MASSACHUSETTS INSTITUTE OF TECHNOLOGY

DEPARTMENT OF NUCLEAR ENGINEERING

Cambridge 39, Massachusetts

THE APPLICATION OF ALTERNATING-DIRECTION IMPLICIT METHODS TO THE SPACE-DEPENDENT KINETICS EQUATIONS

by

Alan L. Wight, Kent F. Hansen

August 1969

MIT - 3903 - 3

MITNE - 106

AEC Research and Development Report

Contract AT(30-1)3903

U.S. Atomic Energy Commission

THE APPLICATION OF ALTERNATING-DIRECTION IMPLICIT METHODS TO THE SPACE-DEPENDENT KINETICS EQUATIONS

by

ALAN LEONARD WIGHT

B.E., University of Saskatchewan, Saskatoon Campus
(1965)
B.A., University of Saskatchewan, Regina Campus
(1966)
S.M., Massachusetts Institute of Technology
(1967)

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

August, 1969

Signature of Author . . . . . .
Department of Nuclear Engineering, August 18, 1969

Certified by . . . . . .
Thesis Supervisor

Accepted by . . . . . .
Chairman, Departmental Committee on Graduate Students

THE APPLICATION OF ALTERNATING-DIRECTION IMPLICIT METHODS TO THE SPACE-DEPENDENT KINETICS EQUATIONS

by

Alan Leonard Wight

Submitted to the Department of Nuclear Engineering of the Massachusetts Institute of Technology on August 18, 1969 in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

ABSTRACT

An approximate solution of the multigroup neutron diffusion kinetics equations with delayed neutrons in two-dimensional geometry can be obtained by matrix splitting methods based on an Alternating-Direction Implicit (ADI) scheme. The method is shown to be consistent and numerically stable. An exponential transformation of the semi-discrete equations reduces the truncation error so that the method becomes useable for practical computations. The results of numerical experiments are presented to illustrate the accuracy and stability of the method. These results indicate that another splitting method based on an Alternating-Direction Explicit scheme is slightly superior.

Thesis Supervisor: Kent F. Hansen
Title: Professor of Nuclear Engineering

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

The author would like to thank Prof. K. F. Hansen for his support and guidance throughout the course of this work.

The author would also like to thank William T. McCormick, Jr., and William H. Reed for their many ideas and numerical results, and Dr. John Yasinsky for the results of the TWIGL code.

The author would also like to thank the National Research Council of Canada for their Scholarship support.

The work was performed under USAEC Contract AT(30-1)-3903. Computation was performed on the IBM 360/65 computer at the MIT Information Processing Center.

The author would like to thank his wife, Gerry, for her support, encouragement, and patience during the course of this work.

The author would also like to thank the typist, Mrs. Esther Grande, for her skill and effort in the preparation of this manuscript.

BIOGRAPHICAL NOTE

Alan Leonard Wight was born on October 11, 1943 in Regina, Saskatchewan, Canada. He received his elementary and secondary education in various communities in the Province, and graduated from Martin Collegiate Institute, Regina, on June 30, 1961.

Mr. Wight attended the College of Engineering, University of Saskatchewan, Saskatoon, Sask., with a Union Carbide Undergraduate Scholarship. He graduated in May 1965 with the degree of Bachelor of Science in Engineering Science (Physics) with Great Distinction. He subsequently enrolled in the College of Arts and Sciences at the University of Saskatchewan, Regina Campus, and received the degree of Bachelor of Arts in May 1966 with Great Distinction as the College's "Most Distinguished Graduate" of that year. He enrolled in the Massachusetts Institute of Technology in Sept. 1966, and was granted the degree of Master of Science in Sept. 1967.

Mr. Wight was married to the former Geraldine Genieve Weidman of Groton, Mass. in Sept. 1968.

Chapter I

INTRODUCTION

The current trend toward very large power reactors, in which the space-time effects can become limiting design considerations (Refs. 1, 2), requires the development of methods of predicting the transient behavior of the neutron flux as a function of position as well as time. An enormous amount of effort has been directed toward this problem.*

In the following dissertation ws shall be concerned with methods of calculating the flux in two-dimensional geometry for reactor transients sufficiently rapid that time derivatives of the flux are not negligible. This eliminates xenon oscillation problems, burnup calculations, and other long period reactor changes.

The methods to be presented are intended to be applicable to a very general class of problems. However, it is intended that they should be useful as "numerical standards" against which other faster, but more approximate, methods can be tested, and in the analysis of reactor accidents.

To obtain the flux, we use the multigroup neutron diffusion kinetics equations with delayed neutrons. The problem can be written in the semi-discrete form (see Appendix A):

$$\frac{d\vec{\Psi}}{dt} = \vec{A}\,\vec{\Psi} \tag{1.1}$$

where $\vec{\Psi}$ is a vector of group fluxes and delayed neutron densities at

---

*See, for example, References 1, 6, 7, 8, 14, 18, 19, 20, 22.

every mesh point in the reactor and $\vec{A}$ is the 'A' matrix which describes the kinetic properties of the reactor. The multigroup equations and the 'A' matrix are discussed in detail in Appendix A.

Formally, the solution of the initial value problem (1. 1) can be written (for $\vec{A}$ constant):

$$\vec{\Psi}(t) = \exp(t\,\vec{A})\,\vec{\Psi}(0), \qquad (1.2)$$

where $\vec{\Psi}(0)$ are the initial conditions. The purpose of this thesis is to investigate methods of computing an approximate numerical solution of Eq. (1. 1) which are based on an Alternating-Direction Implicit (ADI) type of approximation (Ref. 3). These are part of a general class of "Matrix Splitting" methods (Refs. 4, 5).

The problem (1. 1) is difficult to solve because the characteristic times of the system vary from the asymptotic period (order of seconds) to the prompt neutron life time (fractions of microseconds for some systems). Stated another way, the eigenvalues of $\vec{A}$ vary from order +1 inverse second to order -10^9 inverse second. Also, the methods are difficult to analyze mathematically because $\vec{A}$ is not Hermitian, and because the spacial dependence of the problem prevents its being Fourier transformed.

In Chapter II the basic method will be presented and analyzed from a mathematical point of view, then several modifications of the basic method will be presented and discussed. Numerical solutions for several sample problems have been obtained using these methods, as well as using other methods currently under development (Refs. 6, 7) or in use (Ref. 8). In Chapter III these methods will be compared with respect to accuracy

and computing time.

The computer code, STKADI, written in FORTRAN IV for the MIT System/360/65 computer to test the methods is described and listed in Appendix D.

Chapter II

THEORY

1. Properties of the 'A' Matrix

Before discussing methods of obtaining an approximate solution of Eq. (1.1) we shall first review some of the properties of the matrix $\vec{A}$.

$\vec{A}$ is a real, square, irreducible matrix with non-negative off-diagonal elements. This is an "essentially positive" matrix by Def. 8.1, page 257 of Varga (Ref. 9). Thus by Varga's Theorem 8.1, $\exp(t\vec{A})$ is positive for all $t > 0$, and by Theorem 8.2, $\vec{A}$ has a real, simple eigenvalue, ω_o, such that

i) to ω_o there corresponds a positive eigenvector $\vec{e}_o$,

ii) if ω_i is any other eigenvector of $\vec{A}$, then Real (ω_i) is less than ω_o, and

iii) ω_o is increased if any element of $\vec{A}$ increases.

In addition, we know from Theorem 8.3 that $\exp(t\vec{A})$ exhibits asymptotic behavior given by

$$\| \exp(t\vec{A}) \| \sim K \exp(\omega_o t), \tag{2.1}$$

as $t \to \infty$, K some constant independent of t.

This leads us to the following observations about the solution, $\vec{\Psi}$, of Eq. (1.1) for non-negative initial conditions:

i) $\vec{\Psi}(t) > 0$ all $t > 0$ since $\exp(t\vec{A}) > 0$, and $\vec{\Psi}(0) \geq 0$, $\vec{\Psi}(0) \neq 0$,

ii) as t becomes large, $\| \vec{\Psi}(t) \|$ is bounded by $K \exp(\omega_o t)$, and

iii) we can subtract a constant diagonal matrix from $\vec{A}$ to make an $\vec{A}'$ whose largest eigenvalue is zero.

Furthermore, by Theorem 1 in Appendix B, the solution behaves asymptotically like:

$$\vec{\Psi}(t) \sim a \exp(\omega_o t)\, \vec{e}_o \qquad \text{as } t \to \infty, \tag{2.2}$$

where $a = (\vec{e}_o, \vec{\Psi}(0)) \geq 0$.

2. The Alternating-Direction Implicit Method

The solution (1.2) of Eq. (1.1) can be obtained in principle from the convergent series:

$$\exp(t\vec{A})\, \vec{\Psi}(0) = (\vec{I} + t\vec{A} + (t\vec{A})^2/2! + \ldots (t\vec{A})^n/n! + \ldots)\, \vec{\Psi}(0). \tag{2.3}$$

This is generally not feasible in practice for two reasons:

i) the number of terms required, and the number of computations needed for each additional term make the computing time prohibitive, and

ii) round-off error will swamp the solution long before the series converges.

To obtain an approximate solution of Eq. (1.1), we replace the time derivative by a forward difference over a time interval, Δt, and calculate a series of approximate solutions $\vec{\Psi}^j$ at discrete times, t_j, assuming $\vec{A}$ constant over Δt. The algorithm for computing $\vec{\Psi}^{j+1}$ from $\vec{\Psi}^j$ is obtained as follows:

$\vec{A}$ is split into the sum of two parts:

$$\vec{A} = \vec{A}_1 + \vec{A}_2, \tag{2.4}$$

where

$$\vec{A}_1 = \vec{X} + \vec{U}, \qquad \vec{A}_2 = \vec{Y} + \vec{L}.$$

$\vec{X}$ is a symmetric matrix of one half of the diagonal terms of $\vec{A}$ and the off-diagonal stripes associated with diffusion in one direction. $\vec{Y}$ contains one half of the diagonal and off-diagonal stripes associated with diffusion in the perpendicular direction. (See Appendix A.) $\vec{U}$ contains the remaining elements of $\vec{A}$ which appear above the diagonal, and $\vec{L}$ those which appear below. Then with $h = \Delta t/2$, we write

$$\frac{\vec{\Psi}^{j+1/2} - \vec{\Psi}^{j}}{h} = \vec{A}_1\vec{\Psi}^{j+1/2} + \vec{A}_2\vec{\Psi}^{j}$$

$$\frac{\vec{\Psi}^{j+1} - \vec{\Psi}^{j+1/2}}{h} = \vec{A}_1\vec{\Psi}^{j+1/2} + \vec{A}_2\vec{\Psi}^{j+1}, \tag{2.5}$$

or equivalently,

$$(\vec{I} - h\vec{A}_2)\,\vec{\Psi}^{j+1/2} = (\vec{I} + h\vec{A}_1)\,\vec{\Psi}^{j}, \tag{2.6a}$$

$$(\vec{I} - h\vec{A}_1)\,\vec{\Psi}^{j+1} = (\vec{I} + h\vec{A}_2)\,\vec{\Psi}^{j+1/2}, \tag{2.6b}$$

where $\vec{\Psi}^{j+1/2}$ is an intermediate vector which is actually computed.

The name Alternating-Direction Implicit derives from the diffusion term in one direction being handled implicitly in one half step, and explicitly in the next, according to the scheme of Peaceman and Rachford (Ref. 3). The $\vec{L}$ and $\vec{U}$ matrices are treated as in the Gauss-Seidel method, with the $\vec{L}$ and $\vec{U}$ alternately implicit.

The linear systems implied by Eqs. (2.6) can be easily solved by taking advantage of the block structure of the matrices. After the matrix multiply $(\vec{I} - h\vec{A}_1)\,\vec{\Psi}^{j}$ is performed, we are left with a system that looks

like Fig. 2.1 if we divide out h and incorporate $\vec{I}/h$ into the diagonal blocks. The system $\vec{Y}_1 \vec{u}_1 = \vec{v}_1$ can be solved quickly by elimination since $\vec{Y}_1$ is tridiagonal. Starting in group 1, we solve for all $\vec{u}_1$, which is then available for back substitution when solving for $\vec{u}_2$. The process is repeated until all the energy group fluxes are obtained. The delayed groups then require solution of systems. $\vec{\Lambda}_i \vec{u}_{G+i} = \vec{v}'_{G+i}$ which is trivial since $\vec{\Lambda}_i$ is diagonal. For the next half-step the procedure is the same, except the solution proceeds from bottom to top. No iterations are required at any stage of the computation.

By substituting (2.6a) into (2.6b) we obtain a formal expression for the advancement matrix $\vec{B}^j(h)$:

$$\vec{\Psi}^{j+1} = \vec{B}^j(h)\, \vec{\Psi}^j = (\vec{I} - h\vec{A}_2)^{-1} (\vec{I} + h\vec{A}_1)(\vec{I} - h\vec{A}_1)^{-1} (\vec{I} + h\vec{A}_2)\, \vec{\Psi}^j. \tag{2.7}$$

3. Properties of ADI

The advancement matrix $\vec{B}^j(h)$ can be rewritten in a form more convenient for analysis:

$$\vec{B}^j(h) = \vec{I} + 2h(\vec{I} - h\vec{A}_2)^{-1} (\vec{I} - h\vec{A}_1)^{-1} \vec{A}. \tag{2.8}$$

This gives immediately:

Property 1 — For the exactly critical, steady-state reactor, $\dot{\vec{\Psi}} = \vec{A}\, \vec{\Psi} = \vec{0}$, $\vec{\Psi}^1 = \vec{B}'(h)\, \vec{\Psi} = \vec{\Psi}$, which is the exact solution, independent of h.

For the general problem we are more concerned with:

Property 2 — The advancement matrix $\vec{B}^j(h)$ is a consistent approximation to the solution of (1.1).

$$
\begin{vmatrix}
Y_1 & & & & & & & \\
T_{21} & Y_2 & & & & & & \\
T_{31} & T_{32} & Y_3 & & & & & \\
\cdot & \cdot & \cdot & & & & & \\
\cdot & \cdot & \cdot & T_{G,G-1} & Y_G & & & \\
P_{11} & P_{12} & \cdot & \cdot & \cdot & \Lambda_1 & & \\
P_{21} & P_{22} & \cdot & \cdot & \cdot & & \Lambda_2 & \\
 & & \cdot & \cdot & \cdot & & & \\
P_{I1} & P_{I2} & \cdot & \cdot & \cdot & & & \Lambda_I
\end{vmatrix}
\begin{vmatrix} u_1 \\ u_2 \\ u_3 \\ \cdot \\ u_G \\ u_{G+1} \\ u_{G+2} \\ \cdot \\ u_{G+I} \end{vmatrix}
=
\begin{vmatrix} v_1 \\ \cdot \\ \cdot \\ \cdot \\ v_G \\ v_{G+1} \\ \cdot \\ \cdot \\ v_{G+I} \end{vmatrix}
$$

Fig. 2.1 - System to be Solved on First Half-Step

"Consistency" assures us that the method does in fact approximate the solution $\vec{\Psi}(t)$ for "sufficiently small" time steps. The consistency property is proven in Theorem 5, Appendix B.

If we assume h sufficiently small that we can expand $\vec{B}^j(h)$ in a power series, we get

$$\vec{B}^j(h) = \vec{I} + 2h\vec{A} + 2h^2\vec{A}^2 + \ldots . \qquad (2.9)$$

Comparing this to the expansion:

$$\exp(\Delta t\vec{A}) = \exp(2h\vec{A}) = \vec{I} + 2h\vec{A} + (2h\vec{A})^2/2! + \ldots , \qquad (2.10)$$

we arrive at

Property 3 – $\vec{B}^j(h)$ agrees with the expansion of $\exp(\Delta t\vec{A})$ to terms in h^2, and hence is said to be "accurate to order (h^2)."

This property is verified experimentally in Chapter III, Section 2.

Theorem 2, Appendix B, shows that the inverses, $(\vec{I} - h\vec{A}_1)^{-1}$ and $(\vec{I} - h\vec{A}_2)^{-1}$, are non-negative. Since the fundamental eigenvector, $\vec{e}_o$, is positive, we can establish:

Property 4 – If $\omega_o \geq 0$, the component of the solution vector $\vec{\Psi}^j$ in the direction of $\vec{e}_o$ is nondecreasing.

This assures us that the fundamental component will grow when it should, although it does not guarantee that it will grow faster than all other components, nor that it will grow with the correct period.

Equation (2.7) can be rewritten in the form

$$\vec{B}^j(h) = (\vec{I}/h-\vec{A}_2)^{-1}\,(\vec{I}/h+\vec{A}_1)(\vec{I}/h-\vec{A}_1)^{-1}\,(\vec{I}/h+\vec{A}_2). \qquad (2.11)$$

In the limit as h becomes very large, $\vec{I}/h$ becomes negligibly small, and

$$\vec{B}^j(h) \rightarrow (-\vec{A}_2)^{-1}\,(\vec{A}_1)(-\vec{A}_1)^{-1}\,(\vec{A}_2) = \vec{I}.$$

Thus:

Property 5 – As the time step becomes very large, the advancement matrix, $\vec{B}^j(h)$, approaches the identity operator, $\vec{I}$.

This property helps explain the observed tendency for the ADI to underpredict the growth (or decay) of the solution as the time step is increased. More precisely, the growth (or decay) of each eigenvector in the solution is underpredicted, depending on the product $h\omega_i$. Thus, if the initial condition contains a large amount of a component with a large negative eigenvalue, that component will die away very slowly, resulting in considerable error in the computed solution vector, even though the fundamental is well approximated.

The eigenvalues of $\vec{B}^j(h)$ are the solutions of

$$\vec{B}^j(h)\,\vec{v}_i = \xi_i\,\vec{v}_i, \qquad (2.12)$$

or if we let

$$\xi_i = \frac{1 + h\nu_i}{1 - h\nu_i} \qquad (2.13)$$

we can write (2.12) as

$$(1+h\nu_i)(\vec{I} - h\vec{A}_1)(\vec{I} - h\vec{A}_2)\,\vec{v}_i = (1 - h\nu_i)(\vec{I}+h\vec{A}_1)(\vec{I}+h\vec{A}_2)\,\vec{v}_i.$$

This reduces to

$$h\vec{A}\,\vec{v}_i = h\nu_i\,\vec{v}_i + (h\nu_i)\,h^2\vec{A}_1\vec{A}_2\,\vec{v}_i. \tag{2.14}$$

Thus the eigenvalues ν_i and eigenvectors $\vec{v}_i$ are solutions of the characteristic equation:

$$(\vec{I}+h^2\vec{A}_1\vec{A}_2)^{-1}\,\vec{A}\,\vec{v}_i = \nu_i\,\vec{v}_i. \tag{2.15}$$

Comparing (2.15) to the characteristic equation for $\vec{A}$,

$$\vec{A}\,\vec{e}_i = \omega_i\,\vec{e}_i, \tag{2.16}$$

we see that ν_i and $\vec{v}_i$ are approximations to ω_i and $\vec{e}_i$ accurate to order (h^2). Furthermore, the eigenvalues of $\vec{B}^j(h)$ are approximately

$$\xi_i = \frac{1 + h\nu_i}{1 - h\nu_i} = \frac{1 + h\omega_i}{1 - h\omega_i} + 0(h^2)$$

$$= e^{2h\omega_i} + 0(h^2). \tag{2.17}$$

4. Stability

To be useable a numerical method cannot allow some error in the solution vector to grow faster than the correct solution; that is, the method must be stable. To insure stability, we require that the solution should remain bounded for finite time and finite time step. More precisely, we use the definition of stability of Richtmyer and Morton (Ref. 10):

Definition: The advancement matrix, $\vec{B}(\Delta t)$, is stable if there exists a constant, b, such that

$$\|\vec{B}^N(\Delta t)\| \leq b \tag{2.18}$$

for

$$0 < N\Delta t = T, \qquad 0 < \Delta t_j \leq \tau.$$

This condition can be satisfied if the solution grows by no more than a factor $(1+K\Delta t)$, K some constant, with each time step.

We must impose an additional requirement on the definition above. $\vec{B}^j(h)$ contains quantities of the form $h v_g D_g/\Delta x^2$ arising from the approximation to the diffusion operator which become very large as Δx^2 becomes very small. The upper bound in (2.18) must be either independent of Δx^2, or, if this is too strict a requirement, then an upper bound must exist with the ratio $r = h v_g D_g/\Delta x^2$ held fixed.

We first consider the stability of the problem obtained by setting the intragroup transfer terms (including delay-group transfers) to zero to obtain the symmetric matrix, $\vec{a} = \vec{X} + \vec{Y}$. Thus:

$$\vec{\mathcal{B}}^j(h) = (\vec{I} - h\vec{Y})^{-1} (\vec{I}+h\vec{X})(\vec{I} - h\vec{X})^{-1} (\vec{I}+h\vec{Y}). \tag{2.19}$$

Theorem 3 shows that $\vec{\mathcal{B}}^j(h)$ is unconditionally stable if and only if all the eigenvalues of $\vec{X}$ and $\vec{Y}$ are non-positive. Using Gerschgorin's Theorem, we can show that this is true if the net group production term on the diagonal,

$$\sigma_{gg} = \chi_g(1-\beta)(\nu\sigma_f)_g - \sigma_{ag}, \tag{2.20}$$

is negative.

The matrix $\vec{B}^j(h)$ can be written as a sum of (2.19) and a bounded perturbation of order (h). Thus:

$$\vec{B}^j(h) = \vec{\mathcal{B}}^j(h) + h\vec{Q}(h), \tag{2.21}$$

where $\|\vec{Q}(h)\| \leq q$, all h. $\vec{B}^j(h)$ certainly cannot be stable if $\vec{\mathcal{B}}^j(h)$ is not, and conversely if $\vec{\mathcal{B}}^j(h)$ is stable and $\vec{Q}(h)$ bounded, $\vec{B}^j(h)$ must be stable. (See Ref. 10.)

$\vec{B}^j(h)$ can be shown to be stable as follows:

$$\vec{B}^j(h)^N = (\vec{I} - h\vec{A}_2)^{-1}\,(\vec{I}+h\vec{A}_1)(\vec{I} - h\vec{A}_1)^{-1}$$

$$\ldots\ (\vec{I}+h\vec{A}_1)(\vec{I} - h\vec{A}_1)^{-1}\,(\vec{I}+h\vec{A}_2)$$

$$= (\vec{I} - h\vec{A}_2)^{-1}\{(\vec{I}+h\vec{A}_1)(\vec{I} - h\vec{A}_1)^{-1}\ldots$$

$$(\vec{I}+h\vec{A}_2)(\vec{I} - h\vec{A}_2)^{-1}\}\,(\vec{I} - h\vec{A}_2)$$

$$= (\vec{I} - h\vec{A}_2)^{-1}\,\vec{B}'(h)^N\,(\vec{I} - h\vec{A}_2).$$

Now

$$\vec{B}'(h) = \vec{B}_1(h)\,\vec{B}_2(h),$$

with

$$\vec{B}_1(h) = (\vec{I}+h\vec{A}_1)(\vec{I} - h\vec{A}_1)^{-1},$$

and

$$\vec{B}_2(h) = (\vec{I}+h\vec{A}_2)(\vec{I} - h\vec{A}_2)^{-1}.$$

Now we can manipulate $\vec{B}_1(h)$ to obtain

$$\vec{B}_1(h) = (\vec{I} - h\vec{X})^{-1}\,(\vec{I}+h\vec{X}) + 2h(\vec{I} - h\vec{A}_1)^{-1}\,U(\vec{I} - h\vec{X})^{-1}$$

$$= \vec{\mathcal{B}}_1(h) + 2h\vec{Q}_1(h). \qquad (2.22)$$

Now if Condition (2.20) is satisfied, Theorem 4 establishes that $\vec{Q}_1(h)$ is bounded:

$$\|\vec{Q}_1(h)\| \leq q_1 \tag{2.23}$$

Furthermore, all the eigenvalues of $\vec{X}$ are negative, and

$$\rho(\vec{\mathcal{B}}_1(h)) = \|\vec{\mathcal{B}}_1(h)\| < 1.$$

Thus

$$\|\vec{B}_1(h)\| \leq 1 + 2hq_1 \leq e^{\Delta t q_1}. \tag{2.24}$$

Similarly,

$$\|\vec{B}_2(h)\| \leq 1 + 2hq_2 \leq e^{\Delta t q_2}. \tag{2.25}$$

Consequently,

$$\|\vec{B}^j(h)^N\| \leq \|(I - h\vec{A}_2)^{-1}\| \cdot \|\vec{B}'(h)^N\| \cdot \|(I - h\vec{A}_2)\|$$

$$\leq \|(\vec{I} - h\vec{A}_2)^{-1}\| \cdot \|(I - h\vec{A}_2)\| \left(e^{(q_1+q_2)\Delta t}\right)^N$$

$$\leq C\, e^{qT} = b \tag{2.26}$$

since the condition number

$$\|(\vec{I} - h\vec{A}_2)^{-1}\| \cdot \|(\vec{I} - h\vec{A}_2)\|$$

is bounded by a constant for $0 < h \leq \tau < \infty$.

If the time step varies over the computation, or the elements of $\vec{A}$ are functions of time, then we select the maximum $\|\vec{B}_1\|$ and $\|\vec{B}_2\|$, and

perform the same analysis.

The above analysis has not assumed that the reactor is homogeneous, has placed no restriction on the number of neutron or delayed groups, and has placed no restriction on h or T other than that they be real, positive and finite. The only restriction is that the diagonal production term, σ_{gg}, be negative, a condition which is almost always satisfied in practice. Thus the basic method is unconditionally stable.

5. Fractional Step Method

Physically meaningful problems have flux distributions which are everywhere positive. Consequently any negative elements which would appear in the approximate solution $\vec{\Psi}^j$ would render the solution unuseable in practice. Unfortunately, since the matrices $(\vec{I}+h\vec{A}_2)$ and $(\vec{I}+h\vec{A}_1)$ are non-negative only for very small h, the ADI does not necessarily produce a positive solution even when the initial vector is positive. Since we know that the exact solution with positive initial conditions is positive, we seek a method which shares this property.

The advancement matrix

$$\vec{B}^j(\Delta t) = (\vec{I} - \Delta t\vec{A}_2)^{-1}\,(\vec{I} - \Delta t\vec{A}_1)^{-1} \qquad (2.27)$$

is a consistent approximation, and is non-negative. It is accurate to $0(\Delta t)$ and stable for all Δt. Unlike the basic ADI method, it is not exact for the exactly critical problem, nor can the solution be guaranteed to grow when it should. In fact, for very large Δt,

$$\lim_{\Delta t\to\infty} \vec{B}^{j}(\Delta t) = \lim_{\Delta t\to\infty} \frac{1}{\Delta t^2}(\vec{I}/\Delta t-\vec{A}_2)^{-1}\,(\vec{I}/\Delta t-\vec{A}_1)^{-1}$$

$$= \frac{1}{\Delta t^2}(-\vec{A}_2)^{-1}\,(-\vec{A}_1)^{-1}.$$

Thus for large Δt, the solution decreases as Δt increases.

6. Frequency Transformation

Numerical experiments have shown that the basic ADI method of Section 2 is not sufficiently accurate. A large improvement in the accuracy of the method results from a simple change of variable. We define a transformation from the relation

$$\vec{\Psi} = e^{\vec{\Omega}t}\,\vec{\Psi}' \tag{2.28}$$

with $\vec{\Omega}$ a diagonal matrix. The equation for $\vec{\Psi}'$ then becomes

$$\frac{d\vec{\Psi}'}{dt} = e^{-\vec{\Omega}t}\,(\vec{A}-\vec{\Omega})\,e^{\vec{\Omega}t}\,\vec{\Psi}'$$

$$= \vec{A}'\,\vec{\Psi}' \tag{2.29}$$

If $\vec{\Psi}(t)$ has a basically exponential behavior, then $\vec{\Psi}'(t)$ should be a smooth modulation, and hence well approximated by a simple difference.

Now (2.29) is identical in form to (1.1). We thus attempt a solution as before. We first integrate over h to obtain

$$\Psi'(h) - \Psi'(0) = \int_0^h \vec{A}'(t)\,\vec{\Psi}'(t)\,dt$$

$$\approx h\vec{A}'_1(h/2)\,\vec{\Psi}'(h) + h\vec{A}'_2(h/2)\,\vec{\Psi}'(0), \tag{2.30}$$

evaluating $\vec{A}'(t)$ half way between the end points.

By reversing the roles of $\vec{A}_1'$ and $\vec{A}_2'(t)$ on the next half step, and using some algebra, we obtain

$$\vec{\Psi}^{j+1} = \exp\left(\frac{3}{2} h\vec{\Omega}\right)\left(\vec{I} - h\left(\vec{A}_2 - \frac{1}{2}\vec{\Omega}\right)\right)^{-1}$$

$$\left(\vec{I} + h\left(\vec{A}_1 - \frac{1}{2}\vec{\Omega}\right)\right)\left(\vec{I} - h\left(\vec{A}_1 - \frac{1}{2}\vec{\Omega}\right)\right)^{-1}\left(\vec{I} + h\left(\vec{A}_2 - \frac{1}{2}\vec{\Omega}\right)\right)$$

$$\exp\left(\frac{1}{2} h\vec{\Omega}\right)\vec{\Psi}^{j}, \qquad (2.31)$$

after using (2.28) to transform back to $\vec{\Psi}$.

The matrix $\exp(t\vec{\Omega})$ is a diagonal matrix of elements $(e^{t\Omega_{ii}})$ and consequently is simple to evaluate. The remaining terms of (2.31) are the basic ADI method applied to the matrix $(\vec{A}-\vec{\Omega})$. We shall call this the "Transformed ADI" method, and the matrix,

$$\vec{B}(h,\vec{\Omega}) = \exp\left(\frac{3}{2} h\vec{\Omega}\right)\left(I - h\left(A_2 - \frac{1}{2}\vec{\Omega}\right)\right)^{-1}$$

$$\left(I + h\left(\vec{A}_1 - \frac{1}{2}\vec{\Omega}\right)\right)\left(I - h\left(A_1 - \frac{1}{2}\vec{\Omega}\right)\right)^{-1}\left(I + h\left(A_2 - \frac{1}{2}\vec{\Omega}\right)\right)$$

$$\exp\left(\frac{1}{2} h\vec{\Omega}\right), \qquad (2.32)$$

the "Transformed Advancement Matrix." Since $\vec{\Omega}$ has units of $\sec^{-1}$, we shall call it the "Frequency Matrix," and its elements "frequencies."

$\vec{\Omega}$ is to be chosen in such a way as to minimize the error in the solution. It is not obvious a priori how to do this, but a method which has proven extremely successful in practice is to take advantage of information available from the previous step, and compute

$$\Omega_i^j = \frac{\log\left(\Psi_i^j/\Psi_i^{j-1}\right)}{2h}. \tag{2.33}$$

Thus the solution growth over the previous step is used to estimate $\vec{\Omega}$ to be used in the present step.

This requires storing $\vec{\Psi}^{j-1}$ as well as $\vec{\Omega}^j$, so that the transformed ADI requires three times as much storage as the basic ADI method. For the first step $\vec{\Omega} = \vec{0}$ is used.

If $\vec{\Omega}$ were held constant throughout the calculation, and the condition

$$(\sigma_{gg} - \Omega_i) < 0 \tag{2.34}$$

were always satisfied, then the transformed ADI would have the same stability properties as ADI. However, $\vec{\Omega}$ is changed at each step, allowing feedback effects to cause instabilities.

If the solution has become asymptotic, i.e.,

$$\vec{\Psi}^j = e^{\omega\Delta t}\,\vec{\Psi}^{j-1}, \tag{2.35}$$

and

$$\vec{\Omega}^j = \omega\vec{I},$$

then

$$\vec{\Psi}^{j+1} = e^{3/2\omega h}\left[I + 2h\left[\vec{I} - h\left(\vec{A}_2 - \tfrac{1}{2}\omega\right)\right]^{-1}\left[I - h\left(\vec{A}_1 - \tfrac{1}{2}\omega\right)\right]^{-1}(\vec{A} - \omega)\right] e^{1/2\omega h}\vec{\Psi}^j$$

$$= e^{2h\omega}\left[\vec{\Psi}^j + 2h\left[\vec{I} - h\left(\vec{A}_2 - \tfrac{1}{2}\omega\right)\right]^{-1}\left[\vec{I} - h\left(\vec{A}_1 - \tfrac{1}{2}\omega\right)\right]^{-1}(\vec{A} - \omega)\,\vec{\Psi}^j\right].$$

Hence $(\vec{A} - \omega)\,\vec{\Psi}^j = \vec{0}$ since the inverse matrices are non-negative. Thus $\omega = \omega_o$, and $\vec{\Psi}^j \sim \vec{e}_o$. This means that asymptotically the growth is exact

and the solution is proportional to the fundamental eigenvector of the system. Thus we say the solution is "asymptotically exact."

7. Iterative ADI Method

It is not always true that an Order (Δt^2) method is superior to an $0(\Delta t)$ method. To illustrate, consider the following rather simple example:

$$\vec{A} = \begin{bmatrix} -49.5 & 50.5 \\ 50.5 & -49.5 \end{bmatrix},$$

for which

$$\omega_1 = 1, \qquad \vec{e}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

and

$$\omega_2 = -100, \qquad \vec{e}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Note that the eigenvalues are greatly different in order of magnitude. Let us consider two approximations to $\exp(\Delta t \vec{A})$:

$$\vec{B}_1(\Delta t) = (\vec{I} + h\vec{A})(\vec{I} - h\vec{A})^{-1}, \tag{2.36}$$

$$\vec{B}_2(\Delta t) = (\vec{I} - \Delta t \vec{A})^{-1}, \tag{2.37}$$

and an initial vector,

$$\vec{x} = \vec{e}_1 + \vec{e}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 2 \\ 0 \end{bmatrix}. \tag{2.38}$$

If we take time step $\Delta t = .1$, and operate with $\vec{B}_1$, $\vec{B}_2$ and $\exp(\Delta t \vec{A})$ separately on each eigenvector, we get the results shown in Table 2.1. Thus,

Table 2.1. Comparison of $0(\Delta t^2)$ and $0(\Delta t)$ approximation.

	Coefficient of Each Component in Solution		
Component	$\vec{B}_1$ $(0(\Delta t^2))$	$\vec{B}_2$ $(0(\Delta t))$	$\exp(\Delta t \vec{A})$
$\vec{e}_1$	1.106	1.111	1.105
$\vec{e}_2$	-.667	.0909	.000042
$\|x(\Delta t)\|$	1.292	1.113	1.105
error	.187	.008	.0

although $\vec{B}_1$ better approximates the growth of $\vec{e}_1$, $\vec{B}_2$ better approximates the decay of $\vec{e}_2$, and, in the final result, $\vec{B}_2$ gives the better approximation.

Although the above example is contrived, it is not a completely unrealistic example. For real reactor problems, the eigenvalues are separated by many orders of magnitude, such that $|h\omega| \gg 1$ for the largest (in magnitude) eigenvalues. Since one picks time steps such that the fundamental and perhaps a few of the smaller (in magnitude) eigenvalue components of the solution are well approximated, most of the error in the solution by the ADI method comes from the larger eigenvalue components. If these are present in a large amount in the initial vector, which they may well be in a problem with much spacial dependence, the error will be quite large. Thus for some problems an approximation like $\vec{B}_2$ may be preferable to an approximation like the ADI.

To use the advancement matrix

$$\vec{B}(\Delta t) = (\vec{I} - \Delta t \vec{A})^{-1} \tag{2.39}$$

we must solve a system of equations:

$$(\vec{I} - \Delta t\, \vec{A})\, \vec{\Psi}^{j+1} = \vec{\Psi}^{j} \tag{2.40}$$

at each step. This is simply the linear system

$$\vec{G}\, \vec{X} = \vec{Y}, \qquad \vec{G} = \vec{I} - \Delta t\, \vec{A} \tag{2.41}$$

which can be solved by an ADI iterative method as follows: Assuming a starting value of $\vec{X}$ obtained by some method (ADI for example), define a splitting of $\vec{G}$ analogously to the ADI method

$$\vec{G} = \vec{G}_1 + \vec{G}_2 \tag{2.42}$$

$$\vec{G}_1 = \Delta t(\vec{I}/2\Delta t - \vec{A}_1) \tag{2.43}$$

$$\vec{G}_2 = \Delta t(\vec{I}/2\Delta t - \vec{A}_2). \tag{2.44}$$

The iteration scheme then becomes

$$\begin{aligned} (\vec{R}_k + \vec{G}_1)\, \vec{X}^{k+1/2} &= (\vec{R}_k - \vec{G}_2)\, \vec{X}^{k} + \vec{Y} \\ (\vec{R}_k + \vec{G}_2)\, \vec{X}^{k+1} &= (\vec{R}_k - \vec{G}_1)\, \vec{X}^{k+1/2} + \vec{Y}, \end{aligned} \tag{2.45}$$

where $\vec{R}_k$ is some positive diagonal acceleration (or optimization) matrix. If this method is to be used in practice, some scheme for determining optimal $\vec{R}_k$ to speed convergence would have to be invented. However, to test the method the selection

$$\vec{R}_k = \vec{I}/2 \tag{2.46}$$

was made because it was particularly easy to code with the subroutines already available.

An alternative strategy for treating large eigenvalue components is to reduce the time step of the ADI method. In order to be competitive, the Iterative ADI method must employ fewer iterations to achieve the same error reduction than the alternative requires additional steps. In a rather artificial test problem to which the method was applied, it did as well as the ADI, but for the one "practical" problem to which the method was applied, it was very little improvement over the basic method with the same time step, and required four times as much computing. (See Chapter III, Section 5.)

Another iterative method, TWIGL (Ref. 8), uses a much faster iteration scheme, but still appears to require more iterations than other, non-iterative methods require steps, although comparisons are difficult since these problems were not run on the same machine. The TWIGL method is compared with other methods in Chapter III, Section 4. In general, it appears that the longer time step that iterative methods allow costs more in terms of computing time than non-iterative methods.

8. Comparison and Summary

The various ADI methods discussed in this chapter are summarized in Table 2. 2. The requirement that σ_{gg} be negative is not considered as a practical restriction, and is assumed to hold for all methods.

Of the four methods, only the Frequency Transformation can be unstable for some problems. However, the results in Chapter III demonstrate conclusively the great superiority of this method over the others

Table 2.2. Comparison of various methods.

Method	Truncation Error	Eigenvalue	Eigenvector	Advantages	Comments
ADI	$0(\Delta t^2)$	$\frac{1+h\omega_i}{1-h\omega_i} + 0(h^2)$	$\vec{e}_i + 0(h^2)$	Unconditionally stable	Truncation error too high in practice
Fractional Step	$0(\Delta t)$	$\frac{1}{1-\Delta t\omega_1} + 0(\Delta t)$	$\vec{e}_i + 0(\Delta t)$	Advancement matrix is non-negative	Δt's small enough to make this method accurate are also small enough to make ADI non-negative
Iterative ADI	$0(\Delta t)$	$\frac{1}{1-\Delta t\omega_i}$	$\vec{e}_i$	Improved approximation for components having large negative eigenvalues	Requires iteration at each step
Frequency-Transformed ADI	better than $0(\Delta t^2)$	$e^{\omega_o \Delta t}$ (asymptotically)	$\vec{e}_o$	Asymptotically exact, truncation error much better than basic ADI	Unstable for some problems

for a broad class of problems. The other three methods may have some limited application to problems where the Frequency Transformation is unstable, but otherwise it is the method of choice.

Chapter III

RESULTS

1. Introduction

The ADI method and variations described in Chapter II have been coded in FORTRAN IV for the IBM 360/65 computer. The listing and program description of STKADI are in Appendix D. Several trial problems have been run to test the method. The results of these numerical experiments will be discussed in the following sections.

The same problems have also been run on the computer codes LUMAC (Ref. 6) and MITKIN (Ref. 7). LUMAC is a two-dimensional version of GAKIN (Ref. 22) which uses a buckling approximation for the second dimension. MITKIN uses a splitting method very similar to STKADI, except that it is based on an "Alternating-Direction Explicit" approach. The solutions from these two codes will be compared with those from STKADI.

The storage requirements of STKADI are summarized in Table 3. 1. In addition, the program itself requires 11, 500 eight byte words of core storage. Thus a problem of 1000 mesh points, 10 groups, 6 delayed groups and 20 regions requires 120, 000 words.

The observed computing times per step on the 360/65 are listed in Table 3. 2 for various trial problems with and without the frequency transformation, and compared to the reported computing times for the MITKIN and LUMAC codes.

The number of floating-point multiplications (and divisions) in one step of STKADI is given by

Table 3.1. Storage requirements.

Data Stored	Number of Words of Storage Required
Diffusion Coefficients	$4 \times G \times R$
Intragroup Transfer Terms	$G^2 \times R$
Delayed Neutron Production Terms	$G \times I \times R$
Delayed to Neutron Group Transfer Terms	$I \times G \times R$
Delayed Group Decay Constants	I
Flux Vector	$N \times (G+I)$
Frequency Vector	$N \times (G+I)$
Flux Vector for Previous Step	$N \times (G+I)$
Total	$3N(G+I) + RG(G+2I+4) + I$

N = number of mesh points
G = number of energy groups
I = number of delayed groups
R = number of regions

Table 3.2. Computing times.

			Computing time in seconds/step			
N	G	I	STKADI	STKADI (frequencies)	MITKIN	LUMAC
81	2	1	.21	.36	.33	.46
81	4	1	.45	.71	.56	.90
361	2	1	.89	1.51	1.34	1.75
171	4	1	.94	1.49	1.20	1.75
38	2	6	.23	.41		

$$Nf(G,I) = N(186 + 8I + 56I + 2G^2), \tag{3.1}$$

where terms small compared to N are neglected. In addition, the frequency transformation requires two exponential and one logarithm evaluation for each unknown.

Assuming that the total computing time per step is proportional to the number of floating point multiplications, and the additional computing time required by the frequency transformation is proportional to the number of unknowns, the total computing time can be written

$$\tau = \alpha N(f(G,I) + \gamma(I+G)). \tag{3.2}$$

We obtain the constants using the data in Table 3.2. They are

$$\alpha = .41 \times 10^{-4} \text{ sec/step/mesh point}$$

$$\gamma = \begin{cases} 15.2 & \text{with frequencies} \\ 0. & \text{without frequencies.} \end{cases}$$

Thus the 1000 mesh point, 10 group, 6 delayed group problem requires 40 seconds per step or about 70 minutes to do a 100-step problem.

2. CASE1 – Two Group Bare Homogeneous Reactor

CASE1 is a two energy group, one delayed group thermal system. The reactor is a homogeneous square, 200 cm on a side with nine mesh points (ten intervals) in each direction. A positive step change in reactivity of about forty cents is inserted at time zero by a decrease of $.0000369 \text{ cm}^{-1}$ in the thermal capture cross section. The initial conditions correspond to the steady state. Data are given in Appendix C.

The solution was calculated using various time steps with and without the frequency transformation. The results at t = .4 second are shown in Table 3.3. The percentage error is plotted in Fig. 3.1.

From Table 3.3 and Fig. 3.1 it is clear that the frequency transformation is far superior to the basic method. The former achieves less than 1% error with time steps of about .001 second, the latter requires time steps of one-tenth this or smaller to achieve the same error. However, the transformation increases the computing time by about 70% so that the over-all improvement is about a factor of six.

Table 3.3. Results for CASE1 at .4 second.

	Without Frequencies		With Frequencies	
Δt (sec)	Group 1	Group 2	Group 1	Group 2
Exact	1.566	.600	1.566	.600
.0080	1.002 (36.%)	.384 (36.%)	1.061 (32.%)	.407 (32.%)
.0040	1.010 (35.%)	.387 (35.%)	1.389 (11.%)	.532 (11.%)
.0020	1.037 (34.%)	.398 (34.%)	1.647 (-5.2%)	.631 (-5.2%)
.0010	1.127 (28.%)	.432 (28.%)	1.572 (-.4%)	.603 (-.4%)
.0005	1.315 (16.%)	.504 (16.%)	1.566 (-.06%)	.601 (-.06%)
.00025	1.482 (5.3%)	.568 (5.3%)		
.000125	1.544 (1.3%)	.592 (1.3%)		

Note: Numbers in parentheses ( ) are percentage errors.

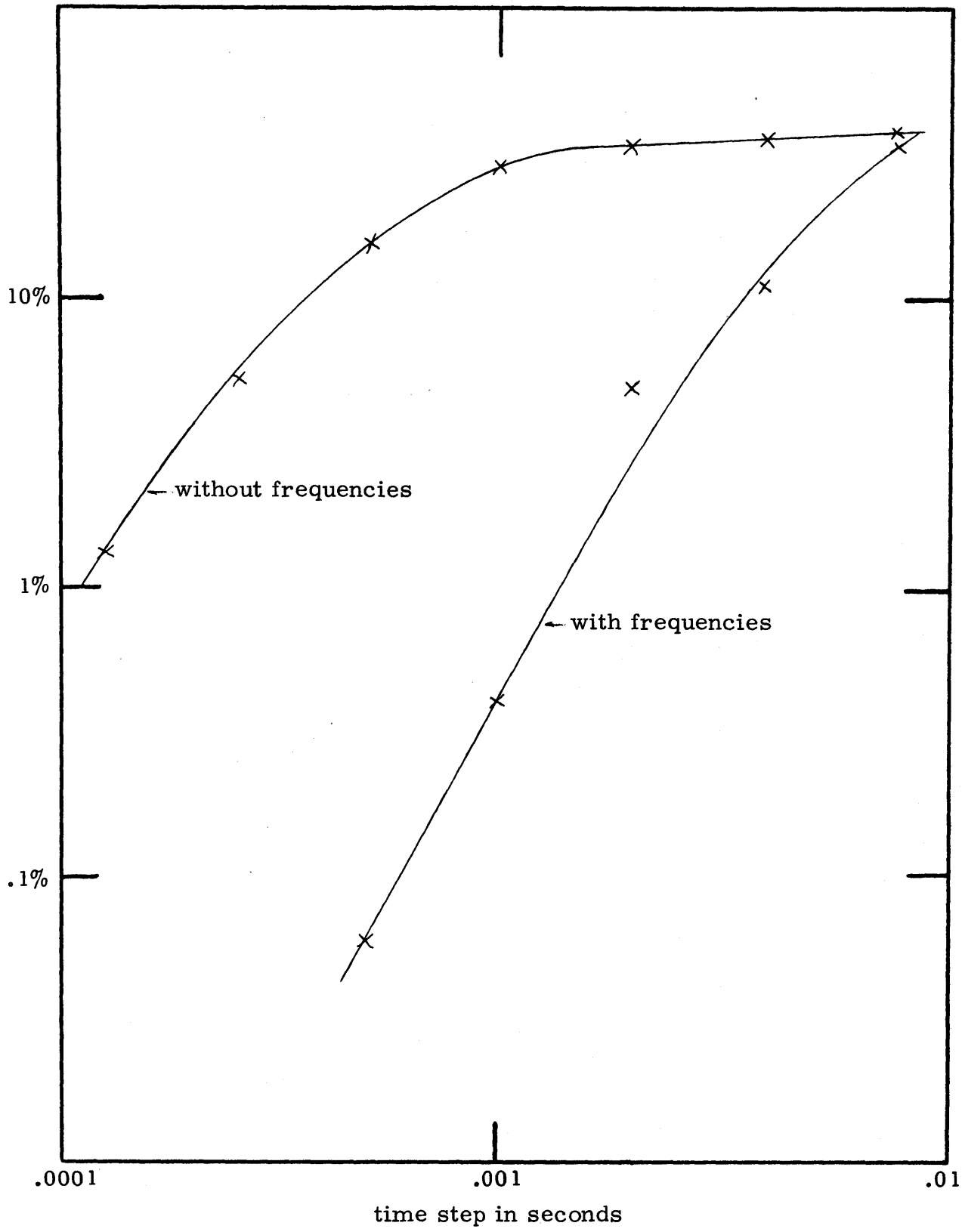


Fig. 3. 1. Error for CASE1.

Note that the thermal flux (without frequencies) for $\Delta t = .008$ and $\Delta t = .004$ differs by only .8%, which in the absence of other information might lead one to conclude that the solution had converged and was accurate to within 1%. Obviously this conclusion would be incorrect. This points out the danger of using the agreement of the solution at two different time steps to establish convergence. However, looking at the differences between the solution at the three largest time steps reveals that the percentage difference actually increases with decreasing time step, revealing that the solution is not converged. This can be used as a quick check on convergence of any method which has a characteristic error curve like Fig. 3.1.

The asymptotic convergence rate of the basic method is $0(\Delta t^2)$ as expected, while the convergence rate for the transformed method is slightly faster.

3. FOURGP – Four Group Bare Homogeneous Fast Reactor

FOURGP is a four energy group, one delayed group fast system. The reactor is a homogeneous square, 150 cm on a side with nine mesh points (ten intervals) in each coordinate direction. A positive step change in reactivity of about 60 cents is inserted by changing the critical value of ν by +.00172. Initial conditions correspond to the steady state of the system. Data are given in Appendix C.

Solutions were obtained with STKADI using the frequency transformation at time steps of $.2 \times 10^{-5}$ and $.4 \times 10^{-5}$ seconds. These are compared with the solution obtained from MITKIN in Table 3.4.

The MITKIN results at a time step of $.4 \times 10^{-5}$ seconds are superior

Table 3.4. FOURGP results – comparison with MITKIN.

Thermal flux, with frequencies.

Time (sec)	Exact	MITKIN (Δt=.4E-5)	STKADI (Δt=.4E-5)	STKADI (Δt=.2E-5)
.00000	.004475			
.00016	.005481	.005473 (-.15%)	.004949 (-9.7%)	.005419 (-1.1%)
.00032	.006381	.006378 (-.05%)	.005927 (-7.1%)	.006352 (-.5%)
.00048	.007149	.007148 (-.01%)	.006931 (-3.0%)	.007139 (-.13%)
.00064	.007804	.007805 (.01%)	.007793 (-.14%)	.007804 (.01%)
.00080	.008362	.008364 (.02%)	.008489 (1.5%)	.008366 (.06%)

Note: Numbers in parentheses ( ) are percentage errors.

to the STKADI results at one-half this time step. Recalling that MITKIN takes less execution time than STKADI, it is apparent that MITKIN is better than STKADI for this problem by a factor of at least two.

4. TWIGL Problems – Two Group Non-Homogeneous Reactor

This series of problems was prepared at Bettis Atomic Power Laboratory (Ref. 11) to test the TWIGL (Ref. 8) Code. The reactor consists of a square core surrounded by a blanket with blanket in the interior as well. It is completely symmetric. The geometry is shown in Fig. 3.2.

The transient is induced by changing the thermal cross section in region 1. Three different transients were studied: a positive step change in reactivity, a positive ramp change in reactivity, and a negative ramp

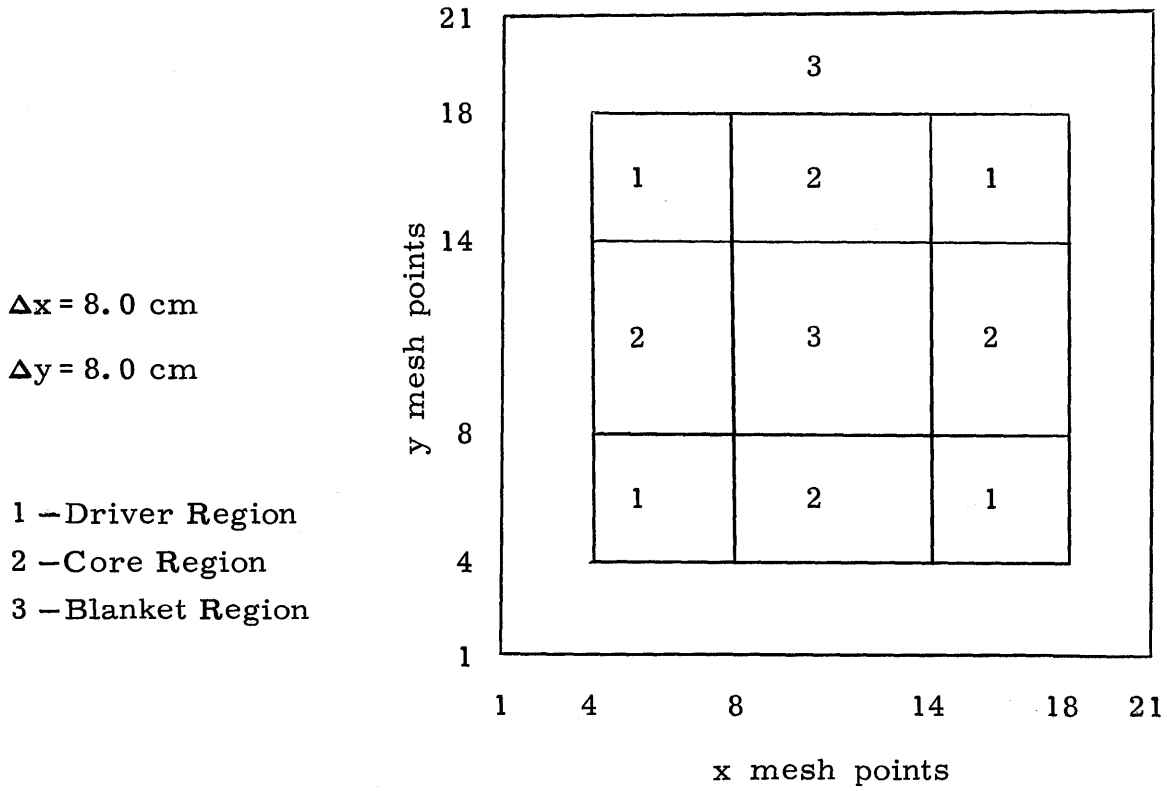


Fig. 3.2. TWIGL geometry.

change. Solutions were obtained from the TWIGL, LUMAC, MITKIN and STKADI codes.

Data are given in Appendix C.

4.1 Positive Step Change

A positive step change in reactivity of about 50 cents is introduced at time zero by reducing the thermal capture cross section in region 1 by .0035 cm^{-1}.

Although the problem is space-dependent, the change in flux shape is quite small. The thermal flux at the mesh point at the exact center of the reactor is used as a basis of comparison. All other points behave similarly.

STKADI solutions were calculated with and without frequency transformation using several different time steps. Results are summarized in Table 3. 5.

Table 3. 5. Results of TWIGL case, positive step change.

Thermal Flux at Center of Core

		Time (sec)			
Δt (sec)	Frequencies	.02	.10	.20	.30
.25 E-04	yes	31.29			
.25 E-04	no	30.71			
.25 E-03	yes	33.21	34.02	34.33	34.64
.25 E-03	no	18.85	25.62	30.05	32.27
.50 E-03	yes	27.56	31.01		
.50 E-03	no	17.18	19.59		
.10 E-02	yes	17.95	53.60		
.10 E-02	no	16.82	17.34		
.20 E-02	yes	16.82	17.61		
.20 E-02	no	16.76	16.84		

Initial flux is 16.75

The great superiority of the transformed method over the untransformed is again evident. The latter is converged at 25 microsecond time steps, while the former is acceptably accurate at a time step ten times as large.

Table 3. 5 shows the tendency of the ADI to approach the identity operator as the time step becomes large. Both methods exhibit this, although

the transformed method approaches unity much more slowly. The result at t = .1 sec, Δt = .0010 seems anomalous. Apparently the frequencies at the beginning of the transient were quite large, and became "locked in" on subsequent steps because the ADI was so close to the identity that the frequencies could not decrease.

The transformed ADI method is compared with the other methods in Table 3.6. The computing times required by STKADI and MITKIN were

Table 3.6. TWIGL results – comparison with other methods.

Thermal Flux at Center of Core

t (sec)	TWIGL (Δt=.001)	LUMAC (EP1=.00008)	MITKIN (Δt=.0002)	STKADI (Δt=.00025)
.00	16.75			
.01	26.70	27.29 (2.2%)	27.32 (2.3%)	27.15 (1.7%)
.02	30.78	31.48 (2.3%)	31.50 (2.3%)	33.21 (7.9%)
.03	32.40	33.06 (2.0%)	33.13 (2.2%)	33.94 (4.8%)
.04	33.15		33.97 (2.5%)	33.90 (2.3%)
.05	33.54		34.63 (3.3%)	33.88 (1.0%)
.10	34.01			34.03 (.05%)
.20	34.31			34.33 (.07%)

Note: Numbers in parentheses ( ) are the percentage deviation from the TWIGL results.

very similar since a larger Δt compensates for STKADI's longer computing time per step. LUMAC takes three times as long for this problem. TWIGL was run on another machine, so that computing times are not comparable.

All the methods agree quite closely so that there appears to be little to choose among the four methods on the basis of accuracy. Hence MITKIN and STKADI are superior to LUMAC because of their shorter running times.

Table 3.7. TWIGL positive ramp – comparison of results.

Thermal Flux in Center of Core

t (sec)	TWIGL (Δt=.01)	LUMAC (EPl=.0008)	MITKIN (Δt=.001)	STKADI (Δt=.00025)
.02			17.26	17.30 (.2%)
.05	18.76		18.79 (.16%)	18.70 (-.33%)
.10	21.74	21.73 (-.03%)	21.75 (.05%)	21.66 (-.37%)
.15	25.96		25.95 (-.02%)	25.84 (-.47%)
.20	32.37	32.49 (.38%)	32.31 (-.16%)	32.15 (-.66%)
.25	34.05		34.12 (.22%)	34.58 (1.58%)
.30	34.24	34.83 (1.7%)	33.57* (-2.0%)	34.29 (.14%)
.35	34.39		33.23 (-3.4%)	34.44 (.15%)
.40	34.54		33.19 (-3.9%)	34.60 (.18%)

*Δt changed to .010 sec.

Note: Numbers in parentheses ( ) are percentage deviations from TWIGL results.

4.2 Positive Ramp Change

A positive ramp change in reactivity of about 2.5 dollars per second is introduced over the time interval $.0 \leq t \leq .2$ seconds by reducing the thermal capture cross section in region 1 by $(.0035)(t/.2)$ cm^{-1}.

The results are compared in Table 3.7 with TWIGL, LUMAC and MITKIN. Agreement is excellent.

4.3 Negative Ramp Change

A negative ramp change in reactivity of about 80 dollars per second is introduced in the interval $.0 \leq t \leq .02$ seconds by increasing the thermal capture cross section in region 1 by $.03\ (t/.02)$ cm^{-1}.

The results are compared in Table 3.8 with TWIGL and MITKIN.

Table 3.8. TWIGL negative ramp – comparison of results.

Thermal Flux at Center of Core

t (sec)	TWIGL ($\Delta t = .001$)	MITKIN ($\Delta t = .0002$)	STKADI ($\Delta t = .00025$)
.000	16.750	16.750	16.750
.010	8.154	7.445 (-8.7%)	8.178 (.3%)
.020	4.594	4.573 (-.5%)	4.605 (.2%)
.030	4.442	4.388 (-1.2%)	4.141 (-.6%)
.040	4.385	4.385 (.0%)	4.377 (-.2%)

Note: Numbers in parentheses ( ) are percentage deviations from TWIGL results.

As with the positive ramp, the agreement is excellent.

5. OBLONG – Non-Homogeneous, Non-Symmetric Reactor

The TWIGL problem is completely symmetric and shows very little change in flux shape over the transient (<5%). A severe test of a space-time method requires a sample problem with no symmetries whatsoever, and a significant change in flux shape. The OBLONG problem was designed for that purpose (Ref. 12).

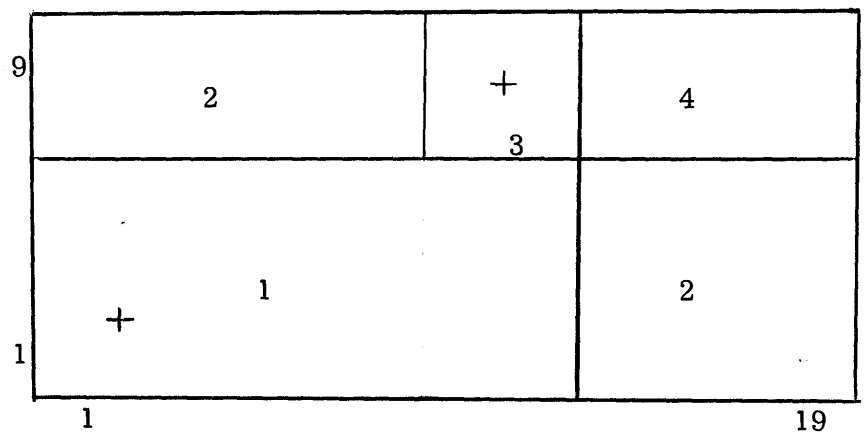
The reactor is a four energy group, one delayed group system. It is a rectangle, 160 by 80 cm divided into four regions as shown in Fig. 3. 3. Region 1 is a driver region, region 2 is a blanket region, and regions 3 and 4 are a water reflector.

The transient is induced by a ramp change in the thermal cross section in region 3 over the time interval $.0 \leq t \leq .2$ seconds. Data are given in detail in Appendix C.

Solutions were obtained from STKADI with the frequency transformation, using various time steps. The results are presented in Tables 3.9 through 3. 12 for the fast and thermal fluxes at the points in regions 1 and 3 shown by crosses in Fig. 3. 3.

The MITKIN results quoted are the most accurate available (Ref. 13). They differ from the MITKIN results at twice and four times the time step by only a few parts in 1000. Also, the good agreement with the LUMAC results, which are converged to within one per cent (Ref. 6), indicates that the MITKIN solution is accurate to better than a fraction of one per cent.

The STAKDI results are in extremely poor agreement. Errors in



Region 1 – driver
Region 2 – blanket
Region 3 – water reflector, perturbation
Region 4 – water reflector
\+ – flux test points
x direction – 19 mesh points
y direction – 9 mesh points

Fig. 3.3. OBLONG geometry.

excess of 5% persist even at a time step of one quarter that used by MITKIN. The results are not even self-consistent – differences between successive STKADI runs are as large as the discrepancy with MITKIN. After the end of the ramp, the solution exhibits a large but damped oscillation. MITKIN also shows this behavior, but to a much lesser extent (Ref. 13).

The failure to converge for reasonable time steps and the excessive oscillations in the solution indicate that the ADI is not a satisfactory method for this problem.

The iterative ADI method (ITRADI) is about 25% efficient for this problem, taking 4 iterations per step to converge. The solution obtained is almost exactly the same as the STKADI solution at the same time step,

Table 3.9. OBLONG results, fast group, region 1.

Group 1 flux at point (2, 8)

t (sec)	MITKIN (Δt=.0005)	LUMAC (EP1=.8E-4)	STKADI (Δt=.000125)	STKADI (Δt=.00025)	STKADI (Δt=.0005)	STKADI (Δt=.0010)	ITRADI (Δt=.0010)
.000	.4463						
.025	.4515		.4468 (-1.0%)		.4463 (-1.2%)	.4462 (-1.2%)	
.050	.4566	.4569 (.06%)	.4525 (-.9%)		.4463 (-2.3%)	.4463 (-2.3%)	
.075	.4620		.4640 (.4%)			.4463 (-3.4%)	
.100	.4677	.4730 (1.1%)	.4781 (2.2%)	.4508 (-3.6%)	.4464 (-4.5%)	.4463 (-4.6%)	.4463 (-4.6%)
.150	.4804	.4830 (.5%)	.4985 (3.8%)			.4463 (-7.1%)	
.200	.4954	.4930 (-.5%)	.5064 (2.2%)	.4918 (-.7%)	.4486 (-9.5%)	.4463 (-9.9%)	.4465 (-9.9%)
.250	.4961		.5225 (5.3%)			.4464 (-10.%)	
.300	.4965	.5123 (3.2%)	.5194 (4.6%)	.5609 (13.%)	.4624 (-6.9%)	.4465 (-10.%)	

Table 3.10. OBLONG results, fast group, region 3.

Group 1 flux at point (11, 2)

t (sec)	MITKIN (Δt=.0005)	LUMAC (EP1=.00008)	STKADI (Δt=.000125)	STKADI (Δt=.00025)	STKADI (Δt=.0005)	STKADI (Δt=.0010)	ITRADI (Δt=.0010)
.000	.1341						
.025	.1363		.1350 (-1.0%)		.1342 (-1.6%)	.1341 (-1.6%)	
.050	.1385	.1385 (.0%)	.1375 (-.7%)		.1346 (-2.8%)	.1342 (-3.1%)	
.075	.1408		.1419 (.8%)			.1343 (-4.6%)	
.100	.1432	.1453 (1.4%)	.1473 (2.8%)	.1394 (-2.7%)	.1371 (-4.3%)	.1346 (-6.0%)	.1352 (-5.6%)
.150	.1488	.1499 (.8%)	.1554 (4.5%)			.1359 (-8.6%)	
.200	.1552	.1551 (-.07%)	.1604 (3.3%)	.1558 (.4%)	.1413 (-8.9%)	.1382 (-11.%)	.1403 (-9.6%)
.250	.1555		.1653 (6.4%)			.1411 (-9.2%)	
.300	.1556	.1605 (3.2%)	.1640 (5.4%)	.1768 (13.%)	.1489 (-4.3%)	.1438 (-7.6%)	

Table 3.11. OBLONG results, thermal group, region 1.

Group 4 flux at point (2, 8)

t (sec)	MITKIN (Δt=.0005)	LUMAC (EP1=.8E-4)	STKADI (Δt=.000125)	STKADI (Δt=.00025)	STKADI (Δt=.0005)	STKADI (Δt=.0010)	ITRADI (Δt=.0010)
.000	.0359						
.025	.0364		.0360 (-1.0%)		.0359 (-1.2%)	.0359 (-1.2%)	
.050	.0368	.0368 (.07%)	.0364 (-.9%)		.0359 (-2.2%)	.0359 (-2.3%)	
.075	.0372		.0374 (.4%)			.0359 (-3.4%)	
.100	.0376	.0381 (1.2%)	.0385 (2.2%)	.0363 (-3.5%)	.0360 (-4.5%)	.0359 (-4.5%)	.0359 (-4.5%)
.150	.0387	.0389 (.6%)	.0401 (3.7%)			.0359 (-7.1%)	
.200	.0399	.0398 (-.2%)	.0408 (2.2%)	.0396 (-.7%)	.0361 (-9.3%)	.0359 (-9.8%)	.0360 (-9.8%)
.250	.0399		.0420 (5.3%)			.0359 (-10.%)	
.300	.0400	.0412 (3.1%)	.0418 (4.6%)	.0451 (13%)	.0373 (-6.7%)	.0360 (-10.%)	

Table 3.12. OBLONG results, thermal group, region 3.

Group 4 flux at point (11, 2)

t (sec)	MITKIN (Δt=.000t)	LUMAC (EPl=.00008)	STKADI (Δt=.000125)	STKADI (Δt=.00025)	STKADI (Δt=.0005)	STKADI (Δt=.0010)	ITRADI (Δt=.0010)
.000	.9684						
.025	1.0101		1.0010 (-.9%)		.9950 (-1.5%)	.9939 (-1.6%)	
.050	1.0540	1.056 (.2%)	1.0474 (-.6%)		1.0255 (-2.7%)	1.0223 (-3.0%)	
.075	1.1013		1.1105 (.8%)			1.0535 (-4.3%)	
.100	1.1525	1.166 (1.2%)	1.1855 (2.9%)	1.1204 (-2.8%)	1.1006 (-4.5%)	1.0873 (-5.6%)	1.0879 (-5.6%)
.150	1.2686	1.278 (.7%)	1.3278 (4.7%)			1.1614 (-8.4%)	
.200	1.4075	1.410 (.2%)	1.4565 (3.5%)	1.4133 (.4%)	1.2914 (-8.2%)	1.2498 (-11.%)	1.2652 (-10.%)
.250	1.4105		1.5023 (6.5%)			1.2726 (-9.8%)	
.300	1.4116	1.451 (2.8%)	1.4920 (5.7%)	1.6094 (14.%)	1.3498 (-4.4%)	1.2889 (-8.7%)	

and is certainly no more accurate. Several other modifications of the basic ADI method were also tried on this problem. They were unstable. Thus it appears that methods based on the splitting (2.4) are inappropriate for this problem.

6. Other Methods

Several variations of the basic method were considered, and rejected. Table 3.13 outlines these variations and the reason for rejection.

Table 3.13. Unusable methods.

	Method	Reason for Rejection
1.	Iterating on the frequency matrix, $\vec{\Omega}$, to obtain an improved approximation	Iteration did not converge for model problem
2.	Selecting $\vec{\Omega}$ at step j from: $\vec{A}\,\vec{x}^{j} = \vec{\Omega}^{j}\,\vec{x}^{j}$	A simple trial problem showed that it was unstable
3.	Leaving group transfer terms entirely explicit	Accurate to only $0(\Delta t)$, requires storing both old and new flux
4.	Using a weighting scheme for the σ_{gg} terms on the diagonal of $\vec{A}$; that is, $\theta\sigma$ was left on the RHS and $(1-\theta)\sigma$ was taken to the LHS	For the OBLONG problem the following results were obtained (with frequencies): $\theta = .0$ – unstable $\theta = .5$ – exactly the same as STKADI $\theta = 1.0$ – unstable
5.	Leaving $\vec{U}$ matrix implicit on both half steps	Unstable for OBLONG problem using frequencies
6.	Weighting schemes for other elements of $\vec{A}$ matrix	A wide variety of ADI schemes were tested at Bettis and found to be unsatisfactory because of high truncation error and instability (Ref. 14)
7.	Recalculating $\vec{\Omega}$ every half step	Unstable for OBLONG problem
8.	Rearranging order of computation to: $(I+hA_1)(I-hA_1)^{-1}$ $(I+hA_2)(I-hA_2)^{-1}$	Unstable for OBLONG problem
9.	Using two or three preceding time steps to calculate the frequency	Either unstable or less accurate than standard method for OBLONG problem, depending on weighting factors used
10.	"Smoothing" the frequency matrix by averaging at each mesh point over four nearest neighbors	Unstable for OBLONG problem

Chapter IV

CONCLUSION

1. Conclusion

The ADI method with the frequency transformation is superior to all other ADI methods considered for the general kinetics problem. However, the Alternating Direction Explicit (ADE) method (Ref. 7) with frequency transformation is comparably accurate in some cases, and far superior in others, to the ADI. Since the ADE is also a faster method, the ADI method is inferior to it for the solution of the general kinetics problem.

The failure to STKADI to treat the OBLONG problem adequately is the decisive factor in this conclusion. To be generally applicable, a space-time method must be able to handle problems with a great deal of spacial dependence. STKADI was not able to do this, and consequently cannot be regarded as a promising method.

2. Recommendations for Further Work

The ADI method handles the spacial differencing by splitting the horizontal and vertical differencing matrices; the ADE by splitting into a lower and an upper triangular matrix. Other differencing schemes (nine point, triangular, etc.) are possible which suggest different splittings, which may lead to even lower truncation error. Work should be continued to find even better splitting methods for the problem (1. 1).

REFERENCES

1. S. Kaplan, A. F. Henry, S. G. Margolis, and J. J. Taylor, "Space-Time Reactor Dynamics," Proc. 3rd U.N. International Conf. on Peaceful Uses of Atomic Energy P/271, 4, 41 (1964).
2. A. F. Henry, "Space-Time Reactor Kinetics," 22.243 Course Notes, MIT (1969), unpublished.
3. D. W. Peacman and H. H. Rachford, Jr., "The Numerical Solution of Parabolic and Elliptic Differential Equations," J. Soc. Ind. Appl. Math. 3, 28 (1955).
4. G. I. Marcuk and N. N. Janenko, "Solving a Multi-Dimensional Kinetics Equation by the Splitting-up Method," Reports of the Academy of Sciences, USSR, Vol 157, No. 6 (pp. 1291-1292) (1964) (in Russian).
5. G. I. Marcuk and U. M. Sultangazin, "About the Convergence of the Splitting-up Method for an Equation of Radiation Transfer," Reports of the Academy of Sciences, USSR, Vol. 161, No. 1 (pp. 66-69) (1965) (in Russian).
6. W. T. McCormick, Jr., "Numerical Solution of the Two-Dimensional Multigroup Kinetics Equations," Ph. D. Thesis, Nuclear Engineering Dept., MIT, MIT-NE-99 (May 1969).
7. W. H. Reed, "Finite Difference Techniques for the Solution of the Reactor Kinetics Equations," Ph. D. Thesis, Nuclear Engineering Dept., MIT, MIT-NE-100 (May 1969).
8. J. B. Yasinsky, M. Natelson, and L. A. Hageman, "TWIGL – A Program to Solve the Two-Dimensional, Two-Group Space-Time Neutron Diffusion Kinetics Equations with Temperature Feedback," WAPD-TM-743 (1968).
9. R. S. Varga, Matrix Iterative Analysis, Prentice Hall, Inc., Englewood, N.J. (1962).
10. R. D. Richtmyer and K. W. Morton, Difference Methods for Initial-Value Problems, Second Edition, Interscience Publishers, New York (1967).
11. J. B. Yasinsky, Personal communication.
12. W. T. McCormick, Jr., Personal communication.
13. W. H. Reed, Personal communication.
14. L. A. Hageman and J. B. Yasinsky, "Comparison of Alternating Direction Time Differencing Methods with Other Implicit Methods for the Solution of the Neutron Group Diffusion Equations," WAPD-T-2203, Bettis Atomic Power Laboratory (March 1969).
15. M. Clark, Jr. and K. F. Hansen, Numerical Methods of Reactor Analysis, Academic Press, New York (1964).

16. J. R. Lamarsh, Introduction to Nuclear Reactor Theory, Addison-Wesley, Reading, Mass. (1966).

17. E. Isaacson and H. B. Keller, Analysis of Numerical Methods, John Wiley and Sons, Inc., New York (1966).

18. K. F. Hansen, "A Comparative Review of Two-Dimensional Kinetics Methods," GA-8169 (August 1967).

19. J. B. Andrews II, "Numerical Solution of the Space-Dependent Reactor Kinetics Equations," Ph. D. Thesis, Nuclear Engineering Dept., MIT (May 1967).

20. J. B. Andrews II, "Numerical Solution of the Time-Dependent Multigroup Diffusion Equations," Nucl. Sci. Eng. 31, 304 (1968).

21. J. N. Franklin, Matrix Theory, Prentice Hall, Inc., Englewood Cliffs, N.J. (1968).

22. K. F. Hansen and S. R. Johnson, "GAKIN, A Program for the Solution of the One-Dimensional Multigroup Space-Time Dependent Diffusion Equations," USAEC Report GA-7543, General Atomic Division, General Dynamics Corp. (1967).

Appendix A

THE NEUTRON DIFFUSION KINETICS EQUATIONS

The time-dependent neutron flux in the multigroup diffusion model is given by

$$\frac{1}{v_g}\frac{\partial \phi_g}{\partial t} = \nabla \cdot D_g \nabla \phi_g + \chi_g \sum_{g'=1}^{G} (\nu \sigma_f)_{g'} (1-\beta)\phi_{g'}$$

$$+ \sum_{g'=1}^{G} \tau_{gg'}\phi_{g'} + \sum_{i=1}^{I} f_{gi}\lambda_i c_i, \qquad \text{(A. 1)}$$

and the delayed neutron precursor density is given by

$$\frac{\partial c_i}{\partial t} = \sum_{g'=1}^{G} \beta_i (\nu \sigma_f)_{g'}\phi_{g'} - \lambda_i c_i, \qquad \text{(A. 2)}$$

where the symbols are defined in Table A. 1 (Refs. 1, 2, 15, 16, 18, 19, 20). We approximate the diffusion term $\nabla \cdot D_g \nabla \phi_g$ for two-dimensional geometry using the five-point central differencing scheme (Refs. 2, 15, 17):

$$\nabla \cdot D_g \nabla \phi_g \approx D_g/\Delta x^2 \{\phi_{g,l,k+1} - 2\phi_{g,l,k} + \phi_{g,l,k-1}\}$$

$$+ D_g/\Delta y^2 \{\phi_{g,l+1,k} - 2\phi_{g,l,k} + \phi_{g,l-1,k}\}, \qquad \text{(A. 3)}$$

shown schematically in Fig. A. 1. We number the fluxes at each mesh point as shown, and place them in a column matrix:

$$\vec{\Phi}_g = \begin{bmatrix} \phi_{g,1} \\ \phi_{g,2} \\ \vdots \\ \phi_{g,N} \end{bmatrix} . \qquad \text{(A. 4)}$$

The central difference operator then becomes a pentadiagonal matrix, $\vec{W}_g$, which couples the flux at each point to the four neighboring points. (Fig. A. 2.) $\vec{W}_g$ can be split into a tridiagonal differencing matrix in one direction, $\vec{Y}_g$, and a tridiagonal differencing matrix in the other direction, $\vec{X}_g$ (Figs. A. 3 and A. 4).

Equations (A. 1) and (A. 2) can now be written in matrix form:

$$\frac{d\Phi_g}{dt} = \vec{v}_g \left\{ \vec{W}_g \vec{\Phi}_g + \sum_{g'=1}^{G} \vec{T}_{gg'} \vec{\Phi}_{g'} + \sum_{i=1}^{I} \vec{F}_{gi} \vec{C}_i \right\} , \quad 1 \leq g \leq G \qquad \text{(A. 5)}$$

$$\frac{d\vec{C}_i}{dt} = \sum_{g'=1}^{G} \vec{P}_{ig'} \vec{\Phi}_{g'} - \vec{\Lambda}_i \vec{C}_i, \quad 1 \leq i \leq I \qquad \text{(A. 6)}$$

where the matrices are defined in Table A. 2. $\vec{\Phi}_g, \vec{C}_i$ are column matrices, $\vec{v}_{g'}, \vec{T}_{gg'}, \vec{P}_{ig}$, and $\vec{\Lambda}_i$ are diagonal matrices.

The system of Eqs. (A. 5, A. 6) can be written in matrix form as

$$\frac{d\vec{\Psi}}{dt} = \vec{A}\, \vec{\Psi}, \qquad \text{(A. 7)}$$

where $\vec{\Psi}$ is the column matrix:

Table A. 1. Definition of symbols – scalars.

g	neutron energy group index
G	total number of neutron energy groups
i	delayed precursor group index
I	total number of delayed groups
k	vertical spacial index
Nk	total number of mesh points in vertical direction
l	horizontal spacial index
Nl	total number of mesh points in horizontal direction
N	total number of mesh points
j	time step index
t	time (seconds)
T	final time, end of transient
Δt	time step (seconds)
h	one-half time step ($h=\Delta t/2$)
v_g	group speed (cm/sec)
ϕ_g	group scalar flux in cm^{-2}/sec
D_g	group diffusion coefficient (cm)
χ_g	group fission yield
ν	number of neutrons per fission (may depend on g)
$(\sigma_f)_g$	macroscopic group fission cross section (cm^{-1})
$\tau_{gg'}$	macroscopic transfer cross section from group g' to group g (for $g'=g$, τ_{gg} is minus the group removal cross section)
f_{gi}	fractional yield of i^{th} group precursors into group g
λ_i	delayed neutron decay constant (sec^{-1})
β_i	delayed group yield fraction
β	total delayed yield
c_i	precursor group concentration
$\nabla\cdot$	divergence operator
∇	gradient operator
$\sum$	summation operator

$$\vec{\Psi} = \begin{bmatrix} \vec{\Phi}_1 \\ \vec{\Phi}_2 \\ \vdots \\ \vec{\Phi}_G \\ \vec{C}_1 \\ \vdots \\ \vec{C}_I \end{bmatrix}, \qquad \text{(A. 8)}$$

and $\vec{A}$ is the 'A' matrix shown in Fig. A. 5. Equation (A. 7) is the multigroup diffusion problem in "semi-discrete" form.

Table A. 2. Definition of symbols – matrices.

$\vec{\Phi}_g$	flux vector
$\vec{C}_i$	delayed precursor vector
$\vec{T}_{gg'}$	intragroup transfer matrix
$\vec{F}_{gi}$	delayed group to energy group transfer matrix
$\vec{P}_{ig}$	delayed group production matrix
$\vec{W}_g$	pentadiagonal diffsuion matrix for two dimensions
$\vec{X}_g$	tridiagonal diffusion matrix in x direction
$\vec{Y}_g$	tridiagonal diffusion matrix in y direction
$\vec{\Lambda}_i$	diagonal delayed precursor decay constant matrix
$\vec{I}$	the identity matrix
$\vec{\Psi}$	total flux vector (all group fluxes plus all delayed precursors)
$\vec{A}$	the 'A' matrix
$\vec{B}^j(h, \Omega)$	the advancement matrix which takes $\vec{\Psi}^j$ into $\vec{\Psi}^{j+1}$
$\vec{\Omega}$	frequency transformation matrix
$\|\cdot\|$	any natural matrix norm

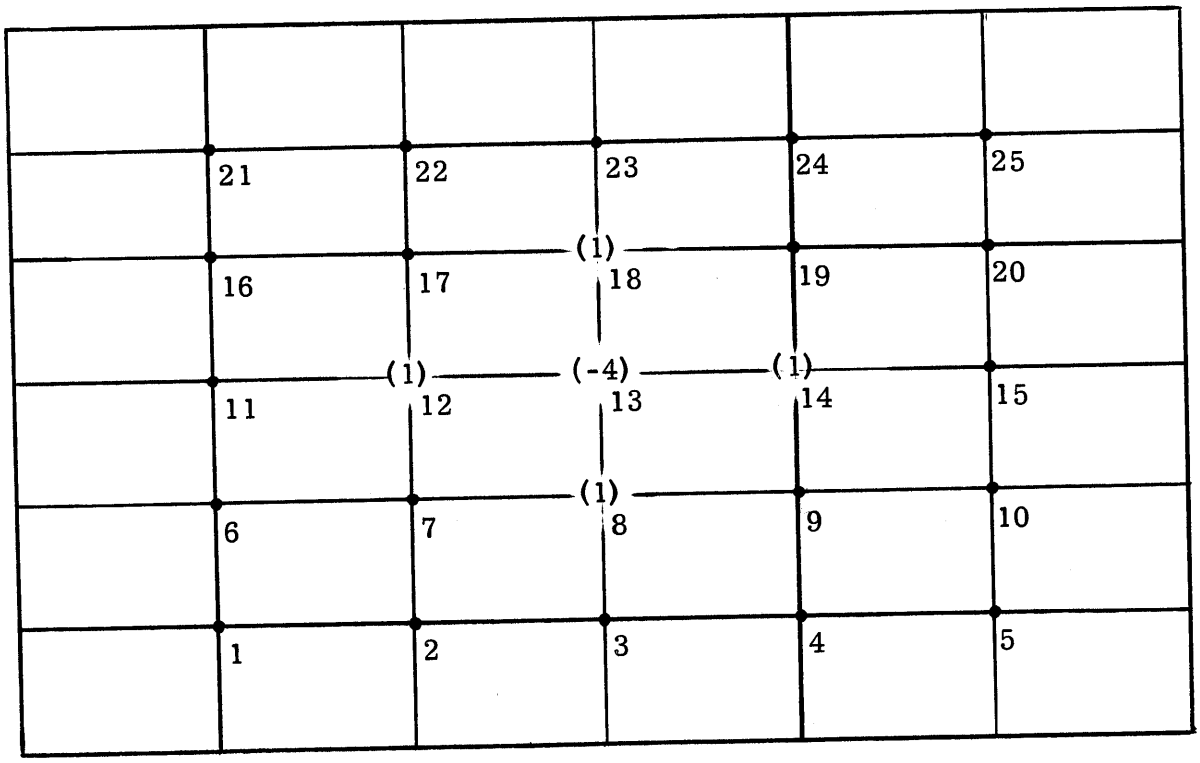


Fig. A. 1. Two-dimensional mesh.

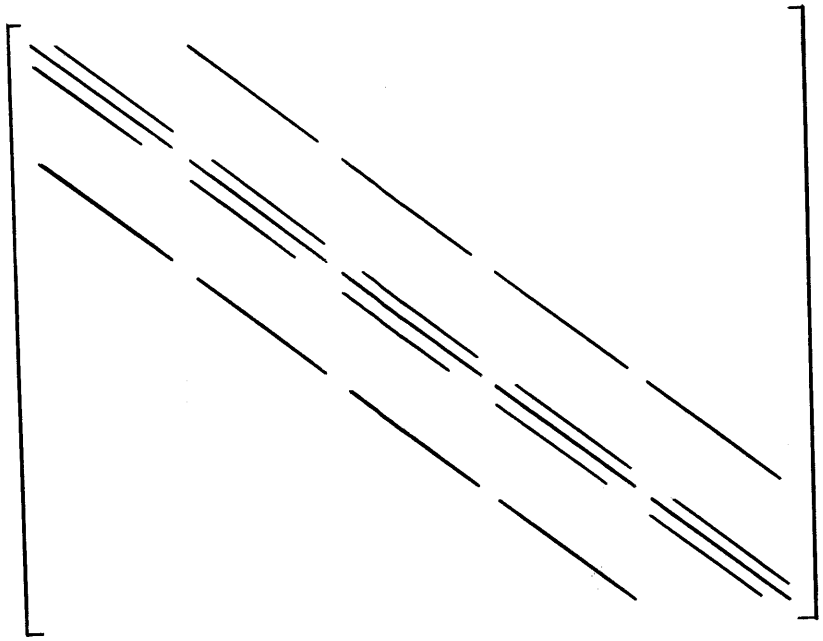


Fig. A. 2. Central differencing matrix, $\vec{W}_g$.

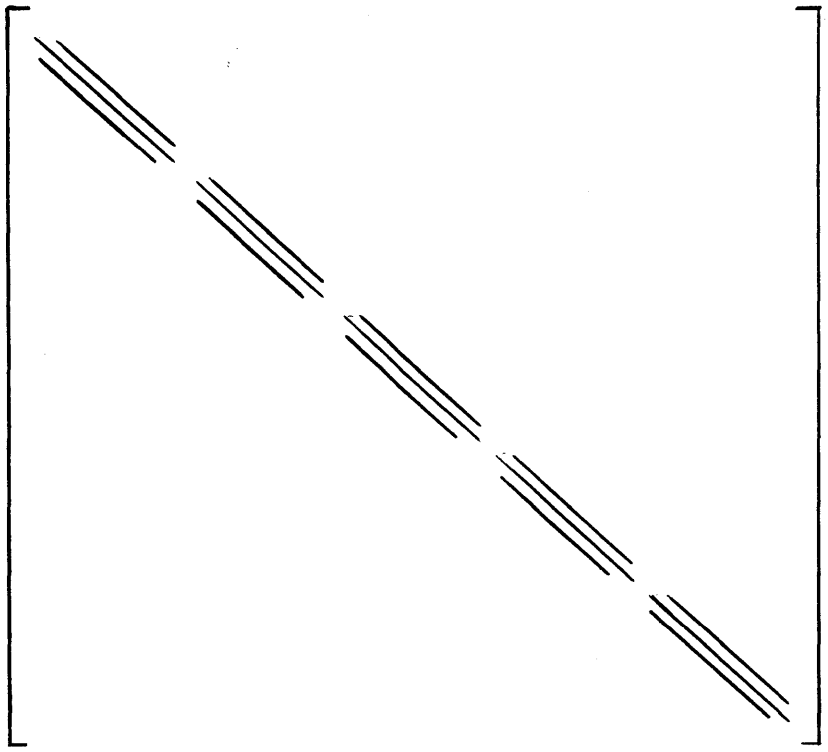


Fig. A.3. Differencing matrix, $\vec{Y}_g$, for 'y' direction.

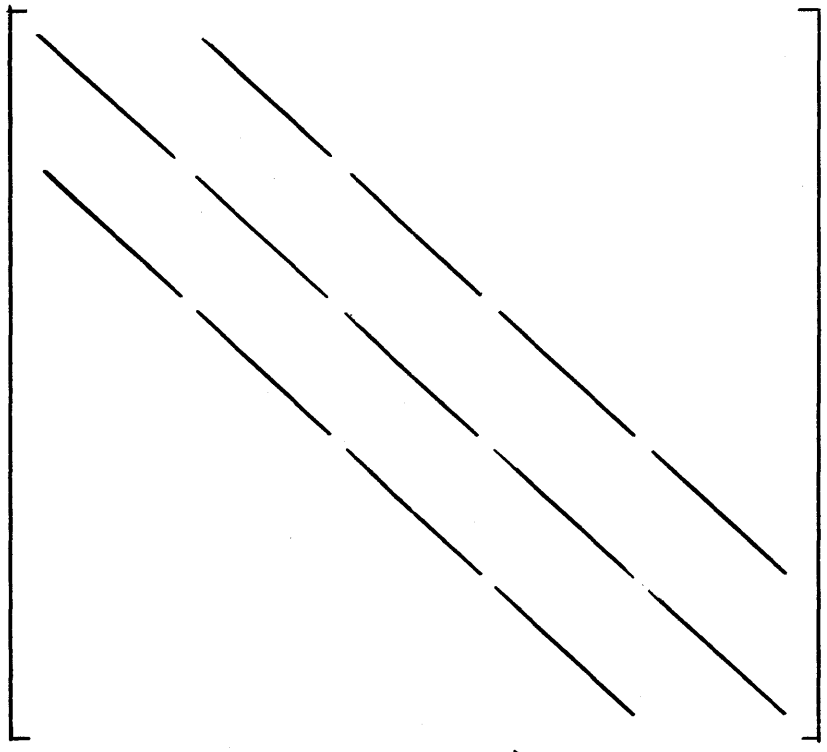


Fig. A.4. Differencing matrix, $\vec{X}_g$, for 'x' direction.

$$
A = \begin{vmatrix}
W_1 & T_{12} & T_{13} & T_{14} & F_{11} & F_{12} & F_{13} \\
T_{21} & W_2 & T_{23} & T_{24} & F_{21} & F_{22} & F_{23} \\
T_{31} & T_{32} & W_3 & T_{34} & F_{31} & F_{32} & F_{33} \\
T_{41} & T_{42} & T_{43} & W_4 & F_{41} & F_{42} & F_{43} \\
P_{11} & P_{12} & P_{13} & P_{14} & \Lambda_1 & & \\
P_{21} & P_{22} & P_{23} & P_{24} & & \Lambda_2 & \\
P_{31} & P_{32} & P_{33} & P_{34} & & & \Lambda_3
\end{vmatrix}
$$

Example shown is for a 4 energy group, 3 delayed group problem.

Fig. A. 5. The 'A' matrix.

Appendix B

THEOREMS

Theorem 1 — As t approaches infinity, the solution vector $\vec{\psi}(t) = \exp(\vec{A}t)\vec{\psi}_o$ approaches $a\, e^{\omega_o t}\, \vec{e}_o$, where $a = (\vec{\Psi}_o, \vec{e}_o)$, and ω_o is the largest eigenvalue of $\vec{A}$.

Proof — We write $\vec{\psi}_o$ as a linear combination of $\vec{e}_o$ and $\vec{v}$, where $(\vec{v}, \vec{e}_o) = 0$, that is, $\vec{\psi}_o = a\,\vec{e}_o + \beta\,\vec{v}$. Now

$$a(\vec{e}_o, \vec{e}_o) + \beta(\vec{e}_o, \vec{v}) = (\vec{e}_o, \vec{\psi}_o),$$

or

$$a = (\vec{e}_o, \vec{\psi}_o),$$

since

$$(\vec{e}_o, \vec{e}_o) = 1.$$

We can now write

$$\vec{\psi}(t) = \exp(\vec{A}t)(a\,\vec{e}_o + \beta\,\vec{v})$$

$$= a\, \exp(\omega_o t)\, \vec{e}_o + \beta\, \exp(\vec{A}t)\, \vec{v}$$

$$= a\, \exp(\omega_o t)(\vec{e}_o + \beta/a\, \exp(\vec{B}t)\, \vec{v}),$$

where

$$\vec{B} = \vec{A} - \omega_o \vec{I}.$$

Note that the largest eigenvalue of $\vec{B}$ is 0, and all the others are given by $\lambda_i = \omega_i - \omega_o$ and have real parts less than zero. (See Sec. 2. 1.) Now we put $\vec{B}$ in Jordan form:

$$J = \vec{S}^{-1}\vec{B}\vec{S} = \begin{bmatrix} \vec{J}_1 & & & \vec{0} \\ & \vec{J}_2 & & \\ & & \vec{J}_3 & \\ & & & \ddots \\ \vec{0} & & & \end{bmatrix}, \qquad \text{(B. 1)}$$

where each of the blocks on the diagonal is of the form

$$J_i = \begin{bmatrix} \lambda_i & 1 & & & \vec{0} \\ & \lambda_i & 1 & & \\ & & \lambda_i & 1 & \\ & & & \ddots & \ddots \\ & \vec{0} & & & \ddots \end{bmatrix}. \qquad \text{(B. 2)}$$

$\vec{J}_i$ is a p_i by p_i matrix where p_i is less than or equal to the multiplicity of the i^{th} eigenvalue, and the λ_i's are arranged in order of nonincreasing real part. $\vec{J}_1$ is a 1×1 matrix since the largest eigenvalue of $\vec{B}$ is simple. Now

$$\exp(\vec{B}t)\,\vec{v} = \exp(\vec{S}^{-1}\vec{J}\,\vec{S}\,t)\,\vec{v}$$

$$= (I + \vec{S}^{-1}(\vec{J}\,t)\,\vec{S} + 1/2!\,\vec{S}^{-1}(\vec{J}\,t)^2\,\vec{S} + \ldots)\,\vec{v}$$

$$= \vec{S}^{-1}\exp(\vec{J}\,t)\,\vec{S}\,\vec{v} = \vec{S}^{-1}\exp(\vec{J}\,t)\,\vec{a}, \qquad \vec{a} = \vec{S}\,\vec{v}. \qquad \text{(B. 3)}$$

But

$$\exp(\vec{J}t) = \begin{bmatrix} 1 & & & 0 \\ & \exp(\vec{J}_2 t) & & \\ & & \exp(\vec{J}_3 t) & \\ & & & \ddots \\ 0 & & & \end{bmatrix}. \qquad \text{(B. 4)}$$

Furthermore, since $\vec{A}$ and $\vec{B}$ share the same eigenvectors, $\vec{e}_o$ is the eigenvector of $\vec{B}$ corresponding to eigenvalue 0, and the transformation $\vec{S}$ also puts $\vec{A}$ into Jordan form. That is,

$$\vec{J}'\vec{S} = \vec{S}\vec{A}, \qquad \vec{S}^{-1}\vec{J}'\vec{S}\vec{e}_o = \vec{A}\vec{e}_o = \omega_o\vec{e}_o, \qquad \vec{J}'\vec{S}\vec{e}_o = \omega_o\vec{S}\vec{e}_o,$$

$$\vec{J}' = \begin{bmatrix} \omega_o & & & \vec{0} \\ & \vec{J}'_2 & & \\ & & \vec{J}'_3 & \\ & & & \ddots \\ \vec{0} & & & \end{bmatrix}, \qquad \text{(B. 5)}$$

Hence

$$\vec{S}\vec{e}_o = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ \end{bmatrix}, \qquad \text{and} \qquad \vec{S} = \begin{bmatrix} \vec{e}_o^{\,T} \\ x \\ x \\ \cdots \end{bmatrix}.$$

Thus

$$\vec{S}\,\vec{v} = \begin{bmatrix} \vec{e}_o^{\,T}\vec{v} \\ x \\ x \\ \cdots \end{bmatrix} = \begin{bmatrix} 0 \\ x \\ x \\ \cdots \end{bmatrix},$$

that is, the first element of $\vec{S}\,\vec{v}$ is zero since $\vec{e}_o$ is orthogonal to $\vec{v}$.

Now

$$\exp(\vec{J}t)\,\vec{S}\,\vec{v} = \begin{bmatrix} 1 & & & \vec{0} \\ & \exp(\vec{J}_2 t) & & \\ & & \exp(\vec{J}_3 t) & \\ \vec{0} & & & \ddots \end{bmatrix} \begin{bmatrix} 0 \\ \vec{a}_2 \\ \vec{a}_3 \\ \vdots \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ \exp(\vec{J}_2 t)\,\vec{a}_2 \\ \exp(\vec{J}_3 t)\,\vec{a}_3 \\ \cdots \end{bmatrix}. \qquad \text{(B.6)}$$

Thus $\|\vec{S}^{-1}\exp(\vec{J}t)\vec{S}\,\vec{v}\| \leq \|\vec{S}^{-1}\| \sum_{i=2}^{n} \|\exp(\vec{J}_i t)\| \cdot \|\vec{a}_i\|$ approaches $\|\vec{S}^{-1}\| \sum_{i=2}^{n} \|\vec{a}_i\| \frac{t^{p_i-1}}{(p_i-1)!} \exp(t\mathrm{Re}\lambda_i)$ as t approaches infinity, using Lemma 8.1 from Varga (Ref. 9). Now, since $\mathrm{Re}\,\lambda_i$ is less than zero, all i greater than 1, this norm goes to zero for large t. Hence $\|\vec{e}_o + \beta/\alpha \exp(\vec{B}t)\,\vec{v} - \vec{e}_o\| = \|\beta/\alpha \exp(\vec{B}t)\,\vec{v}\|$ approaches zero as t approaches infinity, and the vector $\vec{e}_o + \beta/\alpha \exp(\vec{B}t)\,\vec{v}$ approaches $\vec{e}_o$,

completing the proof.

Theorem 2 – The matrix $(\vec{I} - h\vec{A}_1)^{-1}$ is non-negative for all h, provided that the diagonal absorption term, σ_{gg} is negative everywhere for all groups.

First we must prove the following lemma:

Lemma 1 – If $\vec{a}$ is a diagonally dominant, real n by n matrix, and $a_{ij} \leq 0$ for all $i \neq j$, and $a_{ii} > 0$, all i, then $\vec{a}^{-1}$ is non-negative and non-singular.

Proof – Let $\vec{a} = \vec{D} - (\vec{U}+\vec{L})$, then $\vec{D}^{-1}\vec{a} = \vec{I} - \vec{D}^{-1}(\vec{U}+\vec{L})$. Now $\vec{D}$ and $\vec{D}^{-1}$ are non-negative since $\vec{D}$ contains the diagonal terms of $\vec{a}$. Also $\vec{D}^{-1}(\vec{U}+\vec{L})$ is non-negative. Now

$$\rho(\vec{D}^{-1}(\vec{U}+\vec{L})) \leq \max_{i\neq j} \left| \frac{a_{ij}}{a_{ii}} \right| < 1 \qquad \text{(B. 7)}$$

since $\vec{a}$ is diagonally dominant. Let $\vec{B} = \vec{D}^{-1}(\vec{U}+\vec{L})$, then since $\rho(\vec{B}) < 1$, the series

$$(\vec{I} - \vec{B})^{-1} = \vec{I} + \vec{B} + \vec{B}^2 + \ldots$$

converges (see Isaacson and Keller, Ref. 17, p. 15). But since $\vec{B} \geq 0$, $(\vec{I} - \vec{B})^{-1} = \vec{I} + \vec{B} + \vec{B}^2 + \ldots \geq 0$, $(\vec{D}^{-1}\vec{a})^{-1} \geq 0$, and $\vec{a}^{-1}\vec{D} \geq \vec{0}$. Hence $\vec{a}^{-1} \geq \vec{0}$.

We can write the matrix $(\vec{I} - h\vec{A}_1)^{-1}$ as

$$(\vec{I} - h\vec{A}_1)^{-1} = (\vec{I} - h(\vec{X}+\vec{U}))^{-1} = (\vec{I} - h(\vec{I} - h\vec{X})^{-1}\,\vec{U})^{-1}\,(\vec{I} - h\vec{X})^{-1}.$$

Now $\vec{X}$ has positive off-diagonal terms and negative diagonal terms.

$\vec{X}$ is diagonally dominant if $\sigma_{gg} < 0$. Thus $(\vec{I} - h\vec{X})$ satisfies the conditions of the Lemma, and thus $(\vec{I} - h\vec{X})^{-1}$ is non-negative. Now since $\vec{U}$ is non-negative,

$$\vec{V} = h(\vec{I} - h\vec{X})^{-1} \vec{U} \geq \vec{0}.$$

Now $\vec{V}$ is block upper triangular so that $\vec{V}^m = \vec{0}$ where m is the number of energy groups plus the number of delayed groups. Hence the series

$$\vec{I} + \vec{V} + \vec{V}^2 + \vec{V}^3 + \ldots + \vec{V}^{m-1}$$

converges and is equal to $(\vec{I} - \vec{V})^{-1}$. Thus $(\vec{I} - h(\vec{I} - h\vec{X})^{-1} \vec{U})^{-1} \geq \vec{0}$. Thus $(\vec{I} - h(\vec{I} - h\vec{X})^{-1} \vec{U})^{-1} (\vec{I} - h\vec{X})^{-1} \geq \vec{0}$, and the theorem is proved.

Corollary – $(\vec{I} - h\vec{A}_2)^{-1} \geq \vec{0}$, all h, if $\sigma_{gg} < 0$.

Proof – Replace $\vec{A}_1$ by $\vec{A}_2 = \vec{Y} + \vec{L}$ in the proof of the Theorem. The result follows identically.

Extension – $(\vec{I} - h\vec{A}_1)^{-1}$ and $(\vec{I} - h\vec{A}_2)^{-1}$ are non-negative if $h\sigma_{gg} < 1$.

Proof – The matrix $(\vec{I} - h\vec{X})$ is still diagonally dominant if $1 > h\sigma_{gg}$. Hence the result of the theorem still follows. Similarly for $(\vec{I} - h\vec{A}_2)^{-1}$.

Theorem 3 – if $\vec{a} = \vec{a}_1 + \vec{a}_2$ is symmetric, then the advancement matrix $\vec{\mathcal{B}}^j(h)$ is unconditionally stable if and only if all the eigenvalues of $\vec{a}_1$ and $\vec{a}_2$ are less than or equal to zero.

Proof – The advancement matrix $\vec{\mathcal{B}}^j(h)$ is given by

$$\vec{\mathcal{B}}^j(h) = (\vec{I} - h\vec{a}_2)^{-1} (\vec{I} + h\vec{a}_1)(\vec{I} - h\vec{a}_1)^{-1} (\vec{I} + h\vec{a}_2). \qquad \text{(B. 8)}$$

Now the similarity transformation

$$(\vec{I} - h\vec{a}_2)\ \vec{\mathcal{B}}^j(h)(\vec{I} - h\vec{a}_2)^{-1} = (\vec{I}+h\,\vec{a}_1)(\vec{I} - h\,\vec{a}_1)^{-1}\ (\vec{I}+h\,\vec{a}_2)(\vec{I} - h\vec{a}_2)^{-1} \tag{B.9}$$

does not alter the eigenvalues if $(\vec{I} - h\vec{a}_2)$ is nonsingular. Thus

$$\begin{aligned}
\rho(\vec{\mathcal{B}}^j(h)) &= \rho((\vec{I}+h\vec{a}_1)(\vec{I} - h\vec{a}_1)^{-1}\ (\vec{I}+h\,\vec{a}_2)(\vec{I} - h\vec{a}_2)^{-1}) \\
&\leq \|(\vec{I}+h\vec{a}_1)(\vec{I} - h\vec{a}_1)^{-1}\ (\vec{I}+h\,\vec{a}_2)(\vec{I} - h\vec{a}_2)^{-1}\| \\
&\leq \|(\vec{I}+h\,\vec{a}_1)(\vec{I} - h\,\vec{a}_1)^{-1}\| \cdot \|(\vec{I}+h\vec{a}_2)(\vec{I} - h\vec{a}_2)^{-1}\| \\
&= \rho((\vec{I}+h\,\vec{a}_1)(\vec{I} - h\vec{a})^{-1}) \cdot \rho((\vec{I}+h\,\vec{a}_2)(\vec{I} - h\vec{a}_2)^{-1}),
\end{aligned}$$

since $\vec{a}_1$ and $\vec{c}_2$ are Hermitian. Consequently,

$$\rho(\vec{\mathcal{B}}^j(h)) \leq \max_i \left| \frac{1 + ha_i^1}{1 - ha_i^1} \right| \cdot \max_i \left| \frac{1 + ha_i^2}{1 - ha_i^2} \right|, \tag{B.10}$$

where a_i^1, a_i^2 are the eigenvalues of $\vec{a}_1, \vec{a}_2$, respectively.

If all the a's are less than zero, $\rho(\vec{\mathcal{B}}^j(h)) < 1$, and $\vec{\mathcal{B}}^j(h)$ is stable since a matrix is stable if its largest eigenvalue is less than 1. If all $a_i^1 \leq 0$ and all $a_i^2 < 0$, or all $a_i^1 < 0$ and all $a_i^2 \leq 0$, then $\rho(\vec{\mathcal{B}}^j(h)) < 1$ and $\vec{\mathcal{B}}^j(h)$ is again stable.

Furthermore, if the largest a_i^1 and a_i^2 are both zero, then $\rho(\vec{\mathcal{B}}^j(h) \leq 1$, and the possibility of $\rho(\vec{\mathcal{B}}^j(h)) = 1$ exists. If $\rho(\vec{\mathcal{B}}^j(h)) = 1$, then there must exist an eigenvector, $\vec{x}$, of $\vec{\mathcal{B}}^j(h)$ such that

$$\vec{\mathcal{B}}^j(h)\ \vec{x} = \vec{x} = (\vec{I} + 2h(\vec{I} - h\vec{a}_2)^{-1}\ (\vec{I} - h\,\vec{a}_1)^{-1}\ \vec{a})\ \vec{x},$$

or

$$(\vec{I} - h\vec{a}_2)^{-1} (\vec{I} - h\vec{a}_1)^{-1} \vec{a}\,\vec{x} = \vec{0}.$$

Now $(\vec{I} - h\vec{a}_2)^{-1}$ and $(\vec{I} - h\vec{a}_1)^{-1}$ are nonsingular, non-negative matrices, hence we must have $\vec{a}\,\vec{x} = \vec{0}$, which requires that $\vec{a}$ have eigenvalue zero with eigenvector $\vec{x}$. Now since $\vec{\mathcal{B}}^j(h)\,\vec{x}$ is a consistent approximation to $\exp(\vec{a}\,t)\,\vec{x}$, then for sufficiently small h, $\vec{\mathcal{B}}(h)\,\vec{x}$ is arbitrarily close to $\exp(\vec{a}\,t)\,\vec{x}$. But $\vec{\mathcal{B}}^j(h)\,\vec{x} = \vec{x}$ independently of the time step, thus $\exp(\vec{a}\,t)\,\vec{x} = \vec{x} = \exp(\omega_o t)\,\vec{x}$ as t approaches infinity, and the largest eigenvalue of $\vec{a}$ is zero. Thus $\vec{x}$ is the solution vector, and $\vec{\mathcal{B}}^j(h)$ applied to any other vector will decrease. Thus $\vec{\mathcal{B}}^j(h)$ is stable in this case also.

However, if any $a_i^1 > 0$, then for the time step for which $1 - ha_i^1 = 0$, the matrix $(I - h\vec{a}_1)$ is singular, and the solution will exhibit an unbounded growth. Thus the ADI is not unconditionally stable. Similarly, if any $a_i^2 > 0$, the ADI is not unconditionally stable.

Theorem 4 — The matrix $\vec{Q}_1(h) = (\vec{I} - h\vec{A}_1)^{-1}\,\vec{U}(\vec{I} - h\vec{X})^{-1}$ is bounded in norm for positive h if σ_{gg} is negative.

Proof — The matrix $(I - h\vec{A}_1)^{-1}\,\vec{U}$ can be written

$$(\vec{I} - h\vec{A}_1)^{-1}\,\vec{U} = (\vec{I} - (\vec{I} - h\vec{X})^{-1} h\vec{U})^{-1} (\vec{I} - h\vec{X})^{-1}\,\vec{U}$$

$$= (\vec{I} - h\vec{V})^{-1}\,\vec{V} \qquad \text{(B. 11)}$$

where $\vec{V} = (\vec{I} - h\vec{X})^{-1}\,\vec{U}$. Now $(\vec{I} - h\vec{X})^{-1}$ is a block diagonal matrix, and $\vec{U}$ is nonzero only in blocks above the diagonal. Thus $\vec{V}$ is nonzero only in the blocks above the diagonal and $\vec{V}^m = \vec{0}$, where m is the number of

energy plus delayed groups. Consequently the series

$$(I - h\vec{V})^{-1} = (I + (h\vec{V}) + (h\vec{V})^2 + \ldots (h\vec{V})^{m-1}$$

converges. Then

$$(I - h\vec{V})^{-1}\,\vec{V} = (I + (h\vec{V}) + (h\vec{V})^2 + \ldots (h\vec{V})^{m-1})\vec{V}$$

$$= \frac{1}{h}((h\vec{V}) + (h\vec{V})^2 + \ldots (h\vec{V})^{m-1}).$$

Thus

$$\|(I - h\vec{V})^{-1}\,\vec{V}\| \leq \frac{1}{h}\{\|h\vec{V}\| + \|h\vec{V}\|^2 + \ldots \|h\vec{V}\|^{m-1}\}$$

$$= \frac{1}{h}\begin{cases} \dfrac{\|(h\vec{V})\|^m - 1}{\|h\vec{V}\| - 1} & \|h\vec{V}\| \neq 1 \\ \\ m & \|h\vec{V}\| = 1. \end{cases} \qquad \text{(B. 12)}$$

Now,

$$\|h\vec{V}\| = \|(\vec{I} - h\vec{X})^{-1}\,h\vec{U}\|.$$

At this stage we need the following:

Lemma — If the matrix $(r\vec{I} - \vec{B})^{-1}$ is non-negative and nonsingular for $r \geq 0$, then $(r\vec{I} - \vec{B})^{-1}$ is a nonincreasing function of r.

Proof — Let us consider an r_1 and r_2 such that $r_2 - r_1 = \epsilon > 0$, where ϵ is small. Thus

$$(r_1\vec{I} - \vec{B})(r_2\vec{I} - \vec{B})^{-1} = \vec{I} - \epsilon(r_2\vec{I} - \vec{B})^{-1},$$

and

$$(r_2\vec{I} - \vec{B})(r_1\vec{I} - \vec{B})^{-1} = (\vec{I} - \epsilon(r_2\vec{I} - \vec{B})^{-1})^{-1}$$

$$= \vec{I} + \epsilon(r_2\vec{I} - \vec{B})^{-1} + (\epsilon(r_2\vec{I} - \vec{B})^{-1})^2 + \ldots$$

which converges for sufficiently small ϵ to

$$(r_2\vec{I} - \vec{B})(r_1\vec{I} - \vec{B})^{-1} = \vec{I} + \vec{C},$$

with $\vec{C} \geqslant 0$. Thus

$$(r_1\vec{I} - B)^{-1} = (r_2\vec{I} - \vec{B})^{-1} + (r_2\vec{I} - \vec{B})^{-1}\,\vec{C}.$$

Since $(r_2\vec{I} - \vec{B})^{-1}\,\vec{C}$ is non-negative, every element of $(r_1\vec{I} - \vec{B})^{-1}$ must be greater than or equal to the corresponding element of $(r_2\vec{I} - \vec{B})^{-1}$. Hence

$$(r_2\vec{I} - \vec{B})^{-1} \leqslant (r_1\vec{I} - \vec{B})^{-1}. \tag{B. 13}$$

Now since $(r\vec{I} - \vec{B})^{-1}$ is continuous in r, and is nonincreasing for any r_2 and r_1 arbitrarily close together, it must be nonincreasing for all r.

Thus

$$(\vec{I} - h\vec{X})^{-1} \leqslant (h\vec{X})^{-1} \tag{B. 14}$$

and

$$(\vec{I} - h\vec{X})^{-1}\,h\vec{U} \leqslant (-h\vec{X})^{-1}\,h\vec{U}, \tag{B. 15}$$

since multiplication by a non-negative matrix does not affect the inequality.

Also since the matrices are non-negative, the norm must also satisfy the same inequality:

$$\begin{aligned} \|(I - hX)^{-1}\, hU\| &\leq \|(-hX)^{-1}\, hU\| \\ &\leq \|X^{-1}\| \cdot \|U\| \\ &\leq \max_k \left|\frac{1}{\nu_k}\right| \cdot \|U\| \\ &\leq \frac{\|U\|}{\min |v_g \sigma_{gg}/2|} = \beta \end{aligned} \tag{B. 16}$$

since all $|\nu_k| \geq |v_{gg}\sigma_{gg}/2|$ by Gerschgorin's Theorem (Ref. 9).

Furthermore, the function $(x^m - 1)/(x - 1)$ is a nondecreasing function of x, and hence

$$\frac{\|h\vec{V}\|^m - 1}{\|h\vec{V}\| - 1} \leq \frac{\beta^m - 1}{\beta - 1}. \tag{B. 17}$$

Also

$$\|(\vec{I}/h - \vec{X})^{-1}\| \leq \|-\vec{X}^{-1}\| \leq \frac{1}{\min |v_g \sigma_{gg}/2|}. \tag{B. 18}$$

Finally, we have the result that

$$\|\vec{Q}_1(h)\| \leq \frac{1}{\min |v_g \sigma_{gg}/2|} \left(\frac{\beta^m - 1}{\beta - 1}\right) = q_1, \tag{B. 19}$$

where the bound q_1 depends only on reactor properties, and not on $h v_g$ or Δx^2.

Corollary – $\vec{Q}_2(h)$ is bounded under the same conditions as $\vec{Q}_1(h)$.

The proof is identical with the appropriate substitutions.

Theorem 5 – $\vec{B}^j(h)$ is a consistent approximation to the solution of Eq. (1.1).

Proof – We first require the following lemma:

Lemma – If each of the matrices $\vec{C}_1(h)$ and $\vec{C}_2(h)$ are consistent, then the matrix $\vec{B} = \vec{C}_2\vec{C}_1$ is also consistent.

This is proven by Reed as his Theorem 1, page 30 of Ref. 7.

Now from Eq. (2.7) we have

$$\vec{B}^j(h) = (\vec{I} - h\vec{A}_2)^{-1}(\vec{I} + h\vec{A}_1)(\vec{I} - h\vec{A}_1)^{-1}(\vec{I} + h\vec{A}_2)$$

$$= \vec{C}_2(h)\,\vec{C}_1(h) \qquad \text{(B. 20)}$$

Thus $\vec{C}_1(h)$ is

$$\vec{C}_1(h) = (\vec{I} - h\vec{A}_1)^{-1}(\vec{I} + h\vec{A}_2)$$

$$= I + 2h(I - h\vec{A}_1)^{-1}\vec{A}. \qquad \text{(B. 21)}$$

Using one of the common definitions of consistency (Ref. 10),

$$\lim_{\Delta t \to 0} \left\| \left(\frac{\vec{C}_1(\Delta t) - \vec{I}}{\Delta t} - \vec{A} \right) \vec{x} \right\| = 0, \qquad \text{(B. 22)}$$

for $\vec{x}$ a genuine solution of (1. 1), we obtain

$$\lim_{\Delta t \to 0} \left\| \left((I - hA_1)^{-1} \vec{A} - \vec{A} \right) \vec{x} \right\| = \lim_{\Delta t \to 0} \left\| \left[(I - hA_1)^{-1} - I \right] \vec{A}\, \vec{x} \right\|$$

$$= \lim_{\Delta t \to 0} h \left\| \left((I - hA_1)^{-1} \vec{A}_1 \right) \vec{A}\, \vec{x} \right\|$$

$$\leq \lim_{\Delta t \to 0} h \left\| (\vec{I} - h\vec{A}_1)^{-1} \right\| \cdot \left\| \vec{A}_1 \vec{A}\, \vec{x} \right\|. \qquad \text{(B. 23)}$$

Now since all the eigenvalues of $\vec{A}_1$ are negative, the spectral radius of $(\vec{I} - h\vec{A}_1)^{-1}$ is bounded by 1 as h becomes small. Consequently the norm must be bounded by some constant, say β.

However, the term $\left\| \vec{A}_1 \vec{A}\, \vec{x} \right\|$ presents some problems since as h goes to zero, so must Δx^2 and Δy^2, and the norms of $\vec{A}_1$ and $\vec{A}$ become unbounded as Δx^2 and Δy^2 go to zero. However, the norm of $\vec{A}\, \vec{x}$ must remain bounded as Δx^2 and Δy^2 go to zero since $\vec{x}$ is a genuine solution of (1. 1). However, $\vec{y} = \vec{A}\, \vec{x} = \dot{\vec{x}}$ must also be a genuine solution of (1. 1) if $\vec{x}$ is, since $\dot{\vec{y}} = \frac{d}{dt}(\vec{A}\vec{x}) = \vec{A}\frac{d\vec{y}}{dt} = \vec{A}\, \vec{y}$ which is identical to (1. 1). But if $\vec{y}$ is a solution $\left\| \vec{A}\, \vec{y} \right\|$ must be bounded. But since $\vec{A}$ is the sum of $\vec{A}_1$ and $\vec{A}_2$, $\vec{A}_1 \vec{y}$ must also be bounded. So $\left\| \vec{A}_1 \vec{A}\, \vec{x} \right\|$ is bounded for $\vec{x}$ a genuine solution of (1. 1). Thus

$$\lim_{h \to 0} \left\| \frac{C_1(h) - I}{h} - \vec{A} \right\| \leq h\beta \left\| \vec{A}_1 A\, x \right\|$$

$$= 0. \qquad \text{(B. 24)}$$

A similar proof holds for $\vec{C}_2(h)$, and the result of the theorem follows.

Appendix C

DATA FOR TEST PROBLEMS

C. 1 CASE 1 – Two Group Homogeneous Bare Reactor

Perturbation is a uniform step change in the thermal cross section.

Number of neutron energy groups = 2
Number of delayed precursor groups = 1

Geometry: Homogeneous square 200 cm on a side.

$\Delta x = 20$ cm
$\Delta y = 20$ cm

Precursor constants:

$\lambda_1 = .08, \quad \beta_1 = .0064, \quad f_{11} = 1.0, \quad f_{21} = 0.0$

	Group 1	Group 2
v	$.3 \times 10^8$	$.22 \times 10^6$
χ	1.0	0.0

Material properties:

	Group 1	Group 2
D	1.35	1.08
σ_c	.00114	.00137 (.0014069, critical)
ν	2.41	2.41
σ_f	.000242	.00408
$\sigma_{i \to i+1}$	.0023	.00

Initial conditions:

Spacial shape: Cosine
Spectrum: 1.0
.382345
.000347419

C. 2. FOURGP – Four Group Bare Homogeneous Reactor

Perturbation is induced by changing critical value of ν by +.01172.

Number of neutron energy groups = 4
Number of delayed precursor groups = 1

Geometry: Homogeneous square 150 cm on a side.

Δx = 15.0 cm
Δy = 15.0 cm

Precursor constants: $\lambda_1 = .08$, $\beta_1 = .0074$, $f_{11} = .0$, $f_{21} = 1.0$, $f_{31} = .0$, $f_{41} = .0$

	Group 1	Group 2	Group 3	Group 4
v	$.25 \times 10^{10}$	$.5 \times 10^{9}$	$.43 \times 10^{7}$	$.25 \times 10^{6}$
χ	.575	.425	.0	.0

Material properties:

	Group 1	Group 2	Group 3	Group 4
D	2.0291	1.1609	.76965	.35676
σ_c	.00237	.00438	.03266	.1339
ν	3.16578	3.16578	3.16578	3.16578
σ_f	.01316	.00111	.0182	.38769
$\sigma_{i \to i+1}$	.06532	.00481	.00232	.00

Initial condition:

Spacial shape:	Cosine
Spectrum:	1.0000000
	11.2690000
	1.0066000
	0.0044746
	0.0133890

C. 3. TWIGL Reactor – Two Group, Non-Homogeneous System

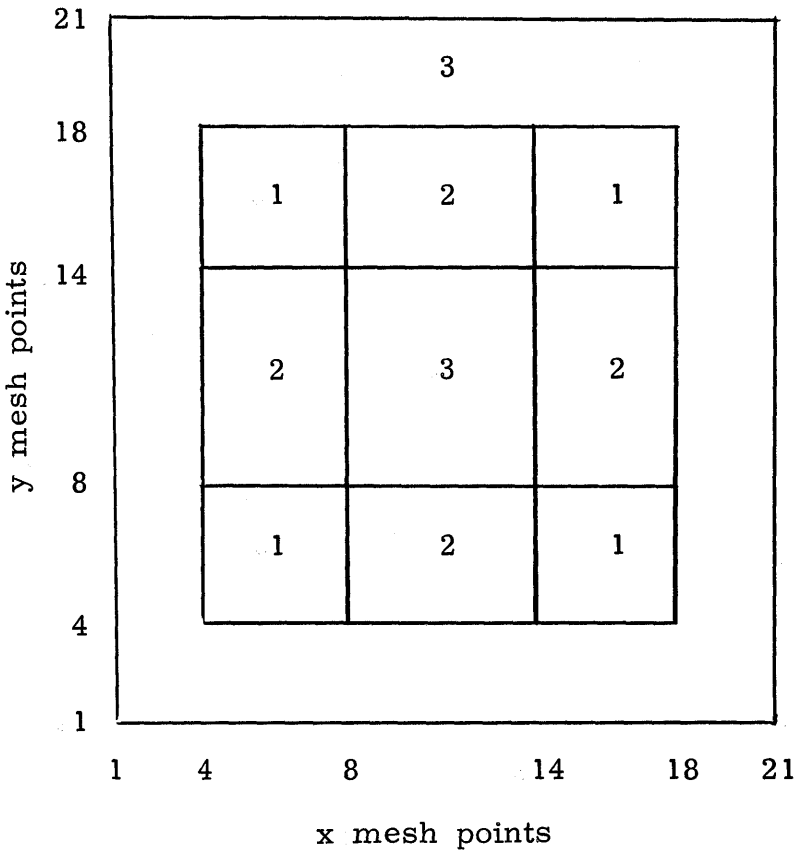
Critical Configuration

Number of neutron groups = 2
Number of precursor groups = 1

Geometry:

$\Delta x = 8.0$ cm

$\Delta y = 8.0$ cm



where the numbers indicate the material composition of that space region.

Delayed constants:

$\lambda_1 = .08, \quad \beta_1 = .0075, \quad f_{11} = 1.0, \quad f_{21} = 0.0$

	Group 1	Group 2
v	$.1 \times 10^8$	$.2 \times 10^6$
χ	1.0	0.0

Material properties:

Material 1

	Group 1	Group 2
D	1.4	0.4
σ_c	.0065	.05
ν	2.1877	2.1877
σ_f	.0035	0.1
$\sigma_{i \to i+1}$	0.01	0.0

Material 2

(same as material 1)

Material 3

	Group 1	Group 2
D	1.3	0.5
σ_c	.0065	0.02
ν	2.1877	2.1877
σ_f	.0015	.03
$\sigma_{i \to i+1}$	.01	0.0

Perturbations which induce transients:

Material 1

		Group 1	
Positive Step:	$\Delta\sigma_c$	-.0035	$t \geq .0$
Positive Ramp:	$\Delta\sigma_c$	-.0035 (t/.2),	$.0 \leq t \leq .2$
		-.0035 ,	$t \geq .2$
Negative Ramp:	$\Delta\sigma_c$	+.03 (t/.02) ,	$.0 \leq t \leq .02$
		+.03 ,	$t \geq .02$

C.4. OBLONG Reactor – Four Group Non-Homogeneous, Non-Symmetric System

Critical Configuration

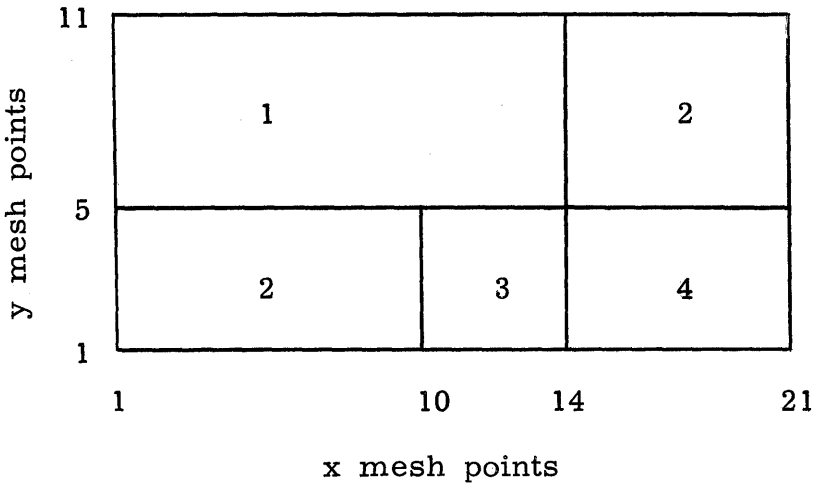
Number of neutron groups = 4

Number of precursor groups = 1

Geometry:

$\Delta x = 8.0$ cm

$\Delta y = 8.0$ cm



where the numbers indicate the material composition of that space region.

Delayed constants:

$\lambda_1 = .08$, $\beta_1 = .0064$, $f_{11} = 0.0$, $f_{21} = 1.0$, $f_{31} = 0.0$, $f_{41} = 0.0$

	Group 1	Group 2	Group 3	Group 4
v	$.1 \times 10^{10}$	$.1 \times 10^{9}$	$.5 \times 10^{7}$	$.2 \times 10^{6}$
χ	.0755	0.245	0.0	0.0

Material properties:

Material 1

	Group 1	Group 2	Group 3	Group 4
D	2.7778	1.0753	.64103	.16260
σ_c	.0013	.001	.0097	.115
ν	1.4507	1.4507	1.4507	1.4507
σ_f	.00136	.00197	.0262	.54
$\sigma_{i \to i+1}$	.0586	.00197	.085	0.0

Material 2

	Group 1	Group 2	Group 3	Group 4
D	3.3333	1.3889	.83333	2.0833
σ_c	.00065	.0005	.0045	.058
ν	1.4507	1.4507	1.4507	1.4507
σ_f	.0007	.0009	.0131	.274
$\sigma_{i \to i+1}$	.0586	.0828	.0850	0.0

Material 3

	Group 1	Group 2	Group 3	Group 4
D	4.1667	2.0833	1.0753	.26247
σ_c	.00077	.00072	.00051	.012
ν	0.0	0.0	0.0	0.0
σ_f	0.0	0.0	0.0	0.0
$\sigma_{i \to i+1}$	.0570	.0822	.0847	0.0

Material 4

(same as material 3)

Perturbation which induces transient:

Material 3

Group 4

Positive Ramp: $\Delta\sigma_c$ $\quad -.003\,(t/.2), \quad .0 < t \leq .2$

$\quad -.003 \quad , \quad t \geq .2$

Appendix D

COMPUTER PROGRAMS

(Only in first five copies)