ANL-6984

327
4/1/65

ANL-6984

Argonne National Laboratory

AN ITERATIVE UNFOLDING METHOD FOR RESPONSE MATRICES

by

Raymond Gold

DISCLAIMER

Printed in USA. Price $2.00. Available from the Clearinghouse for Federal Scientific and Technical Information, National Bureau of Standards, U. S. Department of Commerce, Springfield, Virginia

ANL-6984
Mathematics and Computers
(TID-4500, 37th Ed.)
AEC Research and
Development Report

ARGONNE NATIONAL LABORATORY
9700 South Cass Avenue
Argonne, Illinois 60440

AN ITERATIVE UNFOLDING METHOD FOR RESPONSE MATRICES

by

Raymond Gold

Reactor Physics Division

December 1964

Operated by The University of Chicago
under
Contract W-31-109-eng-38
with the
U. S. Atomic Energy Commission

TABLE OF CONTENTS

LIST OF FIGURES

TABLE

AN ITERATIVE UNFOLDING METHOD FOR RESPONSE MATRICES

by

Raymond Gold

ABSTRACT

The problem of unfolding the output data of a detection system is reviewed. Physical implications of the matrix representation of a detection system are discussed. An iterative method of solution is presented. The properties of the iterative method are investigated, and proof of convergence to the exact solution is given for two-dimensional, triangular, and positive definite type response matrices. The unfolding characteristics of the method are examined by applying the iterative process to the results of a neutron detection system. Limitations of the iterative method are discussed.

I. INTRODUCTION

Since ideal detection systems do not exist in practice, the measurement of a continuous spectrum is often complicated by the problem of finding the actual spectrum from the observed experimental data. The distortion of the actual spectrum usually arises from the inherent limitations of the particular detection system that is used. Obvious examples of such limitations are finite resolving power, multiple interaction modes, and nonlinear detection efficiency. The process of correcting experimental distributions for detector distortion has become known as unscrambling or unfolding.

The present considerations will be restricted to detection systems that measure the intensity distribution of one physical observable. Let $X(\mu)$ represent the actual intensity distribution or spectrum of the observable μ, and $B(\nu)$ represent the distribution arising as a result of the measurement. Then, the distributions $X(\mu)$ and $B(\nu)$ are related by an equation of the form

$$B(\nu) = \int_{\mu_1}^{\mu_2} A(\nu,\mu)\, X(\mu)\, d\mu, \tag{1}$$

where the function $A(\nu,\mu)$ is customarily called the response function of the

detection system. Hence, the unfolding problem is that of solving an integral equation of the above type. It is instructive to briefly review the status of such an equation.

If the interval (μ_1,μ_2) is finite and the prescribed kernel $A(\nu,\mu)$ and prescribed function $B(\nu)$ are well-behaved, then Eq. (1) is called a nonsingular Fredholm equation of the first kind. It is known that no well-behaved solution need exist for this equation when arbitrary well-behaved functions $A(\nu,\mu)$ and $B(\nu)$ are prescribed.[1] Moreover, as pointed out by Eckart,[2] the existence of a solution of Eq. (1) does not imply uniqueness. Namely, if $X_1(\mu)$ is such a solution, then $X_2(\mu) = X_1(\mu) + f(\mu)$ is also a solution, where $f(\mu)$ satisfies the equation

$$0 = \int_{\mu_1}^{\mu_2} A(\nu,\mu)\, f(\mu)\, d\mu. \tag{2}$$

It is evident that subsidiary conditions are required before uniqueness can be guaranteed. Such conditions would play a role analogous to that of boundary conditions in the theory of partial differential equations.

Since our application of Eq. (1) represents the description of a physical process, one might conjecture that the existence of a unique solution is assured by the physical requirements of the problem. However, a closer examination reveals that this is not the case. Indeed, the assumption that Eq. (1) describes a physical process tacitly implies that the function $B(\nu)$ is determined from experimental data and, as such, $B(\nu)$ is not known exactly. As a consequence, one must then consider not Eq. (1), but the modified equation,

$$B(\nu) + E(\nu) = \int_{\mu_1}^{\mu_2} A(\nu,\mu)\, X(\mu)\, d\mu, \tag{3}$$

where $B(\nu)$ now represents the results of some experimental measurement, and $E(\nu)$ is a measure of the error that is inherent in the observation. It has been shown that solutions of Eq. (3) are not only nonunique, but are also subject to severe oscillations.[3] This instability is an inherent characteristic of numerical solutions of the Fredholm integral equation, rather than the particular method of solution that is utilized.[4] It has been recognized[2,3] that only by specifying a criterion of smoothness can a unique (smooth) solution of Eq. (3) be determined. Such a criterion is, of course, equivalent to the proper prescription of boundary data in the aforementioned analogy with partial differential equations.

Let us now turn our attention to the numerical solution of Eq. (3), which usually proceeds by replacing the integral in Eq. (3) by a quadrature formula. In this manner, Eq. (3) is reduced to a system of simultaneous linear equations. In this approximation, Eq. (3) assumes the matrix form,

$$\underline{B} + \underline{E} = A\underline{X}, \tag{4}$$

where $\underline{B}$, $\underline{X}$, and $\underline{E}$ are n-element column vectors and A is an (n x n) matrix. Since the resulting output of the detection system is actually a single vector, $\underline{Y} = \underline{B} + \underline{E}$, Eq. (4) takes the form,

$$\underline{Y} = A\underline{X}. \tag{5}$$

It is customary to call A the response matrix (of the detection system). The vector $\underline{Y}$ will be referred to as the output vector (of the detection system). Consequently, the vector $\underline{Y}$ is not precisely defined, but possesses the inherent error associated with the experimental measurement. In terms of this matrix representation, the input to the detector also takes the form of a vector, namely the input vector (of the detection system). The physical implications of the description imply that for each output vector there must exist an input vector that satisfies Eq. (5).

The order n of this matrix representation is to a certain extent arbitrary. Indeed, the larger one chooses n, the more closely the matrix representation approximates the original integral equation. However, the fact that one is restricted from choosing arbitrarily large n can be seen from the solution of Eq. (4). One has

$$\underline{X} = A^{-1}(\underline{B} + \underline{E}). \tag{6}$$

There is no doubt that as n increases, the contribution of the truncation error in the elements of $\underline{E}$ will decrease. However, there does exist a (nonvanishing) lower bound for the norm of the elements of this error vector. Such a lower bound is usually set by the error that arises from the physical limitations of the measurement and is thereby independent of n. On the other hand, if n is chosen too large, then A may become ill-conditioned, and the matrix A^{-1} will then possess coefficients that become rapidly larger with increasing n.[1] Consequently, there is no reason to believe that the contribution to the exact solution vector $\underline{X}$ arising from the error vector, that is, $A^{-1}\underline{E}$, need become small for arbitrarily large n. Indeed, it is usually found that as n increases, the solutions first become more accurate, but then eventually become worse.

In practice, the minimum order of a response matrix is usually dictated by the range and resolution of the detection system. That is, one naturally chooses n large enough to cover this entire range and, at the same time, retain the resolving power afforded by the instrument.* For detection systems of interest, such a choice implies that the order n will be large. Typically, one finds orders that are larger than n = 20 and quite commonly $n \geq 50$. That such response matrices are ill-conditioned has been

*Otherwise there is little, if anything, to be gained by following any type of unfolding procedure.

demonstrated.[5] Consequently, the exact solution of Eq. (4) or (5), as given in Eq. (6) in terms of A^{-1}, will possess the inherent oscillations that (as has already been mentioned) persist in numerical solutions of the Fredholm equation.

In such circumstances, it is clear that the exact solution can be completely unacceptable. It follows that one does not, in general, desire the exact solution as given by Eq. (6), but more properly seeks a vector that satisfies Eq. (5), together with certain subsidiary conditions imposed by the physical implications of the description. Such a solution will be called an appropriate solution. In other words, appropriate solutions are more than merely members of the set of vectors that satisfy Eq. (5). Appropriate solutions must, in addition, satisfy all relevant physical requirements.

The main features of this problem can be described more precisely in terms of the (infinite) set of vectors that satisfy Eq. (5). This set can be defined as

$$\chi = \{\underline{S} \mid (\underline{Y} - A\underline{S}) \lesssim 0(\underline{E})\}.$$

Thus, a vector $\underline{S}$ is said to satisfy Eq. (5), provided $A\underline{S}$ differs from $\underline{Y}$ by no more than the order of the experimental error. Application of subsidiary conditions will define some subset $\chi' \subset \chi$. Provided the subsidiary conditions have (physical) validity and have been properly applied, one can expect that the input vector will be an element of the subset χ'. On the other hand, one must also anticipate that the subset χ' will generally be infinite. In addition, if some pertinent physical implications have not been utilized or could not be successfully incorporated into the method of solution, then the subset χ' may still possess elements that could not be classified as appropriate solutions. Consequently, there could still exist some uncertainty concerning the nature of the solutions furnished by such a method. In this event, a proper examination of the ability of a given method to provide appropriate solutions could only be ascertained through testing, utilization, and experience with the given method.

Many different unfolding procedures have been proposed and used in an effort to treat this problem. These techniques generally fall into three categories: iterative, least-squares, and smoothing methods. In the iterative approach, one attempts to generate successive approximations which converge to an appropriate solution. Different types of iteration processes have been utilized with some success.[6-9] The least-squares method minimizes the norm of the residual that is constructed from the output vector and an assumed parametric form of the input vector.[10,11] A smoothing method introduces constraints upon the input vector, output vector, or both, in an effort to suppress unwanted oscillatory behavior. For example, the smoothing technique of Phillips[12] has demonstrated some success in dampening out spurious oscillations.

In particular, this report is concerned with a specific iteration method,[7] which has been successfully employed in practical applications.[13-15] The properties of this iteration process will be examined in detail. The main goal of this investigation is to develop a more rigorous foundation for this method than presently exists. Such a foundation will provide the necessary justification for the application of this method.

The general advantages of iterative methods have been discussed by Bodewig.[16] Forsythe[17] has given a general classification for different types of iterative methods. Perhaps the most significant feature of iterative methods, for the unfolding application, lies in the ability to incorporate readily most of the major physical implications of the description. Furthermore, the necessity for an explicit calculation of the inverse matrix, A^{-1}, is avoided. This advantage can be extremely important since most response matrices are ill-conditioned, and thereby the error in the elements of A^{-1} can be prohibitive.

This additional complication does not appear to have been clearly recognized in this application. The error that arises in the elements of A^{-1} is due to two essentially independent sources. First, one has round-off error incurred in the inversion process. Second, one must realize that the elements (a_{ij}) of A also possess inherent error. We shall not dwell upon the subject of the construction of a response matrix, but refer the reader to examples in the literature.[18-22] The error in the matrix elements (a_{ij}) can either be the result of experimental measurement or may arise from certain analytical approximations of the interaction (or detection) process. As a result of these two sources alone, it is apparent that the induced error, which is propagated to the elements of A^{-1}, will increase very rapidly with increasing n. In such circumstances, any reduction in the number of operations required to find an appropriate solution can significantly reduce the propagation of error. The possibility of such a reduction is available when one uses an iterative technique. Moreover, this very feature is often pointed out as one of the general advantages of iterative methods.

Finally, the ability to stop and examine the process, after a given number of iterations, is highly desirable. This affords not only a study of the convergence of the method, but also permits an examination of any member of the sequence of approximations. Consequently, if the iterative method generates successive vectors that possess the characteristics ascribed to an appropriate solution, then one may confidently expect (provided convergence is assured) that the method does furnish appropriate solutions.

The next section reviews the physical implications that arise from the matrix representation of detection systems. The iterative process is then described, and the convergence properties of this method are examined in detail. In Section V, the limitations of this unfolding technique are investigated by applying this method to a Li^6 solid-state detector which is utilized for the measurement of neutron energy spectra.

II. PHYSICAL IMPLICATIONS

For exposition, the matrices that enter into this representation of detection systems, as given in Eq. (5), are displayed below. The vectors $\underline{X}$ and $\underline{Y}$ take the form of column matrices

$$\underline{X} = \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix}; \quad \underline{Y} = \begin{pmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{pmatrix}. \tag{5'}$$

The square (n x n) response matrix A takes the usual form

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdot & \cdot & \cdot & a_{1n} \\ a_{21} & a_{22} & \cdot & \cdot & \cdot & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \cdot & \cdot & \cdot & a_{nn} \end{pmatrix} \tag{5''}$$

Due to physical requirements, Eq. (5) need not be considered in complete generality. For example, one can easily recognize that all elements that enter into Eq. (5) must be restricted to the field of real numbers. In fact, all these elements must be nonnegative real numbers. More specifically, we shall only treat systems that satisfy the following additional conditions:[7,15]

The input vector is nonnegative.	(i)
The output vector $\underline{Y}$ is positive.	(ii)
A is a nonnegative matrix.	(iiia)
$a_{ii} > 0, \quad i = 1,2,\ldots,n.$	(iiib)
A is nonsingular.	(iiic)

The above conditions have by no means been arbitrarily introduced, but follow quite naturally for many detection systems. We outline below the physical justification for these assumptions.

Condition (i) implies that the input vector has positive, or at most vanishing, components. This follows from the fact that the input vector is the discrete approximation of the intensity distribution of some physical observable. Condition (ii), which implies that the vector $\underline{Y}$ must possess

positive components, follows from the finite resolving power of all detection systems. Moreover, Condition (ii), together with Condition (iiic), excludes from consideration the trivial case of a vanishing input vector, which corresponds to the physically uninteresting case in which no observable is present.

Conditions (iiia) and (iiib) are implied by the fact that any interaction between the observable and the detection system will yield a positive or, at most, vanishing result. This follows from the nonnegative character of the interaction cross sections used in the detection system. One might say that these conditions require a response or no response, but never a negative response.

Although we have already implied that A may be ill-conditioned, we shall require Condition (iiic), that A be nonsingular. That this assumption is reasonable follows from the flexible (although not completely arbitrary) order of the matrix representation that has already been considered. In addition, a given matrix representation corresponds to a certain arbitrary subdivision of the ν and μ space of Eq. (1) or (3). Consequently, if a particular mesh of the (μ,ν) space of a given order leads to a singular A matrix, then a different mesh or a different order (or both) may remove this difficulty. Hence it is reasonable to assume that there exists a suitable mesh of suitable order for which the response matrix is nonsingular. We then take Conditions (iiia), (iiib), and (iiic) to define the most general type of response matrix.

We remark that caution need be exercised for the correct interpretation of the above conditions. For example, we note that Condition (ii) does not imply Condition (i). In fact, there may exist infinitely many vectors $\underline{Y}$ satisfying Condition (ii), whose corresponding exact solutions, $\underline{X} = A^{-1}\underline{Y}$, do not satisfy Condition (i). If, however, one is given a vector $\underline{Y}$ that is the result of an experimental measurement, i.e., an output vector, then Condition (i) follows immediately from the physical requirement of the description. In this case, the exact solution $\underline{X} = A^{-1}\underline{Y}$ and the input vector or <u>appropriate</u> solution need not coincide. However, the very fact that $\underline{Y}$ is an output vector implies that the set χ must contain an <u>appropriate</u> solution. The determination of this <u>appropriate</u> solution is just what is desired.

III. THE ITERATION METHOD

It has been stressed that the essential difficulties of the unfolding problem arise from the inexact description inherent in all experimental measurements. On the other hand, a basic understanding of the specific iteration process considered here can only evolve from the application of this method to an exact problem. This assumption is employed throughout the ensuing work on the general characteristics of this iteration method.

In this event, Eq. (5) is exact, and Conditions (i) through (iii) therefore imply the existence of a real, unique, nonnegative solution.* Moreover, in view of Conditions (ii), one can define a real diagonal matrix D, with the unique elements,

$$D_{ii} = x_i/y_i, \quad i = 1,2,\ldots,n. \tag{7}$$

One then has the matrix equation,

$$\underline{X} = D\underline{Y}. \tag{8}$$

It is evident that the matrix D cannot be identified as A^{-1}. In general, $D \neq A^{-1}$, and only for the trivial case wherein A is diagonal (hence A^{-1} is also diagonal) will $D = A^{-1}$.

Hence, a knowledge of the matrix D will also supply the solution of Eq. (5), satisfying Conditions (i) through (iii). The method of successive approximations described herein yields approximate values for the elements of D. One can thereby obtain an approximate solution without the explicit use of A^{-1}.

To initiate the iteration process, one must choose $\underline{X}^{(0)}$, the zero-order approximation of $\underline{X}$. It will become obvious, in sequel, that $\underline{X}^{(0)}$ can be chosen as an arbitrary positive vector. Consequently the zero-order approximations corresponding to Eqs. (7) and (8) are**

$$D_{ii}^{(0)} = x_i^{(0)}/y_i, \quad i = 1,2,\ldots,n; \tag{9a}$$

and

$$\underline{X}^{(0)} = D^{(0)}\underline{Y}. \tag{9b}$$

Using this initial vector $\underline{X}^{(0)}$ in Eq. (5) yields the output vector $\underline{Y}^{(0)}$. That is,

$$\underline{Y}^{(0)} = A\underline{X}^{(0)}. \tag{9c}$$

In terms of this result, one can define $D^{(1)}$, the first-order approximation of D. One has

$$D_{ii}^{(1)} = x_i^{(0)}/y_i^{(0)}, \quad i = 1,2,\ldots,n. \tag{9d}$$

---

*It also follows that the exact solution and the input vector coincide.

**It is often convenient to choose the output vector $\underline{Y}$ as the starting point for the iteration process; i.e., $\underline{X}^{(0)} = \underline{Y}$. In this event, the zero-order approximation of D is simply $D^{(0)} = I$, where I is the identity matrix.

Using Eq. (8) again, one can calculate $\underline{X}^{(1)}$, the first-order approximation of $\underline{X}$. Thus, one has

$$\underline{X}^{(1)} = D^{(1)}\underline{Y}. \tag{9e}$$

Having determined $\underline{X}^{(1)}$, the process may be repeated to obtain $\underline{X}^{(2)}$, the second-order approximation of $\underline{X}$. Namely,

$$\underline{Y}^{(1)} = A\underline{X}^{(1)}; \tag{10a}$$

$$D_{ii}^{(2)} = x_i^{(1)}/y_i^{(1)}, \quad i = 1,2,\ldots,n; \tag{10b}$$

and

$$\underline{X}^{(2)} = D^{(2)}\underline{Y}. \tag{10c}$$

In general, the $(m+1)$ approximation of $\underline{X}$ is found with the following analogous steps:

$$\underline{Y}^{(m)} = A\underline{X}^{(m)}; \tag{11a}$$

$$D_{ii}^{(m+1)} = x_i^{(m)}/y_i^{(m)}, \quad i = 1,2,\ldots,n; \tag{11b}$$

and

$$\underline{X}^{(m+1)} = D^{(m+1)}\underline{Y}. \tag{11c}$$

Equations (11a), (11b), and (11c) determine the recursion relation between successive approximations to the elements of D. One finds

$$D_{ii}^{(m+1)} = \frac{D_{ii}^{(m)} y_i}{\sum_{j=1}^{n} a_{ij} D_{jj}^{(m)} y_j}, \quad i,j = 1, 2, \ldots, n. \tag{12}$$

As an alternate, but fully equivalent, representation of this iteration process, one can also write the recursion relation in the form

$$x_i^{(m+1)} = x_i^{(m)} \cdot y_i/y_i^{(m)} = x_i^{(m)} y_i / \sum_{j=1}^{n} a_{ij} x_j^{(m)}, \quad i,j = 1,2,\ldots,n. \tag{13}$$

At this point, it is relevant to remark upon some aspects of the character of this iteration process. To begin with, one recognizes that all members of the sequence of approximations $\{\underline{X}^{(m)}\}$ are positive vectors.

Hence, convergence to the exact solution, if it occurs at all, must proceed through the positive domain. This implies that the limit (if it exists) must be a positive, or at most nonnegative, vector. From the viewpoint of our application, this behavior requires that Condition (i) be satisfied. Moreover, the simplicity of the recursion relation is evident. The set $\{\underline{Y}^{(m)}\}$ is a sequence of output vectors. Each step of the process can be viewed as a more accurate modification of the actual output vector $\underline{Y}$. Consequently, one may anticipate that each member of the sequence of vectors $\{\underline{X}^{(m)}\}$, generated by this iteration process, possesses smooth behavior. These properties suggest that approximations can be obtained from this method that possess properties similar to that required of an appropriate solution.

IV. CONVERGENCE TO THE EXACT SOLUTION

The domain of applicability of the proposed iteration method must be established. That is, for what class of matrices [satisfying Conditions (iiia), (iiib), and (iiic)] will the iteration method converge to the exact solution? Considered, in sequel, are the two-dimensional, triangular, and positive definite cases.

A. The Simple Two-dimensional System

Our starting point is chosen not only for its simplicity, but also for the insight obtained from the examination of such a system. In this manner, it will serve as a guide for the investigation of the more general n-dimensional case that will follow.

For this simple case, the system and the corresponding solutions are given by

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \tag{14a}$$

$$x_1 = (a_{22}y_1 - a_{12}y_2)/|A|, \tag{14b}$$

and

$$x_2 = (a_{11}y_2 - a_{21}y_1)/|A|, \tag{14c}$$

where $|A|$ represents the determinant of the coefficients of A.

In this event, recursion relations (12) reduce to

$$D_{11}^{(m+1)} = \frac{D_{11}^{(m)}y_1}{a_{11}D_{11}^{(m)}y_1 + a_{12}D_{22}^{(m)}y_2}, \tag{15a}$$

and

$$D_{22}^{(m+1)} = \frac{D_{22}^{(m)}y_2}{a_{22}D_{22}^{(m)}y_2 + a_{21}D_{11}^{(m)}y_1}. \tag{15b}$$

In view of Conditions (i) through (iii), it follows from Eqs. (15a) and (15b) that the elements $D_{ii}^{(m)}$, $i = 1, 2$, are bounded for all m. Moreover, for any m, one has*

*For the more general case of n dimensions, it is evident that Eq. (16) is also valid, $i = 1, 2, \ldots, n$.

$$0 < D_{ii}^{(m)} \leq a_{ii}^{-1}, \quad i = 1,2. \tag{16}$$

It follows from Eqs. (15a) and (15b) that

$$\begin{vmatrix} \left(D_{11}^{(m+1)}a_{11} - 1\right) & D_{11}^{(m+1)}a_{12} \\ D_{22}^{(m+1)}a_{21} & \left(D_{22}^{(m+1)}a_{22} - 1\right) \end{vmatrix} = 0. \tag{17}$$

This equation yields the following two equivalent conditions:*

$$D_{11}^{(m+1)} = \frac{D_{22}^{(m+1)}a_{22} - 1}{|A|D_{22}^{(m+1)} - a_{11}}; \quad a_{12} \neq 0, \quad a_{21} \neq 0; \tag{18a}$$

and

$$D_{22}^{(m+1)} = \frac{D_{11}^{(m+1)}a_{11} - 1}{|A|D_{11}^{(m+1)} - a_{22}}; \quad a_{12} \neq 0, \quad a_{21} \neq 0. \tag{18b}$$

Using Eq. (16), together with Eqs. (18a) and (18b), one has, for all m,

$$a_{11} \geq |A|D_{22}^{(m+1)}; \tag{19a}$$

and

$$a_{22} \geq |A|D_{11}^{(m+1)}. \tag{19b}$$

One can substitute Eqs. (18a) and (18b) into Eqs. (15a) and (15b), respectively, thereby obtaining two recursion relations, each involving a single element. Since the recursion relation for $D_{22}^{(m)}$ follows from the recursion relation for $D_{11}^{(m)}$ upon interchange of indices, one need only consider the latter relation. This is given by (for $m \geq 1$)

$$D_{11}^{(m+1)} = \frac{D_{11}^{(m)}y_1}{a_{11}D_{11}^{(m)}y_1 + a_{12}\left(\frac{1 - a_{11}D_{11}^{(m)}}{a_{22} - |A|D_{11}^{(m)}}\right)y_2}. \tag{20}$$

Equation (20) may be placed in a more convenient form by introducing Eq. (14b). A simple calculation yields

$$\frac{D_{11}^{(m+1)}}{D_{11}^{(m)}} = \frac{1}{1 + \left(\frac{1 - a_{11}D_{11}^{(m)}}{a_{22}/|A| - D_{11}^{(m)}}\right)\left(D_{11}^{(m)} - \frac{x_1}{y_1}\right)}. \tag{21}$$

*It will become apparent that the special case, wherein $a_{12} = 0$ or $a_{21} = 0$, may be treated directly.

We proceed by considering two separate cases, namely $|A| < 0$ and $|A| > 0$. For the former case, we set $|A| = -\rho$, and $\rho > 0$; and Eq. (21) becomes

$$\frac{D_{11}^{(m+1)}}{D_{11}^{(m)}} = \frac{1}{1 - \left(\frac{1 - a_{11}D_{11}^{(m)}}{a_{22}/\rho + D_{11}^{(m)}}\right)\left(D_{11}^{(m)} - \frac{x_1}{y_1}\right)} . \tag{22}$$

We note that for any given m, if $D_{11}^{(m)} > x_1/y_1$, then $D_{11}^{(m+1)} > D_{11}^{(m)} > x_1/y_1$. Conversely, if $D_{11}^{(m)} < x_1/y_1$, it follows from Eq. (22) that $D_{11}^{(m+1)} < D_{11}^{(m)} < x/y$. Hence, we find for $|A| < 0$, that although convergence is assured, the iterative method does not converge to the exact solution.

We shall now prove that convergence to the exact solution is assured, for this simple case, if $|A| > 0$. To this end, we write Eq. (21) in the form

$$D_{11}^{(m+1)} = \frac{D_{11}^{(m)}}{1 + R_{11}^{(m)}\left(D_{11}^{(m)} - x_1/y_1\right)}, \tag{23a}$$

where

$$R_{11}^{(m)} = |A| D_{22}^{(m)} . \tag{23b}$$

Subtracting x_1/y_1 from both sides of Eq. (23a) yields

$$\left(D_{11}^{(m+1)} - x_1/y_1\right) = \left(D_{11}^{(m)} - x_1/y_1\right) \frac{\left(1 - (x_1/y_1)R_{11}^{(m)}\right)}{\left(1 - (x_1/y_1)R_{11}^{(m)} + D_{11}^{(m)}R_{11}^{(m)}\right)} . \tag{24}$$

It is easy to verify that Eqs. (23a) and (23b), and consequently Eq. (24), also hold for the special case $a_{12} = 0$ or $a_{21} = 0$.

Since $|A| > 0$, Eqs. (16) and (23b) imply that

$$0 \leq (x_1/y_1)R_{11}^{(m)} = (x_1/y_1)\,|A|\,D_{22}^{(m)} \leq (x_1/y_1)\,|A|\,a_{22}^{-1} . \tag{25a}$$

Utilizing Eqs. (14b) and (25a), one finds, for all m, that

$$0 \leq (x_1/y_1)R_{11}^{(m)} \leq 1. \tag{25b}$$

Using this result in Eq. (24), one has

$$\left|D_{11}^{(m+1)} - x_1/y_1\right| < \left|D_{11}^{(m)} - x_1/y_1\right| \tag{26}$$

for all m. It follows that the monotone sequence $\{|D_{11}^{(m)} - x_1/y_1|\}$ must converge to zero or some positive lower bound. We shall prove that no such positive lower bound exists. To this end, we shall assume the converse and complete the proof by contradiction.

Thus we set

$$\text{Lim } |D_{11}^{(m)} - x_1/y_1| = \alpha \tag{27}$$

where $\alpha > 0$. In this event, Eq. (24) yields

$$\text{Lim } \left|\frac{1 - (x_1/y_1)R_{11}^{(m)}}{1 - (x_1/y_1)R_{11}^{(m)} + D_{11}^{(m)}R_{11}^{(m)}}\right| = 1. \tag{28}$$

Equation (28) is satisfied if $R_{11}^{(m)} \to 0$ or $D_{11}^{(m)} \to 0$. However, $R_{11}^{(m)} \to 0$ implies that $D_{22}^{(m)} \to 0$. Hence, Eq. (28) is satisfied if either $D_{11}^{(m)} \to 0$ or $D_{22}^{(m)} \to 0$. Let us assume that $D_{11}^{(m)} \to 0$; hence, $D_{22}^{(m)} \to a_{22}^{-1}$. This assumption, together with Eq. (27), implies $\alpha = x_1/y_1$ and thereby

$$\text{Lim } \left(D_{11}^{(m)} - x_1/y_1\right) = -\alpha. \tag{29}$$

Using these results in Eq. (23a), one finds for sufficiently large m

$$D_{11}^{(m+1)} = D_{11}^{(m)}/[1 - \alpha|A|a_{22}^{-1} + \epsilon^{(m)}], \tag{30}$$

where the sequence $\{\epsilon^{(m)}\} \to 0$. Hence for sufficiently large m, $D_{11}^{(m+1)} > D_{11}^{(m)}$. However, this contradicts the assumption that $D_{11}^{(m)} \to 0$.

There remains as yet a consideration of the case $D_{22}^{(m)} \to 0$, hence $D_{11}^{(m)} \to a_{11}^{-1}$. However, it is obvious from the above analysis that $D_{22}^{(m)} \to 0$ implies $x_2 = 0$ and therefore $D_{11}^{(m)} \to a_{11}^{-1}$ provides convergence to the exact solution for this case as well. Hence, for the simple two-dimensional system, convergence to the exact solution is assured if, and only if, $|A| > 0$.

B. The Triangular Case

Let us partition the general n-dimensional system of Eq. (5) in the following manner:

$$\begin{pmatrix} Y' \\ \hline Y_r \end{pmatrix} = \left(\begin{array}{c|c} A' & C_r \\ \hline R_r & A_r \end{array}\right)\begin{pmatrix} X' \\ \hline X_r \end{pmatrix}. \tag{31}$$

One thereby obtains two matrix equations,

$$\underline{Y}' = A'\underline{X}' + C_r\underline{X}_r, \tag{32a}$$

and

$$\underline{Y}_r = R_r\underline{X}' + A_r\underline{X}_r, \tag{32b}$$

which represent (n - 2) dimensional and two-dimensional systems, respectively. That is, the submatrices A' and A_r are square $[(n-2) \times (n-2)]$ and (2×2) matrices, respectively.

On the other hand, we shall need the matrix form of the recursion relation (12), which is given by

$$D^{(m+1)}AD^{(m)}\underline{Y} = D^{(m)}\underline{Y}. \tag{33}$$

This matrix equation may also be partitioned in analogous fashion. One has

$$\left(\begin{array}{c|c} D^{(m+1)'} & \\ \hline & D_r^{(m+1)} \end{array}\right)\left(\begin{array}{c|c} A' & C_r \\ \hline R_r & A_r \end{array}\right)\left(\begin{array}{c|c} D^{(m)'} & \\ \hline & D_r^{(m)} \end{array}\right)\left(\begin{array}{c} Y' \\ \hline Y_r \end{array}\right) = \left(\begin{array}{c|c} D^{(m)'} & \\ \hline & D_r^{(m)} \end{array}\right)\left(\begin{array}{c} Y' \\ \hline Y_r \end{array}\right). \tag{34}$$

The (n - 2) dimensional and two-dimensional systems which result from this partitioning are, respectively,

$$D^{(m+1)'}A'D^{(m)'}\underline{Y}' + D^{(m+1)'}C_rD_r^{(m)}\underline{Y}_r = D^{(m)'}\underline{Y}', \tag{35a}$$

and

$$D_r^{(m+1)}A_rD_r^{(m)}\underline{Y}_r + D_r^{(m+1)}R_rD^{(m)'}\underline{Y}' = D_r^{(m)}\underline{Y}_r. \tag{35b}$$

Let us consider the case where A is triangular. Herein all the nonvanishing elements of A occur either above or below the main diagonal. If these nonvanishing elements occur above the main diagonal, A is called an upper triangular matrix, and if the nonvanishing elements occur below the main diagonal, then A is called a lower triangular matrix. Response matrices of this type often arise in practice.

The proof will be confined to the case wherein A is an upper triangular matrix. It will be evident in what follows that our proof is equally valid for lower triangular matrices.

If A is an upper triangular matrix, then all the elements of the matrix R_r vanish. Moreover, Eqs. (32b) and (35b) become, respectively,

$$\underline{Y}_r = A\underline{X}_r, \tag{36a}$$

and

$$D_r^{(m+1)} A_r D_r^{(m)} \underline{Y}_r = D_r^{(m)} \underline{Y}_r. \tag{36b}$$

However, Eq. (36a) is a two-dimensional system, and Eq. (36b) is just the recursion relation that corresponds to this sytem. Since A_r is also an upper triangular matrix, $|A_r| \neq 0$, and convergence to the exact solution is then assured for this system. We can continue this process by partitioning the (n - 2) dimensional system of Eq. (32a) as well as the associated recursion relation given in Eq. (35a). A knowledge of the convergence to the exact solution for Eq. (36) can then be utilized in this step of the partitioning process. We complete the proof by induction.

Hence, assume that

$$D_{kk}^{(m)} y_k = x_k + \epsilon_k^{(m)}, \quad k = i + 1, \ldots, n; \tag{37a}$$

where

$$\epsilon_k^{(m)} \to 0, \quad k = i + 1, \ldots, n \tag{37b}$$

It follows that the recursion relation for $D_{ii}^{(m)}$ can be written in the form

$$D_{ii}^{(m+1)} = \frac{D_{ii}^{(m)}}{1 + a_{ii}\left(D_{ii}^{(m)} - x_i/y_i\right) + \eta_i^{(m)}}, \tag{38a}$$

with

$$\eta_i^{(m)} = y_i^{-1} \cdot \sum_{k=i+1}^{n} a_{ik} \epsilon_k^{(m)}. \tag{38b}$$

Equation (38a) can be placed in a form analogous to Eq. (23a). One has

$$D_{ii}^{(m+1)} = \frac{D_{ii}^{(m)}}{1 + a_{ii}\left[D_{ii}^{(m)} - \left(x_i/y_i - \gamma_i^{(m)}\right)\right]}, \tag{39a}$$

where

$$\gamma_i^{(m)} = \eta_i^{(m)} \cdot a_{ii}^{-1}. \tag{39b}$$

It follows from our considerations of the simple two-dimensional system that the sequence $\{|D_{ii}^{(m)} - (x_i/y_i - \gamma_i^{(m)})|\}$ converges to zero. Consequently, given an ϵ arbitrarily small, there exists an index m_0 such that

$$\left|D_{ii}^{(m+1)} - (x_i/y_i - \gamma_i^{(m)})\right| \leq \epsilon \tag{40}$$

for all $m > m_0$. Thus, one has

$$\left|\left|D_{ii}^{(m+1)} - x_i/y_i\right| - \left|\gamma_i^{(m)}\right|\right| < \epsilon \tag{41}$$

for all $m > m_0$. However, Eq. (41) implies that either

$$\left|D_{ii}^{(m+1)} - x_i/y_i\right| < \epsilon + \left|\gamma_i^{(m)}\right|, \tag{42a}$$

or

$$\left|D_{ii}^{(m+1)} - x_i/y_i\right| < \left|\gamma_i^{(m)}\right|, \tag{42b}$$

for all $m > m_0$. It follows from the definition of the sequence $\{\gamma_i^{(m)}\}$ that the sequence $\{D_{ii}^{(m)}\}$ converges to x_i/y_i. Consequently, the proof of convergence to the exact solution for triangular systems is complete.

C. The Positive Definite Case

Let us examine the possible implications of our treatment to this point. As a consequence of our two-dimensional considerations, one can show that

$$\begin{vmatrix} a_{kk} & a_{kj} \\ a_{jk} & a_{jj} \end{vmatrix} > 0, \quad \begin{matrix} k < j, \\ k = 1, 2, \ldots, (n-1), \\ j = 2, 3, \ldots, n, \end{matrix} \tag{43}$$

is a necessary condition for convergence to the exact solution. In fact, consider a vector $\underline{X}$ which possesses only two nonvanishing elements, x_k and x_j. Assume further that all $D_{ii}^{(m)}$, $i \neq j, k$, converge to zero. It follows that the recursion relations for the elements $D_{kk}^{(m)}$ and $D_{jj}^{(m)}$ of the n-dimensional system approach those of a two-dimensional system. Indeed, under the present assumptions, one can have the recursion relations for $D_{kk}^{(m)}$ and $D_{jj}^{(m)}$ approximate the recursion relations of a two-dimensional system in an arbitrarily close fashion. However, it has already been proven that convergence to the exact solution for such a system will occur if, and only if, the determinant of the coefficients is positive. Condition (43), that is, the positiveness of all the (2 x 2) principal minors A, then follows from the fact that the choice of nonvanishing elements x_j and x_k of $\underline{X}$ is arbitrary.

Passing to a three-dimensional system, one can show that if $|A| < 0$ [even though Condition (43) is satisfied], convergence to the exact solution is no longer assured. This can be demonstrated by a simple numerical calculation.

These remarks indicate that a necessary condition for convergence to the exact solution in the general case might be that all the principal minors of A be positive. Moreover, one finds this condition is automatically satisfied by triangular response matrices. There exists, however, another important similarity in structure between two-dimensional and triangular response matrices; namely, the eigenvalues of such matrices must be real and positive.

Consequently, one is led to an examination of the following requirements:

(a) All principal minors of A are positive.

(b) All eigenvalues of A are real.

(c) All eigenvalues of A are positive.

Note that Conditions (a) and (b) imply Condition (c). However, Condition (c) obviously implies Condition (b), but does not imply Condition (a). Indeed, one can demonstrate by simple numerical calculations that Condition (c) is not a sufficient condition for convergence to the exact solution. The most obvious class of matrices that are an extension of triangular response matrices, in that Conditions (a) and (b) hold, are the positive definite matrices. Furthermore, it is well known that the system given in Eq. (5) can be transformed by multiplication on the left with $\tilde{A}$, the transpose matrix of A. This yields the system

$$\underline{V} = B\underline{X}, \tag{44a}$$

where

$$B = \tilde{A}A, \tag{44b}$$

and

$$\underline{V} = \tilde{A}\underline{Y}. \tag{44c}$$

Moreover, if A is a response matrix, then B must be a positive definite response matrix. In addition, Eq. (44c) implies that if $\underline{Y}$ is a positive vector, then $\underline{V}$ must also be a positive vector. Hence the above system satisfies Conditions (i) through (iii) and possesses a positive definite response matrix. It follows that one may assume that the matrix A, in Eq. (5), is positive definite without any loss in generality.

To verify that these assumptions provide sufficient conditions for convergence to the exact solution, one can introduce the positive definite quadratic form,

$$F(\underline{Z}) = (\tilde{\underline{Z}} - \tilde{\underline{X}})A(\underline{Z} - \underline{X}). \tag{45}$$

This quadratic form is actually the norm of the vector $(\underline{Z} - \underline{X})$ with respect to A as a metric. If $\underline{Z}$ is some approximation of $\underline{X}$, then $F(\underline{Z})$ is also called the error function (corresponding to the vector $\underline{Z}$). It has been shown that solving Eq. (5) is equivalent to minimizing the error function.[23] Consequently, one need only demonstrate that the iteration method generates a sequence of vectors $\{\underline{X}^{(m)}\}$, which minimize the error function on the nonnegative subspace. To establish this property, it will be necessary to prove several theorems and lemmas that are introduced below.

Theorem I

For positive definite response matrices, the eigenvalues of the matrix $\left(D^{(m)}A\right)$ satisfy*

$$0 < \lambda_i\left(D^{(m)}A\right) \leq 1, \quad i = 1, 2, \ldots, n,$$

for all $m(\geq 1)$.

Since A is a nonnegative matrix, then for any m, the matrix $\left(D^{(m+1)}A\right)$ is also nonnegative. Consequently, the hypothesis of the "Frobenius Theorem" (as applied to reducible matrices) is satisfied.[24] Hence, the matrix $\left(D^{(m+1)}A\right)$ has a maximal nonnegative real eigenvalue μ, such that $\left|\lambda_i\left(D^{(m+1)}A\right)\right| \leq \mu$, $i = 1, 2, \ldots, n$. Moreover, the eigenvector corresponding to this maximal eigenvalue is nonnegative.

Theorem I will be verified with the aid of the following lemma:

Lemma I. If the nonnegative matrix $\left(D^{(m+1)}A\right)$ possesses a positive eigenvector, then the eigenvalue corresponding to this positive eigenvector must be the maximal eigenvalue of $\left(D^{(m+1)}A\right)$.

We set $P^{(m+1)} = D^{(m+1)}A$ and denote the dominant eigenvalue of $P^{(m+1)}$ by μ. It follows that μ is also the dominant eigenvalue of the transposed matrix $\widetilde{P^{(m+1)}}$. Let $\underline{V}_1$ be the eigenvector of $\widetilde{P^{(m+1)}}$ corresponding to the eigenvalue μ. By the "Frobenius Theorem," $\underline{V}_1$ is a nonnegative vector. One has

$$\widetilde{P^{(m+1)}}\underline{V}_1 = \mu\underline{V}_1. \tag{46}$$

*For triangular response matrices, this theorem follows immediately from Eq. (16).

Assume further that there exists a positive eigenvector $\underline{V}_2$ of $P^{(m+1)}$ corresponding to an eigenvalue $\lambda < \mu$. Then,

$$P^{(m+1)}\underline{V}_2 = \lambda \underline{V}_2. \tag{47}$$

Taking the scalar product of Eqs. (46) and (47) with $\underline{V}_2$ and $\underline{V}_1$, respectively, yields

$$\mu(\underline{V}_2, \underline{V}_1) = \lambda(\underline{V}_1, \underline{V}_2).$$

Since $\lambda < \mu$, Eq. (47) can hold only if $(\underline{V}_1, \underline{V}_2) = 0$. However, this is impossible since $\underline{V}_1$ is a nonnegative vector and $\underline{V}_2$ is a positive vector. Hence, the eigenvectors of $P^{(m+1)}$ corresponding to $|\lambda_i| < \mu$ cannot be positive. Conversely, if λ_i does correspond to a positive eigenvector, then λ_i must be the maximal eigenvalue of $P^{(m+1)}$.

We may now apply this lemma. Examination of Eq. (33) reveals $D^{(m)}\underline{Y}$ is a positive eigenvector of the matrix $(D^{(m+1)}A)$ corresponding to an eigenvalue of unity. It follows that

$$|\lambda_i(D^{(m+1)}A)| \leq 1, \quad i = 1, 2, \ldots, n, \tag{48}$$

for all m.

The eigenvalues $\lambda_i(D^{(m+1)}A)$ also satisfy the characteristic equation,

$$|D^{(m+1)}A - \lambda I| = 0, \tag{49a}$$

which can also be written in the form,

$$|A - \lambda D^{-(m+1)}| = 0, \tag{49b}$$

where $D^{-(m+1)}$ is the inverse of $D^{(m+1)}$.* Now $D^{-(m+1)}$ is obviously positive definite for all m, and A is positive definite by assumption. It follows from the theory of quadratic forms,[25] that the eigenvalues found from Eq. (49b), hence those of Eq. (49a), are real and positive. Thus, the proof of Theorem I is complete.

Theorem I will be used to demonstrate that the sequence $\{F(\underline{X}^{(m)})\}$ is monotone decreasing. To this end, one notes that the nonnegative diagonal matrix D of Eq. (8) satisfies the relation,

$$DADY = DY. \tag{50}$$

---

*The abbreviated notation, $S^{-(m)} \equiv [S^{(m)}]^{-1}$, is used throughout.

Let us set

$$\Delta^{(m)} = D^{(m)} - D, \tag{51}$$

where $\{\Delta^{(m)}\}$ must be a sequence of diagonal matrices. The recursion relation for the sequence $\{\Delta^{(m)}\}$ may be obtained by substituting Eq. (51) into Eq. (33) and utilizing Eq. (50). One finds

$$\Delta^{(m+1)}\underline{Y} = \left(I - D^{(m+1)}A\right) \Delta^{(m)}\underline{Y}. \tag{52}$$

It is easy to verify that the matrix $\left(D^{(m+1)}A\right)$ corresponds to a symmetric operator with respect to A as a metric. That is, for arbitrary vectors $\underline{U}$ and $\underline{V}$, one can write

$$\left(\underline{V}, D^{(m+1)}A\underline{U}\right)_A = \tilde{\underline{V}}AD^{(m+1)}A\underline{U}, \tag{53a}$$

and

$$\left(D^{(m+1)}A\underline{V}, \underline{U}\right)_A = \tilde{\underline{V}}\tilde{A}D^{(m+1)}A\underline{U}. \tag{53b}$$

Consequently, for A positive definite, one has

$$\left(\underline{V}, D^{(m+1)}A\underline{U}\right)_A = \left(D^{(m+1)}A\underline{V}, \underline{U}\right)_A. \tag{54}$$

This condition implies that the eigenvectors comprising the modal matrix, $T^{-(m)}$, can be chosen A orthogonal. Proper normalization yields

$$\widetilde{T^{-(m)}}AT^{-(m)} = I. \tag{55}$$

It follows from Eq. (55) that the matrices $T^{-(m)}$ and $T^{(m)}$ exist for all m. Hence, for any m, the matrix $\left(D^{(m+1)}A\right)$ can be diagonalized. One has

$$T^{(m)}D^{(m+1)}AT^{-(m)} = \Lambda^{(m+1)}, \tag{56}$$

where the elements of the diagonal matrix $\Lambda^{(m+1)}$ are just the eigenvalues of $\left(D^{(m+1)}A\right)$.

Let us examine the form of recursion relation (52) in the principal axis system. One can write

$$\underline{Z}^{(m+1)} = \left(I - \Lambda^{(m+1)}\right) \underline{Z}^{(m)}, \tag{57a}$$

with

$$\underline{Z}^{(m+1)} = T^{(m)} \Delta^{(m+1)}\underline{Y}, \tag{57b}$$

and

$$\underline{Z}^{(m)} = T^{(m)}\Delta^{(m)}\underline{Y}. \tag{57c}$$

It is obvious from Theorem I and Eq. (57a) that the norm of $\underline{Z}^{(m+1)}$ is less than the norm of $\underline{Z}^{(m)}$. That is,

$$N\left(\underline{Z}^{(m)}\right) > N\left(\underline{Z}^{(m+1)}\right), \tag{58}$$

where the norm of a vector $\underline{Z}$ is defined in the customary manner as

$$N(\underline{Z}) = \sum_{i=1}^{n} z_i^2. \tag{59}$$

Multiplying Eqs. (57b) and (57c) on the left by $T^{-(m)}$, one finds

$$T^{-(m)}\underline{Z}^{(m+1)} = \Delta^{(m+1)}\underline{Y}, \tag{60a}$$

and

$$T^{-(m)}\underline{Z}^{(m)} = \Delta^{(m)}\underline{Y}. \tag{60b}$$

We proceed by forming the norm of these vectors with respect to A as a metric. Utilizing Eqs. (55) and (45), one finds

$$N\left(\underline{Z}^{(m+1)}\right) = N_A\left(\Delta^{(m+1)}\underline{Y}\right) = F\left(\underline{X}^{(m+1)}\right) \tag{61a}$$

and

$$N\left(\underline{Z}^{(m)}\right) = N_A\left(\Delta^{(m)}\underline{Y}\right) = F\left(\underline{X}^{(m)}\right), \tag{61b}$$

where N_A denotes the norm that has been formed with respect to A as a metric. Using Eqs. (61a) and (61b) in Eq. (58), one has, for all m,

$$F\left(\underline{X}^{(m)}\right) > F\left(\underline{X}^{(m+1)}\right). \tag{62}$$

Since the sequence $F\left(\underline{X}^{(m)}\right)$ is monotone decreasing, it converges. As in the two-dimensional case, convergence to the exact solution is assured unless a positive lower bound exists. However, from the above analysis it follows that convergence to a positive lower bound can occur only if some $\lambda_i\left(D^{(m+1)}A\right)$ converge to zero. Thus, we assume that some $\lambda_i\left(D^{(m+1)}A\right)$ do converge to zero. It follows from Eq. (56) that some subsequence of $\{D_{ii}^{(m)}\}$, $i = 1, 2, \ldots, n$, must converge to zero. Let $\{D_{kk}^{(p)}\}$ be such a subsequence. In this event, it follows from recursion relation (12) that all subsequences of $\{D_{kk}^{(m)}\}$ converge to zero, and hence the sequence $\{D_{kk}^{(m)}\}$ must itself converge to zero. Consequently, one can utilize this property to define two nonvacuous disjoint subsets of the elements of $D^{(m+1)}$. We denote by $\{D_{ii}^{(m+1)'}\}$ the subset containing the elements of $D^{(m+1)}$ that

converge to zero. The remaining elements of $D^{(m+1)}$, which possess a positive lower bound, are denoted by the subset $\{D_{ii}^{(m+1)"}\}$.

This result implies that by choosing m sufficiently large, the recursion relation (33) can be made arbitrarily close to a recursion relation that is representative of a system of lower dimension. This recursion relation is

$$D^{(m+1)"}A"D^{(m)"}\underline{Y}" = D^{(m)"}\underline{Y}", \tag{63a}$$

and the corresponding reduced system obviously is

$$\underline{Y}" = A"\underline{X}". \tag{63b}$$

Herein the matrix A" is positive definite.* Since all the elements of $D^{(m+1)"}$ possess a positive lower bound, then so do all the eigenvalues $\lambda_i\left(D^{(m+1)"}A"\right)$, and convergence to the exact solution for this reduced system is assured. One can write

$$\text{Lim } D^{(m)"} = D_r", \tag{64a}$$

and

$$D_r"A"D_r"\underline{Y}" = D_r"\underline{Y}". \tag{64b}$$

It follows from the assumption of a positive lower bound for all the elements of $\{D^{(m)"}\}$, that the reduced diagonal matrix $D_r"$ must be positive. Consequently, Eq. (64b) may be simplified. One has

$$D_r"\underline{Y}" = (A")^{-1}\underline{Y}". \tag{65}$$

Since the iterative process converges, one can write

$$\lim_{m\to\infty} \underline{X}^{(m)} = \underline{X}^L, \tag{66}$$

or for the individual elements,

$$\lim_{m\to\infty} x_i^{(m)} = x_i^L, \quad i = 1,2,\ldots,n. \tag{67}$$

Concerning the sequence $\{\underline{X}^{(m)}\}$, one has the following theorem.

Theorem II

The convergence of the sequence $\{\underline{X}^{(m)}\}$ to $\underline{X}^L$ is independent of the choice of the (positive) initial vector $\underline{X}^{(0)}$.

*The submatrix A", obtained from A by deleting like rows and columns, is a principal submatrix. Since A is positive definite, it follows that A" must also be positive definite.

Theorem II will be verified with the aid of the following lemma.

Lemma II. The convergence of any element, $x_i^{(m)} \to 0$, $i = 1,2, \ldots, n$, is independent of the choice of the (positive) initial vector.

Taking the derivative of Eq. (13) with respect to $x_j^{(0)}$, one finds

$$\frac{dx_i^{(m+1)}}{dx_j^{(0)}} = \frac{y_i}{y_i^{(m)}}\left(\frac{dx_i^{(m)}}{dx_j^{(0)}}\right) - x_i^{(m)}\frac{y_i}{\left(y_i^{(m)}\right)^2}\left(\frac{dy_i^{(m)}}{dx_j^{(0)}}\right), \quad j = 1,2, \ldots, n. \tag{68}$$

For $x_i^{(m)} \to 0$ and sufficiently large m, Eq. (68) takes the form,

$$\left(\frac{dx_i^{(m+1)}}{dx_j^{(0)}}\right) = \frac{y_i}{y_i^{(m)}}\left(\frac{dx_i^{(m)}}{dx_j^{(0)}}\right) + \epsilon_i^{(m)}, \quad j = 1,2, \ldots, n, \tag{69a}$$

where

$$\lim_{m\to\infty} \epsilon_i^{(m)} = 0. \tag{69b}$$

Taking absolute values on both sides of Eq. (69a), one finds

$$\left|\frac{dx_i^{(m+1)}}{dx_j^{(0)}}\right| \leq \frac{y_i}{y_i^{(m)}}\left|\frac{dx_i^{(m)}}{dx_j^{(0)}}\right| + \left|\epsilon_i^{(m)}\right|, \quad j = 1,2, \ldots, n. \tag{70}$$

Using the assumption $x_i^{(m)} \to 0$ in recursion relation (13), one has $y_i^{(m)} > y_i$ for all m sufficiently large. This fact, together with Eq. (70), implies

$$\left|\frac{dx_i^{(m+1)}}{dx_j^{(0)}}\right| < \left|\frac{dx_i^{(m)}}{dx_j^{(0)}}\right|, \quad j = 1,2, \ldots, n, \tag{71}$$

for all m sufficiently large. Equations (69a), (69b), and (71) imply the convergence of this sequence of derivatives. One can write

$$\lim_{m\to\infty}\left(\frac{dx_i^{(m)}}{dx_j^{(0)}}\right) = \beta_{ij}, \quad j = 1,2, \ldots, n. \tag{72}$$

Using the Taylor series representation of $x_i^{(m)}\left(x_j^{(0)} + \Delta x_j^{(0)}\right)$, one has (to first order)

$$x_i^{(m)}\left(x_j^{(0)} + \Delta x_j^{(0)}\right) = x_i^{(m)}\left(x_j^{(0)}\right) + \left(\frac{dx_i^{(m)}}{dx_j^{(0)}}\right) \cdot \Delta x_j^{(0)}. \tag{73}$$

Consequently, for $m \to \infty$, one finds

$$x_i^{L}\left(x_j^{(0)} + \Delta x_j^{(0)}\right) = x_i^{L}\left(x_j^{(0)}\right) + \beta_{ij}\Delta x_j^{(0)}, \quad j = 1,2, \ldots ,n. \tag{74a}$$

Since by assumption, $x_i^{L}\left(x_j^{(0)}\right) = 0$, one has, for sufficiently small $\Delta x_j^{(0)}$,

$$x_i^{L}\left(x_j^{(0)} + \Delta x_j^{(0)}\right) = \beta_{ij}\Delta x_j^{(0)}, \quad j = 1,2, \ldots ,n. \tag{74b}$$

Hence, for $\Delta x_j^{(0)} > 0$, $\beta_{ij} > 0$; and for $\Delta x_j^{(0)} < 0$, $\beta_{ij} < 0$. Since this is impossible, one must have $\beta_{ij} = 0$, $j = 1,2, \ldots ,n$, which verifies Lemma II.

Thus, for any choice of (the necessarily positive vector) $\underline{X}^{(0)}$, one finds the same reduced system given in Eqs. (63a) and (63b). As the reduced diagonal matrix D_r'' is defined uniquely in terms of this reduced system [viz., Eq. (65)], the iterative method converges to a unique vector, $\underline{X}^{L}$, independent of the choice of $\underline{X}^{(0)}$. This completes the proof of Theorem II.

It follows from Theorem II that the error function assumes a limiting form,

$$F\left(\underline{X}^{L}\right) = \alpha_L \geq 0, \tag{75}$$

independent of the choice of the (positive) initial vector $\underline{X}^{(0)}$. That the limit $\underline{X}^{L}$ minimizes the error function on the nonnegative subspace also follows from Theorem II. Assume that a nonnegative vector $\underline{Z}$ exists such that $F(\underline{Z}) = \alpha_L' < \alpha_L$. One can then contradict Theorem II by choosing $\underline{X}^{(0)} = (\underline{Z} + \underline{\epsilon})$, where $\underline{\epsilon}$ is a positive vector with arbitrarily small elements. Hence, for the positive definite case, the iterative method converges to the exact solution if and only if the exact solution lies in the nonnegative subspace.

In terms of the actual unfolding problem, it is clear that the structure of this iterative method admits the possibility of determining appropriate solutions even when exact solutions are physically meaningless. Standard iterative techniques, which converge to the exact solution, do not offer this possibility. That is, the exact solution $\underline{X}$, for problems of physical interest, is invariably beset with violent oscillations. Consequently, in practical applications, the present iterative method will not converge to $\underline{X}$, but will still minimize the error function on the nonnegative subspace.

V. APPLICATIONS

It has been noted that some success has been demonstrated by this method.[13-15] However, it must be stressed that this limited success, together with the domain of applicability described above, does not imply a general solution of the unfolding problem. First, one should recall that the theoretical treatment given above has been confined to an exact problem. The fact that unfolding problems in practice are not exact, and response matrices may be ill-conditioned, may lead to considerable complications. Secondly, from the viewpoint of practical applications, the rate of convergence of the method can be just as important as the convergence properties of the method. That is, the rate of convergence can be so poor that the iterative method is of little value. In view of Theorem I, it is obvious that the rate of convergence will depend on both the conditioning of the A matrix and the choice of the (positive) initial vector $\underline{X}^{(0)}$.

Hence, there can be no general guarantee that the approximations furnished by this method will be completely satisfactory. Additional physical information and criteria for a given detection system may be required before such a conclusion can be reached. Trial and testing of this method, with the response matrix under consideration and for the simplest physical cases that can arise, would also be advisable. Such a series of tests provides a basis for assessing the ability of the iterative method for a given detection system.

To examine some of the properties and implications of the theoretical treatment, the iteration method has been applied to a Li^6 solid-state neutron spectrometer. In this detection system, the ${}_3Li^6(n,\alpha){}_1H^3$ reaction is employed in conjunction with two separate silicon surface-barrier detectors. Each of the two charged particles, which result from this reaction, impinge on each of the silicon surface-barrier detectors, and the output of these two detectors is summed. The resulting pulse is suitably amplified and fed to a multichannel analyzer for pulse-height analysis.

Since a correction for the finite resolving power of this detector was desired, the response matrix was determined by measurements with a thermal neutron spectrum.* Table I presents the first ten rows and columns of a (33 x 33) response matrix determined from such measurements. This matrix is typical of what arises in practice, since many detectors possess an (approximately) Gaussian representation of the response resolution. Since it is well known that matrices of this type are poorly conditioned,[5] this response matrix should provide a more exacting test for the iteration method than the response matrices that have already been considered.[13-15]

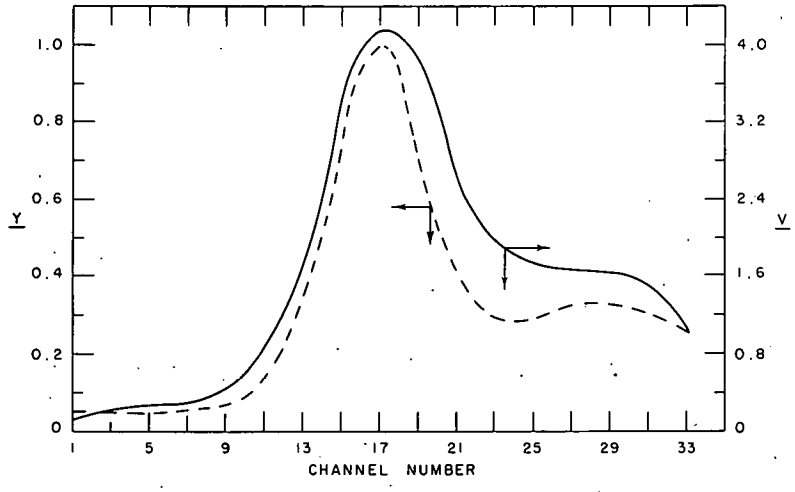
*Since the resolution of this detector was assumed independent of neutron energy, measurements at zero neutron energy (i.e., with thermal neutrons) define the response matrix completely.

Table I

RESPONSE MATRIX FOR THE NEUTRON DETECTOR

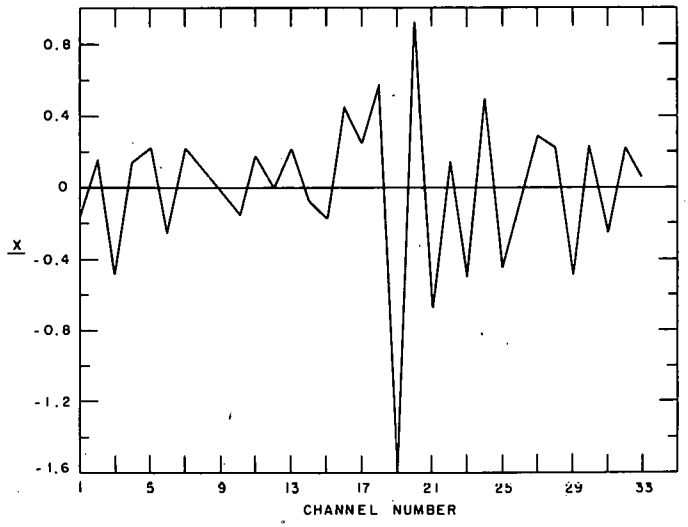
Row \ Col.	1	2	3	4	5	6	7	8	9	10
1	1.0000	0.8142	0.5948	0.3585	0.1843	0.1338	0.0871	0.0622	0.0467	0.0345
2	0.8244	1.0000	0.8142	0.5948	0.3585	0.1843	0.1338	0.0871	0.0622	0.0467
3	0.5083	0.8244	1.0000	0.8142	0.5948	0.3585	0.1843	0.1338	0.0871	0.0622
4	0.2310	0.5083	0.8244	1.0000	0.8142	0.5948	0.3585	0.1843	0.1338	0.0871
5	0.0579	0.2310	0.5083	0.8244	1.0000	0.8142	0.5948	0.3585	0.1843	0.1338
6	0.0209	0.0579	0.2310	0.5083	0.8244	1.0000	0.8142	0.5948	0.3585	0.1843
7	0.0054	0.0209	0.0579	0.2310	0.5083	0.8244	1.0000	0.8142	0.5948	0.3585
8	0.0010	0.0054	0.0209	0.0579	0.2310	0.5083	0.8244	1.0000	0.8142	0.5948
9	0.0019	0.0010	0.0054	0.0209	0.0579	0.2310	0.5083	0.8244	1.0000	0.8142
10	0.0024	0.0019	0.0010	0.0054	0.0209	0.0579	0.2310	0.5083	0.8244	1.0000

The neutron measurements of interest were taken in the center of a fast critical reactor assembly. To unfold the experimental data, the representation had to be symmetrized as described in Eq. (44). The vectors $\underline{Y}$ and $\underline{V}$, which result from the experimental data and the symmetrization process, respectively, are depicted in Fig. 1. As may be anticipated, the exact solution $\underline{X} = B^{-1}\underline{V}$, which is displayed in Fig. 2, proves to be completely unacceptable.



112-3813

Fig. 1. The Li^6 Solid-state Detector Pulse-height Distribution or $\underline{Y}$-vector, and the Symmetrized Experimental Data or $\underline{V}$-vector



112-3808

Fig. 2. The Exact Solution $\underline{X} = A^{-1}\underline{Y}$, Unfolded by the Inverse Response Matrix

Before depicting any iterative approximations, it is perhaps of greater significance to examine the rate of convergence. The simplest index that one can utilize to investigate the rate of convergence is the norm of the residual vector. Since the residual vector is defined as

$$\underline{R}^{(m)} = \underline{V}^{(m)} - \underline{V}, \tag{76}$$

the norm of the residual vector is given by

$$N\left(\underline{R}^{(m)}\right) = \left(\tilde{\underline{V}}^{(m)} - \tilde{\underline{V}}\right)\left(\underline{V}^{(m)} - \underline{V}\right) = \left(\underline{X}^{(m)} - \underline{X}\right) B^2 \left(\underline{X}^{(m)} - \underline{X}\right). \tag{77}$$

Examination of $N\left(\underline{R}^{(m)}\right)$ as a function of m will provide certain insight into the behavior of the iteration method.

To this end, Fig. 3 depicts $N\left(\underline{R}^{(m)}\right)$ as a function of m obtained for three different initial vectors with up to 80 iterations. These initial vectors are, respectively:

$$x_i^{(0)} = C > 0, \quad i = 1,2,\ldots,33; \tag{78a}$$

$$x_i^{(0)} = v_i, \quad i = 1,2,\ldots,33; \tag{78b}$$

$$x_i^{(0)} = 1/v_i, \quad i = 1,2,\ldots,33. \tag{78c}$$

The rate of convergence appears to be only slightly sensitive to the choice of these three initial vectors.

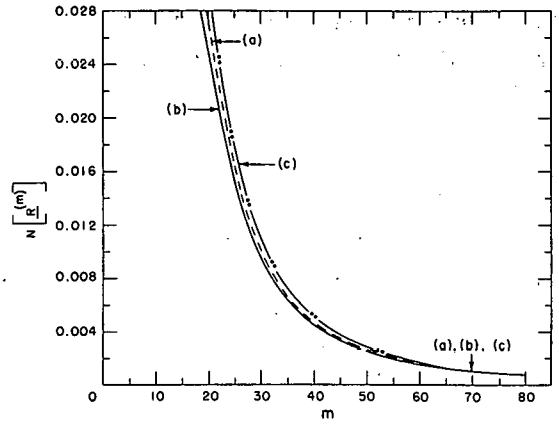


Fig. 3

The Norm of the Residual Vector as a Function of m for up to 80 Iterations with Three Different Initial Vectors: (a) $\underline{X}^{(0)}$ = Constant; (b) $\underline{X}^{(0)} = \underline{V}$; and (c) $x_i^{(0)} = (v_i)^{-1}$, i = 1,2,...,33

112-3809 Rev.

If one utilizes more unusual choices for the initial vector, the resulting effect on the rate of convergence can be pronounced. To demonstrate this behavior, initial vectors of the form

$$x_i^{(0)} = 0.05, \quad i = 1,2,\ldots,33, \quad i \neq 17;$$
$$x_{17}^{(0)} = C > 0, \tag{79}$$

have been used. The results for the initial vector given in Eq. (79), with C = 1, 2, 5, 10, and 100 are displayed in Fig. 4. Here it is apparent that the rate of convergence decreases rather drastically as the constant, C, increases. In addition, for all values of the constant, C, the rate of convergence is considerably reduced when compared with that obtained with the initial vectors of Eq. (78).

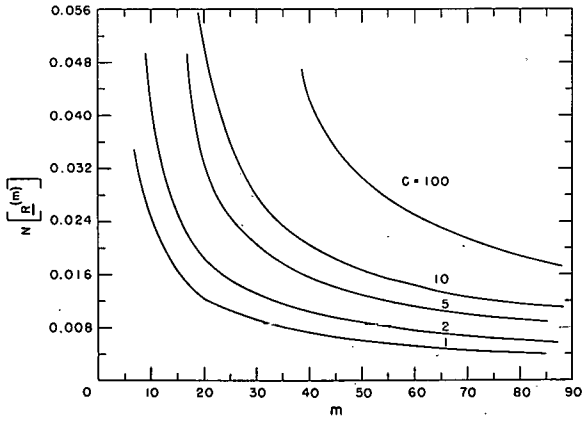


Fig. 4

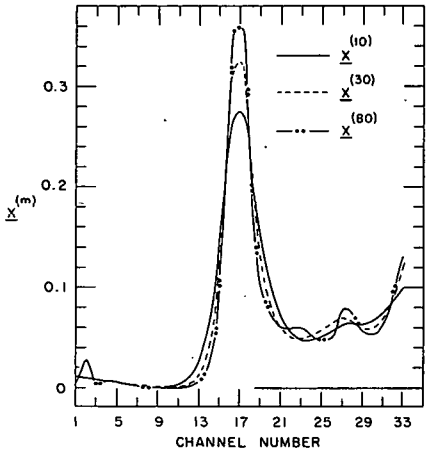
The Norm of the Residual Vector as a Function of m for up to 80 Iterations with Initial Vectors of the Form Given in Eq. (79), Where C = 1, 2, 5, 10, and 100

112-3810 Rev.

Although there can be little justification for an arbitrary initial vector of the type given in Eq. (79), the initial vectors given in Eqs. (78a) and (78b) can be justified on physical grounds. Thus the choice suggested in Eq. (78b) is reasonable when the response matrix introduces small distortions of the input vector. On the other hand, the choice of initial vector given in Eq. (78a) implies no a priori bias for the start of the iteration process. In view of the general influence of $\underline{X}^{(0)}$ upon the rate of convergence, it is apparent that this choice should be based upon pertinent physical implications that arise for the particular detection system under consideration.

Another feature, which is possibly more significant, is revealed in Fig. 4. That is, Fig. 4 also implies that the rate of convergence will be strongly influenced by the conditioning of the response matrix. This conclusion follows from Theorem I, since the eigenvalues that influence convergence, $\lambda_i(D^{(m)}A)$, are characteristic values of the product of the matrix $D^{(m)}$ and the response matrix A. Consequently, the conditioning of the response matrix has important bearing upon the rate of convergence and therefore can affect the adequacy of the method. This behavior can generally be ascertained by examining the approximations generated by the iteration method.

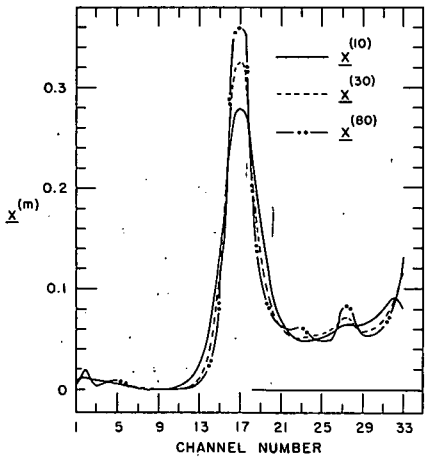
For the present detection system, the iterative approximations of interest have been obtained by employing initial vectors that possess physical justification. As has been mentioned above, these are the first two initial vectors given in Eq. (78). Figures 5 and 6 display the iterative approximations for m = 10, 30, and 80 obtained with these two initial vectors, respectively. The similarity exhibited in Figs. 5 and 6 demonstrates that the iterative approximations are rather insensitive to this particular change of initial vector. However, it is clear from the change in the peak values for these approximations that the rate of convergence is not great enough to properly define an appropriate solution.



112-3802 Rev.

Fig. 5

The Iterative Approximations $\underline{X}^{(10)}$, $\underline{X}^{(30)}$, and $\underline{X}^{(80)}$ of the Neutron Spectrum Obtained with the Initial Vector $\underline{X}^{(0)}$ = Constant



112-3803 Rev.

Fig. 6

The Iterative Approximations $\underline{X}^{(10)}$, $\underline{X}^{(30)}$, and $\underline{X}^{(80)}$ of the Neutron Spectrum Obtained with the Initial Vector $\underline{X}^{(0)} = \underline{V}$.

In this event, as stated earlier, additional criteria must be introduced. Two general conditions can be employed. These conditions are described below.

One can utilize the arresting condition proposed by Skarsgard, Johns, and Green,[8] which terminates the iterative process when the components of the residual vector are reduced to the order of the experimental error. From the data for $N(\underline{R}^{(m)})$, given in Fig. 3, this condition is met when m is roughly 30. Consequently, $\underline{X}^{(30)}$ would be chosen as the appropriate solution by this arresting condition. It is obvious that this arresting condition is precisely the same condition used to define the set χ. Consequently, the satisfaction of this arresting condition only implies that the resulting approximation is a member of χ.

Generally speaking, the arresting condition for iterative methods may only be viewed as a necessary condition. It can become a sufficient condition for a given detection system and iteration process, provided the rate of convergence is sufficiently high. For this to be the case, all the iterates $\{\underline{X}^{(m)}\}$ that are obtained above the arresting point must not differ appreciably from the approximation obtained at the arresting point. Otherwise, as for the present case under consideration (viz., Figs. 5 and 6), additional criteria are necessary to define an appropriate solution.

One further general criterion can be employed. Indeed, it has already been pointed out,[2,3] that a unique smooth solution of the unfolding problem does exist. Consequently, it is reasonable, especially for systems with poor rates of convergence, to choose the iterate $\underline{X}^{(m)}$ (beyond the arresting point) that in some sense possesses the "smoothest" behavior. It is often possible to choose this "smoothest" approximation by simply inspecting the sequence of vectors $\{\underline{X}^{(m)}\}$ generated by the iteration process.

To examine this smoothness criterion in more detail for the present detection system and iteration method, the computations were extended to m = 420 iterations. The approximations which result for m = 120, 180, and 240 are given in Fig. 7. Figure 8 displays similar results for m = 300, 360,

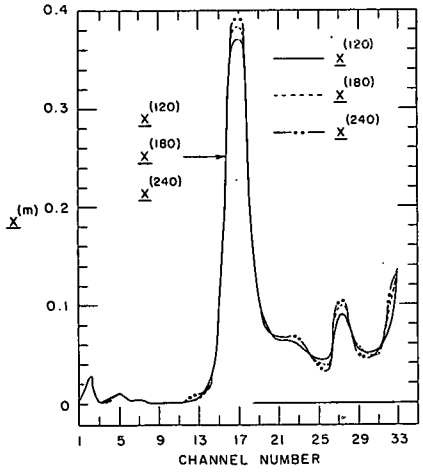


Fig. 7

The Iterative Approximations $\underline{X}^{(120)}$, $\underline{X}^{(180)}$, and $\underline{X}^{(240)}$ of the Neutron Spectrum Obtained with the Initial Vector $\underline{X}^{(0)}$ = Constant

112-3804 Rev. 2

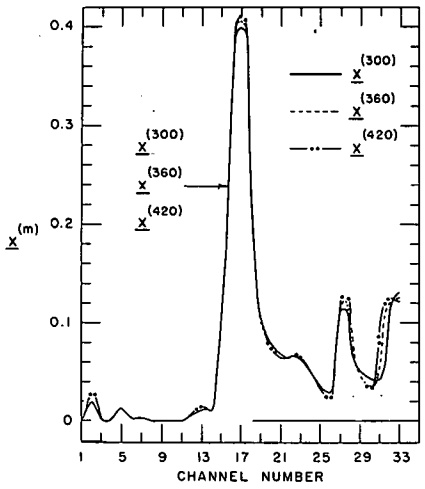


Fig. 8

The Iterative Approximations $\underline{X}^{(300)}$, $\underline{X}^{(360)}$, and $\underline{X}^{(420)}$ of the Neutron Spectrum Obtained with the Initial Vector $\underline{X}^{(0)}$ = Constant

and 420. It is obvious that as m increases, the iterates gradually become more peaked and sharply varying. Hence, in terms of the above criteria, one would choose $\underline{X}^{(30)}$ as the best available approximation of the appropriate solution.

This system also serves as a striking example of how additional physical information can be utilized. The main peak, which arises in the iterative approximations and is centered at channel No. 17, corresponds to thermal neutrons. Therefore, this main peak defines the zero neutron energy point (E_n = 0). The small peak that arises in the neighborhood of channel No. 27 corresponds to the resonance in the ${}_3Li^6(n,\alpha)_1H^3$ cross section at the neutron energy E_n = 0.26 MeV. These two points serve as an energy calibration for the pulse-height response of the detector. In terms of this calibration, one can compare the shape of the small peak at channel No. 27 with the measured features of this cross-section resonance.[26] In particular, the relative full width at half maximum of this resonance is roughly 40%. Inspection of the iterative approximations given in Figs. 5 to 8 reveals that this small peak is barely discernible at m = 10, but attains a reasonable shape at m = 30. For iterations beyond m = 30, this small peak becomes too narrow to be consistent with the known details of this resonance. Consequently, application of this additional physical information not only supports the choice of $\underline{X}^{(30)}$, but also furnishes a specific example of the success of the two general criteria discussed above.

VI. CONCLUSIONS

The application presented above clearly demonstrates that the ability of the iterative method to furnish appropriate solutions depends crucially upon the rate of convergence. If the response matrix is poorly conditioned, the rate of convergence may be too low to afford a proper definition of the appropriate solution. This result is not surprising since

the extent of the set χ is determined chiefly by the conditioning of the response matrix. From this viewpoint, the subset, $\chi' \subset \chi$, determined by the iterative method may not provide an adequate classification of the appropriate solution. In this event, as for the above results with the neutron detection system, one must introduce more subsidiary conditions than have already been employed (in the iteration method) in order to ascertain the appropriate solution.

Two possibilities exist for increasing the general applicability of this iterative method. First, one can consider modifications that could improve the rate of convergence. Perhaps the most apparent modification would be to neglect the components of an iterate $\underline{X}^{(m)}$ that fall below a certain small positive lower bound and continue the iteration process with the reduced system. This corresponds to the assumption that the same identical components actually vanish in the appropriate solution. This technique would keep the eigenvalues, $\lambda_i(D^{(m)}A)$, from becoming too small and thereby increase the rate of convergence.

The second possibility would be to apply additional criteria or constraints directly to the iterative approximations that arise. It would be highly desirable, for example, to incorporate a smoothing technique at every step of the iteration process. Such an approach also suggests that combinations of different unfolding methods, each of which may be successful only to a limited degree, may afford a more adequate solution of the general unfolding problem.

ACKNOWLEDGMENT

The author gratefully acknowledges the assistance of Mr. Norman Jesse who wrote the computer program utilized in these calculations.

REFERENCES

1. L. Fox and E. T. Goodwin, The Numerical Solution of Non-singular Linear Integral Equations, Phil. Trans. Roy. Soc. London, Ser. A 245, 501 (1953).

2. C. Eckart, The Correction of Continuous Spectra for the Finite Resolution of the Spectrometer, Phys. Rev. 51, 735 (1937).

3. G. Kreisel, Some Remarks on Integral Equations with Kernels: $L(\xi_1 - x_1, \xi_2 - x_2, \ldots, \xi_n - x_n; \alpha)$, Proc. Roy. Soc. (London), Ser. A 197, 160 (1949).

4. H. C. van de Hulst, Generalization of Some Methods for Solving an Integral Equation of the First Kind, Bull. Astron. Inst. Neth. 9, 225 (1941).

5. W. R. Dixon and J. H. Aitkens, The Resolution Correction in the Scintillation Spectrometry of Continuous X Rays, Can. J. Phys. 36, 1624 (1958).

6. J. C. Villforth, R. D. Birkhoff, and H. H. Hubbell, Jr., Comparison of Theoretical and Experimental Filtered X-Ray Spectra, ORNL-2529 (1958).

7. R. Gold and N. E. Scofield, Iterative Solution for the Matrix Representation of Detection Systems, Bull. Am. Phys. Soc. 2, 276 (1960).

8. L. D. Skarsgard, H. E. Johns, and L. E. S. Green, Iterative Response Correction for a Scintillation Spectrometer, Radiation Res. 14, 261 (1961).

9. R. P. Uhlig, An Iterative Unfolding Procedure, J. Res. Natl. Bur. Std. (U.S.), A 68A, 401 (1964).

10. A. J. Ferguson, A Program for the Analysis of Gamma-ray Scintillation Spectra Using the Method of Least Squares, CRP-1055 (AECL-1398) (1961).

11. J. I. Trombka, Least-squares Analysis of Gamma-ray Pulse-height Spectra, Applications of Computers to Nuclear and Radiochemistry, NAS-NS 3107, Paper (4-3), OTS, Dept. Commerce, Washington, D. C. (1963).

12. D. L. Phillips, A Technique for the Numerical Solution of Certain Integral Equations of the First Kind, J. Assoc. Computing Machinery 9, 84 (1962).

13. N. E. Scofield, A Technique for Unfolding Gamma-ray Scintillation Spectrometer Pulse-height Distribution, USNRDL-TR-447 (1960).

14. J. F. Mollenauer, A Computer Analysis for Complex Sodium Iodide Gamma Spectra, UCRL-9748 (1961).

15. N. E. Scofield, Iterative Unfolding, Applications of Computers to Nuclear and Radiochemistry, NAS-NS 3107, Paper (3-2), OTS, Dept. Commerce, Washington, D. C. (1963).

16. E. Bodewig, Matrix Calculus (North Holland Publishing Co., Amsterdam, 1956).

17. G. E. Forsythe, Tentative Classification of Methods and Bibliography on Solving Systems of Linear Equations, U. S. Nat. Bur. Std., Applied Math Series 29, 1 (1953).

18. J. W. Motz, Bremsstrahlung Differential Cross-section Measurements for 0.5- and 1.0-MeV Electrons, Phys. Rev. 100, 1560 (1955).

19. N. Starfelt and H. W. Koch, Differential Cross-section Measurements of Thin-target Bremsstrahlung Produced by 2.7- to 9.7-MeV Electrons, Phys. Rev. 102, 1598 (1956).

20. J. H. Hubbell, Response of a Large Sodium-iodide Detector to High-energy X-rays, Rev. Sci. Instr. 29, 65 (1958).

21. J. H. Hubbell and N. E. Scofield, Unscrambling of Gamma-ray Scintillation-spectrometer Pulse Height Distributions, IRE Trans. Nucl. Sci., NS-5, 156 (1958).

22. R. E. Rand, The Analysis of Continuous Spectra Using the Matrix Method, Nucl. Instr. Methods 17, 65 (1962).

23. R. M. Hayes, Iterative Methods of Solving Linear Problems on Hilbert Space, U. S. Nat. Bur. Std., Applied Math Series 39, 71 (1954).

24. F. R. Gantmacher, Applications of the Theory of Matrices (Interscience Publishers, Inc., New York, 1959).

25. F. R. Gantmacher, Theory of Matrices (Chelsa Publishing Co., New York, 1959), Vol. I.

26. J. R. Stehn, et al., Neutron Cross Sections, Vol. I. Z = 1 to 20, BNL-325, 2nd Ed., Suppl. 2 (1964).