

AD-A215 490

4

TECHNICAL REPORT

Office of Naval Research Contract No. N00014-86-K0029

AN ALGORITHM FOR FINDING CHARACTERISTIC
ROOTS OF QUASI-TRIANGULAR MARKOV CHAINS

by

Carl M. Harris
William G. Marchal
and
Richard W. Tibbs

Report No. GMU/22474/111
November 10, 1989

Department of Operations Research and Applied Statistics
School of Information Technology and Engineering
George Mason University
Fairfax, Virginia 22030

DTIC
ELECTE
NOV 24 1989
S E D

This document has been approved
for public release and sale (in
unlimited quantities) at the
estimated price of \$1.00 per copy.

89 11 21 146

TECHNICAL REPORT

Office of Naval Research Contract No. N00014-86-K0029

AN ALGORITHM FOR FINDING CHARACTERISTIC
ROOTS OF QUASI-TRIANGULAR MARKOV CHAINS

by

Carl M. Harris
William G. Marchal
and
Richard W. Tibbs

Report No. GMU/22474/111
November 10, 1989

Department of Operations Research and Applied Statistics
School of Information Technology and Engineering
George Mason University
Fairfax, Virginia 22030

Copy No. 9

This document has been approved for public sale and release; its distribution is unlimited.

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM								
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER								
4. TITLE (and Subtitle) An Algorithm for Finding Characteristic Roots of Quasi-Triangular Markov Chains		5. TYPE OF REPORT & PERIOD COVERED Technical Report								
7. AUTHOR(s) Carl M. Harris William G. Marchal Richard W. Tibbs		6. PERFORMING ORG. REPORT NUMBER GMU / 22474 / 111								
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Operations Research and Applied Statistics George Mason University, Fairfax, VA 22030		8. CONTRACT OR GRANT NUMBER(s) N00014-86-K0029								
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research 800 North Quincy Street Arlington, VA 22247		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Task H B Project 411 8150								
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		12. REPORT DATE November 10, 1989								
		13. NUMBER OF PAGES 26								
		15. SECURITY CLASS. (of this report) Unclassified								
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE								
16. DISTRIBUTION STATEMENT (of this Report)										
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)										
18. SUPPLEMENTARY NOTES										
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)										
<table border="0"> <tr> <td>applied probability,</td> <td>numerical methods</td> </tr> <tr> <td>computational analysis</td> <td>probability</td> </tr> <tr> <td>computational probability,</td> <td>queues</td> </tr> <tr> <td>Markov chains</td> <td>stochastic models</td> </tr> </table>			applied probability,	numerical methods	computational analysis	probability	computational probability,	queues	Markov chains	stochastic models
applied probability,	numerical methods									
computational analysis	probability									
computational probability,	queues									
Markov chains	stochastic models									
20. ABSTRACT (Continue on reverse side if necessary and identify by block number)										
(following)										

Abstract

Numerical rootfinding problems are quite common in stochastic modeling. However, many solutions stop at the presentation of a probability generating function for the state probabilities. But with increasing easy access to computing power, many problems whose answers were typically left in incomplete form or for which there has been a search for alternative solution methods are currently being reexamined. The class of Markov chains whose transition matrices have quasi-triangular layouts (i.e., those having sub- or super-triangular sets of zeros) is a good case in point. They have an especially nice structure which leads to a rather concise representation for the generating functions. But the complete solution then requires the finding of roots. Fortunately, these problems can be shown to have special properties that make accurate rootfinding quite feasible, and we thus supply an efficient numerical procedure for solution.

Accession For	
NTIS GRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

1 INTRODUCTION

As numerous authors have noted (for example, Abolnikov and Dukhovny, 1987, Bailey, 1954, and Powell, 1985), many denumerable discrete-time Markov chains (with particular applications in inventory, dam and queueing modeling) have one of two special transition-matrix structures. These forms have been typically called *quasi-triangular* because of the presence of sub- or super-triangular sets of zeros:

$$A = \begin{bmatrix} a_{00} & a_{01} & a_{02} & a_{03} & \dots \\ a_{10} & a_{11} & a_{12} & a_{13} & \dots \\ a_{20} & a_{21} & a_{22} & a_{23} & \dots \\ \dots & \dots & \dots & \dots & \dots \\ a_{K0} & a_{K1} & a_{K2} & a_{K3} & \dots \\ a_0 & a_1 & a_2 & a_3 & \dots \\ 0 & a_0 & a_1 & a_2 & \dots \\ 0 & 0 & a_0 & a_1 & \dots \\ 0 & 0 & 0 & a_0 & \dots \\ 0 & 0 & 0 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix}$$

and

$$B = \begin{bmatrix} \Sigma_{K-1} & b_{K-1} & b_{K-2} & b_{K-3} & \dots & b_0 & 0 & 0 & 0 & \dots \\ \Sigma_K & b_K & b_{K-1} & b_{K-2} & \dots & b_1 & b_0 & 0 & 0 & \dots \\ \Sigma_{K+1} & b_{K+1} & b_K & b_{K-1} & \dots & b_2 & b_1 & b_0 & 0 & \dots \\ \Sigma_{K+2} & b_{K+2} & b_{K+1} & b_K & \dots & b_3 & b_2 & b_1 & b_0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix},$$

where

$$\Sigma_j = 1 - \sum_{n=0}^j b_n.$$

The structure of these matrices leads to some particularly concise representations for the probability generating functions (PGFs) of the Markov chain equilibrium state process. When the stationary equation for such a

Markov chain is exercised, the PGF of the steady-state probabilities has an algebraic function in its denominator whose roots are critical in the final solution. (We henceforth refer to this denominator function equated to zero as the system's *characteristic equation*).

For the transition matrix $\mathbf{A}$, the characteristic equation (CE) turns out to be (at least for complex z with absolute value ≤ 1)

$$z^K = \sum_{i=0}^{\infty} a_i z^i = \alpha(z), \quad (1)$$

where K is as defined in the matrix representation $\mathbf{A}$ (corresponding, for example, to a constant batch-input module in bulk queues). Under the assumption that $\alpha(z)$ possesses all its derivatives at $z = 1$ (i.e., that the distribution $\{a_i\}$ has all moments), $\alpha(z)$ may be set equal to the Laplace-Stieltjes transform of a distribution function [call that $A(t)$, and set its mean to $1/\mu$] of a nonnegative random variable evaluated at $\lambda(1 - z)$, where λ is an arbitrary positive constant for the time being. Thus we may also write that

$$z^K = A^*[\lambda(1 - z)]. \quad (2)$$

The CE associated with the matrix $\mathbf{B}$ may be written as

$$z^K = \sum_{i=0}^{\infty} b_i z^i = \beta(z), \quad (3)$$

or

$$z^K = B^*[\mu(1 - z)]. \quad (4)$$

In these representations, the constant K is as given in $\mathbf{B}$ (corresponding, for example, to a constant batch-service module in bulk queues), $\beta(z)$ is defined as the PGF of the probabilities $\{b_i\}$, and B^* is the Laplace-Stieltjes transform of a distribution function [call that $B(t)$, and set its mean to $1/\mu$] of a nonnegative random variable evaluated at $\mu(1 - z)$, where μ is an arbitrary positive constant.

Recognize that Equations (1) and (3) are generalizations of the well-known fundamental equation of branching processes, typically written as $z = f(z)$, where f would be the PGF for the number of offspring emanating from one parent. Gross and Harris (1985), for example, provide the details of the root problem for this model. In actuality, it is the **B** problem which is in fact the more direct relative of the branching process, so it is this one on which we comment in detail first.

The CE of the matrix **B** may be rewritten in the standard way by using $z = r \exp(i\theta)$, and we find that

$$r^K \exp(i\theta K) = B^*[\mu(1 - r \exp(i\theta))] \exp(2\pi n i) \quad (5)$$

for $n = 1, 2, \dots, K$. This equation clearly has a root at unity, and by Rouché's theorem, we can show that there are K others inside the unit circle $|z| = 1$ when the chain is ergodic. The condition for ergodicity is that

$$\beta'(1) = \sum_{n=0}^{\infty} n b_n > K$$

or

$$\left(\frac{dB^*[\mu(1-z)]}{dz} \right)_{z=1} > K.$$

These can be shown to be equivalent to the requirement that $K\lambda/\mu < 1$.

For the A-matrix problem, recall that the characteristic equation is

$$z^K = A^*[\lambda(1-z)], \quad (6)$$

where A^* is a Laplace-Stieltjes transform. Ergodicity obtains here when

$$\alpha'(1) = \sum_{n=0}^{\infty} n a_n < K$$

or

$$\left(\frac{dA^*[\lambda(1-z)]}{dz} \right)_{z=1} < K.$$

This is equivalent to requiring that $\lambda/K\mu < 1$.

It is easily shown by Rouché's theorem that (6) has K roots inside and on the unit circle, including the root $z=1$. Abolnikov and Dukhovny (1987) have noted that all the roots on the unit circle are, in fact, simple.

In the prior work of Chaudhry, Harris and Marchal (1989), we have seen under an assumption of infinite divisibility that Equation (1) from matrix A may be rewritten as

$$z^K = [\alpha_1(z)]^K.$$

It then follows that

$$z = \alpha_1(z) \epsilon_{n,K} \quad (n = 1, 2, \dots, K), \quad (7)$$

where $\epsilon_{n,K}$ is an n th (out of K) complex root of unity.

When the original $\alpha(z)$ is infinitely divisible, the function $\alpha_1(z)$ is itself a legitimate PGF with $\alpha_1'(1) = \lambda/K\mu < 1$. The equation $z = \alpha_1(z)$ has only the root $z = 1$; the remaining $K - 1$ roots inside and on the unit circle follow distinctly from the other $K - 1$ equations using Rouché's theorem.

The analyticity of the K th root of $\alpha(z)$ and subsequent use of Rouché's theorem for distinctness also follows when $\alpha(z)$ is nonzero though not necessarily infinitely divisible. (See Chaudhry, Harris and Marchal, 1989.)

Unfortunately, it is not true that all $\alpha(z)$ associated with such models have no zeros inside the unit circle. However, we can feel comfortable knowing that a good number of the problems encountered in practice will have infinitely divisible distributions since many such probability functions are built up by mixtures and convolutions from the infinitely divisible exponential and Erlang. For all other distributions, one should always first try to determine whether $\alpha(z)$ is ever zero, for, if not, the K th-root approach of the infinitely divisible case will work.

The vanishing of $\alpha(z)$ is equivalent to requiring that

$$A^*[\lambda(1-z)] = \int_0^\infty e^{-\lambda(1-z)t} dA(t) = 0.$$

By changing to the polar form $z = r(\cos \theta + i \sin \theta)$ and then separating real and imaginary parts, this is identical to asking that, simultaneously,

$$\int_0^\infty e^{-\lambda(1-r \cos \theta)t} \cos[(\lambda r \sin \theta)t] dA(t) = 0$$

and

$$\int_0^\infty e^{-\lambda(1-r \cos \theta)t} \sin[(\lambda r \sin \theta)t] a(t) dt = 0.$$

Chaudhry, Harris and Marchal (1989) have established firm sufficient conditions for distinctness of the roots for (1) and (2). But neither of their conditions is necessary. Examples are presented in their paper for which $\alpha(z)$ has zeros, but where the characteristic equations still have distinct roots. In fact, they never found an example without distinct roots. But a proof that this is indeed always true is elusive, and the search for a necessary condition will be the subject of future research.

All of these arguments hold in a parallel way for the B-matrix problem, as well, as long as $\beta'(1) > K$, with its K roots strictly inside the unit circle. The B-matrix version of the problem would then have $\beta_1(z)$ in place of $\alpha_1(z)$ in Equation (7).

2 ALGORITHM

The Chaudhry, Harris and Marchal (1989) experiences for both the A- and B-matrix problems yielded a nearly circular pattern of roots on the inside of the unit circle. This is, indeed, perfectly predictable given the above results. This comes directly from writing the right-hand side of the K th root of the characteristic equation as the product of a root of unity and the K th root of the probability generating function. The support of the random variable corresponding to the pgf α_1 is increasingly focused toward 0, so that the net effect of the multiplication is to move the root *slightly* away from the unit circle's boundary toward the center, largely preserving a circle-like appearance. In examples offered in the following, we display a number of diagrams of such patterns. Some possible guidance for the speeding up of the algorithm comes to mind from this formulation. It is to use the n th root (out of K , $n = 1, 2, \dots, K$) of unity as the starting point for each attempted solution of (7). Thus we would set the initial value of the radius at 1 and θ at $2\pi K/n$. We can even do a bit better since we know that the modulus of each root must be less than the positive real root; therefore we use an initial modulus for any particular problem equal to the smallest one obtained so far as n moves to K .

Compare this to the approach used by Chaudhry and colleagues (for example, see Brière and Chaudhry (1987)) to find the roots for the $G/E_K/1$ problem, as in Equations (3) and (4). Their algorithm works on the logarithm of Equation (4), using the real and imaginary parts separately. First, the real equation is solved for r , holding θ fixed. Then the imaginary equation is solved for n , given the derived r and fixed θ . If the result is integral, we are finished; otherwise, the fractional part of n is used and compared to its counterpart related to a larger θ . If the former is smaller, then θ is increased further. Such comparisons are repeated until there is a decrease between the fractional parts of the last two derived n values. This indicates that a root is located at an angle between the last two θ values. Details of the algorithm are given in Brière and Chaudhry (1987).

We have built up *our* algorithm, then, from the intuitively appealing near-circular pattern found for the roots. The roots are ordered according to increasing angle θ , with a starting value for the search for the $(n+1)$ st root using the modulus of the n th root and the angle of the $(n+1)$ st root of unity, $e_{n+1,K}$. We chose to explore the use of a fixed-point algorithm, knowing that a sufficiently well-chosen starting point would lead to convergence of successive substitutions of z into Equation (7). This is particularly easy to formulate in light of the linearity of the left-hand side of (7).

Experimentation with this fixed-point technique has led to success, and detailed examples are presented in the following sections. To start the overall algorithm, it is necessary to provide one of the roots of the problem and the obvious initial choice is to find the real root on $(0, 1)$, since we know the exact angle of this particular root. Any method could be used to solve for this real root, actually, but we chose to use the same fixed-point approach, simply for uniformity in the algorithm. The derivation of a safe starting point for the location of the real root proved to be an interesting problem in its own right. Due to the knowledge that the support for α_1 is toward 0, an initial guess greater than the value of the root would be natural.

At first, an initial value equal to $\alpha'(1)/K$ for the **A**-matrix problem (or equivalently, $K/\beta'(1)$ for the **B**-matrix model) seemed logical, since this is the (single) real root ($= \rho$) for the basic $M/M/1$ queueing problem. While this starting value did work in all cases, the accuracy of the real root found was not ideal because small but bothersome complex components crept into the solution. Nevertheless, all successive roots were found correctly and their values matched those obtained via other rootfinding packages (e.g., the

PACK implementation of the Chaudhry approach).

The progress of convergence of the iteration from starting values outside of the near-circular pattern of roots could suffer from discontinuities in the derivatives of the function $\alpha_1(z)$ brought about by the zeros (if any) of the function α . But it has been observed that the roots of $\alpha(z)$ lie always outside of the locus of the roots of the characteristic equation (7). It has also been observed that the root locus was skewed asymmetrically toward the positive axis in the complex plane, thus making the roots with angles closest to π those of shortest modulus. Thus, if one wished to provide starting values within the basin of attraction of the desired roots, but avoid possible discontinuities in the derivative, one could seek starting values at the angle of $\theta_{K/2} = 2K\pi/[K/2]$ radians, finding the roots in the top half of the complex plane by seeking the roots in the order $2\pi K/n$ with $n = K/2, K/2 - 1, \dots, 0$, thus ending with the real root in $(0, 1)$.

This variation on the technique proved particularly successful, and a starting value of $(r, \theta) = (0, -\theta_{K/2})$ for the first root was used. The effect of this variation was to present a value of slightly less modulus than the desired root, and at an angle of the corresponding root of unity, as the starting value for the fixed point iteration on each root. In general, the numerical impact of the variation was to improve the efficiency of convergence for the fixed point by one to two iterations for each complex root. The real root on $(0, 1)$ was found much more quickly by "coming from the inside" of the root locus, and the complex component of the root was exactly zero.

3 Proof of Convergence

For purposes of this proof, we assume that $\alpha(z)$ is infinitely divisible and thus that any one of its roots is a proper PGF.

Theorem 1 Define the n th characteristic equation for the $G/E_K/1$ and $E_K/G/1$ problems as

$$z_n^K = \alpha(z_n)$$

or

$$z = \alpha(z_n)^{1/K} = f(z_n) = \alpha_1(z_n)\epsilon_{n,K},$$

where $\epsilon_{n,K}$ is the n th of K roots of unity, and the roots $z_{K_2}, z_{K_2-1}, \dots, z_1, z_0$, $K_2 = \lfloor K/2 \rfloor$, are ordered starting with the root of greatest angle in the upper complex half plane, progressing toward the positive real axis. From an initial point $z^{(0)} = (r_0, \theta_0)$ with modulus $r_0 < \{z_{n+1}\}$, $1 \leq n+1 \leq K_2$, the method of successive substitutions,

$$z_n^{(i+1)} = \alpha_1(z_n^{(i)}) \epsilon_{n,K}, \quad (8)$$

applied to the n th characteristic equation for the $G/E_K/1$ and $E_K/G/1$ problems will converge to a unique root.

Proof. Realize first that the characteristic equation is separable into the factors $\epsilon_{n,K}$ and $\alpha_1(z_n) = \sum \alpha_{1,i} z_n^i$ (say). Note that $\alpha_1(z_n)$ is a function of $z_n^{(i)}$ at the i th iteration of successive substitution, but that $\epsilon_{n,K}$ is unchanged by the iteration. We are assured that all the z_n , $n = 1, \dots, K$ are distinct simple roots in or on the unit circle, as previously proven in Chaudhry, Harris and Marchal's Theorems 1 and 4.¹ It thus suffices to show then that the fixed-point approach is guaranteed to converge to a root in $|z| < 1$, since the distinctness of the roots will assure us that there cannot be convergence to the same answer for more than one $\epsilon_{n,K}$. Also note that it has been shown empirically that $|z_{n+1}| < |z_n|$ for $0 \leq n \leq K_2$.

To complete the proof, then, the critical need is to show that the probability generating function $\alpha_1(z)$ is less than $1 - \delta$ (δ an infinitesimal positive constant) for any z chosen within or on the closed and bounded contour $|z| \leq 1 - \delta$. This is indeed true because

$$\begin{aligned} \left| \sum_{i=0}^{\infty} \alpha_{1,i} z^i \right| &\leq \sum_{i=0}^{\infty} |\alpha_{1,i}| |z|^i \leq \sum_{i=0}^{\infty} \alpha_{1,i} (1 - \delta)^i \\ &< (1 - \delta) \sum_{i=0}^{\infty} \alpha_{1,i} = 1 - \delta \end{aligned}$$

since the $\{\alpha_{1,i}\}$ define a legitimate probability distribution under infinite divisibility whenever the traffic intensity < 1 . $\square\square$

¹For certain $\alpha(z)$, this may be due to the fact that the function $\alpha(z)$ is nonzero over the complex unit disk (at least), and for many problems because $\alpha(z)$ is infinitely divisible. In any other case, i.e., if $\alpha(z)$ itself has roots in the complex plane, then the derivatives of $f(z)$ are not continuous everywhere. This alone does not preclude convergence to a fixed point, but may result in erratic performance of the iteration.

4 RESULTS

For the exercise of our algorithm (called RFIX), we chose to apply the technique to a variety of the systems analyzed in Chaudhry, Harris and Marchal (1989) - CHM in the following. The problems covered here are listed by the corresponding table in that referenced paper: Table 1 ($M^{(K)}/M/1$), Table 2 ($PH/E_K/1$), Table 3 ($GH/E_K/1$), Table 5 ($PH/E_K/1$), and Table 7 ($R_n/E_K/1$), all solved as $G/E_K/1$ or B-matrix formulations; and Table 8 ($E_K/GE/1$) solved as an $E_K/G/1$ or A-matrix model. We show our results in tables numbered identically. To the problems of CHM, several $E_L/E_K/1$ and $H_2/E_K/1$ problems unique to this paper were added to illustrate the exercise of our algorithm further. These additional problems appear in this paper as Tables 4 and 6, respectively. For a basis of comparison, the algorithm of Chaudhry and colleagues was exercised on the same systems as our fixed-point algorithm.

The stopping rule for the iteration on each root was that successive root values be less than 10^{-15} in the square of Euclidean distance between them in the complex plane. This resulted in approximately a 10^{-7} error in the real and complex components of the results.

In addition, an implementation of a bivariate Newton's method due to Kahaner, Moler and Nash (1989) was run over all cases. The times shown are those for computing and displaying the roots in the upper half-plane.

To provide a clearer picture of the environments under which our results were obtained, we indicate in Table 0 the nature of the hardware used for each algorithm. The contents of Table 0 apply to all problem types, except for Tables IV and VI, where the Chaudhry algorithm was run on a 16 mHz 386 without a coprocessor. Thus the rootfinding times in Tables IV and VI are without any bias from differing computational platforms.

Table 0:
Computational Platforms

	RFIX	CHAUD	KM&N
machine	16 mHz	12 mHz	16 mHz
on which	386	286	386
algorithm	no	with	no
was run	coprocessor	coprocessor	coprocessor

Table I:
Selected Scenarios from Table 1, CHM, $M/E_K/1$
($\lambda = .1$)

μ	K	intensity $K\lambda/\mu$	RFIX (in sec)	CHAUD (in sec)	KM&N (in sec)
20.0000	10	.05	1	39	5
50.0000	25		2	43	12
200.0000	100		7	51	51
1000.0000	500		30	953	251
1.0526	10	.95	2	4	7
2.6316	25		3	4	15*
10.5263	100		7	10	54*
52.6316	500		30	281	257*

* Found wrong positive, real root less than 1. This happened despite an initial guess = $\text{mod}(z_2) < z_1$.

Table II:
Selected Scenarios from Table 2, CHM, $PH/E_K/1$
($\lambda = 4.21$)

μ	K	<i>intensity</i> $K\lambda/\mu$	RFIX (in sec)	CHAUD (in sec)	KM&N (in sec)
.460	10	.91522	3	21	6 [†]
4.210	10	.10000	3	21	5
.311	15	.90214	4	30	9 [†]
2.806	15	.10000	2	28	7
.155	30	.90214	5	49	16 [†]
1.403	30	.10000	3	49	14

† Found real root = 1.0 instead of $z_1 < z_0 = 1$.

Table III:
Selected Scenarios from Table 3, CHM, $GH/E_K/1$
($\lambda = .85714$)

μ	K	<i>intensity</i> $K\lambda/\mu$	RFIX (in sec)	CHAUD (in sec)	KM&N (in sec)
9.00	10	.95238	5	21	5 [#]
85.00	10	.10084	1	21	4
14.01	15	.91771	4	52*	6 [#]
128.00	15	.10045	2	29	6
29.00	30	.88670	4	90*	12
257.00	30	.10006	3	48	14

* Algorithm missed some complex roots, which were captured in a second pass.

Positive, real root not found (error message = ITERATION NOT MAKING GOOD PROGRESS).

Table IV:
Selected Scenarios for Model $F_{10}/F_K/1$
($\lambda = .1$)

μ	K	intensity $K\lambda/\mu$	RFX (in sec)	CHAUD# (in sec)	KM&N (in sec)
2.8571	20	.7	3	5	20*
6.4286	45		5	8	51*
14.2857	100		10	13	117*
20.0000	20	.1	2	25	23*
45.0000	45		4	28	62*
100.0000	100		9	33	90*
2.2222	20	.9	3	5	22*
5.0000	45		6	7	48*
11.1111	100		10	12	123*

* KM&N became totally lost, finding several roots two or three times, despite initial guesses at unique values. The KM&N subroutines found the negative, real root (a second time) when given a positive, real initial value, for example. Generally, KM&N did all right in the third quarter-plane (negative-negative), becoming confused at or before $\theta = 90^\circ$ for $K=20$ or $\theta = 45^\circ$ for $K=100$. Then previously found roots would be duplicated, with the negative real root (if any) recurring most often.

The Chaudhry software was run on a 16 mHz 386 without a coprocessor.

Table V:
Selected Scenarios from Table 5, CHM, $PH/E_K/1$
($\lambda = 0.2$)

μ	K	<i>intensity</i> $K\lambda/\mu$	RFIX (in sec)	CHAUD* (in sec)	KM&N# (in sec)
20.0	10	.100	2	25	
2.2	10	.909	5	26	
30.0	15	.100	2	35	
3.3	15	.909	5	38	
60.0	30	.100	3	63	
6.6	30	.900	6	66	

* QPACK solution was for slightly different parameter values using the same model. The resultant root values were not identical to those obtained using RFIX and KM&N, but we feel that the above QPACK run times are representative.

KM&N could not perform the rootfinding operation for this problem type. Incorrect values were found for all points.

Table VI:
Selected Scenarios for Model $H_2/E_K/1$
($\lambda_1 = \lambda_2 = .1$)

μ	K	intensity $K\lambda/\mu$	RFIX (in sec)	CHAUD [†] (in sec)	KM&N (in sec)
1.0	15	.100	3	27	10
1.0	30	.100	5	29	19
1.0	95	.100	14	37	60
1.0	500	.100	57	660	
0.2	15	.500	4	7	10
0.2	30	.500	6	8	21
0.2	95	.500	14	17	63
0.2	500	.500	57	333	
.11	15	.909	5	5	11*
.11	30	.909	7	7	21*
.11	95	.909	15	74 [#]	64*
.11	500	.909	58	365 [#]	

* For the real root on the uppersheet at 360° , there was agreement between KM&N and RFIX only to four or so digits.

Missed real roots, so routine was rerun with smaller step size.

† The Chaudhry software was run on a 16 mHz 386 without a coprocessor.

Table VII:
Selected Scenarios for Model $R_n/E_K/1$
($\lambda = 2$)

μ	K	<i>intensity</i> $K\lambda/\mu$	RFIX (in sec)	CHAUD (in sec)	KM&N (in sec)
20.90	10	.10000	1	19	5
2.20	10	.90909	2	20	6
20.00	15	.10000	2	26	7
2.13	15	.93750	2	27	8
20.00	30	.10000	3	46	14
2.06	30	.96774	3	45	15
2.20	100	.90909	8	131	46

Table VIII:
Selected Scenarios from Table 8, CHM, $E_K/GE_3/1$
($\mu = 2$)

λ	K	<i>intensity</i> $\lambda/K\mu$	RFIX (in sec)	CHAUD (in sec)	KM&N in sec)
2	10	.1	2	23	8
18	10	.9	2	22	7
3	15	.1	2	31	9
27	15	.9	2	31	9
6	30	.1	3	51	20
36	30	.9	3	50	21

5 OBSERVATIONS

The performance of the fixed-point algorithm can be seen to be quite effective and stable across a wide range of intensities and root quantities. A unique feature of the fixed-point algorithm is that the number of successive substitution iterations required per root actually goes down as the number of roots grows (i.e., as K grows). This is understandable since the differences between the moduli decrease as K grows, and the starting point for each root is closer to the final point.

The performance in this problem of the other algorithms, namely, CHAUD/ QPACK and KM&N, is indicative of a fundamental difference between equations typically requiring rootfinding in applied probability. In the QPACK and KM&N implementations, the equation $z^K = \alpha(z)$ is analyzed for zeros by testing whether $0 = \alpha(z) - z^K$. This equation represents the usual geometric conceptualization in the complex plane with K zeros inside and on the unit disk. Contrast this with the equation used by RFIX,

$$\begin{aligned} z_n &= \alpha_1(z_n) \epsilon_{n,K} \\ &= \sqrt[K]{r} e^{i(\theta + 2\pi n)/K} \\ &= \sqrt[K]{r} e^{i\theta/n} e^{2\pi ni/K} \end{aligned}$$

for $n = 0, 1, \dots, K - 1$. Recall, as stated in Section 3, that once the root of unity $\epsilon_{n,K}$ is specified, $z_n = \alpha_1(z_n) \epsilon_{n,K}$ is one of the K branches of the multiple-valued function $\alpha(z_n)^{1/K}$. (For example, see Churchill, Brown and Verhey, 1974, p. 88.) Each of the K branches yields an equation of the form $z_n = \alpha_1(z_n) \epsilon_{n,K}$, each of which is a mapping of the domain $\{r > 0, -\pi < \theta < \pi\}$ onto the domain

$$\{ \sqrt[K]{r} = \rho > 0, (2n - 1)\pi/K < \phi < (2n + 1)\pi/K \}, \quad n = 0, \dots, K - 1.$$

Since each of these latter domains can be transformed by an angular rotation to $\{\rho > 0, -\pi/K < \phi < +\pi/K\}$, the RFIX algorithm amounts to solving K completely separate problems with a single root. This is a geometric interpretation of the proof in Section 3 that the algorithm will always

converge to the intended root, whereas the other algorithms are, in fact, solving for the K roots of a single problem, thus allowing for the possibility of "skipped roots" and convergence to undesired roots.

To close out our computational work, we offer pictures of the precise locations of the roots for three of the problems included in Table VIII, namely, Figure 1 for $K = 10, \rho = .1$; Figure 2 for $K = 15, \rho = .9$; and Figure 3 for $K = 30, \rho = .9$.

Figure 1:

*Plot of Roots Inside and On the Unit disk
For $E_{10}/GE_3/1$ Model.*

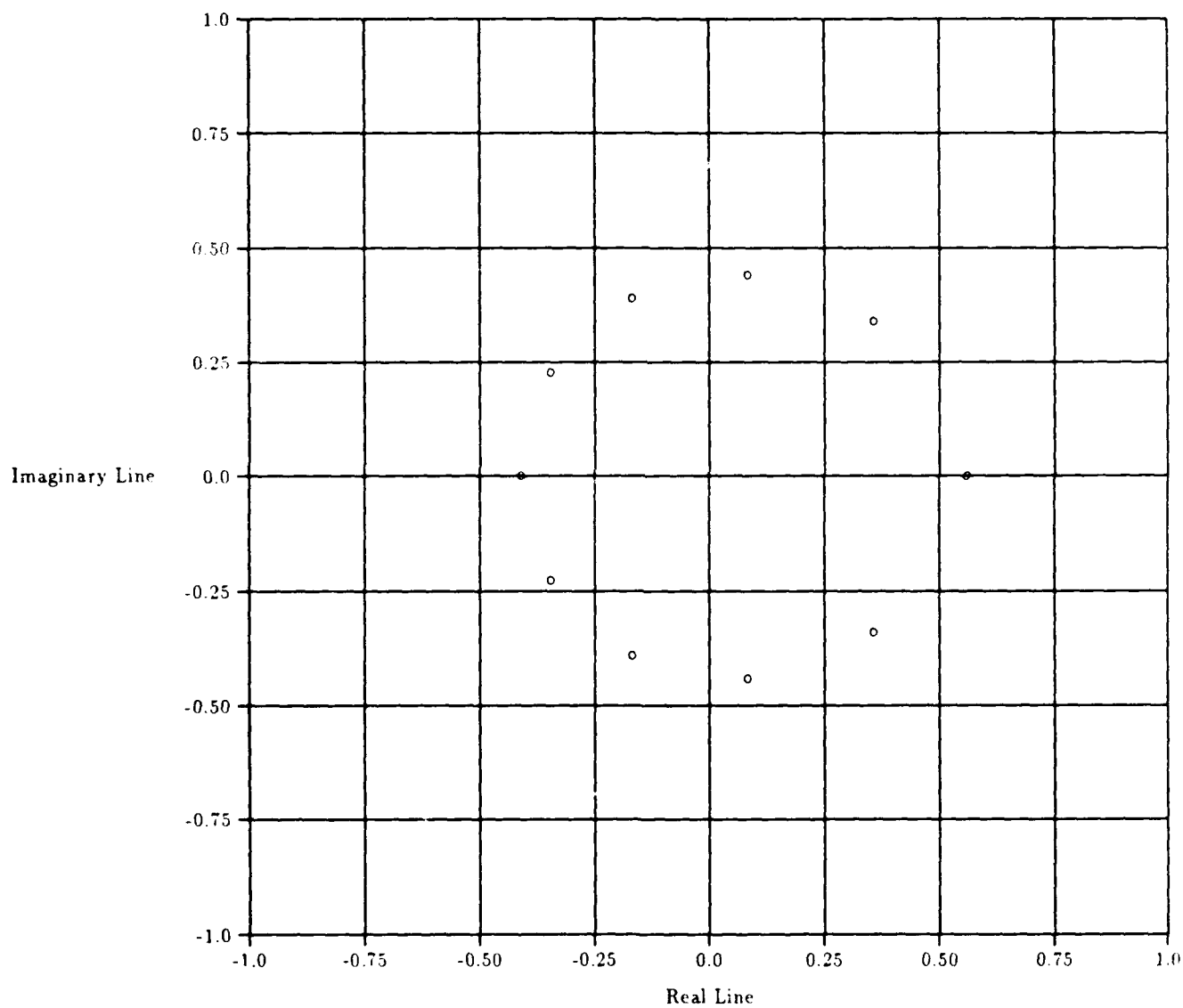


Figure 2:

*Plot of Roots Inside and On the Unit disk
For $E_{15}/GE_3/1$ Model.*

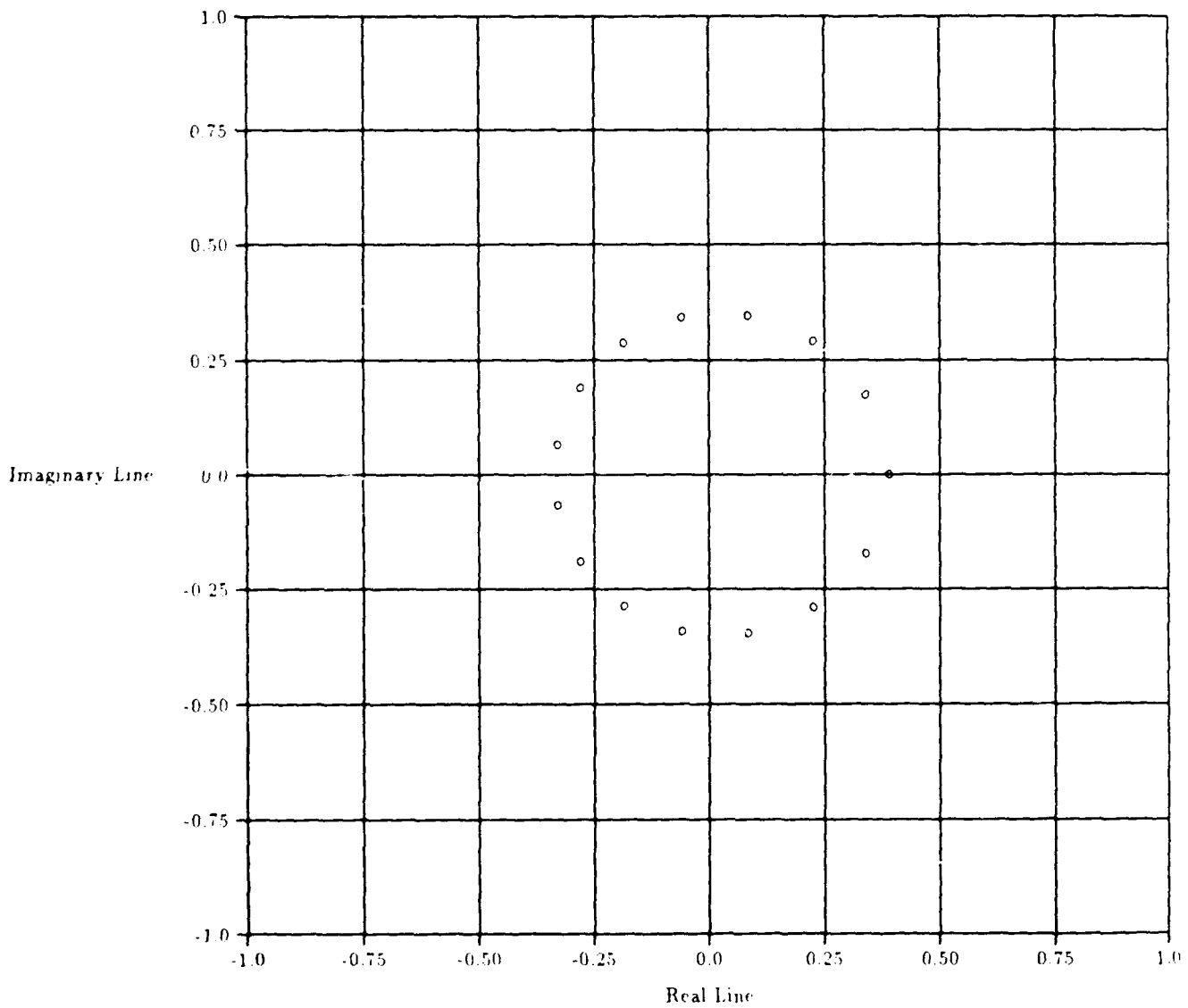
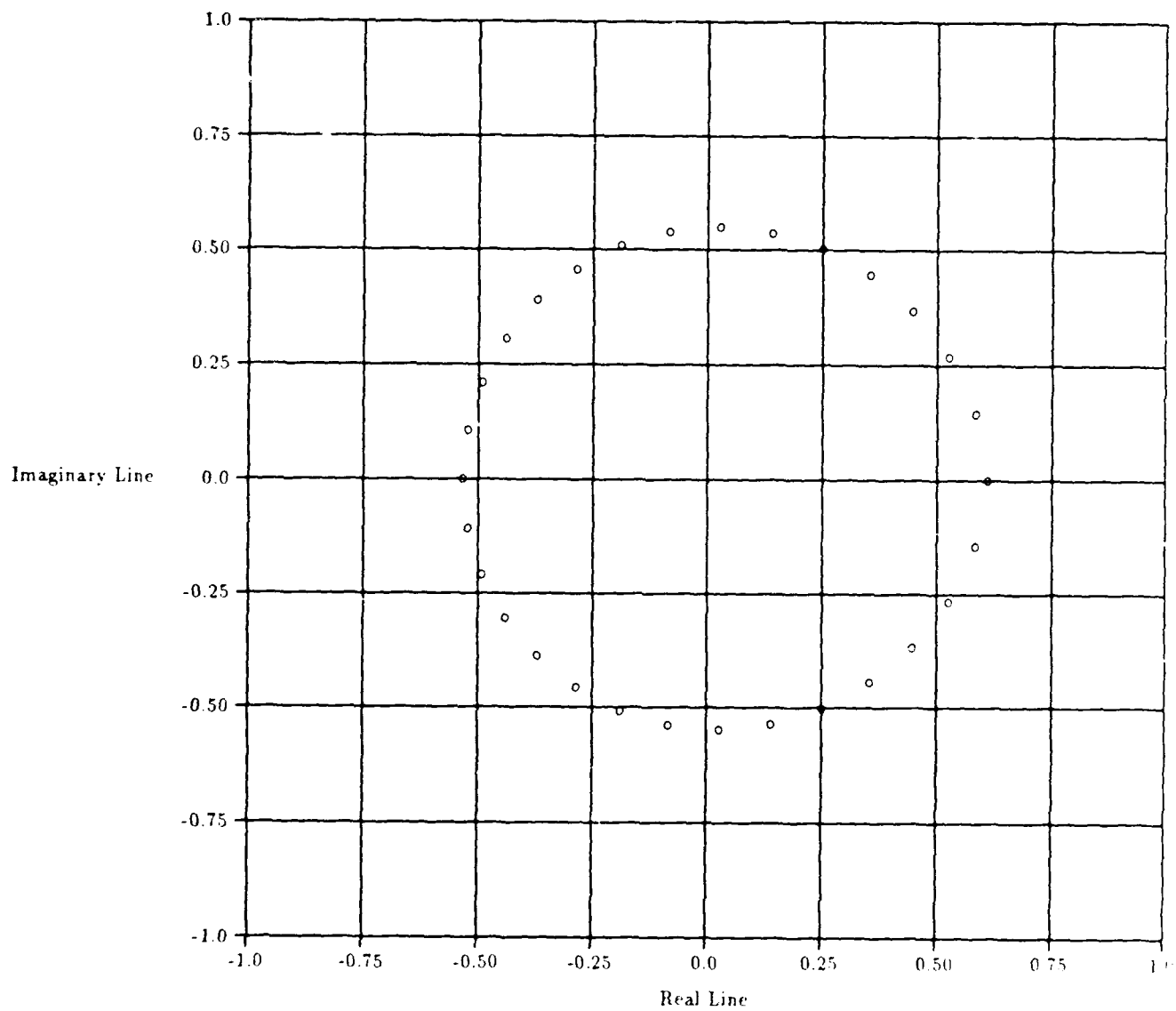


Figure 3:

*Plot of Roots Inside and On the Unit disk
For $E_{30}/GE_3/1$ Model.*



6 CONCLUDING REMARKS

The efficacy of root finding for this problem should help remove the impression that rootfinding in stochastic analysis was fraught with hidden obstacles. Results have indeed been uniformly favorable.

In the process of this work, it was observed over many problems that all three algorithms spent significant time searching for the real root(s) of the problem. In fact, as noted in Table 8, RFIX exhibited convergence problems for queueing systems of the $E_K/G/1$ type. In general, the order of convergence of fixed-point iterations is known to be linear (for example, see Nonweiler and Horwood, 1984), while it is known quadratic for Newton's method and 1.84 (nearly quadratic) for Muller's method.

However, our computational timings suggest a much better general record of performance for our fixed-point approach. As pointed out in Nonweiler and Horwood (1984), the order of convergence only relates to errors and values. But each method (namely, fixed point, Newton's and Muller's) has an associated *computational* burden per iteration, and the simplicity of the fixed-point iteration of RFIX is shown in the illustrative code in Appendix 1. By contrast, the number of operations in the high-level language implementations of M&N's software (Newton's method) and of Chaudhry's QPACK (Muller's method) likely far exceed those of RFIX per iteration; indeed, the sheer size of the KM&N source code (comments excluded) can give some clue as to the fact that execution-time instruction count is several times greater.

What our results show empirically is:

(a) the number of execution-time instructions per iteration is far more dominant than the number of iterations, producing the computational complexity relationship *Fixed Point* < *Newton's Method* < *Muller's Method*; and

(b) the ratio of improvement in speed of convergence (that is, time) between methods is not a constant from problem to problem, particularly with changes in K . That is, if, for a fixed K of any problem type, we look at a row of the appropriate table (namely, I - VIII) and calculate c_1 and c_2 such that $t_{RFX} = c_1 * t_{Newton's}$ and $t_{RFX} = c_2 * t_{Muller's}$, then we will find that over all K in a given problem (that is, table) that c_1 and c_2 would not be constant.

Thus, over all K , we might say that $t_{RFX} = c_1(K) * t_{Newton's}$ and $t_{RFX} = c_2(K) * t_{Muller's}$, where $c_1(K)$ and $c_2(K)$ are nonlinear functions. Thus the

computational stability (that is, consistency of speed of convergence) of our fixed-point method is substantially "more than linear" relative to the other methods in this context.

The fact that this feature was observed over a wide variety of characteristic equations implies that the fixed-point method may be more favorable than other approaches for larger classes of rootfinding problems in probability and statistics.

7 ACKNOWLEDGMENTS

The authors wish to express their sincere thanks to Mohan Chaudhry, Stephen Nash and Jahong Xie for their help in our research.

8 REFERENCES

Abolnikov, L. and Dukhovny, A. (1987). "Necessary and Sufficient Conditions for the Ergodicity of Markov Chains with Transition $\Delta_{m,n}(\Delta'_{m,n})$ -Matrix," *Jl. Appl. Math. & Simul.* 1 13-24.

Bailey, Norman T.J. (1954). "On Queueing Processes with Bulk Service," *Jl. Royal Statist. Soc.* B16 80-87.

Brière, G. and Chaudhry, M.L. (1987). "Computational Analysis of Single-Server Bulk-Arrival Queues," *Queueing Systems* 2, 173-186.

Chaudhry, Mohan L., Harris, C.M. and Marchal, William G. (1989). "Robustness of Rootfinding in Queueing," pending review with *ORSA Jl. on Computing*.

Chaudhry, Mohan L. and Templeton, James G.C. (1983). *A First Course in Bulk Queues*, John Wiley, New York.

Churchill, R.V., Brown, J.W. and Verhey, R.F. (1974). *Complex Variables and Applications*, McGraw-Hill, New York.

Gross, Donald and Harris, Carl M. (1985). *Fundamentals of Queueing Theory*, John Wiley, New York.

Kahaner, D., Moler, C. and Nash, S. (1989). *Numerical Methods and Software*, Prentice-Hall, Englewood Cliffs, NJ.

Nonweiler, T.R.F. and Horwood, Ellis (1984). *Computational Mathematics: An Introduction to Numerical Approximation*. John Wiley, New York.

Powell, Warren B. (1985). "Analysis of Vehicle Holding and Cancellation Strategies in Bulk Arrivals, Bulk Service Queues," *Transportation Science* 19 352-377.

Appendix 1

Illustrative FORTRAN Code for Successive Substitution Algorithm

```
program algor
complex z,e,z1
real e1,e2
r0=.99
pi=3.1415923
do 2 n=0,10
t0=n*pi/10.0
a=r0*cos(t0)
b=r0*sin(t0)
z=cmplx(a,b)
e1=cos(t0)
e2=sin(t0)
e=cmplx(e1,e2)
do 1 i=1,100
z1=e*(1+(20.0/7.0)*(1-z))**(-.5)
write(1,*)z1
a1=real(z1)
a=real(z)
b1=aimag(z1)
b=aimag(z)
if(z.eq.z1)then
goto 3
else
z=z1
endif
1 continue
3 r0=sqrt(a1**2+b1**2)
write(1,*)' '
2 continue
stop
end
```

DISTRIBUTION LIST

Copy No

1	Office of Naval Research 800 North Quincy Street Arlington, VA 22217 Attention: Scientific Officer, Statistics and Probability Mathematical Sciences Division
2	ONR Resident Representative National Academy of Sciences 818 Connecticut Avenue, N.W., Suite 803 Washington, D. C. 20006 - 2702
3 - 8	Director, Naval Research Laboratory Washington, D.C. 20375 Attention: Code 2627
9 - 20	Defense Technical Information Center Building 5, Cameron Station Alexandria, VA 22314
21 - 29	Author
30	GMU Office of Research