

MULTIGROUP TRANSPORT EQUATIONS WITH NONDIAGONAL CROSS
SECTION MATRICES

by

Barton L. Willie

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree
of
DOCTOR OF PHILOSOPHY
in
Physics

APPROVED:

P. F. Zweifel, Chairman

G. Hagedorn

K. Hannögen

W. Greenberg

J. Sluony

S. Bowen

July, 1985
Blacksburg, Virginia

MULTIGROUP TRANSPORT EQUATIONS WITH NONDIAGONAL CROSS
SECTION MATRICES

by

Barton L. Willis

P. F. Zweifel advisor

Department of Physics

(ABSTRACT)

It is shown that multigroup transport equations with nondiagonal cross section matrices arise when the modal approximation is applied to energy dependent transport equations. This work is a study of such equations for the case that the cross section matrix is nondiagonalizable. For the special case of a two-group problem with a noninvertible scattering matrix, the problem is solved completely via the Wiener-Hopf method. For more general problems, generalized Chandrasekhar H equations are derived. A numerical method for their solution is proposed. Also, the exit distribution is written in terms of the H functions.

ACKNOWLEDGEMENTS

First and foremost I would like to thank my thesis advisor Professor Paul Zweifel for his patient guidance, expertise, and professional integrity. I will value my association with the Center for Transport Theory and Mathematical Physics for my entire career. Secondly, I would like to thank Dr. C. V. M. van der Mee who helped me with much of my dissertation research. His unbounded enthusiasm and abilities will not be forgotten. Thirdly, I would like to thank the members of my committee for providing me with a quality education.

I also thank my graduate student colleagues and friends from whom I have learned much about physics and mathematics, especially I thank:

Special thanks goes to
who read my dissertation and made many helpful comments.

Finally, I would like to thank my Blacksburg friends, especially:

I will miss all of them.

TABLE OF CONTENTS

| | |
|--|-----|
| Acknowledgements | iii |
| 1. Introduction | 1 |
| 2. Derivation of the Multigroup Equation | 6 |
| 3. An Equivalent Integral Equation | 12 |
| 4. The Finite Slab Problem | 17 |
| 5. The Half Space Problem | 23 |
| 6. The Wiener-Hopf Method of Solution | 26 |
| Figure 1 | 34 |
| Figure 2 | 35 |
| Figure 3 | 36 |
| 7. Construction of the Wiener-Hopf Factorization | 37 |
| 8. The Exit Distribution | 43 |
| 9. Invertible Scattering Matrices | 52 |
| 10. Invertibility of the Symbol: Invertible C | 54 |
| 11. The Generalized Chandrasekhar H-Equations | 56 |
| 12. A Banach Space Analysis of the H-Equations | 60 |
| 13. References | 69 |
| 14. Vita | 73 |

1. INTRODUCTION

Multigroup transport equations with nondiagonal cross section matrices have been proposed as models of, for example, neutron transport in reactors.^{1,2} Stacey¹ has shown that multigroup equations with nondiagonal cross section matrices arise when the energy dependence of the neutron transport equation is expanded as a finite sum of orthogonal functions, and then the method of weighted residuals is applied to determine equations satisfied by the coefficients of the expansion. The method of weighted residuals is discussed for general problems in Ames,^{3,4} and in Stacey, where different choices of orthogonal functions and weights are considered. Stacey shows that by appropriately choosing the weighting functions it is possible to minimize the error in some particular physical quantity. For example, if the error in the solution is to be minimized at n different spatial locations, he has shown that the appropriate weighting functions are delta functions. It should be noted that the traditional multigroup equation is a special case of this more general procedure. To derive the standard multigroup equations, the orthogonal functions are chosen to be characteristic functions with non-overlapping domains and the weighting functions to be the same set. If this procedure is carried out for the neutron transport equation, then the cross

section matrix will be diagonal. This equation has been studied by numerous authors.⁵⁻⁸ However, for different choices of orthogonal functions and weighting functions, the cross section matrix will in general be nondiagonal, and possibly even nondiagonalizable. In the case of a diagonalizable cross section matrix a similarity transformation applied to the equation will reduce the equation to the equation studied in Refs. 5 and 6, the only difference being that the scattering matrix might no longer be positive, or even real. In such cases it may be easier to work with a nondiagonal cross section matrix directly rather than make the change of variables.⁹ However, the cross section matrix might be nondiagonalizable, in which case no change of independent variables will reduce the transport equation to a form which has already been studied. Furthermore, if the number of groups is large it may be tedious to compute all the eigenvalues of the matrix in order to determine if the matrix is diagonalizable, and if it is diagonalizable it is difficult to determine the correct similarity transformation, since one needs all the eigenvectors. Therefore, this work is study of multigroup transport equations with nondiagonalizable cross section matrices.

There are two methods for solving such transport equations, namely, the Case eigenfunction method,¹⁰ and the

Wiener-Hopf method.¹¹ The Case eigenfunction method has the advantage that, at least at the outset, the formulas resemble the eigenfunction expansions so widely used in quantum mechanics. But unlike quantum mechanics, the operators used in transport theory are in general not self-adjoint, and so it is uncertain that an eigenfunction expansion does in fact exist. For many transport equations, these expansions do exist, but the eigenfunctions can be distributions, i.e. delta functions or principal values. The expansion coefficients satisfy equations which are in general singular integral equations. Singular integral equations have been investigated extensively by mathematicians, and methods for their solution are known.^{12,13} The crucial step in their solution is the construction of the Wiener-Hopf factorization of some function, which is of course the same factorization required by the Wiener-Hopf method for the original equation. Thus, either method of solution is equivalent to constructing a Wiener-Hopf factorization, but the Wiener-Hopf method avoids the principal values and delta functions which appear in the Case eigenfunction method. For this reason, the Wiener-Hopf method will be followed exclusively in this work.

In Sec. 2 a derivation of the multigroup equation is presented. In particular, the reasons behind the possible choices of weight functions are given. In Sec. 3, it is

proven that the multigroup equation is equivalent to an integral equation with a convolution kernel. For the case of a semi-infinite slab, this equation is the vector equivalent of the Wiener-Hopf equation. In Sec. 4, the existence and uniqueness of the integral equation for the finite slab problem is investigated. It is shown that questions of existence and uniqueness can be answered by a simple application of the Fredholm alternative. In Sec. 5, the transport equation will be specialized to a two-group equation, with half range boundary conditions, and a nondiagonalizable cross section matrix. For this problem the Wiener-Hopf factorization can be explicitly constructed provided that the scattering matrix is noninvertible. A noninvertible scattering matrix is not an exceptional case, as Zweifel and Seiwert⁵ have shown that for problems in radiative transfer the scattering matrix has a degenerate form i.e., $C_{ij} = A_i B_j$. In Sec. 6, necessary and sufficient conditions for the existence of the canonical Wiener-Hopf factorization will be found by studying the invertibility of the symbol of the equation. Once the Wiener-Hopf factorization has been found, it is possible to express the exit distribution, that is $\Psi(0, \mu)$, for $\mu > 0$, in terms of the Wiener-Hopf factorization. This will be done in Sec. 8. Rather than restrict the exit distribution solutions to the two-group case, the general N-group equation will be

considered, even though the Wiener-Hopf factorization might not be known. For the N-group equation problem, no general formula exists for the factorization. The best that can be done is to write down equations which the factors must satisfy. These equations are derived in Sec. 11, by making use of Mullikian's work.¹⁴ These equations are nonlinear integral equations which must be solved numerically. If the solutions are to have the analyticity properties required by the Wiener-Hopf factorization, then the equations must be supplemented by constraint equations. In Sec. 12, an iterative method of solution which will automatically satisfy the constraint equations is presented. These solutions are shown to converge in an appropriate Banach space.

2. DERIVATION OF THE MULTIGROUP EQUATION

Consider a one dimensional transport equation with azimuthal symmetry. The density of particles, $\Psi(x, \mu, E)$, is then a function of the one spatial coordinate x , of μ , the cosine of the angle that the velocity vector makes with the x axis, and of E the energy. For radiative transfer problems it is customary to use the wavelength instead of the photon energy. The equation governing the time independent density is

$$\mu \partial_x \Psi + \Sigma(E) \Psi = \int_{-1}^1 \int_0^\infty S(\mu, \mu', E, E') \Psi(x, \mu', E') d\mu' dE'. \quad (1)$$

Here ∂_x is the derivative with respect to x , $\Sigma(E)$ is the total cross section for scattering, which depends only on the energy of the particle, and $S(\mu, \mu', E, E')$ is the rate at which particles with energy E' and direction cosine μ' are scattered into energy E and direction μ . Unless simplifying assumptions are made, solving Eq. (1) is hopelessly difficult. One assumption which greatly reduces the complexity of Eq. (1) is to assume isotropic scattering, i.e., to assume that S is a function of energy only. Even with the assumption of isotropic scattering, the transport equation is still not amenable to analytical solutions, so approximate solutions are sought. One possibility, the so called modal approximation, is to approximate the energy

$$\Psi(x, \mu, E) = \sum_{j=0}^n \Psi_j(x, \mu) \Phi_j(E) . \quad (2)$$

A variety of choices for Φ_j have been considered, for example Laguerre or Hermite polynomials multiplied by Gaussians.^{2, 15} It is now necessary to determine equations for the Ψ_j . Perhaps the most obvious choice is to substitute the approximate solution [Eq. (2)] into the transport equation with the result;

$$\begin{aligned} \sum_{j=0}^n \mu \Phi_j \partial_x \Psi_j + \sum_{j=0}^n \Phi_j \Sigma \Psi_j = \\ \int_{-1}^1 \int_0^\infty S(\mu, \mu', E, E') \sum_{j=0}^n \Phi_j(E') \Psi_j(x, \mu') dE' d\mu' . \end{aligned} \quad (3)$$

If we multiply by $\int_0^\infty dE \Phi_m(E)$; the result is,

$$\begin{aligned} \mu \partial_x \Psi_m + (\Phi_m, \Sigma \Phi_n) \Psi_n = \\ \int_{-1}^1 (\Phi_m, S \Phi_n) \Psi_n d\mu' , \end{aligned} \quad (4)$$

where the Einstein sum convention is in force and

$$(\Phi_m, \Phi_n) = \int_{-1}^1 \Phi_m(x) \Phi_n(x) dx , \quad (5)$$

and S is the integral operator with kernel $S(\mu, \mu', E, E')$.

Now if matrices Σ_{mn} and C_{mn} are defined by;

$$\Sigma_{mn} = (\phi_m, \Sigma \phi_n) \quad (6a)$$

$$C_{mn} = 2(\phi_m, S\phi_n) , \quad (6b)$$

and if we assume isotropic scattering, i.e. S is independent of μ and μ' , then Eq. (3) can be written as a vector equation, namely;

$$\mu \partial_x \Psi + \Sigma \Psi = C/2 \int_{-1}^1 \Psi(x, \mu') d\mu' \quad (7)$$

where Ψ is now a vector with components $(\Psi_0, \Psi_1, \dots, \Psi_n)$. Note that the matrix Σ given by Eq. (6a) is real and symmetric, hence is diagonalizable. Therefore in this case, if U is a matrix which diagonalizes Σ , then $\tilde{\Psi} = U^{-1} \Psi$ satisfies the equation,

$$\mu \partial_x \tilde{\Psi} + \tilde{\Sigma} \tilde{\Psi} = \tilde{C}/2 \int_{-1}^1 \tilde{\Psi}(x, \mu') d\mu' . \quad (8)$$

where $\tilde{\Sigma}$ is the diagonal matrix, $\tilde{\Sigma} = U^{-1} \Sigma U$, and $\tilde{C} = U^{-1} C U$. Equation (8) has been studied under the assumptions of \tilde{C} constant, noninvertible, and $\tilde{C}_{ij} \gg 0$ by Siewert and

Zweifel⁵ in their study of radiative transfer problems, and for invertible \tilde{C} and $\tilde{C}_{ij} \geq 0$ by Sancaktar and Zweifel¹⁶ among others. Note that the matrix \tilde{C} in Eq. (8) is not necessarily positive; however, the analyses given in Refs. 5, and 16 require little modification for negative C matrices. Note that the matrix C is real.

Stacey has pointed out that alternative methods for obtaining equations for the Ψ_n exist, which may be, for some physical problems, more appropriate than the analysis given above. Stacey in Ref. 1, has promoted the method of weighted residuals to determine equations for the Ψ_n . In this method, rather than multiplying Eq. (3) by

$\int_0^\infty dE \phi_m(E)$, the equation is multiplied by $\int_0^\infty dE w_m(E)$,

where the functions w_m , $m=0,1,2\dots N$, are called the weighting functions. More generally, the w_m could be operators. The weighting functions are chosen to minimize the error in the approximate solution [Eq. (2)] in some sense. To illustrate this procedure, it is convenient to rewrite the transport equation [Eq. (1)] in operator form

$$H \Psi = 0, \quad (9)$$

where H is linear, and represent the solution to this

equation as a finite sum plus a remainder,

$$\Psi(x, \mu, E) = \sum_{j=0}^n \phi_j(E) \Psi_j(x, \mu) + \Delta\Psi(x, \mu, E) \quad (10a)$$

$$= \Psi_{\text{approx}} + \Delta\Psi. \quad (10b)$$

Substituting Eq. (10b) into Eq. (9), and making use of the linearity of H , we find

$$H\Psi_{\text{approx}} + H\Delta\Psi = 0. \quad (11)$$

Suppose that it is desired to minimize the error term, i.e. minimize $H\Delta\Psi$, in the least squares sense, that so we require

$$\int dx d\mu dE (H\Delta\Psi)^2 = \text{minimum}. \quad (12)$$

Stacey has shown that the weighting functions (actually in this case they are weighting operators) $w_m(E)$ which do this are given in Dirac notation by

$$w_m(E) = \langle \phi_m | H^* \quad (13)$$

where H^* is the adjoint of H . Stacey gives other choices of weighting operators which minimize the error in the solution

at m spatial points, (by choosing the weights to be delta functions) or the error in the first m -moments of the distribution, etc.

If this procedure is applied to Eq. (3), then it is easy to see that the matrices Σ and C defined by Eqs. (6a)-(6b) will be given by (again in Dirac notation)

$$\Sigma_{mn} = \langle \phi_m | H^* \Sigma | \phi_n \rangle \quad (14a)$$

$$C_{mn} = 2 \langle \phi_m | H^* S | \phi_n \rangle . \quad (14b)$$

The matrix Σ_{mn} is no longer symmetric, in fact it might not even be diagonalizable. The case of a transport equation with a nondiagonalizable cross section matrix and isotropic scattering is the subject of this work.

3. AN EQUIVALENT INTEGRAL EQUATION

As explained in the introduction, the Wiener-Hopf method will be used in this work. In order to use the Wiener-Hopf method, an integral equation equivalent to the integro-differential equation [Eq. (7)] is sought. This integral equation is derived in Theorem 1.

Theorem 1: If Ψ satisfies the integro-differential equation;

$$\mu \partial_x \Psi + \Sigma \Psi = C/2 \int_{-1}^{+1} \Psi(x, \mu') d\mu', \quad x \in [0, \infty) \quad (15)$$

where Σ and C are constant matrices, subject to the boundary conditions:

$$\Psi(x, \mu) \rightarrow 0, \quad x \rightarrow \infty \quad (16a)$$

$$\Psi(0, \mu) = \Phi_+(\mu) \quad \mu \in [0, 1], \quad (16b)$$

$$\Psi \in \bigoplus_{i=1}^n L_2(0, \infty) \otimes L_2(-1, 1), \quad (16c)$$

each of the conditions [Eqs. (16a)-(16b)] holds for where each component of Ψ separately, and we assume that the spectrum of Σ is contained in the right half plane [we denote the spectrum of Σ by $\sigma(\Sigma)$], then G , defined by

$$G(x) = \int_{-1}^{+1} \Psi(x, \mu') d\mu', \quad (17)$$

satisfies the Wiener-Hopf equation

$$G(x) = U(x) + \frac{1}{2} \int_0^{\infty} K(x-y) CG(y) dy, \quad (18)$$

where

$$U(x) = \int_0^1 \exp(-x\Sigma/\mu) \Phi_+(\mu) d\mu \quad (19)$$

and

$$K(x) = \int_0^1 \mu^{-1} \exp(-|x|\Sigma/\mu) d\mu. \quad (20)$$

Proof: Using the definition of G , [Eq. (17)], rewrite Eq. (15) as

$$\mu \partial_x \Psi + \Sigma \Psi = \frac{1}{2} CG. \quad (21)$$

If the right-hand-side of Eq. (21) is known, then one can solve this equation for Ψ by introducing the integrating factor. The integrating factor for this equation is $\exp(x\Sigma/\mu)$, so Eq. (21) can be rewritten as

$$\partial_x [\exp(x\Sigma/\mu) \Psi] = \frac{1}{2\mu} \exp(x\Sigma/\mu) CG(x). \quad (22)$$

Therefore,

$$\exp(x\Sigma/\mu) \Psi(x, \mu) = \frac{1}{2\mu} \int^x \exp(y\Sigma/\mu) CG(y) dy + A(\mu). \quad (23)$$

Here, $A(\mu)$ is an arbitrary constant of integration, and the limits of integration must be chosen to satisfy the boundary

conditions [Eqs. (16a)-(16c)]. In order to satisfy the condition $\Psi(x, \mu) \rightarrow 0$, $x \rightarrow \infty$, for $\mu \leq 0$ it is necessary to choose $A(\mu) = 0$, and the limits of integration to be x and ∞ , therefore;

$$\Psi(x, \mu) = -\frac{1}{2\mu} \int_x^{\infty} \exp[-(x-y)\Sigma/\mu] CG(y) dy, \quad \mu \leq 0. \quad (24a)$$

For $\mu > 0$, the boundary condition Eq. (16b) must be satisfied. To satisfy this condition one must choose $A(\mu) = \Phi_+(\mu)$, and the limits of integration to be 0 and x . So,

$$\Psi(x, \mu) = \exp(-x\Sigma/\mu) \Phi_+(\mu) + \frac{1}{2\mu} \int_0^x \exp[-(x-y)\Sigma/\mu] CG(y) dy \quad (24b)$$

for $\mu > 0$. To derive the integral equation that G must satisfy, multiply both sides of Eq. (24a) by $\int_{-1}^0 d\mu$ and Eq.

(24b) by $\int_0^1 d\mu$, and add these two terms. The result is;

$$\begin{aligned} G(x) = & \int_0^1 \exp(-x\Sigma/\mu) \Phi_+(\mu) d\mu \\ & - \int_{-1}^0 \frac{1}{2\mu} \int_x^{\infty} \exp[-(x-y)\Sigma/\mu] CG(y) dy \\ & + \int_0^1 \frac{1}{2\mu} \int_0^x \exp[-(x-y)\Sigma/\mu] CG(y) dy. \end{aligned} \quad (25)$$

Now interchange the order of integration in the last two

integrals, which is permissible by Fubini's Theorem, due to the condition on the spectrum of Σ , and the boundary condition (16a). Upon simplification, Eq. (18) results.

The integral equation is easily modified to finite slab problems, that is $x \in [0, L]$, where L is the length of the slab measured in mean free paths. For this problem, the boundary condition (16a) must be replaced by $\Psi(L, \mu) = \Phi_-(\mu)$, for $\mu \ll 0$. With this boundary condition it is easy to see that the integral equation for G is replaced by

$$G(x) = U(x) + \int_0^L K(x-y)CG(y) dy \quad (26a)$$

where;

$$U(x) = \int_{-1}^0 \exp(-x\Sigma/\mu) \Phi_-(\mu) d\mu + \int_0^1 \exp(-x\Sigma/\mu) \Phi_+(\mu) d\mu \quad (26b)$$

and $K(\cdot)$ is given by Eq. (20). The existence and uniqueness of solutions to Eq. (26) will be discussed in Sec. 4.

One should not be intimidated by the exponential functions of possibly nonself-adjoint matrices appearing in the above formulas. They are easy to compute if the Jordan decomposition is used. The Jordan decomposition of a matrix

A is defined by

$$A = S + N, \quad SN = NS \quad (27)$$

with S self-adjoint and N nilpotent. Such a decomposition exists for all matrices, and formulas for S and N can be found in Ref. 17. For any sufficiently differentiable function F, $F(A)$ may be computed by

$$F(A) = \sum_{n=0}^M \frac{1}{n!} F^{(n)}(S) N^n \quad (28)$$

where $F^{(n)}$ is the n-th derivative of F, and M is order of nilpotency of N. The matrices $F^{(n)}(S)$ may be computed by making use of the spectral theorem for self-adjoint matrices.

4. THE FINITE SLAB PROBLEM

Quite often in transport theory, it is easy to prove that the equation has at most one solution, but it is considerably more difficult to prove that a solution does in fact exist. Questions of existence and uniqueness are not just academic, as we typically look for time independent solutions to transport equations, and as we know many physical systems have more than one equilibrium (time independent) solution. An example of this is the Vlasov equation. It has infinitely many time independent solutions.¹⁹

If both C and $\Sigma - C$ are positive definite matrices, then it is possible to prove that the transport equation subject to the boundary conditions [Eqs. (16a)-(16b)], has at most one solution, that is, if a solution exists then it is unique. This result holds for either finite or semi-infinite boundary conditions. In the case of a finite slab, the integral operator appearing in Eq. (26a) is compact, so that the Fredholm alternative¹⁸ can be used to prove that the equation has exactly one solution. Recall that the Fredholm alternative states that if K is compact, then either $(I - K)^{-1}$ exists as a bounded operator, or K is not injective. (We denote the identity operator by I .) In other words, if K is injective, then $(I - K)^{-1}$ exists as a bounded operator. A discussion of the applicability of

using the Fredholm alternative to prove the existence of solutions to transport equations can be found in Ref. 20. The uniqueness proof given in this section is a generalization of the proof found in Case and Zweifel.¹⁰ The proof relies on a simple positivity argument.

Theorem 2: The transport equation subject to the boundary conditions (16a)-(16b) has at most one solution if both C and $\Sigma - C$ are positive definite.

Proof: The proof for the finite slab problem will be given; the proof for semi-infinite boundary conditions is essentially the same. For the purpose of contradiction, assume that the transport equation has two solutions, say Ψ_1 and Ψ_2 . Since the transport equation is linear, $\Psi = \Psi_1 - \Psi_2$ is also a solution, but it satisfies the boundary conditions;

$$\Psi(0, \mu) = 0 \quad \mu > 0 \quad (29a)$$

$$\Psi(L, \mu) = 0 \quad \mu < 0, \quad (29b)$$

that is the incident flux is zero. Let Ψ be a column vector, and Ψ^T its transpose. Multiply the transport equation by the operator $\int_0^L dx \Psi^+$. The result is

$$\int_0^L dx \Psi^T (\mu \partial_x + \Sigma) \Psi = \frac{1}{2} \int_0^L dx \Psi^T C \int_{-1}^1 d\mu' \Psi(x, \mu') \quad (30)$$

The first term on the left-hand-side is an exact differential, so Eq. (30) simplifies to

$$\begin{aligned} \frac{1}{2}\mu\langle\P^T\P\rangle(L,\mu) - \frac{1}{2}\mu\langle\P^T\P\rangle(0,\mu) \\ + \int_0^L dx \Psi^T\Sigma\P = \frac{1}{2} \int_0^L dx \Psi^TC\int_{-1}^1 d\mu' \Psi(x,\mu') . \end{aligned} \quad (31)$$

Now, multiply Eq. (31) by the operator $2\int_{-1}^1 d\mu$, and use the boundary conditions (29a)-(29b). We find

$$\begin{aligned} \int_0^1 \mu\langle\P^T\P\rangle(L,\mu) d\mu - \int_{-1}^0 \mu\langle\P^T\P\rangle(0,\mu) \\ + \int_{-1}^1 \int_0^L d\mu dx \Psi^T(x,\mu)\Sigma\P(x,\mu) \\ = \int_{-1}^1 \int_0^L d\mu dx \Psi^T(x,\mu)C\int_{-1}^1 d\mu' \Psi(x,\mu') \end{aligned} \quad (32)$$

The right-hand-side can be analyzed by using the trick in Ref. 10. Consider

$$\int_{-1}^1 \int_{-1}^1 [\Psi^T(x,\mu)-\Psi^T(x,\mu')][C[\Psi(x,\mu)-\Psi(x,\mu')]] d\mu d\mu'$$

$$= 4 \int_{-1}^1 \Psi^T C \Psi(x, \mu) d\mu - 2 \int_{-1}^1 \int_{-1}^1 \Psi^T(x, \mu) C \Psi(x, \mu') d\mu d\mu' . \quad (33)$$

Using the boundary conditions [Eqs. (29a)-(29b)] and Eq. (33), Eq. (32) can be rewritten as

$$\begin{aligned} & \int_0^1 \mu (\Psi^T \Psi)(L, \mu) d\mu - \int_{-1}^0 \mu (\Psi^T \Psi)(0, \mu) d\mu \\ & + \int_0^L \int_{-1}^1 \Psi^T(x, \mu) (\Sigma - C) \Psi(x, \mu) d\mu dx \\ & = - \int_0^L \int_{-1}^1 \int_{-1}^1 [\Psi^T(x, \mu) - \Psi^T(x, \mu')] C [\Psi(x, \mu) - \Psi(x, \mu')] d\mu d\mu' dx \end{aligned} \quad (34)$$

The crucial point is that the integral terms involving $\Psi(0, \mu)$ and $\Psi(L, \mu)$ are greater than zero, so that provided both $\Sigma - C$ and C are positive definite the left-hand side of Eq. (34) is ≥ 0 , while the right-hand-side is ≤ 0 . Therefore, it must be the case that $\Psi = 0$, so uniqueness has been proven. Note that the conditions on C and $\Sigma - C$ reduce to the well known condition $C \in [0, 1)$ in the one-group case.

So far all we have proven is that the transport

equation has at most one solution. Now we must show that the equation has exactly one solution. As previously mentioned, the Fredholm alternative says that uniqueness implies existence provided that the integral operator appearing in Eq. (26a) is compact. If $L < \infty$, then the operator is indeed compact.

Theorem 3: The integral operator defined by

$$(K\psi)(x) = \int_0^L K(x-y)\psi(y) dy \quad 0 < L < \infty \quad (35a)$$

$$K(x) = \int_0^1 \mu^{-1} \exp(-x\Sigma\mu) d\mu \quad (35b)$$

$$\sigma(\Sigma) \subset \{ z \in \mathbb{C} : \operatorname{Re}(z) < 0 \} \quad (35c)$$

is compact on $L_2(0,L) \oplus \dots \oplus L_2(0,L)$.

Proof: It suffices to show that K is Hilbert-Schmidt,¹⁸ which will be true provided

$$\int_0^L \int_0^L [K_{mn}(x-y)]^2 dx dy < \infty \quad (36)$$

holds each element of the matrix K . The verification of inequality (36) is a straightforward calculation, so it will not be presented. Therefore, the conditions;

$$C, \text{ and } \Sigma - C \text{ positive definite} \quad (37a)$$

$$\sigma(\Sigma) \subset \{ z \in \mathbb{C} : \operatorname{Re}(z) > 0 \} \quad (37b)$$

are sufficient to guarantee that the transport equation with finite slab geometry possess a unique solution. Unfortunately, the integral operator with $L = \infty$ [see Eq. (18)] is not compact,²¹ so that in this case uniqueness does not imply existence. Existence for one particular half space problem will be discussed in Sec. 5.

It might be mentioned that it is also possible to prove uniqueness using the equivalent integral equation, rather than the transport equation. To do this, it is sufficient to show that the norm of the integral operator is less than one. This approach might be useful if one wanted to prove uniqueness in a Banach space, say L_1 , rather than in a Hilbert space as was done in Theorem 3. For the one-group problem, this criterion gives the same condition as found by using positivity arguments.

5. THE HALF SPACE PROBLEM

In this Section the transport equation

$$\mu \partial_x \Psi + \Sigma \Psi = \frac{1}{2} C \int_{-1}^1 \Psi(x, \mu') d\mu' \quad (38)$$

subject to the boundary conditions Eqs. (16a)-(16c), will be studied. As previously mentioned, if Σ is diagonalizable then a similarity transformation will reduce Eq. (38) to the problem considered in Ref. 5. More generally, Eq. (38) is also soluble for the case where the matrices Σ and C are simultaneously upper triangularizable. In such a case, a similarity transformation can be applied to the transport equation to reduce it to a tractable form. To see this, suppose that both Σ and C are upper triangular. Then the last component of Ψ , call it Ψ_n , is not coupled to the other components of Ψ . Therefore, Ψ_n satisfies the one-group neutron transport equation; thus its solution can be written in terms of the Chandrasekhar H functions.¹⁰ These functions have been numerically computed to great accuracy. Once the function Ψ_n is known, the $n-1$ 'th component of Ψ obeys an equation involving only Ψ_n and Ψ_{n-1} . Thus Ψ_{n-1} obeys an inhomogeneous one group equation, so it is also possible to solve for Ψ_{n-1} in terms of the H functions. Proceeding in this fashion it is possible to solve for each component of Ψ . Thus we see that whenever Σ and C are simultaneously upper triangularizable, the multigroup

equation is soluble (at least in principle). It would be nice to know necessary and sufficient conditions for two matrices to be simultaneously upper triangularizable. A sufficient condition for two matrices to be simultaneously upper triangularizable is that they commute, but a necessary condition is apparently unknown. In fact, it is not necessary for two matrices to commute in order for them to be simultaneously upper triangularizable; an example will occur in Sec. 7.

In the following, the simplest equation of the form of Eq. (38) will be studied for which Σ is not diagonalizable, and for which Σ and C are not simultaneously upper triangularizable. In particular we consider the two group equation defined by

$$\Sigma = \begin{vmatrix} 1 & \alpha \\ 0 & 1 \end{vmatrix} ; C = \begin{vmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{vmatrix} , \quad (39)$$

with $\alpha \neq 0$ and $c_{21} \neq 0$. It is easily verified that Σ and C satisfy the two previously mentioned conditions. Additionally, in this section we shall assume that the matrix C (the scattering matrix) is noninvertible. Recall that Siewert and Zweifel⁵ have shown that problems in radiative transfer naturally lead to noninvertible scattering matrices. For these problems the scattering matrix has a degenerate form, i.e. there exist constants a_j and b_j so that $c_{ij} = a_i b_j$. It is easy to prove that every

degenerate matrix is noninvertible. For a noninvertible scattering matrix it is possible to construct explicitly the Wiener-Hopf factorization of the symbol of the equivalent integral equation at least for the two-group case, a task which is identical to solving the transport equation. Actually, a similarity transformation can always be applied to set $\alpha = 1$ in Eq. (39), but for bookkeeping purposes it is convenient to leave α as an arbitrary parameter so that the limit $\alpha \rightarrow 0$ is apparent. When $\alpha \rightarrow 0$, Eq. (38) reduces to the multigroup equation with a diagonal cross section matrix. Thus, the limit $\alpha \rightarrow 0$ will serve as a check of the formulas in the next section where the invertibility of the symbol of the Wiener-Hopf equation will be studied. Note that every nondiagonalizable 2×2 matrix is proportional to Σ after an appropriate similarity transformation, so the choice of Σ is not as restrictive as might first be thought.

6. THE WIENER-HOPF METHOD OF SOLUTION

We now proceed to solve the integral equation [Eq. (18)] using the Wiener-Hopf method.^{7,11} Following the standard notation define the functions G^+ , G^- , U^+ , and U^- by;

$$G^+(x) = G(x) \quad \text{for } x > 0 \quad (40a)$$

$$G^+(x) = 0 \quad \text{for } x < 0 \quad (40b)$$

and,

$$G^-(x) = 0 \quad \text{for } x > 0 \quad (40c)$$

$$G^-(x) = G(x) \quad \text{for } x < 0, \quad (40d)$$

and similarly for U^+ and U^- . With these definitions, Eq. (18) can be rewritten as a convolution equation on the entire real line, namely;

$$G^+(x) + G^-(x) = U^+(x) + \frac{1}{2} \int_{-\infty}^{\infty} K(x-y) G^+(y) dy. \quad (41)$$

The function K has been defined in Eq. (20). The Fourier transform of Eq. (41) yields;

$$W(\lambda) \hat{G}^+(\lambda) + \hat{G}^-(\lambda) = \hat{U}^+(\lambda). \quad (42)$$

Here the Fourier transform of a function F is denoted as \hat{F} , where

$$\hat{F}(\lambda) = \int_{-\infty}^{+\infty} \exp(i\lambda x) F(x) dx . \quad (43)$$

The matrix valued function W is called the symbol of the Wiener-Hopf equation, and is given by $I - \hat{K}$ or by

$$W(\lambda) = I - \lambda^{-1}(\tan^{-1}\lambda)C + (1+\lambda^2)^{-1}MC . \quad (44)$$

Here I is the 2×2 identity matrix, and M is the nilpotent part of the matrix Σ . The nilpotent part of the matrix Σ can be computed by making use of Ref. 17, or for this simple case it can be found by inspection. The Jordan decomposition of Σ [see Eq. (27)] is

$$\Sigma = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} + \begin{vmatrix} 0 & \alpha \\ 0 & 0 \end{vmatrix} \quad (45a)$$

$$= I + M . \quad (45b)$$

The crucial step in the Wiener-Hopf method is the construction of the Wiener-Hopf factorization of the symbol. If the symbol is a scalar function, i.e. a 1×1 matrix, then the factorization can always be reduced to quadratures.²² Typically these integrals are not expressible in terms of any standard elementary functions, but nevertheless, they can be numerically evaluated to any desired accuracy. For the one-group case, these

factorizations were first evaluated by Chandrasekhar.²³ Rather than evaluate the integral expressions for the Wiener-Hopf factorizations, Chandrasekhar found approximate solutions to the nonlinear integral equations obeyed by the factors of the symbol. His procedure was to replace the integral terms by quadrature formulas. The resulting equations were tractable algebraic equations. Unfortunately, the Wiener-Hopf factorization of a matrix valued function can only be constructed for special cases when the order of the matrix is greater than one. Therefore, for many cases it will be necessary to resort to a method similar to the one introduced by Chandrasekhar. This will be done in Sec. 12.

A canonical Wiener-Hopf factorization is a pair of functions W^+ and W^- so that

$$W(\lambda) = W^-(\lambda) W^+(\lambda) \quad \lambda \in \mathbb{R}, \quad (46)$$

such that the matrix function W^+ (W^-) is analytic in the open upper (lower) half complex plane, and continuous and invertible in the closed upper (lower) half plane. For the one group case, a factorization of the form Eq. (46) exists only for C less than one. The nonexistence of a canonical factorization implies that the integral equation does not have a unique solution.²⁴ Physically, this means that the

one speed neutron transport equation has a unique solution only if the average number of neutrons liberated per collision is less than one. For radiative transfer problems, C less than one means that the medium absorbs energy on the average from the photons. Since the canonical factorization is by assumption invertible for all real values of the argument of the symbol, a necessary condition for the existence of a canonical factorization is that $W(\lambda)$ must be invertible for $\lambda \in \mathbb{R} \cup \{\infty\}$, that is, $\det W(\lambda) \neq 0$ for $\lambda \in \mathbb{R}_\infty$. For this reason, one must study the zeros of $\det W$. The determinant of the symbol is called the dispersion function. Explicitly, the dispersion function is given by;

$$\det W(\lambda) = 1 - \text{tr } C[\lambda^{-1} \tan^{-1} \lambda] + \alpha c_{21} [1 + \lambda^2]^{-1}. \quad (47)$$

Here, $\text{tr } C$ denotes the trace of C , and the assumption that $\det C = 0$ has been used (i.e. C is noninvertible). It is not surprising that the product αc_{21} plays a special role in Eq. (47), because if either α or $c_{21} = 0$, then all formulas must reduce to the well known results with a diagonal cross section matrix. Observe that the dispersion function has branch points at $\pm i$. We will always choose the branch cuts to be the lines $z = it$, $|t| \geq 1$. Therefore the dispersion function is analytic on the set $\mathbb{C} \setminus \{z \in \mathbb{C} : z = it, |t| \geq 1\}$.

$1, t \in \mathbb{R}$). Note that

$$\lim_{|\lambda| \rightarrow \infty} \det W(\lambda) = 1 \quad (48a)$$

holds inside the region of analyticity. In order to study the zeros of $\det W$, the symmetries;

$$[\det W(\lambda)]^* = \det W(\lambda^*) \quad (48b)$$

$$\det W(-\lambda) = \det W(\lambda), \quad (48c)$$

where the superscript $*$ means complex conjugate, are useful. The symmetries Eqs. (48a)-(48b), are not unusual; they hold whenever the kernel of the integral operator is real and symmetric. These symmetries imply that λ_0 is a zero of the dispersion function if and only if both λ_0^* and $-\lambda_0$ are also zeros of the dispersion function. Therefore the dispersion function must have an even number of zeros. The symmetries [Eqs.(48b)-(48c)] along with the behavior of $\det W$ at infinity [Eq. (48a)] allow one to compute the number of zeros of the dispersion function by computing the change of the argument of $\det W$ along the branch cuts, (the Nyquist method²⁵) just as is done in the one-group case.¹⁰ Recall that the argument principle says that if F is a function of a complex variable which is analytic inside a simple closed positively oriented contour, and F is nonvanishing on the contour, then the change in the argument of F after

completing one complete circuit around the contour is equal to 2π times the number of zeros of F inside of the contour. We apply the argument principle to the contour in Fig. 1. This problem divides into three special cases: (i) $\text{tr } C = 0$, (ii) $\alpha c_{21} = 0$, and (iii) both $\text{tr } C \neq 0$ and $\alpha c_{21} \neq 0$. The case $\text{tr } C = 0$ is solved easily by algebra, and case (ii) is identical to the one-group dispersion function so that the number of zeros is known.¹⁰ These results are summarized in Fig. 2. Case (iii) requires special attention. Unlike the one-group dispersion function, i.e. $\alpha c_{21} = 0$, the dispersion function now has poles at the branch points due to the term $\alpha c_{21} [1+\lambda^2]^{-1}$ [see Eq. (47)]. For this case, the change in the argument when rounding the branch points is now important. For this reason, the change in the argument of the dispersion function (denoted by $\Delta \text{Arg det } W$) along the contour in Fig. 1 will be considered in the limit ϵ and $\delta \downarrow 0$. First study $\Delta \text{Arg det } W$ along the straight lines Γ_ϵ by taking the limit $\epsilon \downarrow 0$ while keeping δ a constant, then study $\Delta \text{arg det } W$ along the circle C_δ by taking the limit $\delta \downarrow 0$. Along the lines Γ_ϵ the real and imaginary parts of the boundary values of $\det W$ are;

$$\text{Re det } (\pm 0 + iy) = (\text{tr } C/2y) \ln \left| \frac{1+y}{1-y} \right| + \alpha c_{21} (1-y^2)^{-1} \quad (49a)$$

$$\text{Im det } W(\pm 0 + iy) = \pm (\pi \text{tr } C)/2y \quad (49b)$$

(Note that Eq. (49b) proves that $\det W$ is nonvanishing on the contour Γ as required by the argument principle.) With these formulas, the Nyquist diagram for the contour Γ_ε can be sketched; for the case $\alpha c_{21} > 0$ and $\text{tr } C > 0$ the result is shown in Fig. 3. The diagrams for the other possible choices of signs of αc_{21} and $\text{tr } C$ are similar. To complete the Nyquist diagrams, the contour C_δ must now be considered. Along the C_δ , the pole term $(1+\lambda^2)^{-1}$ dominates, and the contour approaches a circle at infinity as $\delta \downarrow 0$. On C_δ , the dispersion function can be estimated by, for $\delta > 0$ and $-\pi/2 \ll \theta \ll \pi/2$

$$\det W(i+\delta e^{i\theta}) = -i\alpha c_{21}\delta^{-1}e^{-i\theta} + \text{tr } C(\frac{\theta}{2} + \frac{\pi}{4}) + O(\ln\delta^{-1}), \quad (50a)$$

and for $\pi/2 \ll \theta \ll 3\pi/2$,

$$\det W(i+\delta e^{i\theta}) = -i\alpha c_{21}\delta^{-1}e^{-i\theta} - \text{tr } C(\frac{\theta}{2} + \frac{\pi}{4}) + O(\ln\delta^{-1}) \quad (50b)$$

Therefore, as $\delta \downarrow 0$, the image of C_δ (note C_δ is counterclockwise) approaches a circle at infinity with the opposite orientation, i.e., the image is clockwise. With this information, the Nyquist diagrams can be sketched (see Fig. 3), and the number of zeros of the dispersion function can be deduced. Now that the number of zeros of the

dispersion function is known, the remaining task is to determine whether the zeros are purely real, purely imaginary, or neither. Recall that a necessary condition for the existence of a canonical Wiener-Hopf factorization is that the symbol must be invertible for all real numbers. Therefore, if the dispersion function has a purely real zero, then a canonical factorization does not exist. Fortunately, the graphs of the real and imaginary parts of the dispersion function are easy to sketch, so it is easy to determine if the dispersion function has a real zero. These results are summarized in Fig. 3. Thus we can conclude that $W(\lambda)$, $\lambda \in \mathbb{R}_c$, is invertible for $1 + \alpha c_{21} > \text{tr } C$ and $\text{tr } C < 1$. Note that when the number of groups is reduced to one, the condition $1 + \alpha c_{21} = \text{tr } C$ reduces to the familiar condition $c < 1$.

It might also be mentioned that the line defined by $\alpha c_{21} = \text{tr } C - 1$ (see Fig. 3) is a line of degenerate roots, which of course correspond to degenerate eigenvalues of the transport equation. As this line is approached in the $(\alpha c_{21}, \text{tr } C)$ plane, two nondegenerate zeros collide to form one degenerate root.

We now turn our attention to the construction of the Wiener-Hopf factorization of the symbol.

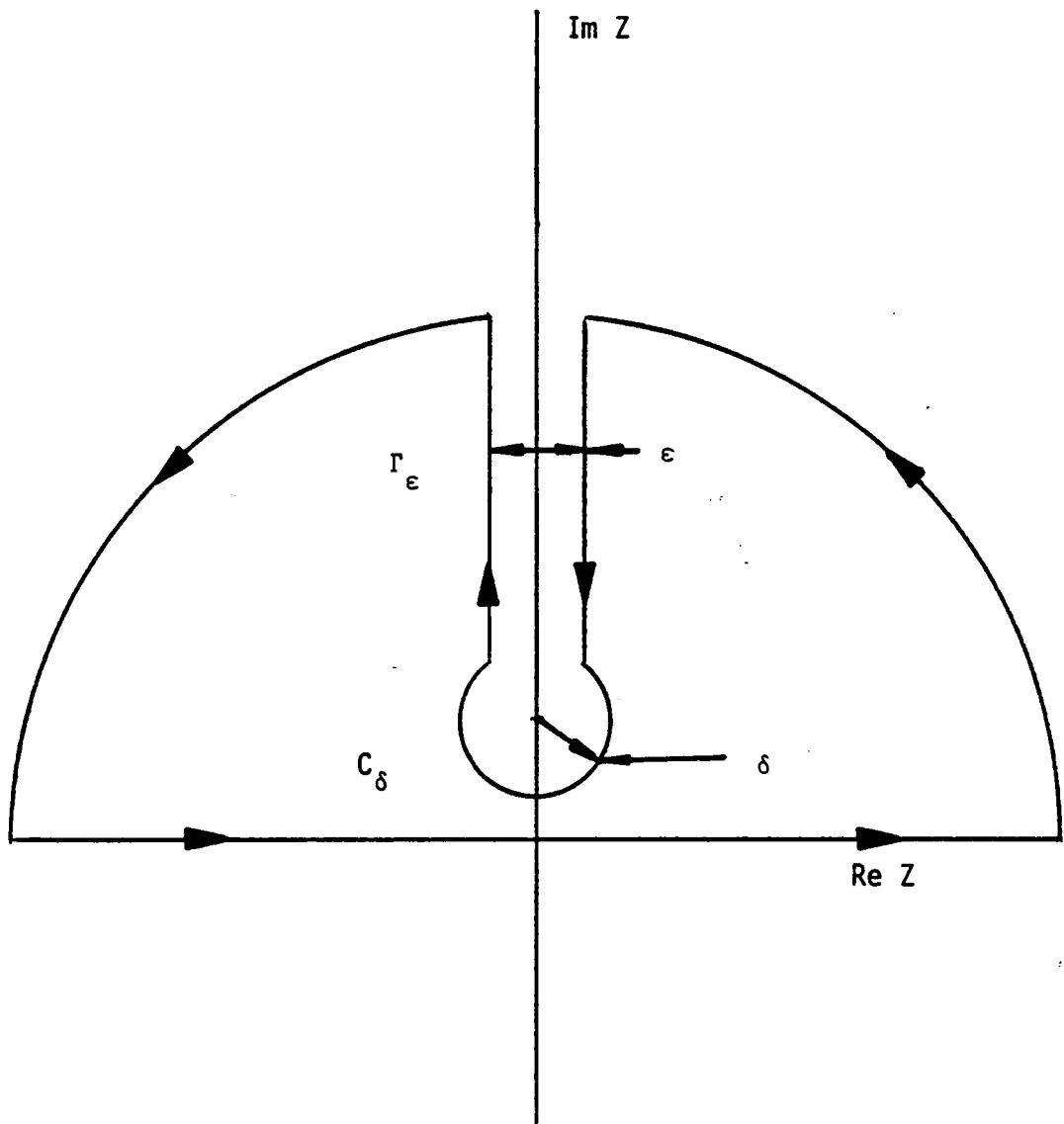


Fig. 1. Contour for computing $\Delta \text{Arg det } W$.

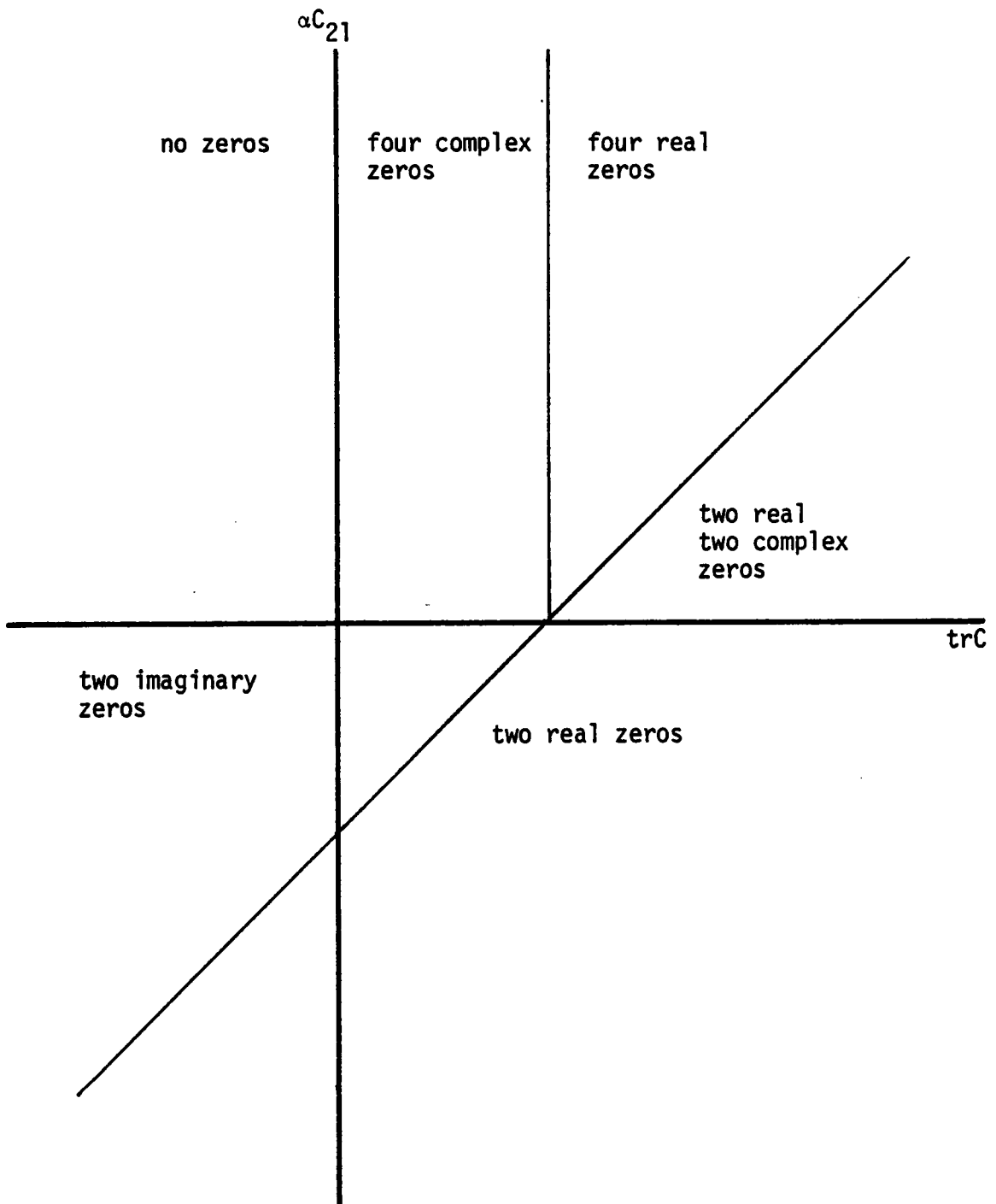


Fig. 2 The zeros of $\det W$ in the $\text{tr}C$, αC_{21} plane.

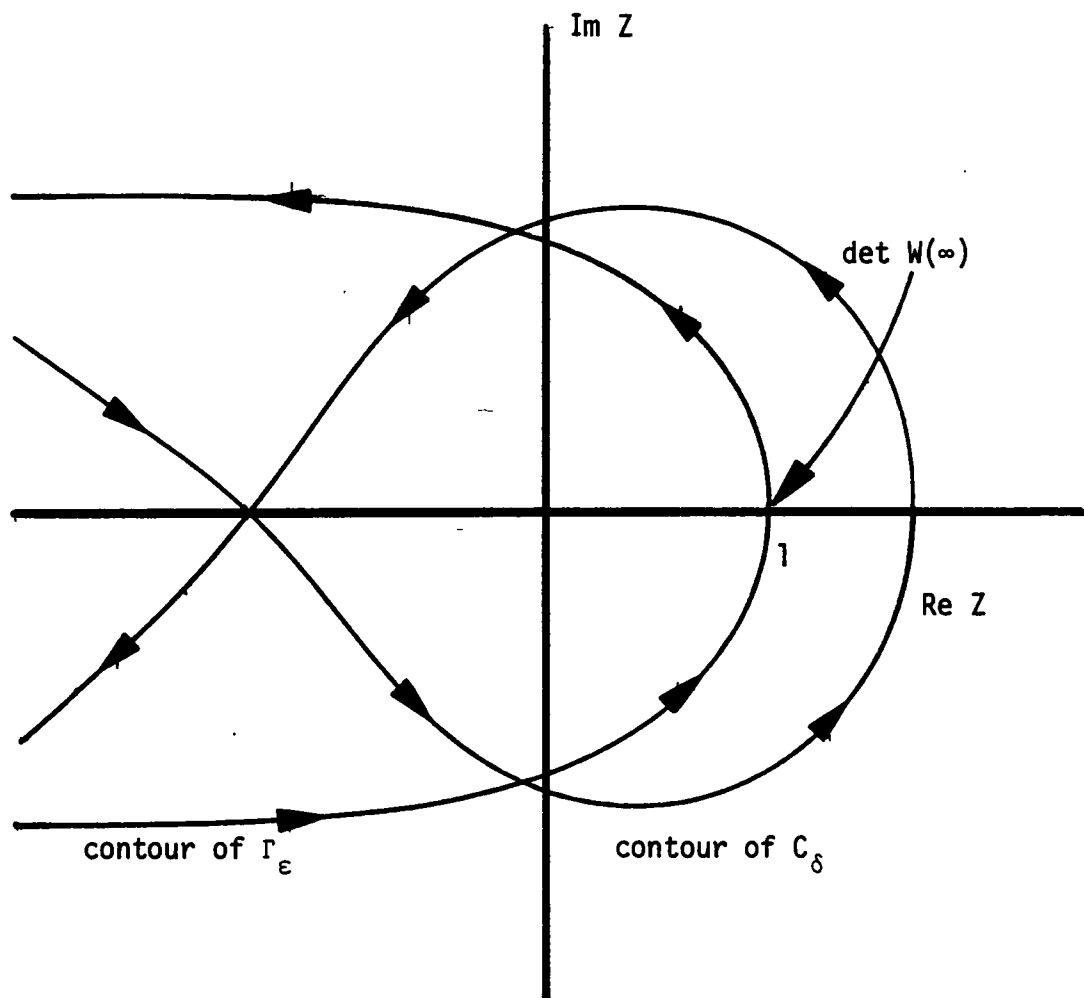


Fig. 3 Nyquist diagram for $\alpha C_{21} > 0$ and $\text{tr} C > 0$.

7. CONSTRUCTION OF THE WIENER-HOPF FACTORIZATION

The matrix valued function to be factored is

$$W(\lambda) = (\lambda^{-1} \tan^{-1} \lambda)C + (1+\lambda^2)^{-1}MC ; \quad (51)$$

the matrix M has been defined in Eq. (45). In general it is not known how to construct the Wiener-Hopf factorization of matrices, but Larsen and Zweifel,²⁶ have shown how to factor any upper triangular matrix. More generally, it is also possible to construct the factorization if the matrix is similar to an upper triangular matrix. Of course, every matrix is similar to an upper triangular matrix, but the similarity transformation must not involve the independent variable if this procedure is to work. That is, the similarity transformation must be constant. The matrix (51) can be upper triangularized by a similarity transformation with constant elements. One possible transformation is given by;

$$S = \begin{vmatrix} -c_{22} & c_{21} \\ c_{21} & 0 \end{vmatrix} \begin{vmatrix} 1 & A \\ 0 & 1 \end{vmatrix}, \quad (52a)$$

and A is

$$A = c_{21}(\text{tr } C)^{-1} \quad \text{if } \text{tr } C \neq 0$$

$$A = 0 \quad \text{if } \text{tr } C = 0. \quad (52b)$$

The matrix S is always invertible because $\det S = -c_{21}^2$ which is nonvanishing by assumption. (Here again we see that c_{21} has a distinguished role.) The choice for the transformation S is not unique; if T is any invertible triangular matrix, then TS will also upper triangularize the symbol. The particular choice [Eq. (52a)] has been made with forethought so that the transformed matrices MC and C are especially simple. Explicitly the transformed matrices are;

$$S^{-1}(I + M)CS = \begin{vmatrix} 0 & 0 \\ 0 & \alpha c_{21} \end{vmatrix}, \quad \text{tr } C = 0, \quad (53a)$$

and for $\text{tr } C \neq 0$

$$S^{-1}(I + M)CS = \begin{vmatrix} 0 & -\alpha c_{21}^2(\text{tr } C)^{-1} \\ 0 & \text{tr } C + \alpha c_{21} \end{vmatrix}. \quad (53b)$$

The transformed matrix $S^{-1}CS$ is given by the same expression but with $\alpha=0$. It is tempting to think that the similarity transformation [Eq. (52a)] applied to the original equation

will result in a similar simplification, but this is not the case. The reason is that although C and MC are simultaneously upper triangularizable, C and Σ are not. Recall that a sufficient condition for two matrices to be simultaneously upper triangularizable is that they commute. This condition is not necessary; an easy calculation shows that the matrices C and MC provide a counterexample.

The Wiener-Hopf factorization can now be computed. If $S^{-1}WS$ is denoted by \tilde{W} , then

$$\tilde{W} = \begin{vmatrix} 1 & K(\lambda) \\ 0 & \det W(\lambda) \end{vmatrix}, \quad (54a)$$

where

$$K(\lambda) = -c_{21} \lambda^{-1} \tan^{-1} \lambda, \quad \text{tr } C = 0, \quad (54b)$$

$$K(\lambda) = -\alpha c_{21}^2 (\text{tr } C)^{-1} (1 + \lambda^2)^{-1}, \quad \text{tr } C \neq 0. \quad (54c)$$

The function \tilde{W} is an upper triangular of second order and the procedure for getting its Wiener-Hopf factorization when it exists has been developed by Cebotarev.²⁷ Here we follow the method of Ref. 26. First we note that the factors of an upper triangular matrix can be taken to be upper triangular, so we set

$$\tilde{W} = X(\lambda)Y(\lambda) \quad (55)$$

with X (Y) analytic and invertible in the lower (upper) half plane. If the elements of the matrices X and Y are denoted by X_{ij} and Y_{ij} respectively, then the following system of equations result when Eq. (55) is substituted into Eq. (51), and corresponding matrix elements are equated,

$$1 = X_{11}Y_{11} \quad (56a)$$

$$1 - (\text{tr } C)\lambda^{-1}\tan^{-1}\lambda + \alpha c_{21}(1+\lambda^2)^{-1} = X_{22}(\lambda)Y_{22}(\lambda) \quad (56b)$$

and,

$$\begin{aligned} &-(c_{21}-A\text{tr } C)\lambda^{-1}\tan^{-1}\lambda - \sigma c_{21}A(1+\lambda^2)^{-1} \\ &= X_{11}(\lambda)Y_{12}(\lambda) + X_{12}Y_{22}(\lambda) . \end{aligned} \quad (56c)$$

These equations do not uniquely determine X and Y , since XU and $U^{-1}Y$ satisfy Eqs. (56a)-(56c) whenever X and Y do, where U is any invertible matrix. It is consistent to impose the conditions

$$X_{ij}(\infty) = Y_{ij}(\infty) = \delta_{ij} . \quad (57)$$

With this condition, Eq. (56a) uniquely determines X_{11} , and Y_{11} to be;

$$X_{11}(\lambda) = Y_{11}(\lambda) = 1, \quad (58)$$

while the solution to Eq. (56b) is given by

$$X_{22}(\lambda) = \exp\left(\frac{1}{2\pi i} \int_{-\infty + i/2}^{+\infty + i/2} \frac{B(z)}{z-\lambda} dz\right), \quad (59a)$$

where

$$B(z) = \ln[1 - (\text{tr } C) z^{-1} \tan^{-1} z + \alpha c_{21} (1+z^2)^{-1}]. \quad (59b)$$

The expression for Y_{22} is the same except that the limits of integration are replaced by $\infty - i/2$ and $-\infty - i/2$. In deriving Eqs. (59a)-(59b), the standard formulas for the factorization of scalar functions have been used. These formulas are proven in many texts, but perhaps the clearest discussion (with many examples) is given in Roos.²² Finally, we must determine Y_{12} and X_{12} . To do this, divide Eq. (56c) by Y_{22} , and define the left-hand side of Eq. (56c) to be $L(\lambda)$. Then

$$\frac{L(\lambda)}{Y_{22}(\lambda)} = \frac{Y_{12}(\lambda)}{Y_{22}(\lambda)} + X_{12}(\lambda). \quad (60)$$

The left-hand-side of this equation is known, while the right-hand-side is the sum of two functions, one analytic in the upper half plane, the other in the lower half plane. To solve for Y_{12} it is only necessary to write LY_{22}^{-1} as the sum of two functions:

$$\frac{L(\lambda)}{Y_{22}(\lambda)} = \frac{Y_{12}(\lambda)}{Y_{22}(\lambda)} + X_{12} = L^+(\lambda) + L^-(\lambda) . \quad (61)$$

with L^+ (L^-) analytic in the upper (lower) half plane. Therefore

$$L^+(\lambda) = \frac{1}{2\pi i} \int_{-\infty - i/2}^{+\infty - i/2} \frac{L(z)/Y_{22}(z)}{z-\lambda} dz \quad (62a)$$

$$L^-(\lambda) = \frac{1}{2\pi i} \int_{-\infty + i/2}^{+\infty + i/2} \frac{L(z)/Y_{22}(z)}{z-\lambda} dz . \quad (62b)$$

Now with the definitions

$$Y_{12}(\lambda) = Y_{22}(\lambda) L^+(\lambda) \quad (63a)$$

$$X_{12}(\lambda) = L^-(\lambda) ; \quad (63b)$$

the matrices X and Y have all the properties required of a factorization.

8. THE EXIT DISTRIBUTION

Once the canonical Wiener-Hopf factorization has been computed, an expression for the exit distribution, that is $F(0, \mu)$ for $\mu < 0$, can be written in terms of the factors of $W(\lambda)$. Unlike the one speed case, the exit distribution will involve derivatives of the factors of $W(1/i\lambda)$. The method followed in this section parallels the one given by van der Meer.²⁸ First, the exit distribution for the two speed problem defined by Eq. (39) will be derived; then the formulas will be generalized to the N-group problem.

Following Gohberg and Krein,²⁹ there exists a resolvent kernel $\gamma(.,.)$ so that the general solution to the Wiener-Hopf equation

$$G(x) = \int_0^{\infty} K(x-y) G(y) dy + U(x) \quad (64a)$$

can be written as

$$G(x) = U(x) + \int_0^{\infty} \gamma(x,y) U(y) dy, \quad (64b)$$

and the general solution to the transposed equation

$$G(x) = \int_0^{\infty} G(y) K(y-x) dx + U(x) \quad (65a)$$

can be written as

$$G(x) = U(x) + \int_0^{\infty} U(y) \gamma(y, x) dy. \quad (65a)$$

Note that the resolvent kernels for Eq. (64a) and Eq. (65a) are identical. Returning to Eq. (24a), the exit distribution can be written in terms of G by the formula

$$F(0, \mu) = \frac{1}{2\mu} \int_0^{\infty} e^{y\Sigma/\mu} C G(y) dy, \quad \mu < 0. \quad (66)$$

Introducing the resolvent kernel $\gamma(.,.)$ this can be rewritten as

$$F(0, \mu) = -\frac{1}{2\mu} \int_0^{\infty} \int_0^{\infty} e^{y\Sigma/\mu} C [\delta(y-z) + \gamma(y, z)] U(z) dz dy. \quad (67)$$

If the expression for $U(z)$ in terms of the incident flux is used in Eq. (67), then

$$F(0, \mu) = -\frac{1}{2\mu} \int_0^{\infty} \int_0^{\infty} \int_0^1 e^{y\Sigma/\mu} C [\delta(y-z) + \gamma(y, z)] e^{-z\Sigma/\mu} \phi(s) ds dz dy. \quad (68)$$

This equation relates the exit distribution to the incident distribution by making use of the resolvent kernel. To write Eq. (68) in terms of the factors of W , it is necessary

to write,

$$\int_0^{\infty} \int_0^{\infty} e^{y\Sigma/\mu} C[\delta(y-z) + \gamma(y,z)] e^{-z\Sigma/s} dz dy, \quad (69)$$

in terms of the factors of W . This will be accomplished in two parts. First

Lemma 1:

$$\begin{aligned} & \int_0^{\infty} \int_0^{\infty} e^{y\Sigma/\mu} C[\delta(y-z) + \gamma(y,z)] e^{-z\Sigma/s} dz dy \\ &= H_1(-\mu) \left[\frac{s\mu}{\mu-s} H_r(s) - s \left(\frac{\mu}{\mu-s} \right)^2 (H_r(s) + (\mu-s)H_r'(s)) M \right], \end{aligned} \quad (70)$$

where,

$$W^{-1}(1/i\mu) = H_1(-\mu)H_r(\mu)$$

is a canonical factorization with H_1 and H_r analytic in the open right half plane and continuous and invertible in the closed right half plane.

Proof: Let $G(x;s)$ be a solution to the matrix Wiener-Hopf equation:

$$G(x;s) = \int_0^{\infty} K(|x-y|) G(y;s) dy + e^{-x\Sigma/s}. \quad (71)$$

In this equation the variable s is considered to be a parameter. Note that the left-hand-side of Eq. (70) is

$$\int_0^{\infty} e^{y/\mu} G(y;s) dy = \hat{G}^+(\mu;s) . \quad (72)$$

If Eq. (71) is extended to the entire real line in the usual way and the Laplace transform is defined by

$$\tilde{G}(\lambda) = \int_{-\infty}^{\infty} dx e^{x/\lambda} G(x) , \quad \text{Re}(\lambda) = 0, \quad (73)$$

and $Z(\lambda)$ is defined by $Z(\lambda) = W(1/i\lambda)$, then the Laplace transform of the integral equation [Eq. (71)] is

$$Z(\lambda) \tilde{G}^+(\lambda) + \tilde{G}^-(\lambda) = \frac{s\lambda}{\lambda-s} I - s\left(\frac{\lambda}{\lambda-s}\right)^2 M . \quad (74)$$

The functions G^+ and G^- have already been defined by Eqs. (40a)-(40d), and the matrix M was introduced in Eq. (45a). Now assume that the factorization of $Z(\lambda)$ is given by

$$Z^{-1}(\mu) = H_l(-\mu)H_r(\mu) , \quad (75)$$

where the functions H_l and H_r are analytic and invertible on the open right half plane, and continuous and invertible on the closed right half plane. Using the factorization [see Eq. (75)], Eq. (74) may be rewritten as

$$\begin{aligned}
& H_1^{-1}(-\mu)\tilde{G}^+(\mu) + H_r(\mu)\tilde{G}^-(\mu) \\
& = H_r(\mu) \left[\frac{s\mu}{\mu-s} I - s \left(\frac{\mu}{\mu-s} \right)^2 M \right] .
\end{aligned} \tag{76}$$

If the right-hand-side of Eq. (76) can be written as the sum of two terms, one analytic and invertible in the right half plane, the other one analytic and invertible in the left half plane, then Liouville's theorem can be invoked to solve for \tilde{G}^+ and \tilde{G}^- . Due to the second order pole in Eq. (76), it is necessary to introduce derivatives of the H-functions into this splitting. By inspection, the splitting is given by the sum of

$$\begin{aligned}
& \frac{s\mu}{\mu-s} [H_r(\mu) - H(s)] \\
& - s \left(\frac{\mu}{\mu-s} \right)^2 [H_r(\mu) - H_r(s) - (\mu-s)H_r'(s)] M ,
\end{aligned} \tag{77a}$$

which is analytic in the right half plane, and the expression

$$\frac{s\mu}{\mu-s} H_r(s) - s \left(\frac{\mu}{\mu-s} \right)^2 [H_r(s) + (\mu-s)H_r'(s)] M , \tag{77b}$$

which is analytic in the left half plane. An application of Liouville's theorem then proves Lemma 1. Note when $M = 0$, Eq. (77b) reduces to the result given in Ref. 28. Using

Lemma 1 it is now possible to write Eq. (69) in terms of the H-functions. To do this it is expedient to define

$$\Gamma(\mu, s) = \text{right hand side of Eq. (70)} . \quad (78)$$

Now substitute the explicit formula for $\exp(-\gamma\Sigma/\mu)$ into Eq. (69). The result is

$$\int_0^\infty \int_0^\infty e^{x/\mu} [1 + (x/\mu)M] C \\ \times [\delta(x-z) + \gamma(x, z)] e^{-z\Sigma/\mu} dz dx . \quad (79)$$

The contribution due to the term $e^{x/\mu}$ gives $C\Gamma(\mu, s)$, while the term $(x/\mu)e^{x/\mu}$ gives rise to derivatives of the function Γ . It is easily checked that

$$\int_0^\infty \int_0^\infty (x/\mu) e^{x/\mu} C [\delta(x-z) + \gamma(x, z)] e^{x-z\Sigma/\mu} dz dx \\ = \mu C \partial_\mu \Gamma(\mu, s) . \quad (80)$$

Therefore,

$$F(0, \mu) = -\frac{1}{2\mu} [1 - \mu \partial_\mu M] C \int_0^1 \Gamma(\mu, s) \Phi(s) ds . \quad (81)$$

It is routine to generalize the exit distribution formula [Eq. (81)] to the N-group problem. If $\Sigma = D + M$ is the Jordan decomposition of Σ with D a diagonal matrix given by $\text{diag}(\sigma_i)_{i=1}^N$, then the right-hand side of Eq. (74) is replaced by

$$\sum_{m=0}^{N-1} (-1)^m \text{diag} \left(s \left(\frac{\mu}{\mu-s} \right)^{m+1} \right)_{i=1}^N M^m. \quad (82)$$

Now it is necessary to write

$$H_r(\mu) \sum_{m=0}^{N-1} (-1)^m \text{diag} \left(s \left(\frac{\mu}{\mu-s} \right)^2 \right)_{i=1}^N M^m \quad (83)$$

as the sum of two terms, just as was done for the two group case. Note that Eq. (83) has poles at $\mu = s/\sigma_i$, which are in the right half plane. Denoting the i -th column of a matrix A by $[A]_{(i)}$ and noting that

$$\begin{aligned} [H_r(\mu) \text{diag} \left(s \left(\frac{\mu}{\mu-s} \right)^{m+1} \right)]_{(i)} \\ = [H_r(\mu)]_{(i)} s \left(\frac{\mu}{\mu-s} \right)^{m+1}, \end{aligned} \quad (84)$$

then Eq. (83) can easily be written as the sum of two terms, one analytic in the right half plane, the other one analytic in the left hand plane. This is accomplished by writing Eq. (84) as the sum of

$$[Q_n^+(\mu, s)]_{(i)} = [H_r(\mu) - \sum_{m=0}^{N-1} \frac{1}{m!} (\mu - \sigma_i/s)^m H_r^{(m)}(s/\sigma_i)]_{(i)} \times s(\frac{\mu}{\mu-s})^{m+1} \quad (85)$$

which is analytic in the right half plane, and

$$[Q_n^-(\mu, s)]_{(i)} = [\sum_{m=0}^{N-1} \frac{1}{m!} (\mu - \sigma_i/s)^m H_1^{(m)}(s/\sigma_i)]_{(i)} s(\frac{\mu}{\mu-s})^{m+1} \quad (86)$$

which is analytic in the left half plane, where $H_r^{(m)}$ and $H_1^{(m)}$ are the m -th derivatives of H_r and H_1 respectively. Therefore the generalization of Lemma 1 to the N -group problem is

$$\Gamma(\mu, s) = H_1(-\mu) \sum_{n=0}^{N-1} Q_n^-(\mu, s) M^n, \quad (87)$$

and the exit distribution $[F(0, \mu)]_{(i)}$ is given by

$$-\frac{1}{2\mu} \sum_{m=0}^{N-1} \frac{1}{m!} (\mu/\sigma_i)^m (\partial_\mu)^m \int_0^1 [M^m \Gamma(\mu/\sigma_i, s)]_{(i)} \phi(s) ds. \quad (88)$$

Not only can the exit distribution be written in terms of the factors of the symbol, but the solution for any value of x can also be written in a similar fashion. This can be

done by making use of Eq. (24b) which relates $F(x,\mu)$ to $G(x)$, and the results of this Section. First note that

$$\hat{G}^+(\mu) = \int_0^1 \Gamma(\mu,s)\phi(s) ds . \quad (89)$$

From this expression it is possible to recover the function G . Now that G is known, the solution $F(x,\mu)$ for $x < 0$ can be computed by making use of Eq. (24b).

9. INVERTIBLE SCATTERING MATRICES

In Sec. 7, a two-group transport equation with half range boundary conditions and a noninvertible scattering matrix was solved. We now turn our attention to problems with invertible scattering matrices which might arise in, for example, energy dependent neutron transport.

The symbol of the Wiener-Hopf equation [see Eq. (51)] is the same for either noninvertible or invertible scattering matrices; however, the dispersion function, that is the determinant of the symbol, gains an additional term proportional to $\det C$ if the scattering matrix is invertible. Recall that the construction of the Wiener-Hopf factorization presented in Sec. 7 was contingent on the simultaneous upper triangularization of the matrices C and MC . It is straightforward to show that C and MC are simultaneously upper triangularizable if and only if C is noninvertible. Therefore, if the scattering matrix is invertible, no general procedure is known for the construction of the factorization; for this case the best that can be done is to derive equations which the factors must satisfy and find an algorithm which solves the equations numerically. These equations are a generalization of Chandrasekhar's H equations. Pragmatically, one often resorts to solving the H equations numerically even when quadrature formulas similar to Eq. (59a) and Eqs. (62a)-(62b)

are known, since iterative methods for solving the H equations often converge rapidly, while the quadrature formulas involve integrations over unbounded sets which may present numerical difficulties.

In Sec. 10, the zeros of the dispersion function are investigated. The generalized Chandrasekhar H equations are derived in Sec. 11, and finally, a numerical method for their solution is given in Sec. 12.

10. INVERTIBILITY OF THE SYMBOL: INVERTIBLE C

If the matrix C is invertible, then the dispersion function is

$$\det W(\lambda) = 1 - \text{tr } C [\lambda^{-1} \tan^{-1} \lambda] + \det C [\lambda^{-1} \tan^{-1} \lambda]^2 + \alpha c_{21} [1 + \lambda^2]^{-1} . \quad (90)$$

For this case, a detailed study of the zeros of $\det W(\lambda)$, similar to the one given in Sec. 6, would be tedious. So for this problem, we give only a sufficient condition for invertibility. Mullikian¹⁴ has shown that a sufficient condition for invertibility of the symbol is that the spectral radius of the matrix B defined by

$$B_{ij} = \int_{-\infty}^{\infty} |K_{ij}(x)| dx \quad (91)$$

is less than one, where K is given by Eq. (20). It is easy to check that the spectral radius of B will be less than one if

$$\|C\| + \|MC\| < 1 \quad (92)$$

where $\| \cdot \|$ is the matrix norm. Not only does this condition give a sufficient condition for the invertibility

of the symbol, but Mullikian¹⁴ has shown that it also gives a sufficient condition for the unique solvability of the integral equation [Eq. (18)]. Mullikian's proof of this fact involves estimating the spectral radius of the integral operator.

11. THE GENERALIZED CHANDRASEKHAR H-EQUATIONS

Mullikian in Ref. 14 has derived nonlinear integral equations which the factors of the symbol satisfy. They are, for $\text{Im } z > 0$,

$$H_r^{-1}(z) = I + \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{H_l(t) \hat{K}(-t)}{t+z} dt \quad (92a)$$

$$H_l^{-1}(z) = I + \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{\hat{K}(t) H_r(t)}{t+z} dt, \quad (92b)$$

where \hat{K} is the Fourier transform of the kernel of the Wiener-Hopf equation, and I is the identity matrix. The matrix valued functions H_l and H_r have the following properties:

(i) H_r and H_l are analytic in the half plane $\text{Im } z > 0$, and continuous for $\text{Im } z \geq 0$.

(ii) H_r and H_l are invertible in the half plane $\text{Im } z \geq 0$

(iii) $I - \hat{K}(z) = H_l^{-1}(z) H_r^{-1}(z)$.

We now specialize Eqs. (92a)-(92b) to the two group problem defined by Eq. (39) and $\det C \neq 0$ by introducing the explicit form of \hat{K} ;

$$\hat{K}(z) = \left(\frac{1}{2} \tan^{-1} z\right) C - (1+z^2)^{-1} M C, \quad (93)$$

where M is the nilpotent matrix given by Eq. (39). Recall

that \hat{K} is analytic on $\mathbb{C} \setminus \{z \in \mathbb{C} : z=it, |t| \geq 1\}$. The integrals appearing in the H-equations [Eqs. (92a)-(92b)] can be rewritten into a more familiar form by making use of Cauchy's Theorem. The calculation for Eq. (92a) will be shown; the procedure for Eq. (92b) is essentially the same. Substitute Eq. (3) into Eq. (1a); the result is

$$H_1^{-1}(z) = I + \frac{1}{2\pi i} \int_{-\infty}^{\infty} \left[\frac{1}{t} \tan^{-1} t C - (1+t^2)^{-1} MC \right] \\ \times \frac{H_r(t)}{t+z} . \quad (94)$$

We analyze each term of the integrand separately. First, consider the term $(1+t^2)^{-1} MC$. Note that $H_r(t)/(t+z)$ is analytic in the upper half plane, and $H_r(t)/(t+z)(1+t^2)$ vanishes at infinity as t^{-2} . Therefore, Cauchy's Theorem yields

$$\frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{H_r(t)}{(1+t^2)(t+z)} = \frac{H_r(1)}{1+z} . \quad (95a)$$

Next, consider the term containing $t^{-1} \tan^{-1} t$. The contour can be completed in the upper half plane if the branch cut $[i, i\infty)$ is avoided. The result is

$$\begin{aligned}
& \int_{-\infty}^{\infty} \frac{\tan^{-1}(t)}{t} \frac{H_r(t)}{t+z} dt \\
&= \int_{i\infty}^i \frac{1}{t} [\tan^{-1}(it+0) - \tan^{-1}(it-0)] \frac{H_r(t)}{t+z} dt, \quad (95b)
\end{aligned}$$

where $\tan^{-1}(it \pm 0)$ are the boundary values of \tan^{-1} along the imaginary axis. Using,

$$\tan^{-1}(it + 0) - \tan^{-1}(it - 0) = -i\pi, \quad |t| \gg 1, \quad (96)$$

the right-hand-side of Eq. (95b) becomes

$$\int_{i\infty}^i \frac{\pi}{it} \frac{H_r(t)}{t+z} dt. \quad (97)$$

A simple change of variables gives

$$\int_{i\infty}^i \frac{\tan^{-1}(t)}{t} \frac{H_r(t)}{t+z} dt = \int_0^1 \frac{iH_r(i/t)}{t(z+i/t)} dt. \quad (98)$$

It is convenient to define functions X and Y by

$$X(z) = H_1(i/z) \text{ and,} \quad (99a)$$

$$Y(z) = H_r(i/z). \quad (99b)$$

The final form of the H equations is derived by substituting Eqs. (95a) and (98) into Eq. (92a). The result is

$$\begin{aligned}
 X^{-1}(z) = I + \frac{1}{2} C \int_0^1 d\mu \frac{z}{z+\mu} Y(\mu) \\
 + \frac{1}{2} MC \frac{z}{z+1} Y(1) . \quad (100a)
 \end{aligned}$$

Simialrly, the H-equation corresponding to Eq. (92b) is

$$\begin{aligned}
 Y^{-1}(z) = I + \frac{1}{2} \int_0^1 d\mu \frac{z}{z+\mu} X(\mu) C \\
 + \frac{1}{2} \frac{z}{z+1} X(1) MC . (100b)
 \end{aligned}$$

Note that if $M = 0$, then Eqs. (100a) and (100b) reduce to the Chandrasekhar H-equations.²³ It is straightfoward to show that every solution of Eqs. (100a)-(100b) provides a factorization of the symbol. A Banach space analysis of these equations is the next topic.

12. A BANACH SPACE ANALYSIS OF THE H EQUATIONS

In Ref. 30, Bowden and Zweifel presented a Banach space analysis of the Chandrasekhar H equations. They showed that the H equations could be solved by iteration. Of course the H equations must be supplemented by constraints which are the analyticity requirements given by conditions (i) and (ii) of Sec. 11. In Ref. 31 Bowden, Menikoff, and Zweifel generalized their results to H equations relevant to multigroup problems. In this Section a similar analysis is given for Eqs. (100a)-(100b). In Refs. 30 and 31, it was shown that the H equations have solutions in the Banach space $L_1(0,1)$ by making use of a contracting mapping principle. For the H equations [Eq. (100a)-(100b)] the terms $X(1)$ and $Y(1)$ prevent one from using a contracting principle in $L_1(0,1)$, because point evaluation is an unbounded operator in $L_1(0,1)$. Thus, instead of $L_1(0,1)$, we must use a Banach space with a supremum norm. Consequently the following analysis parallels the one given by Rall.³²

First, we define some Banach spaces: Let X_0 be the vector space $X_0 = \{ T : T \text{ is a } 2 \times 2 \text{ matrix valued function, with } T_{ij} \text{ continuous on } [0,1] \}$ and define a norm

$$\| U \|_{X_0} = \| \| U \|_m(\cdot) \| \quad (101a)$$

where $\| U \|_m(\cdot)$ is defined for each $s \in [0,1]$ by

$$\|U\|_m(s) = \sup_{\|x\|=1} \|U(s)x\|_2 \quad (101b)$$

where $x \in \mathbb{R}^2$, and $\|\cdot\|_2$ is the usual Euclidian norm on \mathbb{R}^2 . That is, for each fixed s , $\|U\|_m$ is the operator norm of U when U is viewed as an operator on $\mathbb{R}^2 \rightarrow \mathbb{R}^2$. And define another Banach space by

$$X = X_0 \oplus X_0 \quad (102)$$

with norm

$$\|T\|_X = \|T_1\|_{X_0} + \|T_2\|_{X_0} \quad (103)$$

where T_1 and T_2 are the components of T .

We now rewrite the H equations [Eqs. (100a)-(100b)] in a form more suitable for analysis: Postmultiply Eq. (100a) by $X(z)$, and premultiply Eq. (100b) by $Y(z)$. The results are:

$$X(z) = I - \frac{1}{z} \int_0^1 d\mu \frac{z}{z+\mu} CY(\mu)X(z) - \frac{z}{z+1} MCY(1)X(z) \quad (104a)$$

$$Y(z) = I - \frac{1}{z} \int_0^1 d\mu \frac{z}{z+\mu} Y(z)X(\mu)C - \frac{z}{z+1} Y(z)X(1)MC. \quad (104b)$$

Define $\tilde{X} \in X$ to be the ordered pair (X, Y) ; then Eqs. (104a)-(104b) can be rewritten in operator form as

$$\tilde{X} = F\tilde{X} . \quad (105)$$

We hope to show that Eq. (105) can be solved by iteration. That is, we hope to show that the sequence $\{\tilde{X}^{(m)}\}_{m=0}^{\infty}$ defined recursively by

$$\tilde{X}^{(m+1)} = F\tilde{X}^{(m)} \quad (106)$$

will converge to the solution of Eq. (105). We shall always choose $\tilde{X}^{(0)} = I$. First we determine if F is a contraction. Let $\tilde{X} = (X_1, X_2)$ and $\tilde{Y} = (Y_1, Y_2)$. Then

$$F\tilde{X}(z) - F\tilde{Y}(z) =$$

$$\left(\frac{1}{2} \int_0^1 d\mu \frac{z}{z+\mu} C \left[X_2(\mu)X_1(z) - Y_2(\mu)Y_1(z) \right] \right.$$

$$+ \frac{1}{2} MC \frac{z}{z+1} \left[X_2(1)X_1(z) - Y_2(1)Y_1(z) \right] ,$$

$$\frac{1}{2} \int_0^1 d\mu \frac{z}{z+\mu} \left[X_2(z)X_1(\mu) - Y_2(z)Y_1(\mu) \right] C$$

$$+ \frac{1}{2} C \frac{z}{z+1} M \left[X_1(1)X_2(z) - Y_1(1)Y_2(z) \right] \quad (107)$$

The first component can be rewritten as

$$\frac{1}{4} C \int_0^1 d\mu \frac{z}{z+\mu} \left[X_2(\mu) - Y_2(\mu) \right] \left[X_1(z) + Y_1(z) \right]$$

$$\frac{1}{4} C \int_0^1 d\mu \frac{z}{z+\mu} \left[X_2(\mu) + Y_2(\mu) \right] \left[X_1(z) - Y_1(z) \right]$$

$$+ \frac{1}{4} MC \frac{z}{z+1} \left[X_2(1) - Y_1(1) \right] \left[X_1(z) + Y_1(z) \right]$$

$$\frac{1}{4} MC \frac{z}{z+1} \left[X_2(1) + Y_2(1) \right] \left[X_1(z) - Y_1(z) \right]$$

and the second component can be rewritten in a similar form.

Let $B(I; r)$ be the ball $\{T \in X : \|I - T\| \leq r\}$. If

\tilde{X} and $\tilde{Y} \in X$ then

$$\| \tilde{X} + \tilde{Y} \|_X \leq 2(1+r) . \quad (106)$$

Using Eq. (106) and the inequality

$$\left| \int_0^1 d\mu \frac{z}{z+\mu} \right| \leq \ln 2 , \text{ for } z \in [0,1] , \quad (109)$$

it is possible to estimate the norm of $F\tilde{X} - F\tilde{Y}$ by

$$\|F\tilde{X} - F\tilde{Y}\| \leq (1+r) [1n2 \|C\| + \|MC\|] \|\tilde{X} - \tilde{Y}\|. \quad (110)$$

It is convenient to define

$$a = 1n2 \|C\| + \|MC\|. \quad (111)$$

Therefore, F is a contraction operator provided

$$a(1+r) < 1. \quad (112)$$

So F is a contraction operator for sufficiently small a and r . We must now show that r can be chosen large enough so that $F: B \rightarrow B$. For F to map B into B it is sufficient to require

$$r \geq \frac{\|F\tilde{X}^{(0)} - \tilde{X}^{(0)}\|}{1 - a(1+r)}. \quad (113)$$

It is straightforward to estimate $\|F\tilde{X}^{(0)} - \tilde{X}^{(0)}\|$ by

$$\|F\tilde{X}^{(0)} - \tilde{X}^{(0)}\| \leq a. \quad (114)$$

Therefore, inequality (15) can be rewritten as

$$ar^2 + (a-1)r + a \leq 0. \quad (115)$$

The largest value of a which is consistent with inequality (113) is

$$a = \frac{1}{3} \quad (116)$$

For this value of a , r can be determined to be

$$r = 1. \quad (117)$$

All that remains to be checked is inequality (112). With $a = 1/3$ and $r = 1$, we have

$$a(1+r) = 2/3 < 1. \quad (118)$$

Therefore, if

$$\ln 2 \|C\| + \|MC\| < 1/3 \quad (119)$$

then Eq. (105) has a unique solution inside the ball $B(I;1)$. Furthermore, the sequence defined by Eq. (106) converges to this solution if $\tilde{X}^{(0)} \in B(I;1)$.

Finally, we must check that $\tilde{X}^{(m)}$ converges to a solution with the required analyticity and invertibility properties. First,

Lemma 2: If $\tilde{X} = (X, Y)$ is the unique solution to Eq. (105) in the ball $B(I; 1)$, and $a < 1/3$, then

$$\det \left[I + \frac{1}{2} C \int_0^1 d\mu \frac{z}{z+\mu} Y(\mu) + \frac{z}{z+1} M C Y(1) \right] \neq 0$$

for $\operatorname{Re} z \geq 0$.

Proof: Since $(X, Y) \in B(I; 1)$ we have

$$1 \geq \|I - X\| \geq |1 - \|X\||. \quad (120)$$

Thus we have

$$\|X\| \leq 2. \quad (121)$$

Suppose that for some value of z

$$\det \left[I + \frac{1}{2} C \int_0^1 d\mu \frac{z}{z+\mu} Y(\mu) + \frac{z}{z+1} M C Y(1) \right] = 0. \quad (122)$$

This implies that

$$\left\| \frac{1}{2} C \int_0^1 d\mu \frac{z}{z+\mu} Y(\mu) + \frac{z}{z+1} M C Y(1) \right\| \geq 1. \quad (123)$$

Or,

$$\|C\| \int_0^1 \left| \frac{z}{z+\mu} \right| d\mu + \|MC\| \geq 1. \quad (124)$$

But if $\operatorname{Re} z \geq 0$, then

$$\int_0^1 \left| \frac{z}{z+\mu} \right| d\mu \leq 1. \quad (125)$$

So inequality (124) implies

$$\|C\| + \|MC\| \geq 1. \quad (126)$$

But by assumption

$$a = \ln 2 \|C\| + \|MC\| < 1/3.$$

which is a contradiction. Therefore, $\hat{X}(z)$ defined by

$$\hat{X}(z) = I + \frac{1}{z}C \int_0^1 d\mu \frac{z}{z+\mu} Y(\mu) + \frac{z}{z+1}MCY(1) \quad (127)$$

is invertible for $\operatorname{Re} z \geq 0$. The same proof shows that

$$\hat{Y}(z) = I + \frac{1}{z} \int_0^1 d\mu \frac{z}{z+\mu} X(\mu)C + \frac{z}{z+1} X(1)MC$$

is also invertible for $\text{Re } z \geq 0$. Furthermore, \hat{X} and \hat{Y} are analytic on $\mathbb{C}/[-1,0]$. Finally observe that \hat{X} and \hat{Y} satisfy the H equations [Eqs. (100a)-(100b)]. Thus the iteration scheme defined by Eq. (105) converges to the solution of the H equations, and satisfies the constraints.

REFERENCES

¹W.M. Stacey, Modal Approximations, (M.I.T. Press, Cambridge, Mass., 1967).

²J.J. Duderstadt and W.R. Martin, Transport Theory, (John Willey & Sons, New York, 1979), p. 408.

³W.F. Ames, Nonlinear Ordinary Differential Equations in Transport Processes, (Academic Press, New York, 1968), pp. 184-196.

⁴W.F. Ames, Nonlinear Partial Differential Equations in Engineering, (Academic Press, New York, 1972), vol. 2, pp. 146-180.

⁵C.E. Siewert and P.F. Zweifel, "An Exact Solution of Equations of Radiative Transfer for Local Thermodynamic Equilibrium in the Non-Gray Case. Picket Fence Approximation," *Ann. Phys. (N.Y.)* 36, 61 (1966).

⁶C.E. Siewert and P.F. Zweifel, "Radiative Transfer II," *J. Math. Phys.* 7 (11), 2092 (1966).

⁷M.M.R. Williams, "The Wiener-Hopf Technique: An Alternative to the Singular Eigenfunction Method," in Advances in Nuclear Science and Technology, edited by E.J. Henley and J. Lewins (Academic Press, New York, 1973), p. 283.

⁸E.E. Burniston, T.W. Mullikan, and C.E. Siewert, "Steady State Solutions in Two-Group Theory of Neutron

Diffusion," J. Math. Phys. 13 (10), 1461 (1972).

⁹C.E. Siewert (private communication).

¹⁰K.M. Case and P.F. Zweifel, Linear Transport Theory, (Addison-Wesley, Reading, Mass. 1967).

¹¹N. Wiener and E. Hopf, "Über eine Klasse singularer Integralgleichungen," Sitzber. Deutsch Akad. 1931, 696.

¹²N.I. Muskhelishvili, Singular Integral Equations, (Noordhoff, Groningen, Holland 1953).

¹³Gakhov, Boundary Value Problems, (Addison-Wesley, Reading, Mass., 1963).

¹⁴T.W. Mullikian, "Wiener-Hopf Factorizations," Transport Theory Stat. Phys. 3 (4), 215 (1973).

¹⁵S.N. Purohit, "Application of Polynomial Method in the Study of Thermal Neutron Spectrum in a Heavy Crystalline Medium," American Nuc. Soc. 5, 46 (1962).

¹⁶S. Sancaktar, Ph.D. thesis, Virginia Polytechnic Institute and State University, 1975.

¹⁷N. Dunford and J.T. Schwartz, Linear Operators, (Wiley, New York, 1971), Part III.

¹⁸M. Reed and B. Simon, Functional Analysis, (Academic Press, New York, 1975), vol. 1, p. 203.

¹⁹N.A. Krall and A.W. Trivelpiece, Principles of Plasma Physics, (McGraw Hill, New York, 1973), pp. 432-437.

²⁰B.L. Willis, C.V.M. van der Mee, and P.F. Zweifel, "Whither Existence Theory?," to be published in Transport

Theory Stat. Phys.

²¹I.C. Gohberg (private communication).

²²B.W. Roos, Analytic Functions and Distributions in Physics and Engineering, (John Wiley & Sons, New York, 1969).

²³S. Chandrasekhar, Radiative Transfer, (Dover, New York, 1960).

²⁴M.G. Krein, "Integral Equations on a Half Line with Kernel Depending upon the Difference of the Arguments," American Math. Soc. Translations 22, Series 2, 163 (1962).

²⁵E.C. Titchmarsh, The Theory of Functions, 2nd ed. (Oxford University Press, London 1939).

²⁶E.W. Larsen and P.F. Zweifel, "Explicit Wiener-Hopf Factorizations," SIAM J. Appl. Math. 30 (4), 732 (1976).

²⁷C.N. Cebotarev, "Partial Indices for the Riemann Boundary Problem for a Matrix of Second Order," Uspehi Mat. Nauk 11, 144 (1956).

²⁸C.V.M. van der Mee, "Albedo Operators and H-Equations for Generalized Kinetic Models," Transport Theory Stat. Phys. 13, 341 (1984).

²⁹I.C. Gohberg and M.G. Krein, "System of Integral Equations on a Half Line with Kernel Depending on the Difference of Arguments," American Math. Soc. Translations 14, Series 2, 217 (1960).

³⁰R.L. Bowden and P.F. Zweifel, "A Banach space

analysis of the Chandrasekhar H-equation," *Astrophys. J.* 210, 178 (1976).

³¹R.L. Bowden, P.F. Zweifel, and R. Menikoff, "On the solution of nonlinear matrix integral equations in transport theory," *J. Math. Phys.* 17 (9), 1722 (1976).

³²L.B. Rall, Computational Solution of Nonlinear Operator Equations, (Wiley, New York, 1969), pp. 74-85.

**The two page vita has been
removed from the scanned
document. Page 1 of 2**

**The two page vita has been
removed from the scanned
document. Page 2 of 2**