0)$ with $\delta \leq w_i \leq c$.

This establishes the following theorem.

Theorem: For the cubic $f(x) = (x-r)(x^2+2ax+b)$ with $0 < a^2 < b$, the fixed set Σ contained in the line $v = r^2 - ru$ contains no point of stability. There is no fixed set properly containing Σ in any sufficiently small ε-neighborhood of this line.

Boyd's indications [4] of possible stability of the line $v = 0$ appear to arise from the following:

(a) the "partial linearization" inadequately reflects the behavior of the iteration;

(b) any point in G_{k+1} which is not in G_k requires at least k steps to converge;

(c) unavoidable roundoff errors cause differences between the analytical and computational iterations.

V. DISCUSSION

5.1. Incidental Notes

5.1.1.

For many functions, a way of showing them to be admissible is to show that the hypotheses of the Kantorovich theorems are satisfied in some open neighborhood of each zero [1].

5.1.2.

The set $G_0(x_*)$ is, in fact, homeomorphic to an open ball. Certainly each of the sets T_i in the construction of $G_0(x_*)$ is homeomorphic to an open ball. To illustrate that the limiting process involved in taking the union of the T_i's does not create "holes", consider simple connectivity.

Each T_i is simply connected. Suppose $G_0(x_*)$ is not simply connected. Let γ be a loop (or cycle) which cannot be reduced to a point. Let γ_i be the portion of γ not belonging to T_i. The intersect $\cap\gamma_i$ is nonempty by the "nested-interval property" which yields a contradiction.

5.1.3.

Barna's proof of the theorem [2, 3] yields some additional information about the countability of those points not in G. His proof involves showing that, at every point not in $cl(G_1)$, the absolute value of N' is greater than 1. Then an appeal to a mean-value theorem shows that if Δ_0 is any interval on which Newton's method fails to converge, some iterate Δ_i has length greater than any prescribed value.

5.1.4.

Barna's proof that $Z \subset G_0$ involves an interpretation of the inflections of f(x) as local extrema of N(x).

5.1.5.

For $f(z) = z^4 - 2z^2 - 1$, Newton's method fails to converge in the complex plane starting from points near enough to $z = \sqrt{3}/3$.

5.1.6.

The study of the global convergence properties of Bairstow's method is complicated by the fact that there is more than one formulation of the method [6, 7]. The other common formulation takes the result of the Euclidean algorithm as

$$f(x) = p(x,u,v)q(x,u,v)+(x-u)F(u,v)+H(u,v)$$

and applies Newton's method to

$$Q(u,v) = \begin{pmatrix} F(u,v) \\ H(u,v) \end{pmatrix} .$$

For polynomials $f(x)$ for which both $P(u,v)$ and $Q(u,v)$ are admissible, the convergence properties near where $P(u,v) = Q(u,v) = 0$ do not depend on the formulation of Bairstow's method. The difference in global behavior is illustrated by the fact that, in the formulation described here, the crucial equation in Boyd's theorem (Equation 10, [4]) becomes

$$0 = (ru_1-ru_0+v_1-v_0)q(r,u_0,v_0)+(u_1-u_0)F(u_0,v_0) .$$

5.1.7.

Boyd's criterion that (Theorem 2 [4]) $\gamma < 1$ is equivalent to the condition that $3b < 4c$ (in the notation used here).

5.2. Summary

A construct has been developed which is useful in the investigation of the global convergence properties of Newton's method.

The usefulness of the construct has been demonstrated by establishing that Newton's method converges almost globally on a wider class of real polynomials than heretofore known. A conjecture is offered on the widest such class.

The construct has also been used as a tool to describe an exceptional set which arises in a study of Bairstow's method applied to a cubic polynomial.

BIBLIOGRAPHY

1. Anselone, P. M. (editor), Nonlinear Integral Equations (University of Wisconsin Press, 1964) Moore, R. H., "Newton's Method and Variations", pp. 65-98

2. Barna, Béla, "Über das Newtonsche Verfahren zur Annäherung von Wurzeln algebaischer Gleichungen", Publicationes Mathematicae (Debrecen) Vol. 2 (151-52) pp. 50-63

3. Barna, Béla, "Über die Divergenzpunkte des Newtonsche Verfahren zur Bestimmung von Wurzeln algebraischer Gleichungen I, II, III and IV", Publicationes Mathematicae (Debrecen) Vol. 3 (1953) pp. 109-118, Vol. 4 (1956) pp. 384-397, Vol. 8 (1961) pp. 193-207, Vol. 14 (1967) pp. 91-97

4. Boyd, David W., "Nonconvergence in Bairstow's Method", SIAM Journal on Numerical Analysis Vol. 14 (1977) pp. 571-574

5. Dieudonné, J., Foundations of Modern Analysis (Academic Press, 1960)

6. Isaacson, E. and Keller, H., Analysis of Numerical Methods (John Wiley & Sons, 1966)

7. Wilkinson, J., Rounding Errors in Algebraic Processes (Prentice-Hall, 1963)