STATE UNIVERSITY OF NEW YORK AT STONY BROOK

COLLEGE OF ENGINEERING

REPORT No. 116

OPTIMUM LEAST SQUARES APPROXIMATION WITH EXPONENTIALS

by

S. L. Huang and R. D. Joseph

AUGUST, 1968

OPTIMUM LEAST SQUARES APPROXIMATION WITH EXPONENTIALS

by

S. L. Huang and R. D. Joseph
Department of Applied Analysis
State University of New York at Stony Brook
Stony Brook, N. Y.

ABSTRACT

An improved iterative scheme, which optimizes both poles and residues, is developed for approximating a function by a sum of exponentials. In the pole optimization, the damped least squares Taylor method keeps pole increments small, while pole constraints are employed to insure stability. A more direct matrix inversion technique simplifies the residue optimization.

INTRODUCTION

In a recent paper by Chatterjee and Fahmy[1], an iterative scheme was presented for approximating a desired impulse response f(t) by a finite sum of exponentials of the form $\sum_{i=1}^{n} \alpha_i \exp(-q_i t)$ where α_i and q_i are real and $q_i > 0$. The performance index used was the mean square error, and it was minimized by optimizing both the poles (q_i) and the residues (α_i). At the risk of oversimplification, this iterative procedure may be described as follows. First, the least squares Taylor method[2] is utilized to obtain optimum pole locations with residues fixed a priori. Second, the optimum residues are obtained by using the well-known Kautz orthonormal set[3]. Once an initial pole is chosen judiciously, the iterative cycle continues until the preassigned number of poles and iterations are achieved.

Since this iterative technique optimizes both the poles and the residues it yields a lower performance index than that obtained by the Kautz method [3]. However, inherent in this approach is the lack of control over pole increments; hence, for large numbers of poles unstable solutions ($q_i \leqq 0$) may emerge, and assumptions for the least squares Taylor method may be violated. Moreover, a rigorous proof for the convergence of the iterative technique has not been given. The purpose of this correspondence will be to (1) amend the pole optimization so as to preclude intemperate pole increments and unstable solutions, (2) reformulate the residue optimization so that optimal residues are obtained without making use of the rather involved Kautz orthonormal set, and (3) verify the convergence of this technique.

PROBLEM FORMULATION

For comparision purposes, the nomenclature of Chatterjee and Fahmy[1] will be used in this paper. It is desired to employ an iterative process to approximate f(t) on $L_2[0,\infty)$ by the exponential function

$$g_{\ell,m}(t) = \sum_{i=1}^{n} \alpha_i^{(\ell)} \exp(-q_i^{(m)} t), \tag{1}$$

where ℓ and m denote the respective iterations on the residues and poles. The error functional to be minimized is

$$J[g] = \int_0^\infty [f(t) - g(t)]^2 \, dt . \tag{2}$$

POLE OPTIMIZATION:

If

$$q_i^{(k+1)} = q_i^{(k)} + \Delta q_i^{(k)} \tag{3}$$

and the $\Delta q_i^{(k)}$ are small, a linear approximation may be used so that

$$g_{k,k+1}(t) \approx g_{k,k}(t) + \sum_{i=1}^{n} \frac{\partial g_{k,k}(t)}{\partial q_i^{(k)}} \Delta q_i^{(k)} . \tag{4}$$

Using the damped least squares Taylor method[4] to minimize J[g] as well as the pole increments, $\Delta q_i^{(k)}$, the error criterion

$$E = \int_0^\infty \left[f(t) - g_{k,k}(t) - \sum_{i=1}^{n} \frac{\partial g_{k,k}(t)}{\partial q_i^{(k)}} \Delta q_i^{(k)} \right]^2 dt + \lambda^2 \sum_{i=1}^{n} (\Delta q_i^{(k)})^2 \tag{5}$$

will be used, where λ^2 is a positive constant. This insures that the linear approximation in Eq. (4) remains valid during the iteration.

To minimize E with respect to $\Delta q_j^{(k)}$, the equations

$$\frac{\partial E}{\partial(\Delta q_j^{(k)})} = \int_0^\infty 2\left[f(t) - g_{k,k}(t) - \sum_{i=1}^{n} \frac{\partial g_{k,k}(t)}{\partial q_i^{(k)}} \Delta q_i^{(k)} \right] \left[- \frac{\partial g_{k,k}(t)}{\partial q_j^{(k)}} \right] dt$$

$$+ 2\lambda^2 \Delta q_j^{(k)} = 0 \qquad j = 1, \ldots , n \tag{6}$$

must be satisfied. Rewriting Eq. (6) gives

$$\left(\underset{\sim}{A}^{(k)} + \lambda^2 \underset{\sim}{I} \right) \Delta \underset{\sim}{q}^{(k)} = \underset{\sim}{b}^{(k)} , \tag{7}$$

where $\underset{\sim}{I}$ denotes the nxn identity matrix and $\underset{\sim}{A}^{(k)}$ is an nxn real symmetric matrix with elements

$$a_{ji}^{(k)} = \frac{2\,\alpha_j^{(k)}\,\alpha_i^{(k)}}{(q_j^{(k)} + q_i^{(k)})^3} \quad ; \tag{8}$$

$\Delta\underset{\sim}{q}^{(k)}$ and $\underset{\sim}{b}^{(k)}$ are the column vectors described by

$$\Delta\underset{\sim}{q}^{(k)} = \left[\Delta q_1^{(k)}\,\Delta q_2^{(k)} \quad \cdots \quad \Delta q_n^{(k)}\right]^T , \tag{9}$$

and

$$b_j^{(k)} = \frac{\sum_{i=1}^{n} \alpha_j^{(k)} \alpha_i^{(k)}}{(q_j^{(k)} + q_i^{(k)})^2} - \alpha_j^{(k)} \int_0^{\infty} t f(t) \exp(-q_j^{(k)} t)\, dt ,$$

or

$$b_j^{(k)} = \frac{\sum_{i=1}^{n} \alpha_j^{(k)} \alpha_i^{(k)}}{(q_j^{(k)} + q_i^{(k)})^2} + \alpha_j^{(k)} \frac{d}{ds} F(s)\Big|_{s=q_j^{(k)}} \qquad j = 1,\ldots,n, \tag{10}$$

if f(t) is Laplace transformable.

By inspecting its principal minors, it can be shown that $\underset{\sim}{A}^{(k)}$ is a positive definite matrix; thus $\underset{\sim}{A}^{(k)} + \lambda^2 \underset{\sim}{I}$ is positive definite also. Eq. (7) then yields the solution

$$\Delta\underset{\sim}{q}^{(k)} = (\underset{\sim}{A}^{(k)} + \lambda^2 \underset{\sim}{I})^{-1}\, \underset{\sim}{b}^{(k)} . \tag{11}$$

Now let us inspect the $\Delta\underset{\sim}{q}^{(k)}$ given in Eq. (11). Consider the rate of change of E as $\underset{\sim}{q}^{(k)}$ migrates in the direction of $\Delta\underset{\sim}{q}^{(k)}$, i.e.,

$$\frac{\partial E\left(\underset{\sim}{q}^{(k)} + \beta \Delta \underset{\sim}{q}^{(k)}\right)}{\partial \beta}\Bigg|_{\beta=0} = -2\left[\underset{\sim}{b}^{(k)}\right]^T\left[\underset{\sim}{A}^{(k)} + \lambda^2 \underset{\sim}{I}\right]^{-1}\left[\underset{\sim}{b}^{(k)}\right], \qquad (12)$$

where β is a positive scalar[5]. Since $\left[\underset{\sim}{A}^{(k)} + \lambda^2 \underset{\sim}{I}\right]$ is positive definite, $\left[\underset{\sim}{A}^{(k)} + \lambda^2 \underset{\sim}{I}\right]^{-1}$ is also, so it follows from Eq. (12) that E decreases in the direction $\Delta\underset{\sim}{q}^{(k)}$ defined in Eq. (11); the downhill behavior of E is assured. Since $\underset{\sim}{A}^{(k)}$ is positive definite it is an even simpler matter to show that the least squares Taylor method $(\lambda = 0)$[1] also converges provided the $\Delta q_i^{(k)}$ are small. Note that the minimization of the error criterion E given in Eq. (5) indeed guarantees the minimization of the error functional J[g] given in Eq. (2).

To insure the stability of the approximation, the constraints[6]

$$q_{min} < \underset{\sim}{q}^{(k)} + \Delta\underset{\sim}{q}^{(k)} < q_{max} \qquad (13)$$

are imposed on the minimization procedure. For a stable RC realization q_{min} may be set equal to zero and q_{max} some convenient positive number so that the poles are confined to the negative real axis of the complex plane. Such constraints are only used as a check at the end of each iteration. That is, if a pole is found to violate one of its bounds it is set equal to the extreme value allowed and frozen there for a fixed number of iterations.

RESIDUE OPTIMIZATION:

The pole optimization yields

$$g_{k,k+1}(t) = \sum_{i=1}^{n} \alpha_i^{(k)} \exp(-q_i^{(k+1)} t). \tag{14}$$

Let

$$g_{k+1,k+1}(t) = \sum_{i=1}^{n} \alpha_i^{(k+1)} \exp(-q_i^{(k+1)} t) \tag{15}$$

be the best least squares approximation to $f(t) \epsilon L_2 [0,\infty)$ from among the linear combinations of $h_i = \exp(-q_i^{(k+1)} t)$, which are linearly independent. It is well-known[7] that the optimal residues $\alpha_i^{(k+1)}$ are obtained by solving the normal equations:

$$\alpha_1^{(k+1)}(h_1,h_j) + \alpha_2^{(k+1)}(h_2,h_j) + \cdots + \alpha_n^{(k+1)}(h_n,h_j) = (f,h_j)$$

$$j = 1, \ldots, n. \tag{16}$$

Rewriting Eq. (16) gives

$$\underset{\sim}{H}^{(k+1)} \underset{\sim}{\alpha}^{(k+1)} = \underset{\sim}{d}^{(k+1)}, \tag{17}$$

where $\underset{\sim}{H}^{(k+1)}$ is an nxn real symmetric Gram matrix with elements

$$(h_j,h_i) = \int_0^\infty h_j h_i \, dt = \frac{1}{q_j^{(k+1)} + q_i^{(k+1)}} . \tag{18}$$

$\underset{\sim}{\alpha}^{(k+1)}$ and $\underset{\sim}{d}^{(k+1)}$ are the column vectors described by

$$\underset{\sim}{\alpha}^{(k+1)} = \left[\alpha_1^{(k+1)} \; \alpha_2^{(k+1)} \; \ldots \; \alpha_n^{(k+1)} \right]^T, \tag{19}$$

and

$$d_j^{(k+1)} = \int_0^\infty f(t) \exp\left(-q_j^{(k+1)} t\right) dt,$$

or

$$d_j^{(k+1)} = F(s)\Big|_{s = q_j^{(k+1)}} \qquad j = 1, \ldots, n \tag{20}$$

if $f(t)$ is Laplace transformable.

Since $\underset{\sim}{H}^{(k+1)}$ is a positive definite matrix, Eq. (17) yields

$$\underset{\sim}{\alpha}^{(k+1)} = \left[\underset{\sim}{H}^{(k+1)}\right]^{-1} \underset{\sim}{d}^{(k+1)}. \tag{21}$$

Although the resulting $\underset{\sim}{\alpha}^{(k+1)}$ is same as that obtained from the Kautz orthonormal set, Eq. (21) uses a more direct computation.

Thus, the (k+1) th cycle of the iteration is completed and the performance index is evaluated from

$$\begin{aligned} \text{P.I.} &= \int_0^\infty \left[f(t) - g_{k+1,k+1}(t) \right]^2 dt \\ &= \int_0^\infty \left[f(t) \right]^2 dt - 2\left[\underset{\sim}{\alpha}^{(k+1)} \right]^T \underset{\sim}{d}^{(k+1)} + \left[\underset{\sim}{\alpha}^{(k+1)} \right]^T \underset{\sim}{H}^{(k+1)} \left[\underset{\sim}{\alpha}^{(k+1)} \right]. \end{aligned}$$

If the performance index exceeds the specified tolerable error for the preassigned number of iterations, distant poles are successively added to improve the approximation.

CONCLUSION

A flow chart of the algorithm for computing the optimum approximating function is given in Fig. 1. It incorporates all the refinements presented here. A subsequent paper will cover this material in more detail and discuss a method of optimizing the poles and residues simultaneously; various numerical results will also be given.

REFERENCES

[1] R. N. Chatterjee and M. M. Fahmy, "On time-domain approximation," Proc. 11th Midwest Symp. on Circuit Theory, pp. 288-298, May 1968.

[2] M. R. Aaron, "The use of least squares in system design," IRE Trans. Circuit Theory, vol. CT-3, no. 4, pp. 224-231, December 1956.

[3] W. H. Kautz, "Transient synthesis in the time domain," IRE Trans. Circuit Theory, vol. CT-1, no. 3, pp. 29-38, September 1954.

[4] Denneth Levenberg, "A method for the solution of certain nonlinear problems in least squares," Quarterly of Applied Math., vol. 2, no. 2, pp. 164-168, 1944.

[5] John Todd, Survey of Numerical Analysis. New York: McGraw-Hill, 1962, ch. 7.

[6] R. E. Bellman and S. E. Dreyfus, Applied Dynamic Programming. Princeton, N.J.: Princeton University Press, 1962.

[7] P. J. Davis, Interpolation and Approximation. New York: Blaisdell, 1965, ch. 8.

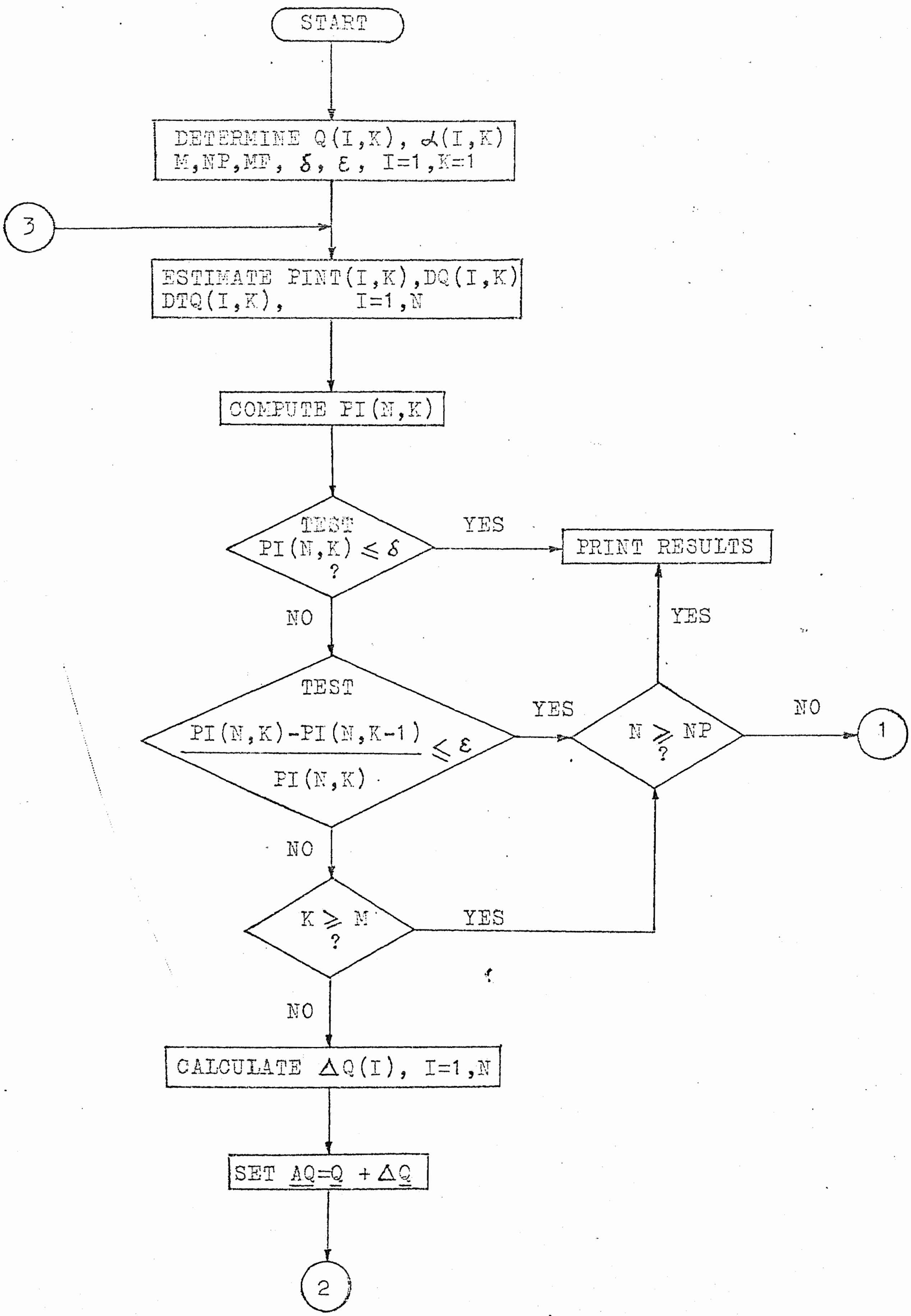
START
DETERMINE Q(I,K), α(I,K)
M,NP,MF, δ, ε, I=1,K=1
3
ESTIMATE PINT(I,K),DQ(I,K)
DTQ(I,K), I=1,N
COMPUTE PI(N,K)
TEST
PI(N,K) ≤ δ
?
YES
PRINT RESULTS
NO
YES
TEST
PI(N,K)-PI(N,K-1) / PI(N,K) ≤ ε
YES
N ≥ NP
?
NO
1
NO
K ≥ M
?
YES
NO
CALCULATE ΔQ(I), I=1,N
SET AQ=Q + ΔQ
2

Fig. 1a

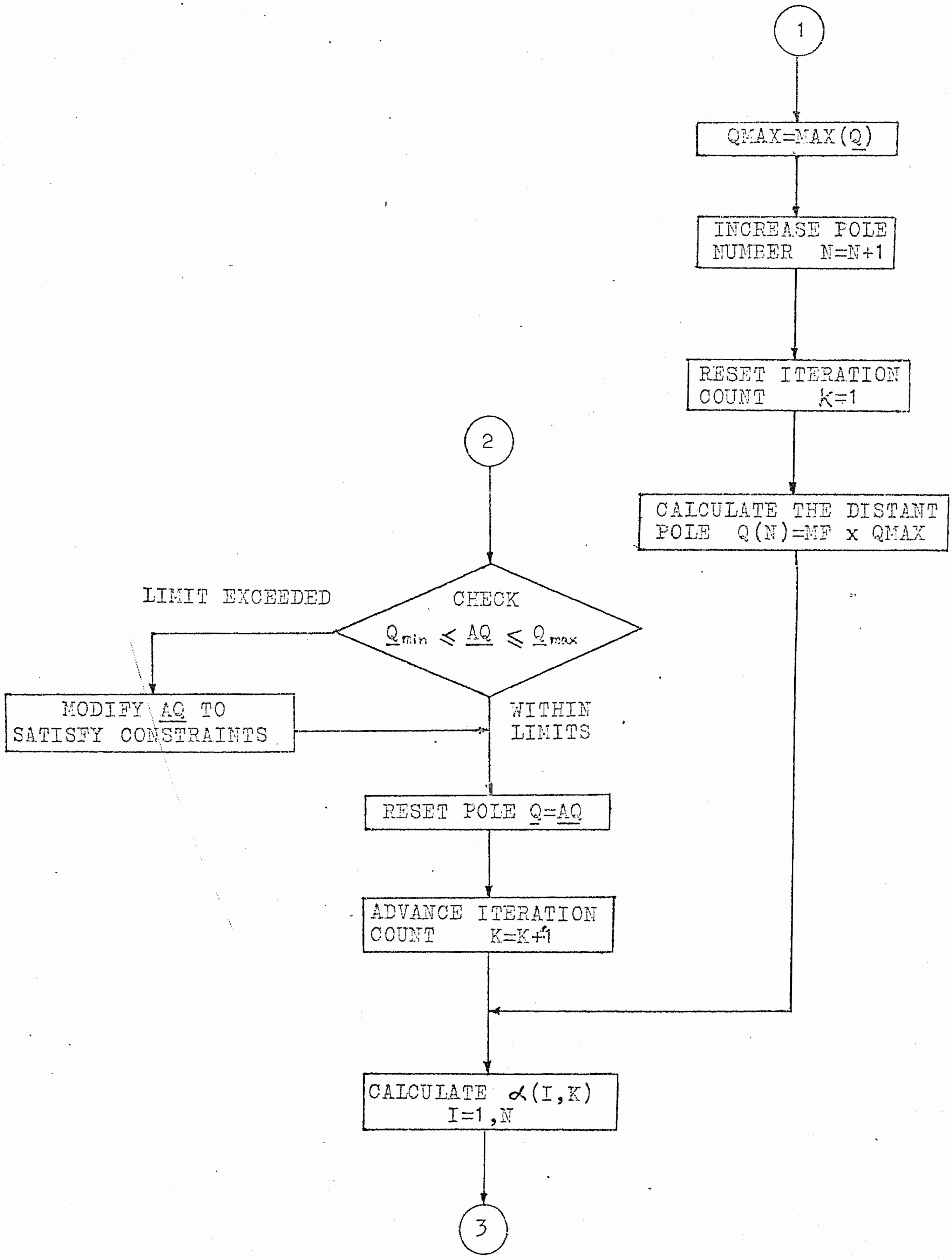
1
QMAX=MAX(Q)
INCREASE POLE NUMBER N=N+1
RESET ITERATION COUNT K=1
CALCULATE THE DISTANT POLE Q(N)=MF x QMAX
2
LIMIT EXCEEDED
CHECK
Qmin ≤ AQ ≤ Qmax
MODIFY AQ TO SATISFY CONSTRAINTS
WITHIN LIMITS
RESET POLE Q=AQ
ADVANCE ITERATION COUNT K=K+1
CALCULATE α(I,K) I=1,N
3

Fig. 1b

FLOW CHART LEGEND

Q = Pole

α = Residue

M = Maximum number of iterations to be performed

NP = Maximum number of poles to be tried

MF = Multiplying factor

δ = Tolerable error

ϵ = Ditto

PI = Performance index

I = Pole count

K = Iteration count

ΔQ = Pole increment

$$PINT = \int_0^\infty \left[f(t) \right]^2 dt$$

$$DQ(I) = \int_0^\infty f(t) \exp(-q_i t)\, dt$$

$$DTQ(I) = \int_0^\infty t f(t) \exp(-q_i t)\, dt$$

Footnotes:

Manuscript received ____________________;

Revised ________________________________.

Work of second author supported by Faculty Research Fellowship from The Research Foundation of the State University of New York.

Department of Applied Analysis

State University of New York

Stony Brook, New York 11790