Point reactor kinetics using Galerkin weighted residuals

by

Joseph Kirk Gasper

A Thesis Submitted to the
Graduate Faculty in Partial Fulfillment of
The Requirements for the Degree of
MASTER OF SCIENCE

Major Subject: Nuclear Engineering

Iowa State University
Ames, Iowa

1971

TABLE OF CONTENTS

I. INTRODUCTION AND LITERATURE SURVEY

An optimal control problem presented by Weaver (1) is investigated. The power level of a reactor with temperature feedback is to be changed a stated amount in a fixed time interval in a manner which would minimize the control rod motion. The analysis is performed utilizing the point reactor kinetics model with one delayed precursor. The temperature feedback is assumed to reduce the reactivity of the system by an amount proportional to the neutron density. Weaver utilized an analog computer to generate a solution for a specific constant of proportionality. This thesis studies the effect of varying this constant of proportionality.

The problem is formulated, as given in Weaver (1) and outlined in Section II part A, by applying Pontryagin's maximum principle to the point kinetics model and results in a coupled set of four first-order nonlinear differential equations with two initial conditions and two final conditions. One of the techniques employed in this thesis to solve this nonlinear boundary value problem is a linearization technique utilized by Kenneth and McGill (2). This technique, which will be referred to as quasilinearization, reduces the set of nonlinear differential equations to a set of linear differential equations which must be solved iteratively to obtain the nonlinear solution. Kenneth and McGill (2) solved linearized problems using numerical

integration techniques. The investigation reported in this thesis uses Galerkin weighted residuals to obtain a linear set of algebraic equations which can be solved by conventional methods. This technique is one of a broad class of methods known as "the method of weighted residuals", which is reviewed in depth by Finlayson and Scriven (3). For the purposes of this thesis, the treatment given in Crandall (4) for the approximate solution of an ordinary linear differential equation is sufficient and is outlined in the following example taken from Crandall.

The equation is

$$\frac{dX(t)}{dt} = -X(t) \qquad 0 < t \leq 1$$

$$X(0) = 1 \quad . \tag{1}$$

A trial solution

$$X(t) = C_0 + C_1 t + C_2 t^2 \tag{2}$$

is proposed with trial functions t and t^2. The parameters C_0, C_1, and C_2 are to be determined. The initial condition is satisfied if $C_0 = 1$ and Equation 2 becomes

$$X(t) = 1 + C_1 t + C_2 t^2 \quad . \tag{3}$$

The next step is to develop a singular criterion for obtaining the "best" approximation with the trial solution.

The substitution of Equation 3 into Equation 2 yields the residual,

$$R(t) = \frac{dX(t)}{dt} + X(t) = 1 + C_1(1+t) + C_2(2t+t^2) \quad . \quad (4)$$

For the exact solution, the residual vanishes over the entire interval $0 \leq t \leq 1$. For the approximate solution the parameters C_1 and C_2 are adjusted so that the residual stays close to zero over the interval $0 \leq t \leq 1$. According to Galerkin's method the weighted averages of the residuals over the desired interval should vanish. The weighting functions are the trial functions that were used to construct the trial solution. Defining the constants in this manner has the effect of making the residuals orthogonal to the weighting functions over the interval of interest. In this case the equations obtained by Galerkin's method are

$$\int_0^1 t\, R(t)dt = \frac{1}{2} + \frac{5}{6}C_1 + \frac{11}{12}C_2 = 0$$

$$\int_0^1 t^2\, R(t)dt = \frac{1}{3} + \frac{7}{12}C_1 + \frac{9}{20}C_2 = 0 \quad . \quad (5)$$

Solving Equation 5 for C_1 and C_2 yields

$$X(t) = 1 - 0.9143t + 0.2857t^2$$

as the approximate solution. Improved approximations are

obtained by increasing the number of trial functions and repeating the procedure. The extension of this technique to handle coupled sets of linear differential equations with mixed boundary conditions is taken up in Section II part B.

Quasilinearization and weighted residual techniques have been applied separately in the field of nuclear engineering. The quasilinearization technique with numerical integration has been used by Stacey (5) to solve a Xenon oscillation problem. Kaplan (6, 7, 8) has applied the method of weighted residuals to flux synthesis problems.

In addition to problem considerations, the investigation includes a study of the solution technique itself. Techniques to extend the range of convergence of the algorithm are considered. Various metrics are compared to determine which provides the earliest and most reliable indication of convergence.

II. THEORETICAL DEVELOPMENT

A. Problem Development

The reactor control problem is formulated so that a power change can be made with minimum control rod motion. The development presented here follows that described by Weaver (1).

The equations which describe the point reactor kinetics model with one delayed precursor group are given by

$$\frac{dx_1(t)}{dt} = \frac{P(t) - \beta}{\Lambda} x_1(t) + \lambda x_2(t) \tag{6}$$

$$\frac{dx_2(t)}{dt} = \frac{\beta}{\Lambda} x_1(t) - \lambda x_2(t) \tag{7}$$

where

$x_1(t)$ = neutron density,

$x_2(t)$ = precursor concentration,

β = precursor fraction,

Λ = mean neutron lifetime,

λ = average precursor decay constant,

$P(t)$ = reactivity.

The reactivity is expressed as

$$P(t) = \rho(t) - \alpha x_1(t) \tag{8}$$

where

$\rho(t)$ = externally applied reactivity,

$\alpha x_1(t)$ = reactivity feedback due to temperature.

Combining Equation 8 and Equation 6 yields

$$\frac{dx_1(t)}{dt} = \frac{1}{\Lambda}[\rho x_1(t) - \alpha x_1^2(t) - \beta x_1(t)] + \lambda x_2(t) \quad . \tag{9}$$

The boundary conditions to be satisfied are given by

$$x_1(0) = n_0 \tag{10}$$

$$x_1(T) = 10n_0 \tag{11}$$

$$x_2(0) = \frac{\beta n_0}{\Lambda\lambda} \tag{12}$$

$$\dot{x}_1(T) = 0 \quad . \tag{13}$$

Equations 10 and 11 arise from the desire to increase the power level by a factor of 10 and the assumption that the power level is proportional to the neutron density. Equation 12 represents the fact that the reactor is at steady state prior to the power change. It is desired that the reactor should remain at the final power level following this power change. This condition is satisfied by Equation 13 provided some reactivity adjustments are made for $t > T$ in order to allow the precursor concentration to attain equilibrium. This can be accomplished, reference (1), with subsequent external reactivity control given by

$$\rho(t) = (\rho_T - \alpha n_T)e^{-\lambda(t-T)} + \alpha n_T \quad t > T \tag{14}$$

where

$\rho_T = \rho(T)$ from nonlinear solution

$n_T = 10n_0$.

Equation 14 represents the solution to Equation 9 and Equation 7 which reduce to an algebraic equation and an ordinary differential equation for $t > T$.

The power change is to be made so that the performance index defined by

$$J = \int_0^T \rho^2(t)dt \tag{15}$$

is minimized. The performance index of Equation 15 places greater weight on larger values of $\rho(t)$ and thus penalizes large control rod movements. Power level changes with minimum disturbances of the neutron flux distribution can be of importance in power reactors where local overheating and possible Xenon oscillations are a problem.

To obtain the control which satisfies Equations 7, 9, and 15 and the boundary conditions given by Equations 10-13, Pontryagin's maximum principle is applied. The Pontryagin H-function is given by

$$H[x(t),\rho(t),u(t)t] = \sum_{j=1}^{n} u_j f_j[x(t),\rho(t)t]$$

$$+ L[x(t),\rho(t),t] \qquad (16)$$

where u(t) is the Lagrange multiplier and n represents the number of equations. For this problem n = 2 and $f_1[x(t),\rho(t),t]$ and $f_2[x(t),\rho(t),t]$ are given by the right hand side of Equation 9 and Equation 7 respectively. $L[x(t),\rho(t),t]$ is given by the integrand of Equation 15. Substituting these expressions into Equation 16 yields

$$H(x,\rho,u,t) = \frac{1}{\Lambda}(u_1\rho x_1 - u_1\alpha x_1^2 - u_1\beta x_1) + u_1\lambda x_2 + u_2\frac{\beta}{\Lambda}x_1$$

$$- u_2\lambda x_2 + \rho^2 \qquad (17)$$

where the explicit time dependence has been dropped for notational simplicity. The next step is to minimize $H(x,\rho,u,t)$ with respect to all admissible control vectors,

$$\frac{\partial H(x,\rho,u,t)}{\partial \rho} = \frac{1}{\Lambda} u_1 x_1 + 2\rho = 0$$

and form the optimal control

$$\rho^0 = \frac{-u_1 x_1}{2\Lambda} \quad . \qquad (18)$$

The optimal H-function is given by

$$H^0(x,u) = H(x,u,\rho^0) \quad . \qquad (19)$$

Hence,

$$H^0(x,u) = \frac{1}{\Lambda}\left(\frac{-u_1^2x_1^2}{4\Lambda} - u_1\alpha x_1^2 - u_1\beta x_1\right) + u_1\lambda x_2 + u_2\frac{\beta}{\Lambda}x_1 - u_2\lambda x_2 \quad . \tag{20}$$

The set of 2n equations

$$\dot{x}_1 = \frac{\partial H^0(x,u)}{\partial u_1} = \frac{1}{\Lambda}\left(\frac{u_1x_1^2}{2\Lambda} + \alpha x_1^2 + \beta x_1\right) + \lambda x_2 \quad , \tag{21}$$

$$\dot{x}_2 = \frac{\partial H^0(x,u)}{\partial u_2} = \frac{\beta}{\Lambda}x_1 - \lambda x_2 \quad , \tag{22}$$

$$\dot{u}_1 = \frac{\partial H^0(x,u)}{\partial x_1} = \frac{1}{\Lambda}\left(\frac{u_1x_1}{2\Lambda} + 2\alpha x_1 + \beta\right)u_1 - \frac{\beta}{\Lambda}u_2 \quad , \tag{23}$$

$$\dot{u}_2 = \frac{\partial H^0(x,u)}{\partial x_2} = -\lambda u_1 + \lambda u_2 \quad , \tag{24}$$

is obtained by taking the appropriate partial derivatives of Equation 20. The final step is the substitution of the solutions of Equations 21-24 into Equation 18 to obtain the optimal control.

The specific problem to be solved in this investigation involves a power change from 10 kw to 100 kw in the period of one second. The reactor physics constants are

$\beta = 0.0064$

$\lambda = 0.1 \text{ sec}^{-1}$

$\Lambda = 10^{-3} \text{ sec}$

and α is varied between 10^{-5} and 9×10^{-5}. The boundary

conditions are given by

$$x_1(0) = 10.0 \tag{25}$$

$$x_1(1) = 100.0 \tag{26}$$

$$\dot{x}_1(1) = 0.0 \tag{27}$$

$$x_2(0) = 640.0 \quad . \tag{28}$$

B. Solution Technique Development

Equations 21-24 along with the boundary conditions given by Equations 25-28 form a nonlinear two point boundary value problem. The solution technique is derived for a general nonlinear two-point boundary-value problem and the results of the application of this method to the optimal control problem are given in the next section. The development that follows will treat a restricted class of problems in which the first I/2 variables, in a set of I equations, are specified at the initial and final time to simplify notation. The procedure for treating more general boundary conditions will then be given.

The general set of I nonlinear ordinary differential equations

$$\dot{X}_i(t) = f_i(X_1,\ldots,X_I,t) \quad , \quad i = 1,\ldots,I \tag{29}$$

are so ordered that the first I/2 variables have known initial and final boundary conditions.

$$X_i(0) = X_{i0} \qquad i = 1,\ldots,I/2$$

$$X_i(T) = X_{iT} \qquad i = 1,\ldots,I/2 \tag{30}$$

Equation 29 is solved by selecting a set of functions $X_i^0(t)$ which may not satisfy either Equation 29 or Equation 30, and successively improving upon these functions. If, on the nth step of this process, the functions $X_i^n(t)$ have been obtained, the improved functions $X_i^{n+1}(t)$ which satisfy

$$\dot{X}_i^{n+1}(t) = f_i(X_1^{n+1},\ldots,X_I^{n+1},t) \qquad i = 1,\ldots,I \tag{31}$$

are approximated by linearizing the right hand side of Equation 30 about $X_i^n(t)$ to obtain

$$\dot{X}_i^{n+1}(t) = f_i(X_1^n,\ldots,X_I^n,t) + \sum_{j=1}^{I} \left.\frac{\partial f_i}{\partial X_j}\right|_{X=X^n} (X_j^{n+1} - X_j^n)$$

$$i = 1,\ldots,I \quad . \tag{32}$$

This equation can be rearranged to form

$$\dot{X}_i^{n+1}(t) = \sum_{j=1}^{I} S_{ij}(X^n,t)X_j^{n+1}(t) + V_i(X^n,t) \tag{33}$$

where

$$S_{ij}(X^n,t) = \left.\frac{\partial f_i}{\partial X_j}\right|_{X=X^n}$$

$$V_i(X^n,t) = f_i(X_1^n,\ldots,X_I^n,t) - \sum_{j=1}^{I} \left.\frac{\partial f_i}{\partial X_j}\right|_{X=X^n} X_j^n \quad .$$

The procedure for obtaining the solution to the set of linear nonhomogeneous differential equations, given in Equation 33, subject to the boundary conditions given in Equation 30, is now considered.

A trial solution X(t) is proposed whereby each component, $X_i(t)$ i=1,...,I, in the trial solutions is composed of a linear combination of trial functions, $\phi_j(t)$ j=1,...,M, plus a constant term. The trial solution is thus of the form

$$\begin{aligned}
X_1(t) &= C_{10} + C_{11}\phi_1(t) + C_{12}\phi_2(t) + \ldots + C_{1M}\phi_M(t) \\
X_2(t) &= C_{20} + C_{21}\phi_1(t) + C_{22}\phi_2(t) + \ldots + C_{2M}\phi_M(t) \\
&\;\;\vdots \\
X_I(t) &= C_{I0} + C_{I1}\phi_1(t) + C_{I2}\phi_2(t) + \ldots + C_{IM}\phi_M(t)
\end{aligned} \qquad (34)$$

where the coefficients from the nth approximation are known and those of the n + 1st approximation are to be determined.

The first I equations of the algebraic system used to determine the unknown coefficients are obtained by requiring the trial solution satisfy the boundary conditions given by Equation 30. The remaining M x I equations are obtained by requiring the residuals given by

$$R_i(t) = \dot{X}_i^{n+1}(t) - S_i(X_i^n,t)X_i^{n+1}(t) - V_i(X_i^n,t) \quad 1 = 1,\ldots,I \tag{35}$$

be orthogonal to each weighting function defined as

$$w_j = \emptyset_j(t) \quad j = 1,\ldots,M$$

on the interval $0 \leq t \leq T$. Mathematically this is accomplished by requiring

$$\int_0^T w_j(t)R_i(t)dt = 0 \qquad \begin{matrix} j = 1,\ldots,M \\ i = 1,\ldots,I \end{matrix} \quad . \tag{36}$$

The physical interpretation that the average of the weighted residual over the entire interval is zero still applies. This set of (I + 1) x M equations is solved for the coefficients of the $X_i^{n+1}(t)$ approximation.

In the computer program that was written to implement this solution technique the integration of the weighted residuals is done by a sixteen point Gaussian quadrature. The solution of the linear system of (I + 1) x M equations is obtained by a Gaussian elimination routine with full pivoting.

The iterative process which evolves from these considerations can be summarized as:

Step 1. Guess a set of functions $X_i^n(t)$, i = 1,...,I.

Step 2. Form the $S_i(X^n,t)$ and $V_i(X^n,t)$ terms of

Equation 33.

Step 3. Form the first I equations of the linear system using the boundary conditions.

Step 4. Form the remaining I x M equations by evaluating the weighted residuals.

Step 5. Solve the (I + 1) x M equations to obtain the coefficients for the new approximation $X_i^{n+1}(t)$.

Step 6. Using the solution obtained in Step 5 repeat Steps 2 through 6. Continue until the process converges.

This summary gives the essential steps of the Galerkin weighted residual (GWR) algorithm which is used to solve the reactor optimal control problem.

Regular polynomials, Legendre polynomials, and Chebyshev polynomials are used as trial functions in the investigation. When regular polynomials are used $\phi_j(t) = t^j$, when Legendre polynomials are used $\phi_j(t) = P_j(t)$ and when Chebyshev polynomials are used $\phi_j(t) = T_j(t)$.

The technique for handling general boundary conditions is now developed. The general form of the boundary condition is

$$h_i(X,\dot{X},t) = 0 \quad , \tag{37}$$

which is evaluated at the proper time. The left hand side of Equation 37 is linearized about X^n to obtain

$$h_i(X^{n+1},\dot{X}^{n+1},t) = h_i(X^n,\dot{X}^n,t) + \sum_{j=1}^{I} \left.\frac{\partial h_i}{\partial X_j}\right|_{X=X^n} (X_j^{n+1} - X_j^n)$$

$$+ \sum_{j=1}^{I} \left.\frac{\partial h_i}{\partial \dot{X}_j}\right|_{\dot{X}=\dot{X}^n} (\dot{X}_j^{n+1} - \dot{X}_j^n) = 0 \quad . \tag{38}$$

The rearranged form of Equation 38

$$\sum_{j=1} \left.\frac{\partial h_i}{\partial X_j}\right|_{X=X^n} X_j^{n+1} + \sum_{j=1}^{I} \left.\frac{\partial h_i}{\partial \dot{X}_j}\right|_{X=X^n} \dot{X}_j^{n+1} = -h_i(X^n,\dot{X}^n,t)$$

$$+ \sum_{j=1}^{I} \left.\frac{\partial h_i}{\partial X_j}\right|_{X=X^n} X_j^n + \sum_{j=1}^{I} \left.\frac{\partial h_i}{\partial \dot{X}_j}\right|_{X=X_n} \dot{X}_j^n \quad , \tag{39}$$

represents a linear equation. If $h_i(X,\dot{X},t)$ is a linear function of X and $\dot{X}$, the boundary condition is satisfied exactly. However, if it is a nonlinear function then the boundary condition is only approximately satisfied until $X_i^{n+1} = X_i^n$ and the method has converged.

A study of the convergence of quasilinearization methods is given in Leondes and Paine (9, 10). This study includes a technique to increase the range over which the algorithm will converge by decreasing the size of the step taken between X^{n+1} and X^n approximations. This technique forms the new $X_i^{n+1}(t)$ approximation from the $\tilde{X}_i^{n+1}(t)$ and $X_i^n(t)$ approximations which are computed by the usual process. The new approximation is

$$X_i^{n+1}(t) = X_i^n(t) + \varepsilon[\tilde{X}_i^{n+1}(t) - X_i^n(t)] \tag{40}$$

where epsilon is varied between zero and one. Physically the quantity in the brackets represents the direction of search and epsilon represents the step size. The reduction of epsilon can be viewed as an attempt to insure the validity of the linearization process. Leondes and Paine (9, 10) have shown theoretically and practically that this method extends the range of convergence. This modification adds another step to the iterative process.

Step 5.5. Form the modified $X_i^{n+1}(t)$ approximation according to Equation 40.

An alternative method of extending the range of convergence is given by Bryson and Ho (11). They suggest a formulation that replaces Equation 31 with

$$\dot{X}_i^{n+1}(t) - f_i(X_1^{n+1},\ldots,X_I^{n+1},t)$$

$$= (1-\varepsilon)[\dot{X}_i^n - f_i(X_1^n,\ldots,X_I^n,t)] \qquad (41)$$

where $X^{n+1}(t)$ comes closer to satisfying Equations 29 and 30 than did $X^n(t)$ and $0 < \varepsilon \leq 1$. The reduction of epsilon can again be viewed as reducing the step size to insure the validity of the linearization process. If $\varepsilon = 1$ Equation 41 reduces to Equation 31 and a full step is taken. Defining

$$X_i^{n+1}(t) = X_i^n(t) + \delta X_i^{n+1}(t)$$

and linearizing $f_i(X_1^{n+1},\ldots,X_I^{n+1},t)$ about $X_i^n(t)$ yields

$$\frac{d}{dX}\left[\delta X_i^{n+1}(t)\right] - \sum_{j=1}^{I} \left.\frac{\partial f_i}{\partial X_j}\right|_{X=X^n} \delta X_i^{n+1}(t)$$

$$= -\varepsilon\left[\dot{X}_i^n(t) - f_i(X_1^n,\ldots,X_I^n,t)\right] \quad . \tag{42}$$

The use of this formulation requires the initial approximation to satisfy the problem boundary conditions exactly. The boundary conditions, corresponding to Equations 25-28, used in the iterative process with this method of quasi-linearization are

$$\delta x_1(0) = 0$$

$$\delta x_1(1) = 0$$

$$\delta \dot{x}_1(1) = 0$$

$$\delta x_2(0) = 0 \quad . \tag{43}$$

When this formulation is used Equation 42 replaces Equation 33 in Step 2 of the iterative process and epsilon is lowered from one until the process converges.

III. RESULTS AND DISCUSSION

A. Problem Results

Solution approximations to the reactor optimal control problem given by Equations 21-28 with varying amounts of temperature feedback are obtained with the aid of a computer program written to implement the GWR algorithm. Solutions obtained from a quasilinearization algorithm similar to that suggested by Kenneth and McGill (2) but modified to handle generalized boundary conditions in the manner suggested by Lewallen (12) comprise an independent check of the approximations obtained. These solutions will be referred to as the accepted solutions in the discussion which follows.

The neutron density and external reactivity approximations obtained with the GWR algorithm using regular polynomial functions are shown in Figures 1 and 2 for several values of alpha. The accepted solutions are also shown and the GWR approximations agree with these solutions to four significant figures. The apparent effect of increasing alpha is to increase the nonlinearity of the problem. This is illustrated by the fact that ten trial function approximations are needed to obtain the stated agreement with the accepted solutions for $\alpha = 8 \times 10^{-5}$ and $\alpha = 9 \times 10^{-5}$; whereas eight trial functions or less are sufficient for smaller values of alpha.

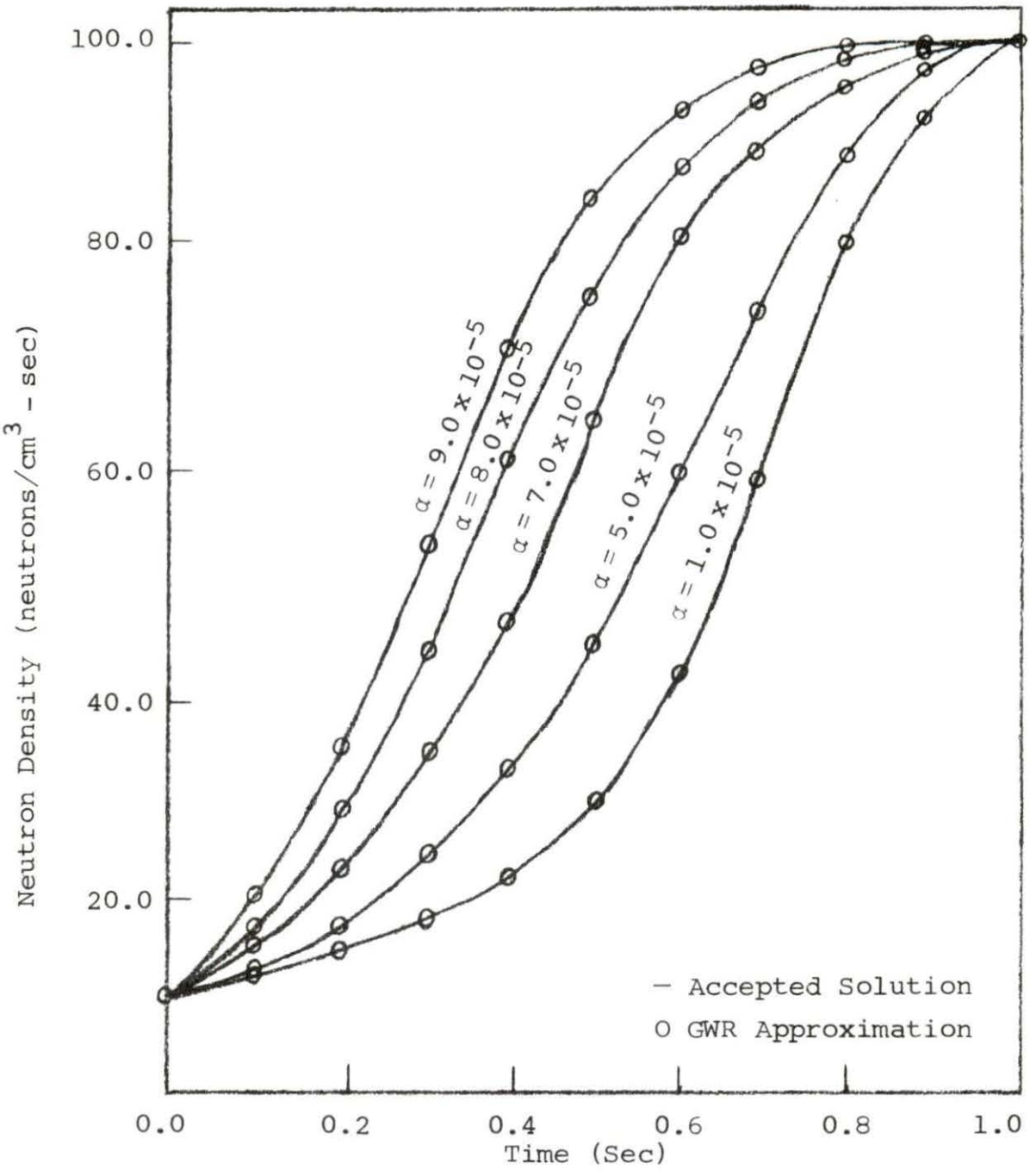

Figure 1. Neutron density time histories for specific amounts of temperature feedback

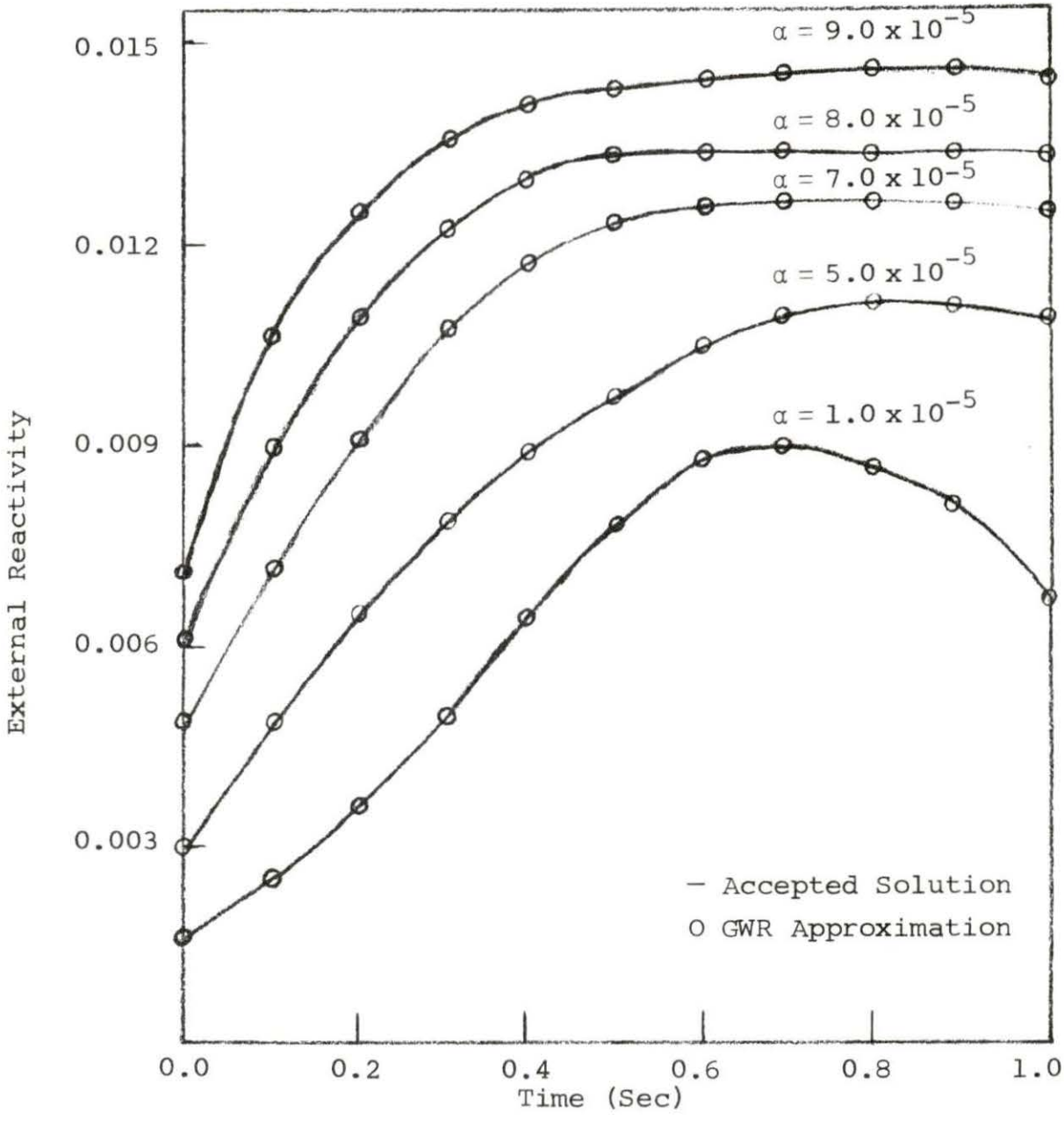

Figure 2. External reactivity time histories for specific amounts of temperature feedback

Figure 2 shows that as alpha is increased the reactivity must be increased more rapidly and to a higher level to compensate for the increased temperature feedback. This in turn causes the neutron density to rise more rapidly as shown in Figure 1. The lack of a pronounced peak in the external reactivity curves for higher values of alpha is explained by considering the solution for $t > 1$ second when the neutron density is maintained at a constant final level and the system becomes linear. The $\alpha = 9 \times 10^{-5}$ neutron density curves approach this constant final level much more rapidly than do the $\alpha = 10^{-5}$ neutron density curves. Thus the reactivity approaches its asymptotic behavior, given by Equation 14, more rapidly. The total reactivities, given by Equation 8, corresponding to these power level changes are shown in Figure 3. As alpha increases the total reactivity peaks earlier and the slope of the curve approaches zero as t approaches 1.0. It is interesting to note that the total reactivity goes through a minimum at approximately $\alpha = 5 \times 10^{-5}$.

The solution approximations for $\alpha = 10^{-5}$ obtained by these two algorithms and the plots of the analog solution obtained by Rosztoszy, Sage and Weaver (13) are shown in Figures 4 and 5. The results obtained by the GWR algorithm and Kenneth and McGill quasilinearization algorithm are in agreement but they differ from Weaver's solution. The reason for this discrepancy is not known.

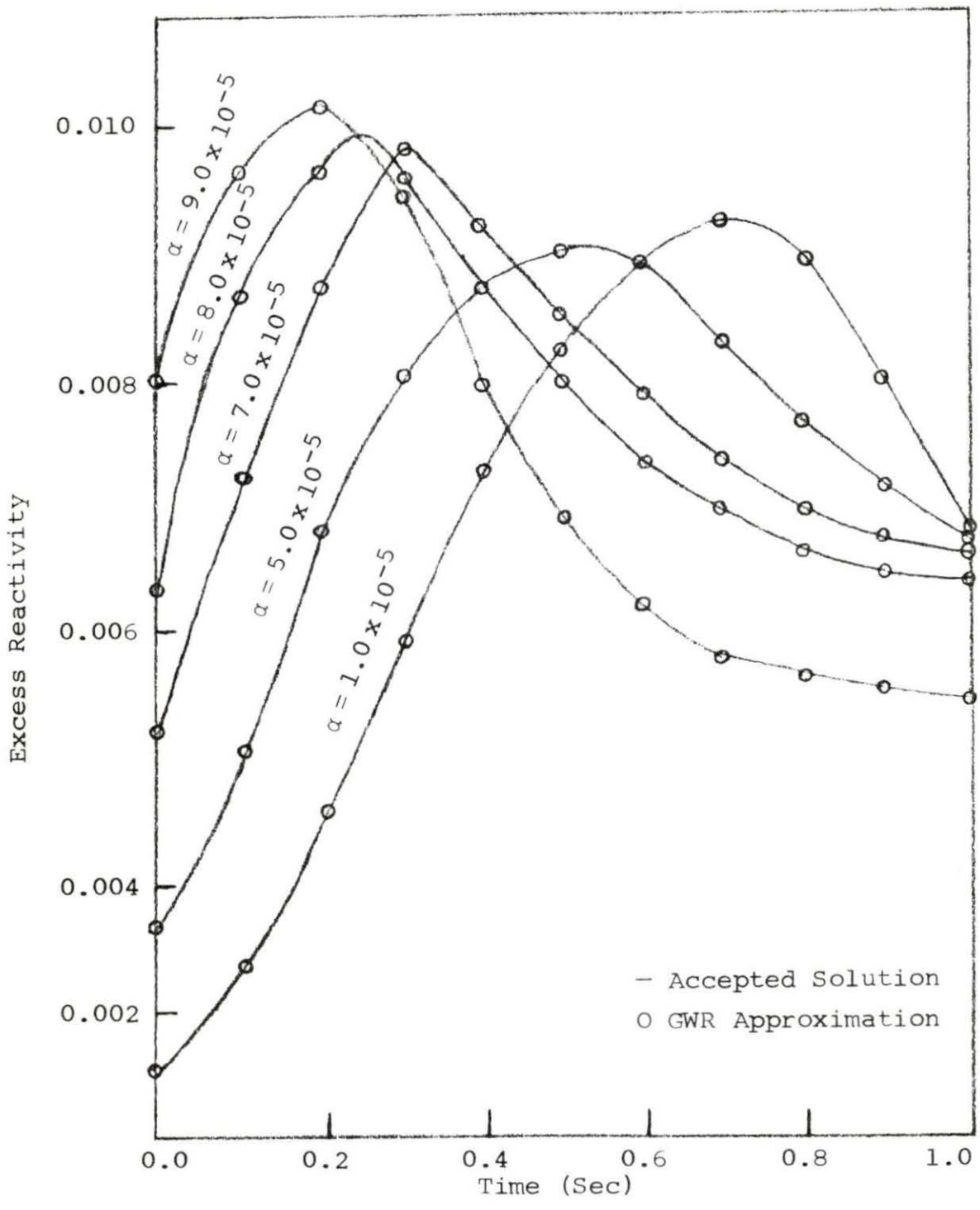

Figure 3. Excess reactivity time histories for specific amounts of temperature feedback

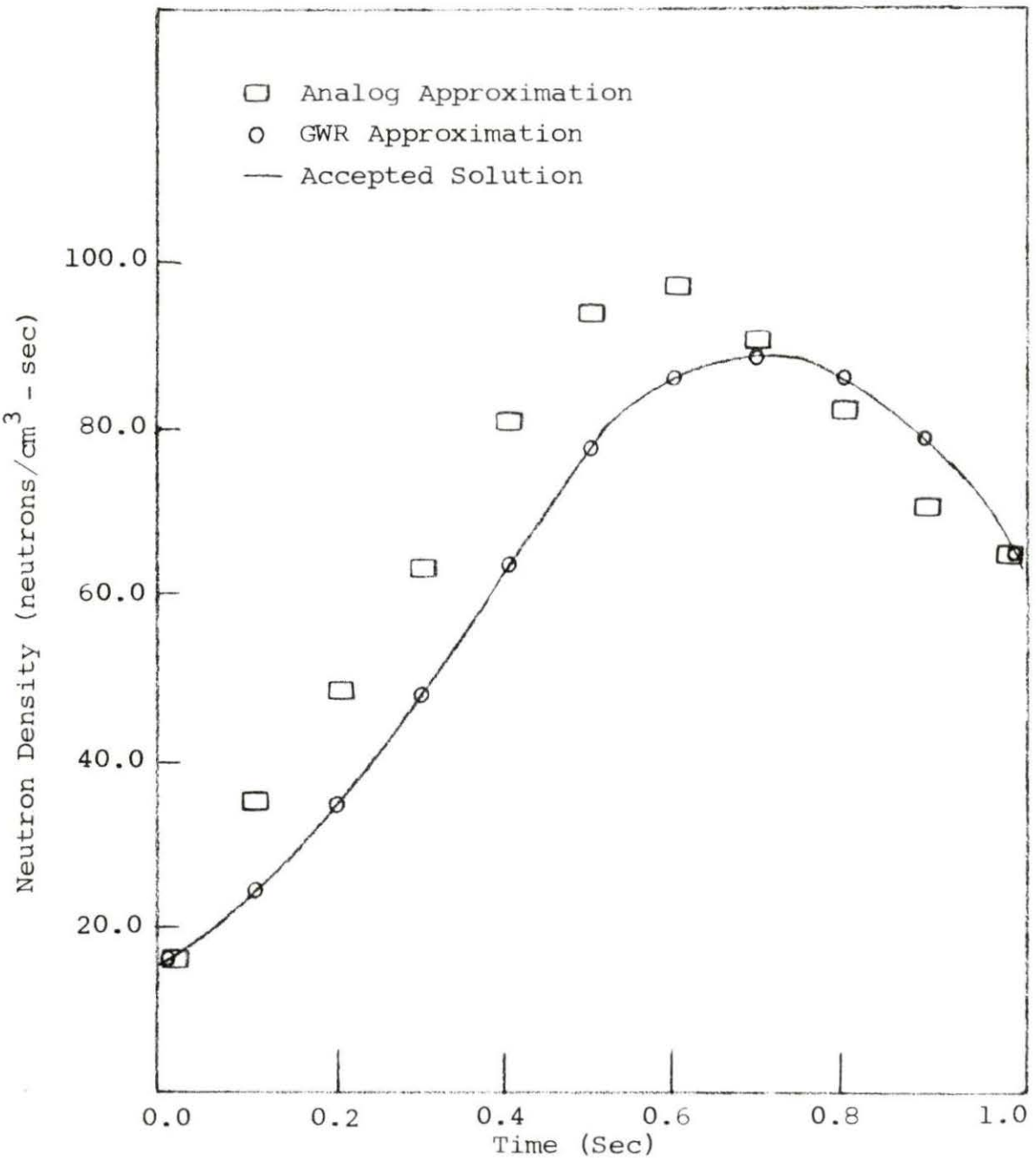

Figure 4. The GWR approximation, accepted solution, and the approximation given in Weaver (1) for the neutron density with $\alpha = 10^{-5}$

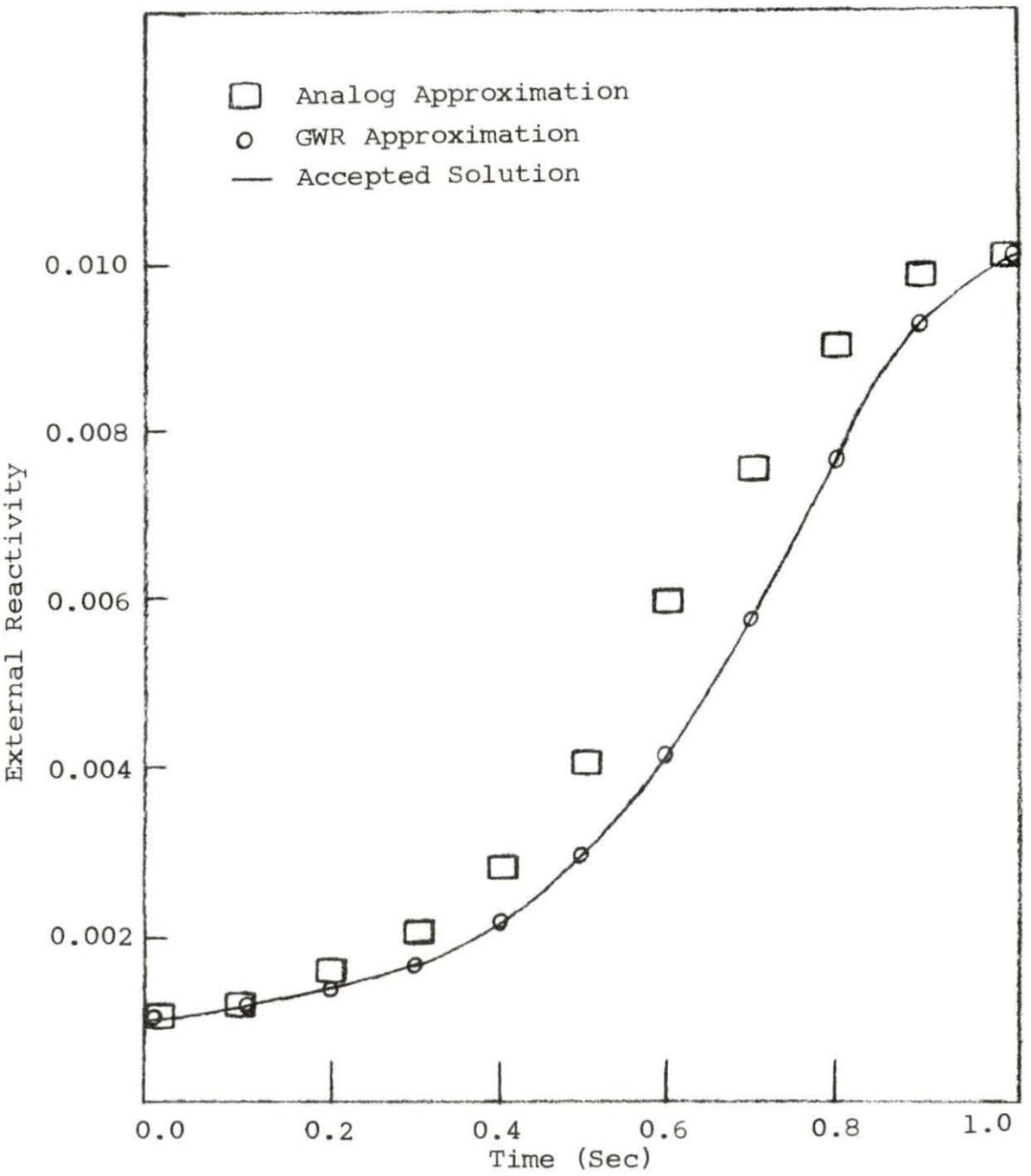

Figure 5. The GWR approximation, accepted solution, and the approximation given in Weaver (1) for the external reactivity with $\alpha = 10^{-5}$

The GWR approximations in Figures 1 and 2 are obtained from a four trial function $\alpha = 10^{-5}$ approximation, which is started with initial trial solutions

$$x_1 = 10 + 180t - 90t^2$$

$$x_2 = 640$$

$$u_1 = -10^{-6}$$

$$u_2 = -10^{-7} \quad .$$

Neither two nor three trial function approximations for $\alpha = 10^{-5}$ are accurate enough to be used to start higher number of trial function approximations. The time histories of the four trial function approximation are used, in turn, to start a five trial function $\alpha = 10^{-5}$ approximation and this process is continued until the eight trial function approximation is obtained. The results of each step of this process are shown in Figure 6. As expected, the approximations improve as the number of trial functions increases.

The four-trial function $\alpha = 10^{-5}$ approximation can also be used to start a four-trial function $\alpha = 2 \times 10^{-5}$ approximation. This in turn can be used to start a four-trial function $\alpha = 3 \times 10^{-5}$ and four trial function approximations up to $\alpha = 8 \times 10^{-5}$ can be obtained by using this process. Larger alpha value approximations cannot be obtained using four trial functions. The four-trial function

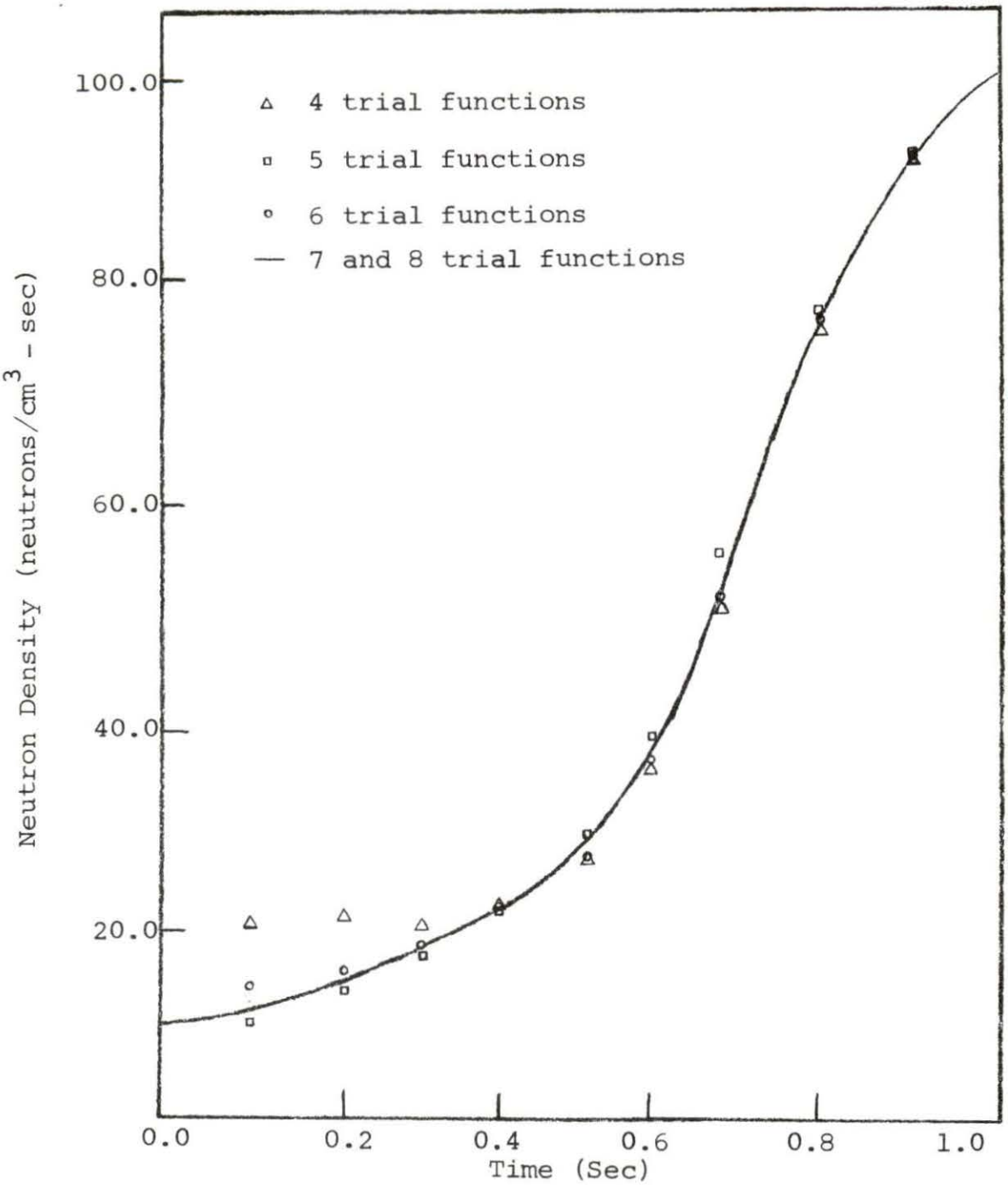

Figure 6. The neutron density GWR approximations for $\alpha = 10^{-5}$ with 4, 5, 6, 7 and 8 trial functions

approximations can be extended to eight trial functions for values of alpha up to 5×10^{-5}. The four-trial function $\alpha = 6 \times 10^{-5}$ approximation is not close enough to the accepted solution to start larger trial function approximations. This problem is circumvented by using the eight-trial function $\alpha = 5.0 \times 10^{-5}$ approximation to start an eight-trial function $\alpha = 5.2 \times 10^{-5}$ approximation and this process of stepping out the previous solution by $.2 \times 10^{-5}$ is continued until the $\alpha = 9.0 \times 10^{-5}$ approximation is obtained. The ten-trial function $\alpha = 8 \times 10^{-5}$ and $\alpha = 9 \times 10^{-5}$ approximations are obtained by stepping out the appropriate eight-trial function approximation.

B. Solution Technique Studies

1. Range of convergence

The discussion in the previous section indicated that it is desirable to extend the range over which the GWR algorithm will converge. Methods suggested by Bryson and Ho (11) and by Leondes and Paine (9,10) (Section 2, part B) are applied to the problem in this section. These techniques used the parameter ε to slow down the rate of convergence and to extend the range of convergence. The solutions obtained in the part A of this section were obtained with an epsilon of one and the algorithm was taking whole steps between iterations. Table 1 shows the effect

Table 1. Four-trial function solutions for various alpha values using $\alpha = 10^{-5}$ for a starting solution

ε	$\alpha \times 10^{-5}$	Number of Iterations
1.0	2	3
1.0	3	4
1.0	4	4
1.0	5	Diverged
0.9	4	8
0.9	5	9
0.9	6	9
0.9	7	Diverged
0.8	6	10
0.8	7	12
0.8	8	Diverged
0.7	7	13
0.7	8[a]	15

[a] A four-trial function solution could not be found for $\alpha = 9 \times 10^{-5}$.

of varying epsilon in the Bryson and Ho quasilinearization technique for extending a four trial function $\alpha = 10^{-5}$ approximation to larger values of alpha. These results show that this formulation does extend the range of convergence by reducing epsilon from one. It should be noted that the

largest value of epsilon that will allow convergence usually required the fewest iterations.

Although the results shown in Table 1 are obtained by the Bryson and Ho technique, similar results are obtained using the standard quasilinearization method and the Leondes and Paine method of extending the range of convergence. This technique, as programmed, requires less storage space and less time per iteration than did the Bryson and Ho technique. For this reason the Bryson and Ho technique is not used for any other work in this investigation.

Figure 7 shows the approximate convergence bounds of the algorithm using the Leondes and Paine formulation for an eight-trial function solution. The ordinate is the maximum absolute difference in the neutron density time histories between the accepted solution for a given alpha value, and the accepted solution for $\alpha = 10^{-5}$ which is used as the starting solution for the larger alpha values. The abscissa is the maximum absolute difference in the neutron density time histories between x_1^5 and x_1^4. This method greatly increases the range of convergence for this problem. An epsilon value of .25 has a convergence range approximately seven times greater than the convergence range of an epsilon value of 1.

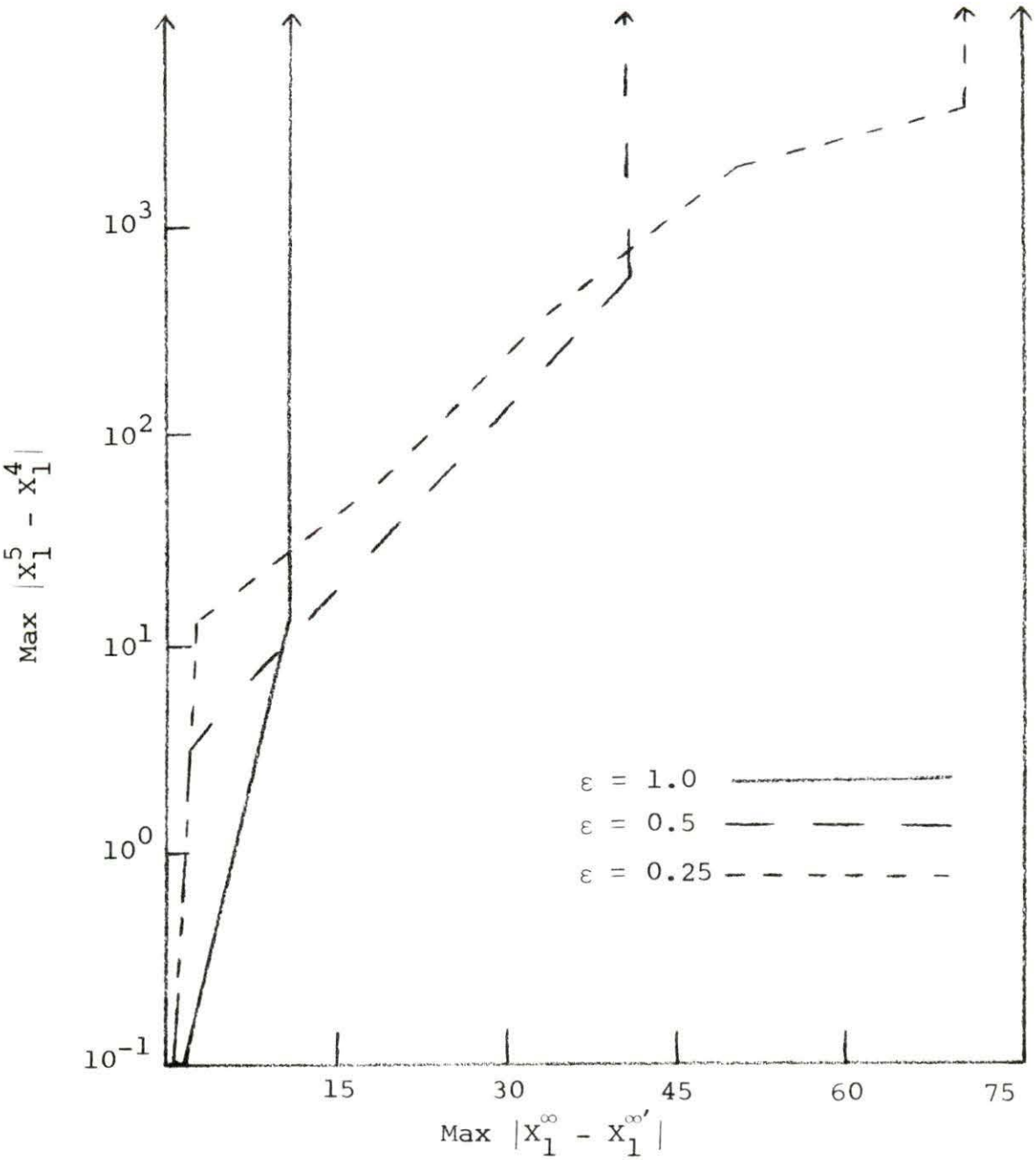

Figure 7. The convergence bounds of the GWR algorithm for eight-trial function approximations. (X_1^{∞} is the accepted solution for a specific value of alpha and $X_1^{\infty'}$ is the accepted solution for $\alpha = 10^{-5}$)

2. Convergence metrics

The development of a convergence metric that will provide an early indication of the convergence of the GWR algorithm is an important part of this investigation. This metric should tell the algorithm user if the algorithm can be expected to converge after a few iterations. A metric is needed because it is possible to determine very little about the convergence of the algorithm by viewing successive time histories.

A comparison of four metrics is made to determine one that has the desired characteristics. The first metric is the sum of the squares of the differences of the time histories between successive iterations:

$$P_1 = \sum_{j=1}^{I} \sum_{k=1}^{16} (X_{jk}^{n+1} - X_{jk}^{n})^2$$

where the subscript j refers to the variable number, the subscript k refers to the sixteen Gaussian quadriture points and the superscript n refers to the iteration number. The second metric

$$P_2 = \sum_{j=1}^{I} \sum_{k=1}^{16} \left(\frac{X_{jk}^{n+1} - X_{jk}^{n}}{X_{jk}^{n+1}}\right)$$

can only be used if X is not zero at any point in the interval. The next metric

$$P_3 = \sum_{j=1}^{I} \int_0^T [\dot{X}_j^n - f(X_j^n, t)]^2 \, dt$$

can be viewed as a measure of the degree to which X^n satisfies the differential equation. It should be noted that $\dot{X}^n$ is computed directly by this algorithm; whereas other quasilinearization algorithms must use numerical techniques to determine $\dot{X}^n$. The final metric

$$P_4 = \sum_{j=1}^{I} \max |X_j^{n+1} - X_j^n|$$

is similar to the one used by Leondes and Paine (9) in their proof of the convergence of quasilinearization methods.

Figure 8 shows a comparison of the four metrics for a representative run of the GWR algorithm. The plots of all these metrics eventually form a straight line when the value of the metric is plotted versus iteration number on a semi-logarithmic plot and this phenomena will be referred to as exponential convergence.

The metrics P_1 and P_2 both show exponential convergence during the later iterations. However, P_1 has very large values during the initial iterations and rapidly reduces to a value similar to the other metrics. This reduction may occur at any time during the iterative process and limits the usefulness of this metric as a predictive parameter.

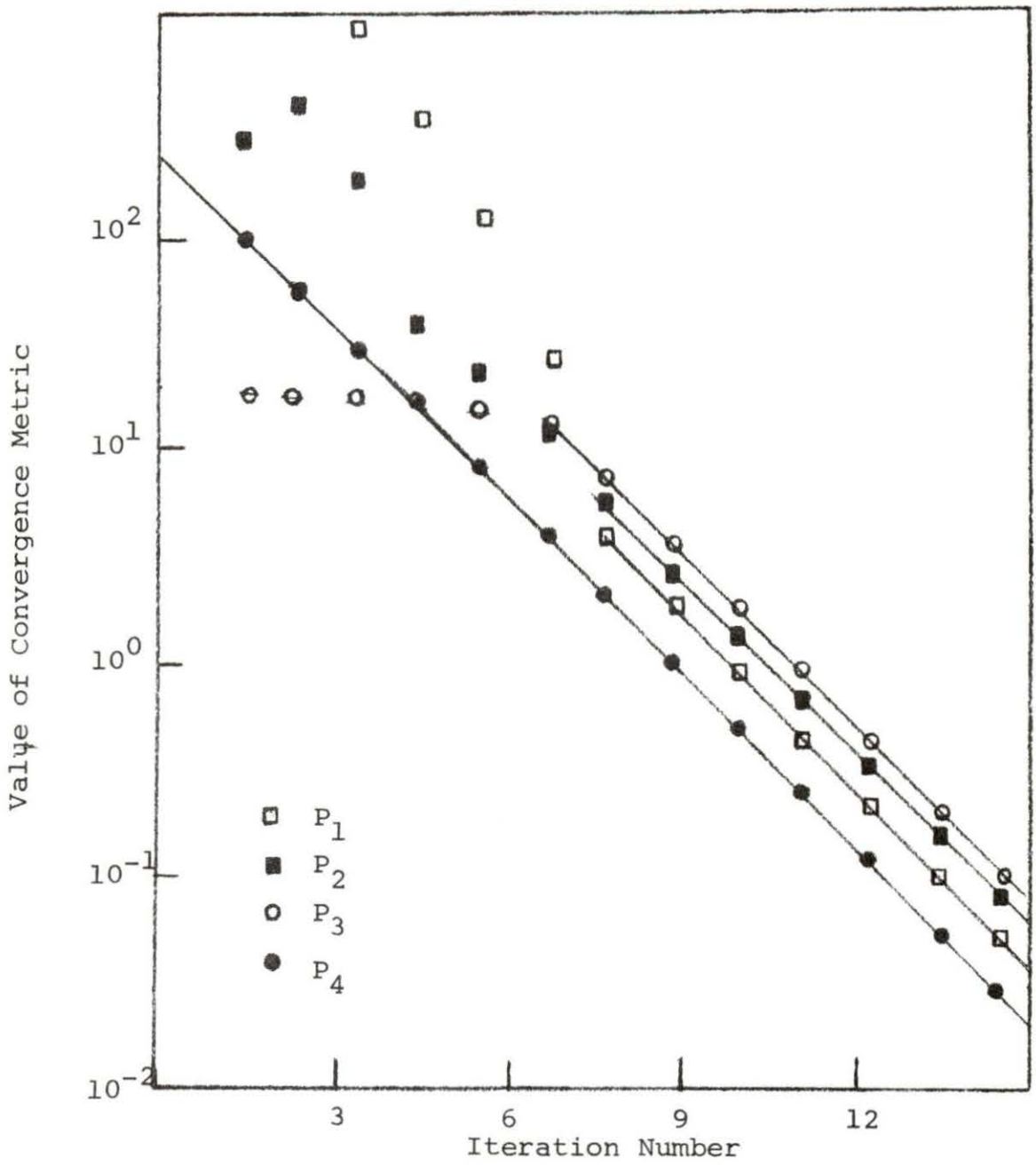

Figure 8. The four metrics during the convergence of the GWR algorithm

The metric P_2 is also limited in its usefulness as a predictive parameter because it is too erratic during the initial iterations.

The metric P_3 also shows exponential convergence during the later iterations but changes very little during the initial iterations. P_3 does show a steady decline, though small, during the initial iterations and if P_3 is monotonically decreasing the algorithm is usually converging. However, P_3 may fluctuate during the initial iterations and it may be difficult to determine if P_3 is decreasing.

P_4 shows exponential convergence in the first few iterations and continues this pattern until the algorithm converges. This indicates that the algorithm is converging toward a solution after a few iterations. Other runs indicate that this metric might fluctuate during the initial iterations but that it would start converging exponentially after a few iterations if the problem is going to converge to a solution.

The results of this study show that the metric P_4 shows the best predictive qualities for this problem. All of the metrics do show the desired predictive characteristics but P_4 consistently shows these characteristics earlier and appears to be the most reliable metric for this problem.

3. Miscellaneous considerations

Two additional modifications of the algorithm are briefly considered. A method of scaling the variables so they are on the interval (-1, 1) is considered and the use of Legendre and Chebyshev polynomial trial functions is discussed.

In order to improve the solution of the system of linear algebraic equations it is desirable to keep the variables on the (-1, 1) interval. This can be done by choosing appropriate scale factors. If the approximate maximum absolute value of each variable can be obtained from physical considerations the scaling factors are easily defined. However, in this problem there are no physical constraints on the Lagrangian multipliers and constant scaling factors are more difficult to estimate. A variable scaling technique which uses the X^n approximation to determine the scale factors for X^{n+1} approximation is used with this problem. Scaling the problem enables the Gaussian elimination routine to obtain a better inversion of the coefficient matrix, but there is no noticeable difference between the time histories obtained by scaled and unscaled versions of the algorithm. Williams (14) has shown that keeping the variables on the (-1, 1) interval increases the range of convergence of numerical integration routines. For the GWR algorithm no differences in the convergence

characteristics of scaled and unscaled versions are observed. However, this study did not do sufficient research in this area to justify any conclusions.

Regular polynomials were used in the GWR algorithm because they were simple to program. However, Legendre and Chebyshev polynomials have properties that could be useful when used with this algorithm. Legendre polynomials were tried as trial functions in attempt to better condition the algebraic system. However, a column of zeros was produced in the coefficient matrix and this problem could not be overcome without substantially altering the computer program. Chebyshev polynomials were tried because they approximate a function with fewer terms than any other series of polynomials. The solutions obtained with Chebyshev polynomial trial functions were identical to those obtained by the algorithm using an equal number of regular polynomial trial functions. The algorithm which used Chebyshev polynomials required more time per iteration and for this reason Chebyshev polynomials were not used.

IV. SUMMARY AND CONCLUSIONS

This investigation shows that the GWR algorithm can be a useful tool for the solution of coupled sets of nonlinear differential equations. The solutions to the reactor optimal control problem obtained with this algorithm agreed to four significant figures with the accepted solutions obtained by a standard quasilinearization technique. This algorithm's ability to handle derivative boundary conditions directly may be advantageous in the solution of some problems. The algorithm has the disadvantage of requiring large amounts of computer space. There is not enough information in this study to establish the usefulness of the GWR algorithm with respect to other quasilinearization algorithms.

The study of various additions to the original algorithm show that some of these additions improve the algorithm. The convergence range study shows that the range of the algorithm can be substantially improved by slowing the rate of convergence. This increase in the range of convergence was observed with both the Leondes and Paine technique, and the Bryson and Ho quasilinearization method. However, the algorithm, as programmed, required more computer storage space when the Bryson and Ho method was used. The convergence metric study showed that while all the metrics had some

predictive capabilities the metric

$$P_4 = \sum_{j=1}^{I} \max |X_j^{n+1}(t) - X_j^n(t)|$$

most reliably predicted the convergence of the GWR algorithm for this problem. It is possible that the other metrics may be more reliable for other problems. Legendre and Chebyshev polynomial functions showed no advantage over regular polynomial trial functions for this particular problem.

V. TOPICS FOR FURTHER STUDY

Several problems which need further investigation have been mentioned in previous sections. A more thorough study of the Bryson and Ho method is needed to determine whether or not it can handle some problems better than the regular formulation. A study could be made on the convergence characteristics of the GWR algorithm for smaller epsilon values to see if the range of convergence is further extended. The variable scaling technique should be studied to see if it extends the range of convergence.

A study of the usefulness of the GWR algorithm could be made. This algorithm and other quasilinearization algorithms could be used to solve several problems and the results and abilities of the algorithms compared. The possibility exists of applying the method of weighted residuals directly to the nonlinear differential equations and solving the resulting set of nonlinear algebraic equations.

The optimum time interval for a power level change could be determined by varying the final time and evaluating the performance index given by Equation 15 for each interval. The GWR algorithm could be used to solve any problem in nuclear engineering which reduces to a coupled set of nonlinear first-order differential equations.

VI. LITERATURE CITED

1. Weaver, Lynn E. Reactor dynamics and control. New York, N.Y., American Elsevier Publishing Company, Inc. 1968.

2. Kenneth, Paul and McGill, Robert. Solution of variational problems by means of a generalized Newton-Raphson operator. AIAA Journal 2: 1761-1766. 1964.

3. Finlayson, B. A. and Scriven, L. E. The method of weighted residuals - A review. Applied Mechanics Reviews 19: 735-748. 1966.

4. Crandall, Stephen H. Engineering analysis. New York, N.Y., McGraw-Hill Book Company, Inc. 1956.

5. Stacey, Weston M., Jr. Control of xenon spatial oscillations. Nuclear Science and Engineering 38: 229-243. 1969.

6. Kaplan, S. Some new methods of flux synthesis. Nuclear Science and Engineering 13: 22-31. 1962.

7. _____. Variational methods in reactor physics. Trans. Am. Nucl. Soc. 5: 412-423. 1962.

8. _____. Application of synthesis approximation to three-dimensional depletion calculation and to cell theory. Trans. Am. Nucl. Soc. 6: 254-255. 1963.

9. Leondes, C. T. and Paine, G. Extensions in quasilinearization techniques for optimal control. Journal of Optimization Theory and Applications 2: 316-330. 1968.

10. _____. Computational results for extensions in quasilinearization techniques for optimal control. Journal of Optimization Theory and Applications 2: 395-410. 1968.

11. Bryson, Arthur E., Jr. and Ho, Yu-Chi. Applied optimal control. Waltham, Mass., Blaisdell Publishing Company. 1966.

12. Lewallen, Jay M. A modified quasilinearization method for solving trajectory optimization problems. AIAA Journal 5: 962-965. 1967.

13. Rosztoszy, J. R., Sage, A. P., and Weaver, L. E. Application of Pontryagin's maximum principle to flux-state changes in nuclear reactors, in reactor kinetics and control. AEC Symposium Series, No. 2: 265-277. 1964.

14. Williams, S. D. Normalization and convergence acceleration for indirect optimization methods. Journal of Spacecraft and Rockets 7: 100-102. 1970.

VII. ACKNOWLEDGMENTS

The author wishes to express his appreciation to Drs. Michael R. Ringham, Glenn Murphy, William D. Leech, and Dale D. Grosvenor who formed his graduate committee. Special thanks are due Dr. Michael R. Ringham for a number of enlightening consultations during the course of this investigation.

The author wishes to thank the National Science Foundation for the traineeship under which the work was accomplished.