Open access • Journal Article • DOI:10.1287/MNSC.22.5.576

Line Search Techniques Based on Interpolating Polynomials using Function Values Only — Source link

Arie Tamir

Institutions: Northwestern University

Published on: 01 Jan 1976 - Management Science (INFORMS)

Topics: Difference polynomials, Degree of a polynomial, Convergence tests, Polynomial and Rate of convergence

DISCUSSION PAPER NO. 79

LINE SEARCH TECHNIQUES BASED ON INTERPOLATING POLYNOMIALS USING FUNCTION VALUES ONLY

by

Arie Tamir*

April 9, 1974

* Northwestern University
Graduate School of Management

ABSTRACT

In this study we derive the order of convergence of some line search techniques based on fitting polynomials, using function values only. It is shown that the rate of convergence increases with the degree of the polynomial. If viewed as a sequence, the rates approach the Golden Section Ratio when the degree of the polynomial tends to infinity.

LINE SEARCH TECHNIQUES BASED ON INTERPOLATING POLYNOMIALS USING FUNCTION VALUES ONLY *

Introduction

Most efficient methods for unconstrained minimization utilize a one-dimensional search along directions generated by the method. If P is the function to be minimized, X the current vector of decision variables, and S the search direction, then the one-dimensional search problems is to choose $\alpha > 0$ yielding the first local minimum of $P(X + \alpha S)$. A significant portion of the total computational effort is expended in this search. The problem can be particularly difficult when P is an interior or exterior penalty function. This is a situation of great practical importance because penalty functions are widely used.

The most popular one-dimensional search procedures for use in unconstrained minimization utilize quadratic [2,7] or 2 point cubic [1,5,7] interpolation of P. When applied to penalty functions these interpolation approaches have serious deficiencies. Quadratic interpolation has the drawback that its order of convergence is approximately 1.3, significantly less than that of 2 point cubic interpolation, which is 2 [8]. The 2-point cubic, however, requires the computation of ∇P. This is usually time consuming and is often difficult to code. In some cases ∇P may not be available analytically.

A one-dimensional search based on quadratic and cubic interpolations using functions values only, is studied in [3]. The performance of this procedure on several test problems involving penalty functions has been significantly better than that of competing methods. The algorithm in [3] has motivated this study on order of convergence of related search techniques based on fitting polynomials, using functions values only.

*This research was partially supported by the Office of Naval Research under Contract No. N00014-67-A-0404-0010.

The algorithm studied in this paper is as follows.

Let x be a scalar variable, and $f(x)$ the function to be minimized, assumed differentiable. An isolated minimum of f is assumed to occur at α, where

$$f'(\alpha) = 0 \tag{1}$$

Let n be a fixed integer greater than 1. If $x_i, x_{i-1}, \ldots, x_{i-n}$ are $n+1$ approximations to α, and $P_n(x)$ is the unique polynomial of degree less than or equal to n which satisfies

$$P_n(x_{i-j}) = f(x_{i-j}) \ , \quad j = 0,1,\ldots,n \tag{2}$$

then the new approximation to α, x_{i+1}, is chosen to satisfy

$$P_n'(x_{i+1}) = 0 \quad . \tag{3}$$

The procedure is repeated, fitting the next polynomial to $x_{i+1}, x_i, \ldots, x_{i-(n-1)}$. This algorithm is henceforth referred to as the Sequential Polynomial Fitting Algorithm (SPFA).

We note that the SPFA is different from the algorithm discussed in [3] , in that the points through which the polynomial passes need not bracket a minimum of f. However, the bracketing algorithms do not lend themselves to the difference equation approach used in most convergence rate derivations. Further, the two procedures are closely related. Other authors [4], [8] give intuitive arguments that the rate of convergence of sequential and bracketing algorithms are the same, then proceed to analyze the convergence rate of the SPFA for the special case $n = 2$. We know of no proof that the convergence rates are the same, although the conjecture seems reasonable.

Convergence and Convergence Rates

In this work speed of convergence of line search methods is measured in terms of the following concepts. (See [8], [9]).

Definition

Let the sequence $\{e_k\}$ converge to 0. The rate of convergence of $\{e_k\}$ is defined as the supremum of the nonnegative numbers p satisfying

$$0 \le \overline{\lim_{k\to\infty}} \frac{|e_{k+1}|}{|e_k|^p} < \infty .$$

(The case o/o is regarded as finite). The average order of convergence is the infinum of the numbers $p > 1$ such that

$$\overline{\lim_{k\to\infty}} |e_k|^{1/p^k} = 1 .$$

The order is infinity if the equality holds for no $p > 1$.

Let

$$J = \{x| \ |x - \alpha| \le L\} \tag{4}$$

throughout this section, f is assumed to satisfy the following conditions. (The notation $f^{(i)}(x)$ denotes the i^{th} derivative of f).

Assumption 1.

1. $f^{(2)}(x) \ne 0$ for all $x \in J$. Note that this is equivalent to $f^{(2)}(x) > 0$ for all $x \in J$, since $f^{(2)}(\alpha) \ne 0$ implies $f^{(2)}(\alpha) > 0$.

2. $f^{(n+1)}(\alpha) \ne 0$.

3. $f^{(n+2)}$ is continuous on J.

4. If we define constants M_0, M_1 and M_2 such that, for all $x \in J$

$$|f^{(2)}(x)| \ge M_0, \ | f^{(n+1)}(x)/(n+1)!| \le M_1, \ | f^{(n+2)}(x)/(n+2)!| \le M_2 \tag{5}$$

then the interval width L in (4) is small enough to satisfy

$$\Gamma = L[\frac{2}{M_0}(M_2L + (n+1)\ M_1)]^{1/(n-1)} < 1 \tag{6}$$

and

$$\Gamma_1 = (1/M_0L)(M_2(2L)^{n+1} + M_1(n+1)(2L)^n) < 1/2 \tag{7}$$

We note that if the constants M_0, M_1 and M_2 are defined as the sharpest possible bounds for a given L, then M_1 and M_2 are nonincreasing in L and M_0 is non-decreasing in L. Since Γ and Γ_1 are decreasing functions of L, approaching zero as L approaches zero, an L satisfying (6 - 7) can always be found.

We also require the assumption that the convergence rate of the sequence $\{e_i\}$ is at least of order 1, where

$$e_i = x_i - \alpha\ . \tag{8}$$

Assumption 2*

$$\lim_{k\to\infty} \frac{e_{k+1}}{e_k} = \beta \tag{9}$$

where β is finite and is not a root of the polynomial

$$\Psi(x) = \prod_{j=1}^{n} (x^{j+1}-1) + \sum_{k=1}^{n} (x^{k+1}-x^k) \prod_{\substack{j=1\\ j\neq k}}^{n} (x^{j+1}-1)\ .$$

It is easy to verify that $\beta = 0$, i.e. superlinear convergence satisfies Assumption 2.

The main result of this section is

Theorem 1. Under assumptions 1 and 2, the order of convergence of the SPFA, using polynomials of degree n, is equal to the unique positive root, σ_n, of the polynomial

(*) The stronger version of this assumption, i.e., superlinear convergence, is also made for the case n = 2 in [8, p. 143].

$$C_n(x) = x^{n+1} - \sum_{j=1}^{n} x^{n-j} . \tag{10}$$

The sequence of roots, $\{\sigma_n\}$ is increasing, approaching the Golden Section Ratio $\tau = (1 + \sqrt{5})/2 = 1.618$ as n approaches infinity.

A table of positive roots of $C_n(x)$ is given below

n	root, σ_n	σ_n/τ
2	1.324	.81
3	1.465	.90
4	1.534	.94
5	1.570	.97
6	1.590	.98

Cubic polynomials (n=3) yield 90% of the maximum attainable convergence rate, and the ratio σ_n/τ increases slowly for $n > 3$. Given the added complexity of dealing with polynomials of degree greater than 3, there is little reason for considering such polynomials in practical interpolation schemes.

In the remainder of this section, we give a number of results leading to a proof of Theorem 1. The following two theorems, proved in appendix A, insure that the sequence $\{x_i\}$ is well defined, and converges to the minimal point α.

Theorem 2. Define $J = \{x \mid |x-\alpha| \leq L\}$ and suppose that α is the unique minimum of f in J. Let $x_i, x_{i-1}, \ldots, x_{i-n}$ in J define the polynomial $P_n(x)$ of degree $\leq n$ satisfying

$$P_n(x_{i-j}) = f(x_{i-j}) \qquad j = 0,1,2,\ldots,n .$$

If f and J satisfy Assumption 1 then $P_n'(x)$ has a real root in J.

Theorem 3. Suppose that the conditions of Theorem 2 hold and let x_{i+1} in J be a real root of the derivative of the interpolatory polynomial $P_n(x)$ determined by x_i, $x_{i-1},\ldots,x_{i-n}$. Then the sequence $\{x_k\}$ converges to α and

$$|x_k - \alpha| \le K\Gamma^{r(n,k)} \tag{11}$$

for some constant K. $\Gamma < 1$ (defined in (6)), and

$$r(n,k) = n^{k/(n+1)} . \tag{12}$$

Hence the sequence $\{e_k\}$ converges to zero with average order of convergence greater than or equal to $n^{1/(n+1)}$.

We now derive results on the (stepwise) order of convergence of the SPFA. In Appendix A, it is shown that

$$P_n'(x) = f'(x) - \frac{f^{(n+1)}(\xi(x))}{(n+1)!} \sum_{k=0}^{n} \prod_{\substack{j=0\\ j\neq k}}^{n} (x-x_{i-j}) - \frac{f^{(n+2)}(\eta(x))}{(n+2)!} \prod_{j=0}^{n} (x-x_{i-j}) \tag{13}$$

where $\xi(x)$ and $\eta(x)$ are in the interval determined by $x_i, x_{i-1},\ldots,x_{i-n},x$. Substituting $x = x_{i+1}$ into (13), and using the relations

$$P_n'(x_{i+1}) = 0 , \quad (x_{i+1}-x_{i-j}) = (e_{i+1}-e_{i-j})$$

and

$$f'(x_{i+1}) = e_{i+1} f^{(2)} (\theta(x_{i+1}))$$

where $\theta(x_{i+1})$ is in the interval $[x_{i+1},\alpha]$, yield

$$e_{i+1} f^{(2)}(\theta(x_{i+1})) = \frac{f^{(n+1)}(\xi(x_{i+1}))}{(n+1)!} \sum_{k=0}^{n} \prod_{\substack{j=0\\ j\neq k}}^{n} (e_{i+1}-e_{i-j}) + \frac{f^{(n+2)}(\eta(x_{i+1}))}{(n+2)!} \prod_{j=0}^{n} (e_{i+1}-e_{i-j}) .$$

Suppose that $e_{i-j} \neq 0$, $j = 0,1,\ldots,n$.

$$e_{i+1}f^{(2)}(\theta(x_{i+1})) = \prod_{j=1}^{n} e_{i-j} \left\{ \frac{f^{(n+1)}(\xi(x_{i+1}))}{(n+1)!} \left[\prod_{j=1}^{n} \left(\frac{e_{i+1}}{e_{i-j}} - 1\right) \right.\right.$$

$$\left. + \sum_{k=1}^{n} \left(\frac{e_{i+1}}{e_{i-k}} - \frac{e_i}{e_{i-k}}\right) \prod_{\substack{j=1 \\ j\neq k}}^{n} \left(\frac{e_{i+1}}{e_{i-j}} - 1\right) \right] + \frac{f^{(n+2)}(\eta(x_{i+1}))}{(n+2)!} (e_{i+1} - e_i) \,.$$

$$\left. \prod_{j=1}^{n} \left(\frac{e_{i+1}}{e_{i-j}} - 1\right) \right\} \tag{14}$$

By Assumption 2 (i.e. $e_{i+1}/e_i \to \beta$), the ratios e_{i+1}/e_{i-j} in (14) approach β^{j+1} as $i \to \infty$, while $(e_{i+1}-e_i)$ approaches zero. If we define A_{i+1} by

$$e_{i+1} = A_{i+1} \prod_{j=1}^{n} e_{i-j} \tag{15}$$

and let $i \to \infty$

$$A_{i+1} \to \Psi(\beta)\ f^{(n+1)}(\alpha)/(n+1)!\ f^{(2)}(\alpha) \equiv A \,. \tag{16}$$

By Assumption 1 (conditions 1, 2) and Assumption 2, $A \neq 0$, so by (15), for i sufficiently large, $e_{i-j} \neq 0$, $j = 0,1,\ldots,n$ implies $e_k \neq 0$ for all $k \geq i$. For i sufficiently large and $\epsilon > 0$, (15) yields

$$(|A|-\epsilon) \prod_{j=1}^{n} |e_{i-j}| \leq |e_{i+1}| \leq (|A| + \epsilon) \prod_{j=1}^{n} |e_{i-j}| \tag{17}$$

or, defining

$$\delta_i = |e_i|(|A| + \epsilon)^{1/(n-1)} \,, \quad \gamma_i = |e_i|(|A| - \epsilon)^{1/(n-1)}$$

$$\delta_{i+1} \leq \prod_{j=1}^{n} \delta_{i-j} \,, \quad \gamma_{i+1} \geq \prod_{j=1}^{n} \gamma_{i-j} \tag{18}$$

Since $e_i \neq 0$ for i sufficiently large, we may take logs of the inequalities (18) yielding the difference inequalities

$$d_{i+1} \leq \sum_{j=1}^{n} d_{i-j} \,, \quad c_{i+1} \geq \sum_{j=1}^{n} c_{i-j} \tag{19}$$

where

$$d_i = \ln \delta_i \,, \quad c_i = \ln \gamma_i \,. \tag{20}$$

We apply the following theorem, due to Ostrowski [10, P. 98], to (19).

Lemma 1.

Consider the equation

$$x^n - \sum_{j=0}^{n-1} \rho_j x^j = 0$$

with $\rho_j \geq 0$, $j = 0,\dots,n-1$, having positive root σ, and the infinite sequence $\{u_i\}$ satisfying the difference inequality

$$u_{i+n} - \sum_{j=0}^{n-1} \rho_j u_{i+j} \geq 0 \ , \quad i = 1,2,\dots \tag{21}$$

where $u_1,\dots,u_n$ are positive. Then we have

$$u_i \geq \gamma \sigma^i \qquad i = 1,2,\dots$$

where

$$\gamma = \min_{1 \leq j \leq n} \frac{u_j}{\sigma^j} > 0 \ .$$

By following Ostrowski's proof, it is easy to verify that, if the reverse inequality holds in (21) then

$$u_i \leq \delta \sigma^i \ , \ \delta = \max_{1 \leq j \leq n} \frac{u_j}{\sigma^j} \geq \gamma$$

Since $|e_i| \to 0$ we can assume without loss of generality that $d_1,\dots,d_{n+1}$ and $c_1,\dots,c_{n+1}$ are negative. Then, applying lemma 1 to the sequences $\{-c_i\}$, $\{-d_i\}$ yields

$$-d_i \geq \gamma \sigma^i \ , \ -c_i \leq \delta \sigma^i$$

where

$$\gamma = \min_{1 \leq j \leq n+1} \frac{-d_j}{\sigma^j} \ , \quad \delta = \max_{1 \leq j \leq n+1} \frac{-c_j}{\sigma^j} \ ,$$

and σ is the unique positive root of the polynomial $C_n(x)$ in (10) as shown in appendix B. Thus, using (20) and (18)

$$(|A|-\varepsilon)^{-1/(n-1)} \exp(-\delta\sigma^i) \leq |e_i| \leq (|A|+\varepsilon)^{-1/(n-1)} \exp(-\gamma\sigma^i) \tag{22}$$

Hence

$$\frac{|e_{i+1}|}{|e_i|^t} \leq g_1 \exp\{\sigma^i(\delta t - \gamma\sigma)\} \quad , \quad g_1 < \infty \tag{23}$$

In appendix B we show that $\sigma > 1$. Hence the right hand side of (23) is finite for all i if $(\delta t - \gamma\sigma) \leq 0$, i.e., if $t \leq \gamma\sigma/\delta$. Again using (22) we obtain

$$\frac{|e_{i+1}|}{|e_i|^t} \geq g_2 \exp\{\sigma^i(\gamma t - \delta\sigma)\} \quad , \quad g_2 > 0$$

which approaches infinity as $i \to \infty$ if $\gamma t - \delta\sigma > 0$, i.e., if $t > \delta\sigma/\gamma$. Hence the order of convergence of the SPFA is less than or equal to $\delta\sigma/\gamma$ and greater than or equal to $\gamma\sigma/\delta$.

To show that the order of convergence is exactly σ, we use the following lemma [10, P. 92].

Lemma 2.

Consider the linear difference equation

$$u_{i+1} = k_{i+1} + \sum_{j=0}^{n} a_j u_{i-j} \quad , \qquad i = n, n+1, \ldots$$

where the a_j are constants and $\{k_i\}$ is a specified sequence. The associated characteristic polynomial is

$$Q(x) = x^{n+1} - \sum_{j=0}^{n} a_j x^{n-j} .$$

Let $r_1, \ldots, r_{n+1}$ be the roots of $Q(x)$, with $|r_1| \geq |r_2| \geq \ldots \geq |r_{n+1}|$.

Assume that $|r_1| > 1 > |r_2|$ and, for some s, $0 < s < |r_1|$

$$k_i = 0(s^i)$$

which means $|k_i|/s^i \to c$ for some constant c as $i \to \infty$. Then there exists α_1 such that, as $i \to \infty$

$$\frac{u_i}{r_1^i} \to \alpha_1 .$$

In addition, if $s > |r_2|$

$$u_i = \alpha_1 r_1^i + O(s^i) \quad .$$

If $s = |r_2|$ and m is the maximum multiplicity of all zeros of $Q(x)$ with modulus $|r_2|$ then

$$u_i = \alpha_1 r_1^i + O(i^m |r_2|^i) \ .$$

A careful examination of the proof in [10] shows that Lemma 2 is true even if the condition $|r_1| > 1 > |r_2|$ is replaced by the weaker condition

$$|r_1| > 1, |r_1| > |r_2| \quad .$$

Taking absolute values and logs of (15), and defining

$$d_i = \ell n\ |e_i|,\ B_i = \ell n\ |A_i|$$

we obtain

$$d_{i+1} = B_{i+1} + \sum_{j=1}^{n} d_{i-j} \ , \qquad i = n, n+1, \ldots \ .$$

Further defining

$$u_i = \frac{d_i}{\ell n|A|+S} \ , \qquad k_i = \frac{B_i}{\ell n|A|+S}$$

where $S = -1$ if $|A| < 1$ and $S = 1$ otherwise, yields

$$u_{i+1} = k_{i+1} + \sum_{j=1}^{n} u_{i-j} \ , \qquad i = n, n+1, \ldots \qquad (24)$$

where, for i sufficiently large

$$|k_{i+1}| < 1 \ . \qquad (25)$$

The characteristic polynomial of (24) is $C_n(x)$ in (10). Consider first the case where n+1 is odd. It is shown in appendix B that, in this case, the roots of $C_n(x)$ satisfy $|r_1| > 1 > |r_2|$. By (25), we can apply Lemma 2 with $s = 1$ to obtain

$$u_i = \alpha_1 r_1^i + O(1)$$

implying

$$|e_i| = \exp\{-\beta_1 r_1^i + 0(1)\}$$

where $\beta_1 > 0$ since $|e_i| \to 0$. This implies that

$$\frac{|e_{i+1}|}{|e_i|^t} = \exp\{\beta_1 r_1^i(t-r_1) + 0_1(1) + t0_2(1)\}$$

which implies that the order of convergence of the sequence $\{e_i\}$ is r_1. Suppose now that n+1 is even. Then, from appendix B, $r_1 > 1$ and $r_2 = -1$. The comment following Lemma 2 justifies its use in this circumstance and, using $s = |r_2| = 1$ we obtain

$$u_i = \alpha_1 r_1^i + 0(i^m |r_2|^i) \ .$$

As shown in appendix B, m = 1, so

$$u_i = \alpha_1 r_1^i + 0(i)$$

which implies

$$|e_i| = \exp\{\gamma_1 r_1^i + 0(i)\} \ . \tag{26}$$

Since $|e_i| \to 0$, $\gamma_1 \leq 0$. If $\gamma_1 = 0$ then $|e_i| = \exp\{0(i)\}$, which contradicts (11). Hence $\gamma_1 < 0$. It is then easily verified that (26) implies that the order of convergence of the sequence $\{e_i\}$ is again r_1. Theorem 1 follows from the preceding discussion and Appendix B.

A concluding remark is in order. The above discussion depends substantially on the assumption that $f^{(2)}(\alpha) \neq 0$. In fact we can weaken this assumption as follows. Suppose that $f^{(r)}(\alpha) = 0$, $r = 1,\ldots,k-1$, and $f^{(k)}(\alpha) \neq 0$, where $n+1 > k \geq 2$. The minimality of α implies that k is even and $f^{(k)}(\alpha) > 0$. It is easy to verify that Theorem 2 is still valid if M_0 is the minimum of $f^{(k)}(\alpha)$ on J and (7) is replaced by

$$[(k-1)!/M_0L^{k-1}]\ (M_2(2L)^{n+1} + M_1(n+1)(2L)^n) < \frac{1}{2}\ .$$

Theorem 1 is also valid if σ_n is replaced by θ_n where $\theta_n > 1$ and is the unique positive root of the polynomial

$$(k-1)x^{n+1} - x^{n-1} - x^{n-2} - \dots - 1\ .$$

The sequence of roots $\{\theta_n\}$ is increasing and converges to $\theta = \dfrac{1 + \sqrt{1 + \frac{4}{k-1}}}{2}$.

We also note that, τ, the bound on the convergence rates of interpolating polynomials using function values only, is easily exceeded when derivative values are incorporated to define the polynomial. For example, the quadratic obtained using the False Position method converges with rate equal to the Golden Section Ratio τ, (see [8]). This method utilizes values of the function and its first derivative only. Newton's method uses second derivatives and has rate equal to 2.

Acknowledgement

The author is indebted to Professor L.S. Lasdon for his suggestions and assistance during the research work.

Appendix A

Existence Theorem of a Zero of the Derivative of the Interpolation Polynomial

In this appendix we prove Theorems 2 and 3, assuring that the sequence of roots $\{x_i\}$, generated by the algorithm, is well defined in the neighborhood of α, and coverges to α.

Proof of Theorem 2. Since $f^{(n+1)}(x)$ is continuous it is well known (e.g. [12, P. 61]) that

$$f(x) = P_n(x) + \frac{f^{(n+1)}(\xi(x))}{(n+1)!} \prod_{j=0}^{n} (x-x_{i-j}) \tag{A.1}$$

where $\xi(x)$ lies in the interval determined by $x_i, x_{i-1}, \ldots, x_{i-n}, x$. To derive an expression for $P_n'(x)$ we apply a result due to Ralston [11], which states that

$$\frac{1}{(n+1)!} \frac{d}{dx} f^{(n+1)}(\xi(x)) = \frac{1}{(n+2)!} f^{(n+2)}(\eta(x)) \tag{A.2}$$

where $\eta(x)$ is again a mean value in the interval of interpolation. Differentiating (A.1) and using (A.2) yield

$$P_n'(x) = f'(x) - \frac{f^{(n+1}(\xi(x))}{(n+1)!} \sum_{k=0}^{n} \prod_{\substack{j=0 \\ j\neq k}}^{n} (x-x_{i-j}) - \frac{f^{(n+2)}(\eta(x))}{(n+2)!} \prod_{j=0}^{n} (x-x_{i-j}) \tag{A.3}$$

We now show that under the assumptions of the theorem $P_n'(x)$ has a zero in J. Note first that $f^{(2)}(x) > 0 \;\; \forall \;\; x \in J$ since α is a minimum point and hence $f^{(2)}(\alpha) \geq 0$. The theorem follows when we prove that $P_n'(\alpha-L) < 0$ and $P_n'(\alpha+L) > 0$. $f'(\alpha) = 0$ implies

$$f'(x) = f'(x) - f'(\alpha) = (x-\alpha)\, f^{(2)}(\gamma(x))$$

where $\gamma(x)$ is in J.

Substituting $x = \alpha - L$ in (A.3) yields

$$P_n'(\alpha-L) = -L\, f^{(2)}(\gamma(\alpha-L)) - \frac{f^{(n+1)}(\xi(\alpha-L))}{(n+1)!} \sum_{k=0}^{n} \prod_{\substack{j=0\\ j\neq k}}^{n} (\alpha-L-x_{i-j})$$

$$- \frac{f^{(n+2)}(\eta(\alpha-L))}{(n+2)!} \prod_{j=0}^{n} (\alpha-L-x_{i-j})$$

$P_n'(\alpha-L)$ is negative if

$$T = \frac{1}{Lf^{(2)}(\gamma(\alpha-L))}\left[-\frac{f^{(n+1)}(\xi(\alpha-L))}{n+1)!}\sum_{k=0}^{n}\prod_{j\neq k}(\alpha-L-x_{i-j}) - \frac{f^{(n+2)}(\eta(\alpha-L))}{(n+2)!}\prod_{j=0}^{n}(\alpha-L-x_{i-j})\right] < 1$$

$$\text{But} \quad T \leq |T| \leq \frac{M_2}{M_0}\frac{(2L)^{n+1}}{L} + \frac{M_1}{M_0}(n+1)\frac{(2L)^n}{L} < 1 \quad .$$

Similar arguments lead to the conclusion that $P_n'(\alpha+L) > 0$, and hence the theorem follows.

<u>Proof of Theorem 3</u>. Substituting $x = x_{i+1}$ in (A.3) we obtain

$$f'(x_{i+1}) = \frac{f^{(n+1)}(\theta_1)}{(n+1)!}\sum_{\ell=0}^{n}\prod_{\substack{j=0\\ j\neq \ell}}^{n}(x_{i+1} - x_{i-j}) + \frac{f^{(n+2)}(\theta_2)}{(n+2)!}\prod_{j=0}^{n}(x_{i+1} - x_{i-j})$$

where

$$\theta_1 = \xi(x_{i+1}) \,,\ \theta_2 = \eta(x_{i+1}) \ .$$

Defining $e_k = x_k - \alpha$, $k = 1,2,\ldots$ and noting that

$$f'(x_{i+1}) = e_{i+1}\, f^{(2)}(\theta_3) \ ,\quad \theta_3 = \gamma(x_{i+1})$$

yield

$$M_0|e_{i+1}| \leq M_1 \sum_{\ell=0}^{n}[|e_{i+1}|L^{n-1}(2^n-1) + \prod_{j\neq\ell}|e_{i-j}|] +$$

$$+ M_2 [|e_{i+1}|L^n(2^{n+1}-1) + \prod_{j=0}^{n}|e_{i-j}|] \quad .$$

Hence,

$$|e_{i+1}| \leq \left\{\frac{M_1}{M_0}\frac{(2L)^n}{L}(n+1) + \frac{M_2}{M_0}\frac{(2L)^{n+1}}{L}\right\}|e_{i+1}| + \frac{M_1}{M_0}(n+1)\max_{0\leq j\leq n}|e_{i-j}|^n + $$

$$+ \frac{M_2}{M_0}\max_{0\leq j\leq n}|e_{i-j}|^{n+1} .$$

By Assumption 1

$$\frac{M_1}{M_0}\frac{(2L)^n}{L}(n+1) + \frac{M_2}{M_0}\frac{(2L)^{n+1}}{L} < \frac{1}{2} .$$

Thus,

$$|e_{i+1}| \leq C \max_{0\leq j\leq n}|e_{i-j}|^n \tag{A.5}$$

where

$$C = 2\left(\frac{M_1}{M_0}(n+1) + \frac{M_2}{M_0}L\right) .$$

Define $d_{i+1} = |e_{i+1}|\ C^{1/(n-1)}$. Then (A.5) yields

$$d_{i+1} \leq \max_{0\leq j\leq n} d_{i-j}^n \tag{A.6}$$

We show that if $k = t(n+1) + \ell$, $t \geq 1$, $\ell = 0,1,\dots,n$ then (A.6) implies that $d_k \leq \Gamma^{n^t}$ where $\Gamma = L\,C^{1/(n-1)}$. The proof is by induction on k.

Let $t = 1$, $\ell = 0$ and consider d_{n+1}, then

$$d_{n+1} \leq \mathrm{Max}\{d_0^n, d_1^n, \dots, d_n^n\} \leq \Gamma^n .$$

Let $k = t(n+1) + \ell$ and suppose that the result holds for indices smaller than k. If $\ell = 0$, then

$$d_k = d_{t(n+1)} \leq \max_{0\leq \ell\leq n}[d_{(t-1)(n+1)+\ell}]^n \leq \Gamma^{n^{t-1}n} = \Gamma^{n^t} .$$

Let $\ell > 1$, then

$$d_k = d_{t(n+1)+\ell} \leq \mathrm{Max}\{\max_{0\leq j<\ell}[d^n_{t(n+1)+j}], \max_{\ell\leq j\leq n}[d^n_{(t-1)(n+1)+j}]\}$$

$$\leq \mathrm{Max}\{\Gamma^{n^t}, \Gamma^{n^{t+1}}\} \leq \Gamma^{n^t} .$$

Hence $d_k \leq \Gamma^{n^t}$ where $k = t(n+1) + \ell$, $t \geq 1$, $\ell = 0,1,\ldots,n$, so

$$|e_k| = d_k \ C^{-1/(n-1)} \leq C^{-1/(n-1)}\Gamma^{n^t} .$$

$t = \frac{k}{n+1} - \frac{\ell}{n+1}$ and $\Gamma < 1$, (Assumption 1), imply that

$$|e_k| \leq C^{-1/(n-1)} \Gamma^{r(n,k)}$$

and the theorem follows.

Appendix B

The Roots of the Indicial Equation

In this appendix we study the properties and roots of the polynomial

$$C_{k-1}(z) = z^k - z^{k-2} - z^{k-3} - \dots - 1 \;. \tag{B.1}$$

We will show that $C_{k-1}(z)$, $k \geq 3$, has a unique simple positive root, σ_{k-1}, with modulus > 1, and that all other roots are also simple with moduli less than or equal to 1. In fact, it will be proved that if k is odd σ_{k-1} is the only real root and that the other k-1 roots are inside the unit disc. If k is even $z = -1$ and σ_{k-1} are the only real roots and the other k-2 roots have moduli less than 1. It is also shown that the sequence $\{\sigma_k\}$, $k = 2,3,\dots$ is increasing and tends to the Golden Section ratio, τ, (i.e., $\tau^2-\tau-1=0$, $\tau>1$).

Lemma B.1.

Let $C_{k-1}(z)$, $k \geq 3$, be defined by (B.1). $C_{k-1}(z)$ has a unique simple positive root, σ_{k-1}, and $1 < \sigma_{k-1} < \tau$, where τ is the Golden Section ratio. If k is odd σ_{k-1} is the only real root, and if k is even $z = -1$ is the only other real root of $C_{k-1}(z)$ and is simple.

Proof.
$$C_{k-1}(z) = z^k - \frac{z^{k-1} - 1}{z - 1} = \frac{1}{z-1}[z^{k-1}(z^2-z-1)] + \frac{1}{z-1}$$

Let τ be the Golden Section ratio, i.e., $\tau^2 - \tau - 1 = 0$, $\tau > 1$.
$C_{k-1}(\tau) = \frac{\tau^{k-1}}{\tau-1}(\tau^2-\tau-1) + \frac{1}{\tau-1} = \frac{1}{\tau-1} > 0$. It is easy to verify that for $k \geq 3$, $C_{k-1}(1) < 0$, and hence there exists a positive root $1 < \sigma_{k-1} < \tau$. To see that σ_{k-1} is simple and also the unique positive root observe first that

$$C_{k-1}(z) = (z - \sigma_{k-1})(z^{k-1} + a_2 z^{k-2} + a_3 z^{k-3} + \ldots + a_k z^0)$$

$$a_2 = \sigma_{k-1}$$

$$a_i = \frac{1}{\sigma_{k-1}} (a_{i+1} + 1) \qquad k > i \geq 3$$

$$a_k = \frac{1}{\sigma_{k-1}} \quad .$$

Thus $a_i > 0$, $i = 2,3,\ldots,k$ and the result follows. Suppose that k is even, then $C_{k-1}(z) = (z+1)\ (z^{k-1} - z^{k-2} - z^{k-4} - \ldots - 1)$. Hence $z = -1$ is a simple root. It is easily verified that $C_{k-1}(t)$ is negative for $-1 < t \leq 0$ and positive for $t < -1$. Hence $z = -1$ is the unique nonpositive real root. Suppose now that $k \geq 3$ is odd.

$$C_{k-1}(z) = z^k - (z+1)(z^{k-3} + z^{k-1} + \ldots + 1)$$

$$= \frac{z^{k-1}}{z-1} (z^2 - z - 1) + \frac{1}{z-1} \ .$$

From the first expression we see that $C_{k-1}(t) < 0$ for $-1 \leq t \leq 0$. Now let $t < -1$. Clearly $t^2 - t - 1 > 0$ and $C_{k-1}(t) < 0$. Hence $C_{k-1}(t) < 0$ for all nonpositive t and σ_{k-1} is the unique real root.

Lemma B.2.

All the roots of $C_{k-1}(z)$ are simple.

Proof. Define $D_{k-1}(z) = (z-1)\ C_{k-1}(z) = z^k(z-1) - z^{k-1} + 1$.

If $z \neq 1$ is a multiple root of $C_{k-1}(z)$ it is also a multiple root of $D_{k-1}(z)$ and $D'_{k-1}(z) = 0$

$$(k+1)z^k - k\ z^{k-1} - (k-1)\ z^{k-2} = 0$$

$z = 0$ is not a root and we have

$$(k+1)\, z^2 - kz - (k-1) = 0$$

which implies that z is real. The preceding lemma assures that real roots are simple and the lemma follows.

The following lemma shows that the sequence $\{\sigma_k\}$ is an increasing one.

<u>Lemma B.3</u>.

$\{\sigma_k\}$, $k = 2,3,\ldots$ is an increasing sequence and $\lim_k \sigma_k = \tau$.

<u>Proof</u>. To show the monotonicity property we prove that $C_k(\sigma_{k-1}) < 0$. Lemma B.1 then assures that $\sigma_k > \sigma_{k-1}$.

$$C_k(z) = \frac{z^{k+1}(z-1) - z^k + 1}{z - 1} = z\left(C_{k-1}(z) - \frac{1}{z-1}\right) + \frac{1}{z-1}$$

$$C_k(\sigma_{k-1}) = \sigma_{k-1}\left(0 - \frac{1}{\sigma_{k-1} - 1}\right) + \frac{1}{\sigma_{k-1} - 1} = -1 \quad .$$

The sequence $\{\sigma_k\}$ is a bounded increasing sequence and hence $\lim_k \sigma_k = \beta$ exists.

$$\sigma_{k-1}^{k-1}\,(\sigma_{k-1}^2 - \sigma_{k-1} - 1) = -1\ , \quad 1 < \sigma_k < \tau$$

$$\Rightarrow \beta^2 - \beta - 1 = 0 \quad \text{and } \beta = \tau\text{, the Golden Section root.}$$

To prove that the $(k-1)$ roots of $C_{k-1}(z)$ that differ from σ_{k-1} have moduli ≤ 1, we introduce the following two results.

<u>Theorem B.1</u>. (Traub, [12, p. 51])

Let $f_k(z) = z^k - a(z^{k-1} + z^{k-2} + \ldots + 1)$, $ka > 1$ and $k \geq 2$. Then $f_k(z)$ has one positive simple root, γ_k, and $\max(1,a) < \gamma_k < 1+a$. All other roots are also simple with moduli less than 1.

Lemma B.4.

(Ostrowski, [10, p. 222]).

Let B be a closed region in the z-plane, the boundary of which consists of a finite number of regular arcs, and let f(z) and h(z) be regular on B. Assume that for no value of the real parameter t, running through the interval $a \leq t \leq b$, the function $f(z) + th(z)$ becomes zero on the boundary of B. Then the number N(t) of the zeroes of $f(z) + th(z)$ inside B is independent of t for $a \leq t \leq b$.

We are now ready to prove the main result.

Theorem B.2. If k is odd the k-1 roots of $\frac{C_{k-1}(z)}{z - \sigma_{k-1}}$ have moduli < 1. If k is even the k-2 roots of $\frac{c_{k-1}(z)}{(z-\sigma_{k-1})(z+1)}$ have moduli < 1.

Proof. Let $\varepsilon > 0$ be arbitrary small and $k \geq 3$ and consider the polynomial $C_{k-1}(z) - tz^{k-1}$ for $t \in [\varepsilon,1]$ where

$$C_{k-1}(z) = z^k - z^{k-2} - z^{k-3} - \ldots - 1 .$$

We show that $C_{k-1}(z) - t\,z^{k-1} \neq 0$ for all z in $\{z|\ |z| = 1\}$. Since $C_{k-1}(1) - t < 0 \quad \forall \quad t \in [\varepsilon,1]$ it is sufficient to show that $(z-1)\{C_{k-1}(z) - tz^{k-1}\} \neq 0$ for all $z \neq 1$ and $|z| = 1$. Suppose $(z-1)\{C_{k-1}(z) - t\,z^{k-1}\} = 0$ for some $z \neq 1$ and $|z| = 1$. Then

$$\{z^{k-1}[z^2 - z(t+1) - (1-t)] + 1\} = 0$$

$$|z^{k-1}|\,|z^2 - z(t+1) - (1-t)| = |-1| \Rightarrow |z^2 - z(t+1) - (1-t)| = 1 .$$

If $z = \cos\theta + i\sin\theta$, then

$$[\cos 2\theta - (t+1)\cos\theta - (1-t)]^2 + [\sin 2\theta - (t+1)\sin\theta]^2 = 1 ,$$

which yields

$$-2(1-t)\cos^2\theta - (t + t^2)\cos\theta + (1+t^2) + (1-t) = 0 .$$

Let $y = \cos\theta$ then it is clear that $y = 1$ is one root of the quadratic

$$2(1-t)y^2 + (t+t^2)y - (t^2-t+2) = 0 . \quad \text{(B.2)}$$

For $t = 1$, $y = 1$ is the only root and we obtain $\cos\theta = 1$ which contradicts the assumption $z \neq 1$. Let $t \in [\varepsilon,1)$, then the second root of (B.2), is $y(t) = \frac{-(t^2-t+2)}{2(1-t)}$, $y(t) < \frac{2t - 2}{2(1-t)} = -1$. Thus we have the contradiction $\cos\theta < -1$. Observing that for $t = 1$, $C_{k-1}(z) - tz^{k-1}$ yields the polynomial $f_k(z)$ with $a = 1$, discussed in Theorem B.1, we apply Lemma (B.4) to conclude that for any positive t arbitrarily close to zero the polynomial $C_{k-1}(z) - tz^{k-1}$ has $k-1$ roots <u>inside</u> the disc $\{z| \ |z| \leq 1\}$. Continuity arguments (see for example [12, Appx. A]) lead to the conclusion that $C_{k-1}(z)$ has $k-1$ roots in $\{z| \ |z| \leq 1\}$. By substituting $t = 0$ in (B.2) we easily verify that the only possible root of $C_{k-1}(z)$ on the boundary of the disc is $z = -1$ which is a root if and only if k is even. Hence the theorem is proved.