27
8-25
25 NTIS
IS-3684

USING THE ARGUMENT PRINCIPLE
IN A PROXIMITY TEST FOR ZEROS
OF REAL POLYNOMIALS

Charles W. Schelin

AMES LABORATORY, USERDA
IOWA STATE UNIVERSITY
AMES, IOWA

Date Transmitted: August 1975

PREPARED FOR THE U. S. ENERGY RESEARCH AND DEVELOPMENT ADMINISTRATION
UNDER CONTRACT W-7405-eng-82

MASTER

DISTRIBUTION OF THIS DOCUMENT UNLIMITED

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

USING THE ARGUMENT PRINCIPLE IN A PROXIMITY
TEST FOR ZEROS OF REAL POLYNOMIALS

Charles W. Schelin

Ames Laboratory, ERDA
Iowa State University
Ames, Iowa 50010

NOTICE
This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Energy Research and Development Administration, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness or usefulness of any information, apparatus, product or process disclosed, or represents that its use would not infringe privately owned rights.

Date Transmitted: August 1975

PREPARED FOR THE U.S. ENERGY RESEARCH AND
DEVELOPMENT ADMINISTRATION UNDER CONTRACT NO. W-7405-eng-82

DISTRIBUTION OF THIS DOCUMENT UNLIMITED

—NOTICE—

This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Energy Research and Development Administration, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product or process disclosed, or represents that its use would not infringe privately owned rights.

Available from: National Technical Information Service
U. S. Department of Commerce
P.O. Box 1553
Springfield, VA 22161

Price: Microfiche \$2.25

TABLE OF CONTENTS

	PAGE
ABSTRACT	iv
0. INTRODUCTION	1
1. CAUCHY INDICES AND THE ARGUMENT PRINCIPLE	2
2. DEVISING A STURM SEQUENCE	10
3. DIVIDING CHEBYSHEV POLYNOMIALS	18
4. COMPARISON WITH THE SCHUR-COHN TEST	26
5. REFERENCES	32

ABSTRACT

In this paper a technique for testing a given circle for zeros of a given real polynomial is presented. The test technique consists of an algorithm for the numerical evaluation of the argument principle. Using this algorithm the exact number of zeros, including multiplicities, possessed by any real polynomial in the unit circle may be determined. After casting the argument principle in terms of Cauchy indices, a Sturm sequence of Chebyshev polynomials, whose length is at most half the degree of the given polynomial, is constructed to determine these indices. The method is found to compare favorably with the Schur-Cohn test.

Distribution List

Chicago Patent Office	1
Oak Ridge	27
Ames Laboratory Library	5
Charles W. Schelin	<u>7</u>
Total	40

0. Introduction

Many of the iterative schemes used in solving polynomial equations, (e.g. Newton's method), are locally convergent and hence require sufficiently close starting values. Polyalgorithms, or composite algorithms, constructed to solve large classes of polynomial equations, usually have an initial stage, consisting of a search method and proximity test, in which these starting values are determined. (See [3], [4], or [13], p. 236). A frequently used search method, developed by Lehmer, [6], and modified by Stewart, [10], isolates polynomial zeros by systematically transforming suspect disks to the unit disk, then testing with the Schur-Cohn algorithm [2], [9]. The Schur-Cohn algorithm, which amounts to a clever application of Rouché's theorem, may be used to decide if a given polynomial has a zero in the unit disk, and in some cases may be used to determine the exact number of zeros a polynomial possesses in the unit disk. Knowledge of the precise number of zeros in a given disk may be useful in polynomial deflation as well as zero isolation. If a sufficiently small disk is found to contain k zeros, the center may be treated as a k -fold zero, yielding a reduced polynomial of degree k less, as opposed to one less, than the original polynomial.

In this paper a method for the numerical evaluation of the argument principle is proposed to determine the precise number of zeros a real polynomial possesses in the unit circle. The procedure involves the construction of a Sturm sequence of Chebyshev

polynomials of the second kind, starting with polynomials of degree $\lfloor n/2 \rfloor$ and $\lfloor n/2 \rfloor - 1$, or less, where n is the degree of the given polynomial. The number of computational steps and possible instabilities of this technique are compared to the corresponding steps and instabilities of the Schur-Cohn method, which itself is a consequence of the argument principle via Rouché's theorem. The technique presented here is found to require fewer multiplications and divisions in general, and fewer total arithmetic operations for polynomials of large degree, ($n \geq 20$).

In section 1, classical results are reviewed and modified, giving rise to a Sturm sequence which is devised in section 2. Computational details and instability indicators are examined in section 3, and finally, operation counts and a numerical example are presented in section 4.

1. Cauchy Indices and the Argument Principle

Upon review of the connection between Cauchy indices and the principle of the argument we find that the number of zeros a real polynomial possesses in the unit circle may be expressed as the Cauchy index of a rational function of Chebyshev polynomials on the interval $[-1, 1]$. For our purposes we consider the following specialized version of the argument principle.

THEOREM 1.1

If $\Gamma: z=z(\tau)$, $\alpha \leq \tau \leq \beta$, is a positively oriented Jordan curve containing no zeros of the polynomial f , and if φ is a continuous argument function, then the number, N , of zeros of f in the interior of Γ , including multiplicities, is given by

$$N = \frac{1}{2\pi} (\varphi[\beta] - \varphi[\alpha]), \quad (1.1)$$

where $\varphi[\tau]$ denotes $\varphi(f(z(\tau)))$.

In case a subdivision $\alpha = \tau_0 < \tau_1 < \dots < \tau_h = \beta$ of $[\alpha, \beta]$ can be found so that $|\varphi[\mu] - \varphi[\tau]| < \pi/2$ for all $\mu, \tau \in [\tau_{i-1}, \tau_i]$, for $i = 1, 2, \dots, h$, then a method due to Henrici [5], for computing winding numbers may be used to obtain (1.1). We are concerned with the case where $z(\theta) = e^{i\theta}$, $0 \leq \theta \leq 2\pi$, and $f(z) = \sum_{k=0}^n a_k z^k$, a_k real for $k = 0, 1, \dots, n$, thus

$$f(z(\theta)) = \sum_{k=0}^n a_k \cos k\theta + i \sum_{k=1}^n a_k \sin k\theta. \quad (1.2)$$

It is not clear how the required subdivision could be obtained in this case.

For any point $z(\tau)$ of the curve Γ ,

$$\tan \varphi(z(\tau)) = \operatorname{Im} z(\tau) / \operatorname{Re} z(\tau), \quad (1.3)$$

so if $\alpha < \tau \leq \beta$ and $|\varphi(z(\tau)) - \varphi(z(\alpha))| \geq \pi$, then there exists ξ , $\alpha \leq \xi \leq \tau$ with $\operatorname{Re} z(\xi) = 0$. i.e., ξ is a pole of $\tan \varphi(z(\tau))$.

Cauchy indices are used to describe changes in a rational function at its poles and thus in view of (1.3) may be used to record the increase of $\varphi(z(\tau))$ along a Jordan curve.

DEFINITION 1.1

Suppose g is a real function whose only singularities are poles. The left Cauchy index of g at a pole ξ , denoted $\text{LIg}(\xi)$, is given by

$$\text{LIg}(\xi) = \begin{cases} 1/2 & \text{if } \lim_{x \rightarrow \xi^-} g(x) = -\infty \\ -1/2 & \text{if } \lim_{x \rightarrow \xi^-} g(x) = \infty \end{cases} \quad (i)$$

The right Cauchy index of g at a pole ξ , denoted $\text{RIg}(\xi)$, is given by

$$\text{RIg}(\xi) = \begin{cases} 1/2 & \text{if } \lim_{x \rightarrow \xi^+} g(x) = \infty \\ -1/2 & \text{if } \lim_{x \rightarrow \xi^+} g(x) = -\infty \end{cases} \quad (ii)$$

The Cauchy index of g at the pole ξ is given by

$$\text{Ig}(\xi) = \text{LIg}(\xi) + \text{RIg}(\xi), \quad (iii)$$

and the Cauchy index of g over an interval $[\alpha, \beta]$, where $\alpha < \beta$ and neither α nor β is a pole of g , is defined as

$$I_{\alpha}^{\beta} g(x) = \sum \text{Ig}(\xi), \quad (iv)$$

the sum being taken over all poles ξ , $\alpha < \xi < \beta$.

As an immediate consequence of this definition, if $\alpha < \gamma < \beta$ and γ is not a pole of g , then

$$I_{\alpha}^{\beta} g(x) = I_{\alpha}^{\gamma} g(x) + I_{\gamma}^{\beta} g(x) \quad (1.4)$$

Another more subtle consequence, again reminiscent of the definite integral, is

$$I_{\alpha}^{\beta} g(x) = -I_{\beta}^{\alpha} g(x) . \quad (1.5)$$

To establish (1.5), make the change of variables $x = \alpha + \beta - y$. Then $I_{\alpha}^{\beta} g(x) = I_{\beta}^{\alpha} h(x)$, where $h(x) = g(\alpha + \beta - x)$. Now poles ξ of g are in one-to-one correspondence with poles $\eta = \alpha + \beta - \xi$ of h . For $\epsilon > 0$, $x = \xi + \epsilon > \xi$, $y = \alpha + \beta - \xi - \epsilon = \eta - \epsilon < \eta$, and so $\lim_{x \rightarrow \xi^{+}} g(x) = \lim_{y \rightarrow \eta^{-}} h(y)$. From (i) and (ii) above, we have $RIg(\xi) = -LIh(\eta)$. Similarly $LIg(\xi) = -RIh(\eta)$, thus (1.5) holds.

Note that for poles ξ of even order, $Ig(\xi) = 0$. Also, for the function $g(\theta) = \tan \theta$ whose poles are of the form $\xi = (2k+1)\pi/2$, k an integer, we have

$$LI \tan \xi = -1/2, \quad RI \tan \xi = -1/2, \quad \text{and thus } I \tan \xi = -1 . \quad (1.6)$$

With (1.6) in mind, we present the following result which may be viewed as a somewhat more general version of the usual Cauchy index theorem, (See Marden [7], p. 129).

THEOREM 1.2

Suppose $\Gamma: z=z(\tau)$, $\alpha \leq \tau \leq \beta$, is a closed curve and φ is a continuous argument function. Let $\varphi[\tau]$ denote $\varphi(z(\tau))$ for points of Γ and assume $\varphi[\alpha] \neq (2k+1)\pi/2$ for any integer k . Assume further that the set $V = \{\tau \in (\alpha, \beta) \mid \varphi[\tau] = (2k+1)\pi/2 \text{ for some integer } k\}$ is finite. Then

$$\pi I_{\alpha}^{\beta} \tan \varphi[\tau] = \varphi[\alpha] - \varphi[\beta] . \quad (1.7)$$

Proof

Let the elements of V be ordered as $\tau_1 < \tau_2 < \dots < \tau_s$.

If $s = 0$, then $I_\alpha^\beta \tan \varphi[\tau] = 0$ and $|\varphi[\alpha] - \varphi[\beta]| < \pi$ as φ is continuous. Γ is closed so $\varphi[\alpha] - \varphi[\beta] = 2j\pi$ for some integer j . Thus $j = 0$ and the theorem holds.

For $s > 0$, $|\varphi[\alpha] - \varphi[\tau_1]| < \pi$ and $|\varphi[\tau_s] - \varphi[\beta]| < \pi$ by the continuity of φ . The rays from the origin through $z(\tau_1)$ and $z(\tau_s)$ either coincide or form a straight angle. Including the ray from the origin through $z(\alpha) = z(\beta)$, (when this point is not itself the origin), we have a pair of angles which are either coincident or supplementary. Thus the quantity

$$Q = (\varphi[\alpha] - \varphi[\tau_1]) + (\varphi[\tau_s] - \varphi[\beta]), \quad (1.8)$$

assumes one of the three values $\pm \pi$, 0 . If $Q = -\pi$, then $\varphi[\alpha] < \varphi[\tau_1]$ and $\varphi[\tau_s] < \varphi[\beta]$. Again appealing to the continuity of φ , we find $\varphi[\tau] \rightarrow \varphi[\tau_1]^-$ as $\tau \rightarrow \tau_1^-$, and $\varphi[\tau] \rightarrow \varphi[\tau_s]^+$ as $\tau \rightarrow \tau_s^+$. Recalling (1.6), we have then for the case $Q = -\pi$,

$$Q = \pi(LI \tan \varphi[\tau_1] + RI \tan \varphi[\tau_s]). \quad (1.9)$$

Similar reasoning establishes (1.9) in the other cases as well. For if $Q = \pi$, apparently $\varphi[\alpha] > \varphi[\tau_1]$ and $\varphi[\tau_s] > \varphi[\beta]$, resulting in $\varphi[\tau] \rightarrow \varphi[\tau_1]^+$ as $\tau \rightarrow \tau_1^-$ and $\varphi[\tau] \rightarrow \varphi[\tau_s]^-$ as $\tau \rightarrow \tau_s^+$. Examining (1.6) we now find $LI \tan \varphi[\tau_1] = RI \tan \varphi[\tau_s] = 1/2$ and (1.9) is true. Likewise, should $Q = 0$, then $\varphi[\alpha] - \varphi[\tau_1] = -(\varphi[\tau_s] - \varphi[\beta])$ and $\varphi[\tau]$ approaches

$\varphi[\tau_1]$ and $\varphi[\tau_s]$ from the same side as $\tau \rightarrow \tau_1^-$ and $\tau \rightarrow \tau_s^+$, respectively. Thus $LI \tan \varphi[\tau_1] = -RI \tan \varphi[\tau_s]$, and (1.9) holds. Now if $s = 1$, $\varphi[\alpha] - \varphi[\beta] = Q$ and the theorem is true.

For $s \geq 2$, use (1.8) and write

$$\varphi[\alpha] - \varphi[\beta] = Q + \sum_{i=2}^s (\varphi[\tau_{i-1}] - \varphi[\tau_i]) . \quad (1.10)$$

Now for each i , $i = 2, \dots, s$, the quantity $(\varphi[\tau_{i-1}] - \varphi[\tau_i])$ takes on one of the three values $\pm \pi, 0$. Considering cases, as above, it is an easy matter to show that

$$\varphi[\tau_{i-1}] - \varphi[\tau_i] = \pi(RI \tan \varphi[\tau_{i-1}] + LI \tan \varphi[\tau_i]) \quad (1.11)$$

for $i = 2, 3, \dots, s$. Combining (1.8), (1.9), (1.10), and (1.11), we have

$$\begin{aligned} \varphi[\alpha] - \varphi[\beta] &= Q + \sum_{i=2}^s (\varphi[\tau_{i-1}] - \varphi[\tau_i]) \\ &= Q + \pi \sum_{i=2}^s (RI \tan \varphi[\tau_{i-1}] + LI \tan \varphi[\tau_i]) \\ &= Q + \pi(RI \tan \varphi[\tau_1] + LI \tan \varphi[\tau_s]) \\ &\quad + \pi \sum_{i=2}^{s-1} (LI \tan \varphi[\tau_i] + RI \tan \varphi[\tau_i]) \\ &= \pi \sum_{i=1}^s I \tan \varphi[\tau_i] = \pi I_{\alpha}^{\beta} \tan \varphi[\tau] . \end{aligned}$$

As a consequence of theorems 1.1 and 1.2 we obtain the following result.

THEOREM 1.3

If $f(z) = \sum_{k=0}^n a_k z^k$, with a_k real, $k = 0, 1, \dots, n$, and if
 if $f(z) \neq 0$ whenever $|z| = 1$, then the number, N , of zeros of f
 inside the unit circle, including multiplicities, is given by

$$N = -I_0^\pi \frac{v(\theta)}{u(\theta)}, \quad (1.12)$$

where $u(\theta) = \sum_{k=0}^n a_k \cos k\theta$ and $v(\theta) = \sum_{k=1}^n a_k \sin k\theta$.

Proof

We immediately have

$$\begin{aligned} N &= \frac{1}{2\pi} (\varphi[\beta] - \varphi[\alpha]) = \frac{1}{2\pi} (-\pi I_0^{2\pi} \tan \varphi[\theta]) \\ &= -\frac{1}{2} I_0^{2\pi} \frac{v(\theta)}{u(\theta)}. \end{aligned}$$

Now, π is not a singularity of $\frac{v(\theta)}{u(\theta)}$ as

$$u(\pi) = \sum_{k=0}^n a_k \cos k\pi = \sum_{k=0}^n a_k (-1)^k = f(-1) \neq 0,$$

so we may write $I_0^{2\pi} \frac{v(\theta)}{u(\theta)} = I_0^\pi \frac{v(\theta)}{u(\theta)} + I_\pi^{2\pi} \frac{v(\theta)}{u(\theta)}$. Using trigonometric

addition formulas, $u(2\pi - \gamma) = u(\gamma)$ and $v(2\pi - \gamma) = -v(\gamma)$. Thus

replacing θ with $2\pi - \gamma$ we find $I_\pi^{2\pi} \frac{v(\theta)}{u(\theta)} = I_\pi^0 \frac{-v(\gamma)}{u(\gamma)} = -I_0^\pi \frac{v(\gamma)}{u(\gamma)}$,

the last equality following from (1.5). Hence $I_0^{2\pi} \frac{v(\theta)}{u(\theta)} = 2I_0^\pi \frac{v(\theta)}{u(\theta)}$

and the theorem holds.

We conclude this section by stating a classical result which may be used to compute the Cauchy index of a function over an interval, namely Sturm's theorem. The proof in Henrici [5] may be easily extended to this slightly more general case.

THEOREM 1.4

Suppose $f_0, f_1, \dots, f_m$ is a sequence of real continuous functions on the interval $[\alpha, \beta]$ satisfying:

$$f_0(\alpha)f_0(\beta)f_m(\alpha)f_m(\beta) \neq 0, \quad (i)$$

and

$$\text{if } 1 < j < m \text{ and } f_j(\xi) = 0 \text{ for some } \xi \in [\alpha, \beta], \quad (ii)$$

$$\text{then } f_{j-1}(\xi)f_{j+1}(\xi) < 0 .$$

For $x \in [\alpha, \beta]$, let $\Psi(x)$ denote the number of variations in sign in the sequence $f_0(x), f_1(x), \dots, f_m(x)$. Then

$$\Psi(\alpha) - \Psi(\beta) = I_{\alpha}^{\beta} \frac{f_1(x)}{f_0(x)} + I_{\alpha}^{\beta} \frac{f_{m-1}(x)}{f_m(x)} . \quad (1.13)$$

A sequence satisfying (i) and (ii) of theorem 1.4 is termed a Sturm sequence. When f_m is of constant sign on $[\alpha, \beta]$, (typically a non-zero constant function), then $I_{\alpha}^{\beta} \frac{f_{m-1}(x)}{f_m(x)} = 0$, simplifying (1.13).

2. Devising a Sturm Sequence

Given polynomials f_0 and f_1 , a Sturm sequence may be formed using the Euclid algorithm for greatest common divisors. Divide f_0 by f_1 and take f_2 to be the negative of the resulting remainder. Proceeding in this fashion the degree of f_j is always less than that of f_{j-1} and ultimately a constant function is obtained. Should the zero function be obtained, all previous functions possess a common factor which may be removed by division, the resulting sequence being a Sturm sequence whose last term is a non-zero constant. By modifying (1.12) we find starting functions f_0 and f_1 , which are linear combinations of Chebyshev polynomials. Then a Sturm sequence is formed using an analogue of Euclid's algorithm.

We begin by noting that

$$I_0^\pi \frac{v(\theta)}{\sin \theta u(\theta)} = I_0^\pi \frac{v(\theta)}{u(\theta)}, \quad (2.1)$$

where $u(\theta)$ and $v(\theta)$ are as in theorem 1.3. First, $\sin \theta > 0$ on $(0, \pi)$ so the two ratios have the same poles and signs in the open interval. Secondly, poles are not introduced at the endpoints as

$$\frac{v(\theta)}{\sin \theta} = \sum_{k=1}^n a_k \frac{\sin k\theta}{\sin \theta}, \quad (2.2)$$

which takes on the values $\sum_{k=1}^n k a_k$ and $\sum_{k=1}^n (-1)^{k+1} k a_k$, at 0 and π ,

respectively. Now, the ratio $\frac{\sin k\theta}{\sin \theta}$ in (2.2) amounts to $U_{k-1}(\cos \theta)$,

U_k being the k th Chebyshev polynomial of the second kind [8]. Furthermore, $\cos k\theta = T_k(\cos \theta)$ where T_k is the k th Chebyshev polynomial of the first kind. Combining (1.12) and (2.1), and using (1.5) with the change of variables $x = \cos \theta$ we have:

THEOREM 2.1

If $f(z) = \sum_{k=0}^n a_k z^k$, with a_k real, $k = 0, 1, 2, \dots, n$, and if $f(z) \neq 0$ whenever $|z| = 1$, then the number, N , of zeros of f inside the unit circle, including multiplicities, is given by

$$N = \frac{1}{2\pi} \int_{-1}^1 \frac{f_1(x)}{f_0(x)} dx, \quad (2.3)$$

where

$$f_0(x) = \sum_{k=0}^n a_k T_k(x), \text{ and } f_1(x) = \sum_{k=1}^n a_k U_{k-1}(x), \quad (2.4)$$

T_k and U_k being Chebyshev polynomials of the first and second kind, respectively.

Throughout the two phase construction which follows, we take f , f_0 , and f_1 , to be defined as in theorem 2.1. In the first phase the functions $f_2, f_3, \dots, f_s, f_{s+1}$, where $s = [n/2]$, are formally defined but coefficients are not actually computed. The variations in sign of this sequence at ± 1 can be determined from $f_0(\pm 1)$ and $f_1(\pm 1)$. Each step of phase 2 amounts to a division of polynomials in the second Chebyshev form.

For the first phase we need the following multiplication formulas, which are easily obtained from trigonometric addition formulas.

Taking $U_{-1} \equiv 0$, we have

$$T_1 U_{k-1} = T_k + U_{k-2}, \text{ for } k \geq 1, \quad (2.5)$$

and

$$U_1 U_{k-j} = U_{k-j+1} + U_{k-j-1}, \text{ for } k \geq j. \quad (2.6)$$

With (2.5) in mind define f_2 according to

$$f_2 = T_1 f_1 - f_0. \quad (2.7)$$

$$\text{Then } f_2(x) = \sum_{k=1}^n a_k [T_k(x) + U_{k-2}(x)] - \sum_{k=0}^n a_k T_k(x)$$

$$= \sum_{k=1}^n a_k U_{k-2}(x) - a_0 T_0(x)$$

$$= (a_2 - a_0) U_0(x) + \sum_{k=3}^n a_k U_{k-2}(x),$$

as $U_0(x) = T_0(x) \equiv 1$, and $U_{-1}(x) \equiv 0$.

All conditions for a Sturm sequence, except perhaps $f_2(\pm 1) \neq 0$, are satisfied by the functions f_0 , f_1 , and f_2 , on $[-1, 1]$. First

$$f_0(\pm 1) = \sum_{k=0}^n a_k T_k(\pm 1) = f(\pm 1) \text{ which is non-zero by assumption. The}$$

second equality depends on the fact that $T_k(1) = 1$, and $T_k(-1) = (-1)^k$, for all non-negative integers k . Secondly, should $f_1(\xi) = 0$, (2.7) insures that $f_2(\xi)f_0(\xi) = -f_0^2(\xi) \leq 0$. Equality cannot hold here or else we have a quantity ξ , $-1 < \xi < 1$, where $f_0(\xi) = f_1(\xi) = 0$. This would give rise to a value θ , $0 < \theta < 2\pi$, where $f(e^{i\theta}) = u(\theta) + iv(\theta) = 0$, contrary to hypothesis.

For $j = 2, 3, \dots, s$, where $s = [n/2]$, define

$$f_{j+1} = U_1 f_j - f_{j-1}. \quad (2.8)$$

Using (2.6) successively for each j , we find that

$$f_j(x) = \sum_{k=j}^n a_k U_{k-j}(x) - \sum_{k=0}^{j-2} a_k U_{j-2-k}(x) \quad (2.9)$$

or

$$f_j(x) = \sum_{k=0}^{j-2} [a_{j+k} - a_{j-k-2}] U_k(x) + \sum_{k=j-1}^{n-j} a_{j+k} U_k(x), \quad (2.10)$$

for $j = 2, 3, \dots, s+1$. As before, except perhaps for the condition that $f_{s+1}(\pm 1) \neq 0$, the functions $f_0, f_1, \dots, f_{s+1}$ form a Sturm sequence on $[-1, 1]$.

Examination of (2.10) reveals that the leading $n-2j$ coefficients of f_j are not altered in passing to f_{j+1} . The leading coefficients $a_n, a_{n-1}, \dots, a_j$, merely slide to the next Chebyshev polynomial of lower degree, while the lower order coefficients, $a_0, \dots, a_{j-2}$, now

occurring with negative signs, slide to the next Chebyshev polynomial of higher degree. Definition (2.8), if used for $j = s+1$, produces a polynomial f_{s+2} whose degree is larger than $n - [n/2] - 1$, an upper bound on the degree of f_{s+1} . Clearly this is a poor direction to take if we are to end our sequence with a constant polynomial.

Before proceeding to phase two of the construction we determine the sign variations which have occurred thus far. If $\Psi_1(x)$ represents the number of sign variations in the sequence $f_0(x), f_1(x), \dots, f_{s+1}(x)$, while $\Psi_2(x)$ denotes the sign variations in the sequence $f_{s+1}(x), f_{s+2}(x), f_{s+3}, \dots, f_m$, (f_m a constant function), to be determined in phase two, then the number, N , of zeros of f in the unit circle may be written as

$$N = [\Psi_1(-1) - \Psi_1(1)] + [\Psi_2(-1) - \Psi_2(1)]. \quad (2.11)$$

The quantity $[\Psi_1(-1) - \Psi_1(1)]$ may be quickly obtained from the following result.

THEOREM 2.2

Assuming f_0 and f_1 are as in (2.4), the functions $f_2, f_3, \dots, f_{s+1}$, where $s = [n/2]$, are defined by (2.7) and (2.8), and $\Psi_1(x)$ is as defined above, let

$$L_0 = \frac{f_1(-1)}{f_0(-1)} = \frac{\sum_{k=1}^n (-1)^{k+1} k a_k}{\sum_{k=0}^n (-1)^k a_k} \quad (2.12)$$

and

$$R_0 = \frac{f_1(1)}{f_0(1)} = \frac{\sum_{k=1}^n k a_k}{\sum_{k=0}^n a_k}. \quad (2.13)$$

Then

$$\psi_1(-1) = \begin{cases} s+1, & \text{if } L_0 < -s, \\ s, & \text{if } L_0 = -s, \text{ or } L_0 > 0, \text{ or } L_0 \text{ is} \\ & \text{a non-integer in } (-s, 0). \\ s-1, & \text{if } L_0 \text{ is an integer in } [1-s, 0]. \end{cases} \quad (2.14)$$

and

$$\psi_1(1) = \begin{cases} 1, & \text{if } R_0 < s \text{ but } R_0 \text{ is not a} \\ & \text{non-negative integer,} \\ 0, & \text{otherwise.} \end{cases} \quad (2.15)$$

Proof

The sequence $f_0(-1), f_1(-1), \dots, f_{s+1}(-1)$ has at most $s+1$ sign variations. The presence, or lack, of a variation between $f_0(-1)$ and $f_1(-1)$ is indicated by the sign of L_0 . For the later terms, noting that $T_1(-1) = -1$ and $U_1(-1) = -2$, we have

$$f_2(-1) = -[f_1(-1) + f_0(-1)],$$

and

$$f_3(-1) = -2f_2 - f_1 = [f_1 + 2f_0],$$

which gives rise to the general formula

$$f_j(-1) = (-1)^{j-1} [f_1(-1) + (j-1)f_0(-1)], \quad j = 1, 2, \dots, s+1. \quad (2.16)$$

No variation in sign occurs between $f_j(-1)$ and $f_{j+1}(-1)$ if the product $f_j(-1)f_{j+1}(-1)$ is non-negative. From (2.16) and (2.12) we find

$$f_j(-1)f_{j+1}(-1) = -f_0^2(-1)[L_0 + j - 1][L_0 + j], \quad (2.17)$$

for $j = 1, 2, \dots, s$. Using (2.17), $f_j(-1)f_{j+1}(-1) \geq 0$ if and only if

$$0 \leq L_0 + j \leq 1. \quad (2.18)$$

Now L_0 is fixed, and at most 2 integers j could satisfy (2.18). The range of j is $1 \leq j \leq s$ for the products (2.17), thus there are two values of j which satisfy (2.18) only if $L_0 = -1, -2, \dots, 1-s$, whence $\Psi_1(-1) = s+1 - 2 = s-1$. For $L_0 = 0$, only $j = 1$ satisfies (2.18) so $\Psi_1(-1) = s+1 - 1 - 1 = s-1$. When $L_0 = -s$, or L_0 is a non integer but $-s < L_0 < 0$, exactly one j satisfies (2.18) so $\Psi_1(-1) = s$. If $L_0 < -s$ no j in the given range can satisfy (2.18) so $\Psi_1(-1) = s+1$, and if $L_0 > 0$, $L_0 + j > 1$ for all acceptable j , thus $\Psi_1(-1) = s$ and (2.14) is established.

In a similar fashion we verify (2.15). Using $T_1(1) = 1$ and $U_1(1) = 2$, we obtain

$$f_j(1) = f_1(1) - (j-1)f_0(1), \quad j = 1, 2, \dots, s+1, \quad (2.19)$$

Products of consecutive terms may be written

$$f_j(1)f_{j+1}(1) = f_0^2(1)[R_0-j+1][R_0-j], \quad (2.20)$$

for $j = 1, 2, \dots, s$. The product (2.20) will be negative if and only if

$$0 < j - R_0 < 1. \quad (2.21)$$

Clearly for a given R_0 , at most one j can satisfy (2.21), and if R_0 is itself an integer no such j exists. In fact, (2.21) can be satisfied by one j in the range $1 \leq j \leq s$ precisely when R_0 is a non-integer such that $0 < R_0 < s$. For this case, and for the case $R_0 < 0$, $\Psi_1(1) = 1$. Otherwise, $\Psi_1(1) = 0$, and the theorem holds.

In summary, we find from (2.14) and (2.15) that $[\Psi_1(-1) - \Psi_1(1)]$ may take on one of the four values $s+1, s, s-1, s-2$. The sequence to be constructed in phase 2 begins with f_{s+1} which is of degree not greater than $n-s-1$, i.e., $s-1$ when n is even and s when n is odd, and thus $|\Psi_2(-1) - \Psi_2(1)|$ cannot exceed s , (or $s-1$ in the even case). If we are concerned only with the existence of a zero, rather than the precise number, in the unit circle, the construction may be terminated should $[\Psi_1(-1) - \Psi_1(1)]$ dominate s , (or $s-1$). In such a case we have computed only L_0 and R_0 .

If $f_{s+1}(\pm 1) \neq 0$, which is easily checked using (2.16) and (2.19), an alternative to phase 2 is possible. Rather than continuing the present sequence until a constant function is obtained, evaluate

$$I_{-1}^1 \frac{f_s(x)}{f_{s+1}(x)}. \quad \text{In view of theorems 1.4 and 2.1, where } N \text{ is defined,}$$

we have

$$N = [\Psi_1(-1) - \Psi_1(1)] - I_{-1}^1 \frac{f_s(x)}{f_{s+1}(x)} . \quad (2.22)$$

The evaluation of $I_{-1}^1 \frac{f_s(x)}{f_{s+1}(x)}$ can be achieved using a Sturm sequence whose starting functions are f_{s+1} and f_s , or some cleverly chosen pair whose ratio has the same Cauchy index on $[-1,1]$. Yet another way to

find $I_{-1}^1 \frac{f_s(x)}{f_{s+1}(x)}$, when s is not too large, is to solve $f_{s+1}(x) = 0$

on $[-1,1]$, using a method like bisection which depends on a sign change at a zero. Only zeros ξ of odd multiplicity are then obtained, and by evaluating $f_s(\xi)$, the Cauchy indices are determined. In addition to computational effort, this approach has the disadvantage of missing poles ξ of odd multiplicity when ξ occurs as a zero of f_{s+1} with even multiplicity m , and as a zero of f_s with odd multiplicity less than m . This situation will not occur if f has no zeros on the unit circle, but in practice the existence of zeros on the unit circle will be determined only as an algorithm proceeds.

3. Dividing Chebyshev Polynomials

In forming each remaining function of the Sturm sequence, from f_{s+1} to a constant function, a polynomial division is required. This is not a new turn of events, as (2.8) is such a division. However, in (2.8) the quotient is U_1 and the new coefficients can be easily

determined from the original polynomial. In phase 2 the quotient and new coefficients are determined from coefficients of the previous two polynomials via division and inner product formation, allowing rounding error propagation and possible instabilities.

A new function f_{j+2} , is formed from the previously determined functions f_{j+1} and f_j according to

$$f_{j+2} = Q_{j+1} f_{j+1} - f_j, \quad (3.1)$$

where Q_{j+1} is to be determined so that the degree of f_{j+2} is less than that of f_{j+1} . Letting

$$f_{j+1}(x) = \sum_{k=0}^d c_k U_k(x) \text{ and } f_j(x) = \sum_{k=0}^m b_k U_k(x), \quad (3.2)$$

we note that for $j = s$ and n even, (i.e., $n = 2s$), $m = s$ and from (2.10) we obtain the initial c 's and b 's:

$$c_k = a_{s+k+1} - a_{s+k-1}, \quad k = 0, 1, \dots, s-1; \quad (3.3)$$

$$b_s = a_n, \quad b_{s-1} = a_{n-1}, \quad b_k = (a_{s+k} - a_{s-k-2}), \quad k = 0, 1, \dots, s-2. \quad (3.4)$$

The degrees of f_{j+1} and f_j must satisfy $d \leq m - 1$, with the strict inequality occasionally holding. This situation is encountered in forming f_{s+2} if and only if n is even and $a_n = a_0$, ($c_s = a_n$ when n is odd), in which case f may be shown to have a zero in the unit disk,

using a well known family of zero bounds, (See [13], p. 203).

As an extreme case, f_{s+1} could be identically zero. Consider the example

$$f(z) = 6z^4 - 35z^3 + 62z^2 - 35z + 6 \quad (3.5)$$

given by Lehmer [6], where Schur-Cohn test fails. Phase 1 would yield

$$f_0(x) = 6T_4(x) - 35T_3(x) + 62T_2(x) - 35T_1(x) + 6,$$

$$f_1(x) = 6U_3(x) - 35U_2(x) + 62U_1(x) - 35,$$

$$f_2(x) = 6U_2(x) - 35U_1(x) + 56,$$

$$f_3(x) \equiv 0.$$

Thus $f_0(-1) = 144$, $f_0(1) = 4$, $f_1(-1) = -288$, $f_1(1) = 8$, giving $L_0 = -2$, $R_0 = 2$, and so $\Psi_1(-1) - \Psi_1(1) = 2$. Now, $f_2(x) = 24x^2 - 70x + 50$, which is of constant sign on $[-1, 1]$, so we may think of the Sturm sequence as terminating with f_2 , and conclude that f has 2 zeros in the unit circle. Anytime $f_{s+1} \equiv 0$, f_s must be of constant sign for $f_s(\xi) = 0$ would force $f_{s+1}(\xi)f_s(\xi) < 0$, an impossibility. Thus extreme handmade examples such as (3.5), which are "accidental", as Lehmer puts it, will be easily disposed of in phase 1.

To form Q_{j+1} of (3.1) we use the following multiplication formula, again a case of trigonometric addition formulas,

$$2T_k U_j = U_{k+j} + U, \text{ where } U = \begin{cases} U_{j-k}, & \text{if } j \geq k, \\ 0, & \text{if } j = k - 1 \\ -U_{k-j-2}, & \text{if } j < k - 1. \end{cases} \quad (3.6)$$

Now, letting

$$Q_{j+1}(x) = 2 \sum_{k=0}^{m-d} q_k U_k(x), \text{ and } f_{j+2}(x) = \sum_{k=0}^{d-1} p_k U_k(x), \quad (3.7)$$

we find the $m+1$ unknowns, $q_{m-d}, q_{m-d-1}, \dots, q_0, p_{d-1}, p_{d-2}, \dots, p_0$, by expanding (3.1) and equating coefficients. First, using (3.6), the product $Q_{j+1} f_{j+1}$ becomes

$$\sum_{k=0}^{m-d} q_k \sum_{j=0}^d c_j U_{k+j} + \sum_{k=0}^{m-d} q_k \sum_{j=k}^d c_j U_{j-k} - \sum_{k=2}^{m-d} q_k \sum_{j=0}^{k-2} c_j U_{k-j-2},$$

and upon further rearrangement we have,

$$\begin{aligned} Q_{j+1} f_{j+1} &= \sum_{k=0}^m \left(\sum_{j=m_0}^{m_1} q_j c_{k-j} \right) U_k + \sum_{k=0}^d \left(\sum_{j=k}^{m_2} q_j c_{j-k} \right) U_k \\ &\quad - \sum_{k=0}^{m-d-2} \left(\sum_{j=k+2}^{m_3} q_j c_{j-k-2} \right) U_k, \end{aligned} \quad (3.8)$$

where $m_0 = \max(0, k-d)$, $m_1 = \min(m-d, k)$, $m_2 = \min(k+d, m-d)$, and $m_3 = \min(m-d, k+d+2)$, with the last sum taken as zero in the case $d = m - 1$. To insure that f_{j+2} is of degree less than d , we take

$$c_d q_{m-d} = b_m, \quad (3.9)$$

$$c_d q_{m-d-1} = b_{m-1} - c_{d-1} q_{m-d},$$

$$c_d q_{m-d-2} = b_{m-2} - c_{d-1} q_{m-d-1} - c_{d-2} q_{m-d-2},$$

$$c_d q_{m-d-j} = b_{m-j} - \sum_{k=m_4}^{j-1} q_{m-d-k} c_{d-j+k} + \delta \sum_{k=2}^{m_5} q_{m-j+k} c_{k-2}$$

$$- \epsilon \sum_{k=d}^{m_6} q_k c_{k-d}, \text{ for } 3 \leq j \leq m-d,$$

where $m_4 = \max(0, j-d)$, $m_5 = \min(j-d, d+2)$, $m_6 = \min(m-d, 2d)$,

and $\delta = 0$ if $j < d+2$, 1 otherwise; $\epsilon = 0$ unless $j = m-d$ and

$m-d > d$, in which case $\epsilon = 1$. So each q can be found by forming

an inner product and dividing by c_d . Equations (3.9) are complicated

by their inclusion of the unlikely cases where $d < m-d-2$ or $d < m-d$.

Typically, $m-d = 1$, or at least $d \geq m-d$, and (3.9) is a triangular

system with coefficient matrix $C = (c_{ij})$, where

$$c_{ij} = \begin{cases} 0 & \text{if } j > i \\ c_{d-k} & \text{if } k = i-j; i, j = 1, 2, \dots, m-d. \end{cases} \quad (3.10)$$

Phase 2 amounts to solving a rather simple reducible linear system of order $m+1$. Having obtained the $m-d+1$ q 's, we use them in finding the p 's. From (3.1) we obtain

$$p_j = \sum_{k=0}^{m_7} q_k c_{j-k} + \sum_{k=j}^{m_8} q_k c_{k-j} - \sum_{k=j+2}^{m_9} q_k c_{k-j-2} - b_j, \quad (3.11)$$

for $j = 0, 1, \dots, d-1$, where $m_7 = \min(m-d, j)$, $m_8 = \min(j+d, m-d)$, and $m_9 = \min(m-d, d+j+2)$, with the last sum taken as zero in case $j > m-d-2$.

The values $f_{j+2}(\pm 1)$ are found by forming

$$Q_{j+1}(\pm 1) = 2 \sum_{k=0}^{m-d} (\pm 1)^k q_k, \quad (3.12)$$

then using (3.1) with the previous values of $f_{j+1}(\pm 1)$ and $f_j(\pm 1)$.

In case $|f_{j+2}(\pm 1)|$ is small, and its sign is questionable, the alternate form

$$f_{j+2}(\pm 1) = \sum_{k=0}^{d-1} (\pm 1)^{k+1} k q_k \quad (3.13)$$

might be used as a check on this sign. However, (3.13) usually requires more arithmetic operations, and the p 's are formed from the q 's allowing further chance for rounding errors.

When the sign of $f_{j+2}(\pm 1)$ is questionable in the presence of round-off noise, say $|f_{j+2}(\pm 1)| < 2^{2-t}$ when t binary digits are used in computing, the best strategy seems that of Lehmer [6], or Stewart [10]; terminate the present calculations and test a different circle! Another source of difficulty is determining the exact degree of the polynomial f_{j+2} , should $|p_{d-1}|$, (or even $|p_{d-k}|$, $k = 1, 2, \dots, j$, $j \leq d$), be small, prior to solving for f_{j+3} .

That the leading coefficient of the present polynomial is crucial in determining the new polynomial is clearly evident in (3.9). Considering

the matrix C of (3.10), and using the estimate $\text{cond}(A) \geq \|A\|_\infty / (\min_i |a_{ii}|)$, (See Conte [1], p. 151), we find that the condition number of C may be estimated by

$$\text{cond}(C) \geq 1 + \sum_{j=1}^{m-d} |\hat{c}_{d-j}|, \quad (3.14)$$

in the usual case of $d \geq m-d$, where $\hat{c}_k$ denotes c_k/c_d . Should any of the $|\hat{c}_k|$ be large, say of order 2^{t-2} , the system is highly ill-conditioned with respect to the accuracy used. Although Wilkinson [12], has found that for a wide class of triangular systems accurate results can be obtained, (independent of the condition number), at present it seems that a large $|\hat{c}_j|$ should be taken as an indicator of possible instabilities and a different circle should be tested.

To summarize, starting with $j = s$, and assuming for simplicity that $d \geq m-d$, proceed with phase 2 as follows:

given $b_m, b_{m-1}, \dots, b_0$ and $c_d, c_{d-1}, \dots, c_1, c_0$, (i)

form $\hat{c}_{d-k} = c_{d-k}/c_d$ and $\hat{b}_{m-k} = b_{m-k}/c_d$ for

$k = 1, 2, \dots, m-d$; if any $\hat{c}_{d-k}$ is too large, in the

sense of (3.14), or if forming any $\hat{b}_{m-k}$ causes overflow,

cease computation, otherwise,

$$\text{form } q_{m-d-j} = \hat{b}_{m-d-j} - \sum_{k=0}^{j-1} q_{m-d-k} \hat{c}_{d-j+k}, \quad (\text{ii})$$

for $j = 0, 1, \dots, m-d$, the sum being zero when

$j = 0$; next,

form $Q_{j+1}(\pm 1)$, as in (3.12), then $f_{j+2}(\pm 1)$; (iii)

if either $|f_{j+2}(1)|$ or $|f_{j+2}(-1)|$ is too small

cease computation, otherwise, record possible
sign variations, and

form p_j , $j = 0, 1, \dots, d-1$, as in (3.11); next, (iv)

when $d = 1$ or $\max(\{j \mid 0 \leq j \leq d-1 \text{ and } p_j \neq 0\}) = 0$, (v)

p_0 represents the final (constant) function in

the sequence; if $|p_0|$ is small, f may have

a zero on the unit circle, otherwise record

possible sign variations, in either case cease
computation; finally,

set $m = d$, $d = \max(\{j \mid 0 \leq j \leq d-1 \text{ and } p_j \neq 0\})$, (vi)

$b_k = c_k$, $k = 0, 1, \dots, m$, and $c_k = p_k$, $k = 0$,

$1, \dots, d$, then return to (i).

The possible difficulties outlined above are not unique to the test method presented here. In the next section we find similar cases where the Schur-Cohn test fails. The advantage of the present method is that the first sequence element to be constructed, (excluding the trivial computations for L_0 , R_0 , f_s , and f_{s+1}), is of degree less than $n - [n/2] - 1$, as opposed to being of degree less than n .

4. Comparison with the Schur-Cohn Test

The Schur-Cohn test requires forming a sequence of polynomials

$$f_j(z) = \sum_{k=0}^{n-j} a_k^{(j)} z^k, \quad j = 1, 2, \dots, n, \text{ from the given, (perhaps complex),}$$

$$\text{polynomial } f_0(z) = \sum_{k=0}^n a_k^{(0)} z^k, \text{ according to}$$

$$f_{j+1}(z) = \overline{a_0^{(j)}} f_j(z) - a_{n-j}^{(j)} f_j^*(z), \quad j = 0, 1, \dots, n-1, \quad (4.1)$$

where f_j^* , the reciprocal polynomial of f_j , is given by $f_j^*(z) = \sum_{k=0}^{n-j} \overline{a_{n-k}^{(j)}} z^k$.

Thus coefficients may be generated according to

$$a_k^{(j+1)} = \overline{a_0^{(j)}} a_k^{(j)} - a_{n-j}^{(j)} \overline{a_{n-j-k}^{(j)}}, \text{ for} \quad (4.2)$$

$k = 0, 1, \dots, n-j-1, j = 0, 1, \dots, n-1$. With definitions (4.1)

and (4.2), the Schur-Cohn test may be stated: if $f_j(0) \neq 0$ for

$j = 0, 1, \dots, n$, and if $f_{j_k}(0) < 0$ for $j_1 < j_2 < \dots < j_m$, then

the number, N , of zeros of f_0 in the unit circle, including multiplicities, is given by

$$N = \begin{cases} \sum_{k=1}^m (-1)^{j_k}, & \text{if } m \text{ is even,} \\ n+1 + \sum_{k=1}^m (-1)^{j_k}, & \text{if } m \text{ is odd.} \end{cases} \quad (4.3)$$

This test fails if $f_j(0) = 0$ for some j , although we may conclude that f_0 has at least one zero in the unit disk in this case. Whenever f_j is self-reciprocal, i.e., its zeros can be grouped in pairs that are symmetric with respect to the unit circle, as in example (3.5) above, we have $f_j(0) = 0$. In theory, the test given in this paper, which applies only to real polynomials, will never fail. In the reality of numerical computation this unfortunately is not the case.

The difficulties in determining a sign when $|f_{j+2}(\pm 1)|$ is small correspond exactly to difficulties when $|f_j(0)|$ is small. At first glance, the possible instability associated with dividing by c_d in (3.9) has no analogue in the Schur-Cohn test. However, the modified form of this test, shown by Stewart, [11], to be less sensitive to underflows or overflows, involves a similar division. Stewart's modification of (4.1) is

$$f_{j+1} = f_j - m_{j+1} f_{j-1} \quad (4.4)$$

where $m_{j+1} = a_{n-j}^{(j)} / \bar{a}_0^{(j)}$. Note that the divisions required in (3.9) could be removed by replacing (3.1) with

$$f_{j+2} = Q_{j+1} f_{j+1} - r_j f_j \quad (4.5)$$

where $r_j = |c_d|$, or even $r_j = c_d^2$.

The number of operations required in forming the test sequences is fewer for the method of this paper than for the Schur-Cohn method.

For each j , using (4.4), we must divide once in forming m_j , then multiply $n-j+1$ times, yielding

$$\sum_{j=0}^{n-1} n-j+2 = \frac{n^2}{2} + \frac{5n}{2} \quad (4.6)$$

multiplications and divisions for the Schur-Cohn sequence. Taking the case $d = m-1$ in (3.2) for each j , we form $q_1 = \hat{b}_m, \hat{b}_{m-1}, \hat{c}_{d-1}$, then $q_0 = \hat{b}_{m-1} - \hat{c}_{d-1} \hat{b}_m$ requiring 3 divisions and a multiplication. (One fewer division is required if we take $q_1 = b_m/c_d$, then $q_0 = [b_{m-1} - c_{d-1}q_1]/c_d$). Now $p_k = q_1(c_{k+1} + c_{k-1}) + q_0c_k - b_k$ requires 2 multiplications and 2 multiplications are required to update the new functions at ± 1 for each step, except the last. (With the Schur-Cohn test, in Stewart's form, $|m_{j+1}|$ need only be compared to 1). Finally, if multiplications are used in forming $f_1(\pm 1)$, $n-1$ are necessary, and L_0, R_0 , and $f_s(\pm 1)$ require 4 more for a total of

$$n+1 + \sum_{k=0}^{n-[n/2]-2} 2k+8 = \begin{cases} \frac{n^2+14n-20}{4}, & \text{if } n \text{ is even,} \\ \frac{n^2+16n-9}{4}, & \text{if } n \text{ is odd.} \end{cases} \quad (4.7)$$

If additions and subtractions are included, the total number of arithmetic operations required becomes

$$\begin{aligned} & (5n^2+98n-168)/8, \text{ when } n \text{ is even,} \\ & (5n^2+108n-121)/8, \text{ when } n \text{ is } n \text{ is odd.} \end{aligned} \quad (4.8)$$

The Schur-Cohn test requires $n^2 + 4n$ total arithmetic operations.

Thus the present method requires fewer multiplications and divisions when $n > 3$, and fewer total operations when $n \geq 20$.

A situation where one would suspect numerical difficulties is the case where the given polynomial has all its zeros near the unit circle. In such a case the difficulty is sensed early in both tests. The product of the moduli of the zeros equals $\left| \frac{a_0}{a_n} \right|$, which means m_0 is near 1 in the Schur-Cohn test, and c_d may be near 0 when starting phase 2 of the present test. We conclude with an example of this type.

Consider the polynomial

$$\begin{aligned} & x^9 - 1.1x^8 + 1.12x^7 - 1.232x^6 + 2.084x^5 - 2.2924x^4 + 1.075648x^3 \\ & - 1.1832128x^2 + 1.07910544x - 1.187015984, \end{aligned} \quad (4.9)$$

which has zeros $\pm .7 \pm .7i$, $\pm .5 \pm .9i$, and 1.1. All computing was done with 10 significant decimal digits. The results given here are rounded to 4 digits for convenience. First we find $f_0(-1) = -13.35$, $f_0(1) = -.6359$, $f_1(-1) = 59.29$, and $f_1(1) = 3.838$, so $L_0 = -4.440$ and $R_0 = -6.036$, giving $\psi_1(-1) - \psi_1(1) = 5 - 1 = 4$. The coefficients and function values computed in phase 2 are listed in table A below. The calculations for the Schur-Cohn test follow in table B.

Table A: Phase 2 Coefficients

j	U_0	U_1	U_2	U_3	U_4	U_5	q_1	q_0	$f_j(-1)$	$f_j(1)$
4	-1.109	1.005	-.04498	1.120	-1.100	1.000			-19.23	1.930
5	1.008	-.04878	.04089	.08702	1.000		1.000	-1.187	5.881	14.72
6	-.1365	.1103	.02511	-.1743			-5.736	-1.325	78.90	-190.3
7	-1.460	.5416	.2850				-6.117	1.250	189.3	-19.88
8	-.6970	.8544					.3336	.9060	-33.68	195.00
9	1.114								1.114	1.114

Counting sign variations for $j = 5, 6, 7, 8, 9$, we find $\Psi_2(-1) - \Psi_2(1) = 2 - 2 = 0$.

Thus $N = 4$ as required.

Table B: Schur-Cohn Coefficients

j	m_j	$a_0^{(j)}$	$a_1^{(j)}$	$a_2^{(j)}$	$a_3^{(j)}$	$a_4^{(j)}$	$a_5^{(j)}$	$a_6^{(j)}$	$a_7^{(j)}$	$a_8^{(j)}$	$a_9^{(j)}$
0		-1.187	1.079	-1.183	1.075	-2.292	2.089	-1.232	1.12	-1.1	1
1	-.8424	-.3446	.1524	-.2397	.03775	-.5325	.1578	-.3758	.1232	-.1909	
2	-.5541	-.2388	.0842	-.05915	-.04966	-.2375	.1369	-.1930	.0388		
3	-.1623	-.2325	.05282	-.03693	-.08821	-.2455	.1273	-.1794			
4	.7715	-.09412	-.04536	.1525	-.02016	-.2170	.08651				
5	-.9191	-.01461	-.2448	.1340	.1200	-.2587					
6	17.71	4.567	2.370	2.238	4.456						
7	.9756	.2200	-.1861	.07374							
8	.3353	.1952	-.1237								
9	-.6336	.1168									

Note that $|m_j| > 1$ only for $j = 6$, thus by (4.3) we find $N = 9+1-6 = 4$.

5. References

1. S. D. Conte and C. de Boor, Elementary Numerical Analysis, 2nd ed., McGraw Hill, New York, 1972.
2. A. Cohn, "Über die Anzahl der Wurzeln einer Algebraischen Gleichung in einem Kreise," Math. Z., v. 14, 1922, pp. 110-148.
3. L. M. Delves and J. N. Lyness, "A Numerical Method for Locating the Zeros of an Analytic Function," Math. Comp., v. 21, 1967, pp. 543-560.
4. D. K. Dunaway, "Calculation of Zeros of a Real Polynomial Through Factorization Using Euclid's Algorithm," SIAM J. Num. Anal., v. 1, 1963, pp. 581-591.
5. P. Henrici, Applied and Computational Complex Analysis, v. 1, Wiley and Sons, New York, 1974.
6. D. H. Lehmer, "A Machine Method for Solving Polynomial Equations", J. Assoc. Comput. Mach., v. 8, 1961, pp. 151-162.
7. M. Marden, "The Geometry of Zeros of a Polynomial In a Complex Variable", Amer. Math. Soc. Surveys No. 3, New York, 1949.
8. T. J. Rivlin, The Chebyshev Polynomials, Wiley and Sons, New York, 1974.
9. J. Schur, "Über Algebraische Gleichungen, die nur Wurzeln mit Negativen Realteilen Besitzen," Z. Angew. Math. Mech., v. 1, 1920, pp. 307-311.
10. G. W. Stewart III, "On Lehmer's Method for Finding the Zeros of a Polynomial," Math. Comp., v. 23, 1969, pp. 829-835.
11. G. W. Stewart III, Some Topics in Numerical Analysis, Oak Ridge National Laboratory Report ORNL-4303, 1968.
12. J. H. Wilkinson, Rounding Errors In Algebraic Processes, Prentice-Hall, Englewood Cliffs, N.J., 1963.
13. D. M. Young and R. T. Gregory, A Survey of Numerical Mathematics, v. 1, Addison-Wesley, Reading, Mass., 1972.