REFLECTION AND TRANSMISSION COEFFICIENTS IN MULTI-WAVE INHOMOGENEOUS MEDIA

by

JOHN LESLIE DOANE

B. E., Yale University
(1964)
S. M., Massachusetts Institute of Technology
(1965)
E. E., Massachusetts Institute of Technology
(1966)

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
June, 1970

Signature of Author: ______________________
Department of Electrical Engineering, May 27, 1970

Certified by: ______________________
Thesis Supervisor

Accepted by: ______________________
Chairman, Departmental Committee on Graduate Students

REFLECTION AND TRANSMISSION COEFFICIENTS
IN MULTI-WAVE INHOMOGENEOUS MEDIA

by

JOHN L. DOANE

Submitted to the Department of Electrical Engineering on May 27, 1970, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

ABSTRACT

The propagation of single-frequency waves in media with continuously-varying parameters is studied theoretically, with particular application to wave conversions in microwave magnetoelastic delay lines. The theory involves analytic solutions of linear ordinary differential equations with variable coefficients.

A procedure previously developed is extended to provide a systematic method of deriving WKB quasi-normal mode solutions for systems describable by coupled first-order, coupled second-order, or single n^{th} order differential equations. Previous WKB solutions for the fourth-order system describing "magnetostatic"-to-"exchange" spin wave conversion at a turning point are shown to be in error by explicit calculation of the correct WKB solutions. A fourth-order differential equation has been solved previously by the WKB method and applied to various cases in the plasma literature to match the WKB solutions on either side of the turning point. The same equation is solved here by the WKB method in a manner particularly suited for wave conversion problems. By explicit calculation of the first error terms in the asymptotic expansions, it is shown that certain reflections have been overlooked in the earlier solutions.

A systematic procedure is developed for finding exact solutions of certain differential equations of arbitrary order describing reflection and transmission of waves in media with a bounded region of inhomogeneity. This procedure involves solution by generalized hypergeometric functions. It is applied to find reflection coefficients for spin wave conversion at a fourth-order turning point near which the static magnetic field varies monotonically. Solution for a variation with a single valley is also outlined. A "source equation" method is described which allows a broader class of differential equations to be solved exactly using the above procedure. This method is applied to a fourth-order equation to find the reflections and transmission efficiencies at a magnetoelastic crossover point. Comparison with previous solutions shows that successive approximations schemes are not reliable for calculating reflections. Coupled second-order equations are derived and combined into one sixth-order equation simultaneously describing interactions between "magnetostatic," "exchange," and elastic waves in the presence of both a turning point and a crossover point. An expression for the overall conversion efficiency is found in terms of sines of complex arguments.

Hypergeometric functions can also be used to solve certain coupled mode problems in which reflections are neglected. Here they are used to find analytic expressions for conversion efficiencies at a magneto-elastic coupling point. Previous solutions required successive approximations or numerical integration. Solutions for the reflections on certain nonuniform transmission lines are also described.

Thesis Supervisor: Frederic R. Morgenthaler

Title: Professor of Electrical Engineering

ACKNOWLEDGMENTS

The author wishes to express his gratitude to Professor F. R. Morgenthaler for his work in supervising this research and suggesting the topic of investigation. Particular thanks are in order for the timeliness of his suggestions and advice, and for his encouragement and interest in the author. He would also like to thank Professors H. A. Haus and D. J. Epstein for their suggestions as thesis readers.

It is a pleasure to acknowledge my colleagues A. Platzker and H. L. Hu, for helpful discussions.

Mrs. D. Radcliffe is to be commended for her patience, diligence and skill in typing this manuscript.

A note of particular thanks is reserved for the Fannie and John Hertz Foundation of Los Angeles, California, for their patience in providing financial support for my graduate education at M. I. T.

To my wife Daryl Ann goes gratitude for her constant encouragement, optimism, devotion, and suggestions.

TABLE OF CONTENTS

Page

Chapter

APPENDICES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Survey of this work.

(a) Motivation and summary of major results.

This work arose from a desire to understand how certain quantities such as loss and field gradients affect the various wave-excitation processes which occur in magnetoelastic delay lines at microwave frequencies (Chapter 4). All of these processes require inhomogeneous static magnetic fields, such as occur in rods or slabs. These processes include the excitation of medium-wavelength "magnetostatic" modes in magnetized yttrium iron garnet by relatively long wavelength electromagnetic radiation from a fine wire antenna, the conversion from such "magnetostatic" spin waves to short-wavelength exchange-dominated spin waves and subsequently to elastic waves. The inverses of these processes also occur, in reverse order (see Fig. 3). Magnetoelastic delay lines are of interest because the overall group delay of a pulse from an antenna to an elastic wave and back can be changed by varying the applied magnetic field. As with elastic delay lines, the basic delay is relatively long due to the slow speed of sound and spin waves as compared to electromagnetic waves.

Particular interest at the start of this work was centered on the conversion from medium wavelength waves to the short wavelength spin waves. Recent evidence had indicated that these medium wavelength "magnetostatic waves" acted as an intermediary between the electromagnetic waves and the exchange-dominated spin waves (part 4.3(a)1). Furthermore, VASILE and LAROSA 1968a have derived a fourth-order differential equation which could model this conversion process. (See parts 4.1(b) and 4.3(b) for a differential derivation, valid for cylindrical symmetry as well as for slabs.) However, they indicated that the conversion would always be essentially complete in the cases of interest for practical devices. There was no suggestion of how losses and the steepness of the static magnetic field gradient inside the delay line might affect this medium-to-short wavelength conversion. Certainly one would expect an incident wave to generate reflected waves of the same type if the gradient is too steep. The presence of such reflected waves had in fact been indicated by recent experimental evidence (see part 4.3(a)1). Losses could also be expected to influence the conversion and perhaps generate such reflections because losses cause a splitting in the dispersion relation similar to the splitting caused by, for example, magnetoelastic coupling (see Figs. 4 and 9).

The analysis of conversion by VASILE and LAROSA 1968a was discovered in the course of this work to be invalid due to an incorrect construction of the WKB quasi-normal mode approximate solutions of the fourth-order differential equation. Hence a systematic procedure was developed for deriving such solutions in general. The approach extends the results of KELLER and KELLER 1962 by showing how the "WKB amplitudes" may be calculated explicitly, regardless of whether the problem is formulated in terms of coupled first-order, coupled second-order, or single n^{th} order differential equations. (See part 2.2(a) in connection with Appendices 1 and 2.) The WKB solutions break down strongly, however, near the "turning point" where the group velocities of the two types of spin waves vanish and the power in each wave is divided equally between the two types (see 4.3(b)2 and 4.3(b)3).

Hence a successive approximations method of calculating the conversion efficiency is not useful. Since the validity of the WKB approximate solutions is closely related to the validity of geometrical optics and ray theory, these latter approaches also break down near the turning point.

When WKB solutions break down at a critical point, it is common to apply the WKB method to match the solutions on both sides of that point, and hence find the reflection and transmission coefficients. When only two waves are strongly interacting, a phase-integral method is sometimes used, based on solutions of the second-order Airy equation (see parts 2.4(b)2 and 4.3(c)). Usually, however, asymptotic solutions of the complete n^{th} order differential equation are necessary (part 2.4(a)). Since such solutions of a fourth-order equation had been previously applied to similar situations in the plasma literature (see part 2.4(d)), they were applied here to study the spin-wave conversion at the turning point (part 4.3(d)). Again, however, complete conversion was predicted. Such a conclusion is invalid in this case because the error terms in the asymptotic expansions indicate the presence of other reflected waves. A warning that results of the WKB method may be invalid in certain cases for this reason was given by HEADING 1967b and 1968. Therefore, the leading error terms were calculated explicitly in the present work (see 4.3(d)). The basic mathematical solutions of the relevant fourth-order equation by WASOW 1950 and RABENSTEIN 1958 included only estimates of the order of these errors. Furthermore, their treatments were not well suited for wave-conversion studies.

In an effort to find better solutions than the WKB method could provide, it was discovered that certain fourth- and even higher-order differential equations could be transformed into equations whose solutions are generalized hypergeometric functions (sections 3.3 and 3.4). In addition, the reflection coefficients can emerge from such solutions in a very natural way (see section 3.1 for a formulation of the properties of generalized hypergeometric functions particularly well adapted for this purpose). The discovery of such solutions resulted from combining the ideas of EPSTEIN 1930b and HEADING and WHIPPLE 1952. When the solutions were applied to conversion at the spin-wave turning point, it was found that reflections from an incident wave into the wave of the same type traveling in the opposite direction become significant precisely when the error terms in the WKB method become significant (see part 4.3(e)).

Later a "source equation" method was discovered which allows a broader class of differential equations to be solved using generalized hypergeometric functions (see parts 3.3(a) and 3.3(f)). This method was found upon investigation of the solution for a derivative field when the field itself can be found from a hypergeometric function solution (see 3.2(a)6). Solutions were thus obtained for the complete fourth-order equation describing reflections and transmissions at the magneto-elastic coupling point (part 4.2(c)1). Previously, only successive approximations and numerical integration solutions were available (see SCHLÖMANN and JOSEPH 1964 and KIRCHNER et al. 1966, respectively). A complete solution was desirable to understand whether discrepancies between recent experimental results and the previous theories were due to limitations of the theories or of the experiments (see 4.2(c)2). Furthermore, the present work shows how to solve the relevant second-order system, in which reflections are neglected, by using hypergeometric functions instead of numerical integration of the two first-order "coupled mode" equations (see 4.2(b)). This discovery then led to an indication of how to solve analytically certain coupled-mode problems in homogeneous media (3. 2(c)) and certain cases of nonuniform transmission lines (3. 2(b)).

The solutions of the fourth-order spin-wave turning point and magnetoelastic crossover point problems are still somewhat unsatisfactory, however. The solutions of the former problem depend on the location of of the crossover point (see the discussions in 4. 3(d) and 4. 3(e)), while the solutions of the latter clearly depend on the influence of the turning point (see the discussion in 4. 2(c)2). Therefore, it was decided to try to find a solution of the combined problem. First of all, the relevant sixth-order differential equation was derived (see 4. 4(a)). Then it was shown how generalized hypergeometric functions could again be used to find solutions. In this way, the overall conversion efficiency was expressed in terms of sums and products of hyperbolic sines (part 4. 4(b)). More work, however, is needed to interpret the result.

(b) Relation to previous work.

Much of the material in section 1. 2, Chapter 2, section 3. 1 and part 3. 2(a) is based on information which has been in various places in the literature for a fairly long time. It is included primarily as a background for the new results presented here and also as an introduction to the concepts involved. However, the sum of this material may provide a better understanding of the overall picture than can be obtained easily from the literature. See also BUDDEN 1961 for an earlier summary. Section 1. 2 of the present work gives a survey of the basic properties of the wave-coupling which can occur in multi-wave media even in homogeneous regions of space. Chapter 2 describes the origin of the new types of couplings which can occur in media with continuously-varying parameters, and outlines various methods of solution for reflection and transmission coefficients. Section 2.1 summarizes and compares these methods. Included then in section 3.1 and part 3.2(a) are descriptions of the techniques applied originally by EPSTEIN 1930b to the solution of such problems using hypergeometric functions. Reference to more recent related work occurs primarily in parts 2.2(a), 2.3, 2.4(c), 2.4(d), 3.2(d), 4.2(a)1, 4.2(c)2, and 4.3(a)1. See also the bibliography for various groupings of representative work in the literature.

For clarity, it is helpful to note here some of the important related problems which this work does not consider. First of all, by "inhomogeneous" media we mean only those whose parameters vary continuously along the direction of propagation. For a survey of other examples of "inhomogeneities," see KARBOWIAK 1967. Such examples include nonuniformities perpendicular to the propagation direction, as occur in surface wave problems and certain waveguide problems, and multiple discrete layers. Since we consider only ordinary (one-dimensional) differential equations, we can treat transverse variations only through constants resulting from the separation of variables in a partial differential equation. Thus we do not include any calculations of ray paths, for example. Similarly, we do not consider directly the coupling of energy from localized sources such as fine-wire antennas. Nonlinear effects and the propagation of pulses or transients are also beyond the scope of this work. Note, however, the work of PRICE 1965 who obtained information on pulse propagation by calculating inverse transforms of single-frequency reflection coefficients resulting from solutions of second-order equations by hypergeometric functions. Note finally that we consider no numerical methods of solution. ALTMAN and CORY 1969a and 1969b may be consulted for a summary of such methods which are applicable.

(c) Conclusions.

Several conclusions from this work can be mentioned. First of all, the discussions in part 4.2(c)2 and 4.3(b)3 show that successive approximations methods are not useful in calculating reflection coefficients. For slowly-varying media in which the WKB quasi-normal mode solutions are approximately valid, however, transmission coefficients may be obtained by such methods. Furthermore, oscillations in the reflection or transmission coefficients as a function of gradients in the medium parameters are likely to occur only when these parameters have discontinuities, or discontinuities in their gradients, etc., with respect to distance. See part 4.2(c)2 and also BREKHOVSKIKH 1960, Fig. 78. For magnetoelastic coupling, the basic conversion efficiency depends only on the local gradient of the magnetic field at the crossover point and not on the shape of the field variation elsewhere. The net conversion efficiency may depend on such "global" variations only to the extent that reflections are generated by rapid field variations or approaches to cutoff (see Fig. 7).

Secondly, the results of the WKB method are unreliable for cases where the unperturbed wavenumber of an extraordinary electromagnetic wave ("magnetostatic" wave) becomes very large. This statement applies to treatments of second-order equations as well as the corresponding complete fourth-order equations (see 2.4(d) and 4.3(d)). At finite distances from the critical point, and for finite gradients of the medium parameters, the error terms in the asymptotic expansions mask the presence of certain reflected waves. Evaluation of the reflection coefficients at finite distances is necessary because other wave-coupling processes will occur and because the differential equation solved by the WKB method is a valid approximation only in a limited region of space. The masked reflections become important in spin wave conversion at a turning point when the magnetic field gradient exceeds a critical value, which probably depends most strongly on the wavenumber at the turning point (see 4. 3(e)).

Thirdly, there need be no particular danger that nonlinear instabilities will occur at the turning point in "magnetostatic"-to-"exchange" spin-wave conversion. Note from part 4. 3(b)3 that the total power flow in each quasi-normal mode remains constant as the wave approaches the turning point as long as the mode amplitudes are approximately constant. The group velocity approaches zero which implies that the energy density may then be very large, leading to nonlinear effects. The quasi-normal mode amplitudes do change rapidly, however. Furthermore, 4. 3(21) shows that the two types of power in each wave become equal in magnitude at the turning point. As long as the dynamic magnetic field remains finite and reasonably constant, the net power flow vanishes with the same dependence on wavenumber as the group velocity. Recall from 4. 3(b)3 that the dynamic magnetic field increases rapidly only if the amplitude of the WKB quasi-normal mode remains constant, which does not happen.

Fourthly, the dispersion relation plays a fundamental role in the solutions for reflection and transmission coefficients by either the WKB method or the generalized hypergeometric equation method. In the WKB method, the dispersion relation appears in the determination of the location of saddle points in the plane of integration versus real distance (see 4. 3(d)). Also, in order to transform an equation to a generalized hypergeometric equation it must have the same form as the dispersion relation in homogeneous regions (see 3. 3(a)). The parameters of the transformed equation are then simply related to the solutions of the dispersion relation outside the inhomogeneous regions(see 3. 3(c)).

Fifthly, it is possibly to solve for the reflection and transmission coefficients from certain useful differential equations of arbitrary order (section 3. 3). Even if the results do not correspond directly to physical situations, the exact solutions can be a very valuable check on any solutions by numerical methods. Since generalized hypergeometric functions are expressed in terms of an integral resembling an inverse transform (equation 3. 1(24)), it is also at least conceivable that problems in pulse and transient propagation might be solvable directly without first finding the single-frequency reflection and transmission coefficients.

Finally, solutions of differential equations by hypergeometric functions may be applicable to other problems of coupled modes in continuously-varying media (see 3. 2(c) and 4. 2(b)). A representative sample of physical systems for which such solutions may prove useful is listed in the bibliography.

Note in addition that the solution by generalized hypergeometric functions for spin wave conversion at the turning point (4. 3(e)) indicates that reflections of incident "magnetostatic" waves are not significantly increased by the presence of loss. Such reflections were postulated by KEDZIE 1968 to result from the split in the real part of the dispersion relation due to loss (see Fig. 9).

1.2 Waves in Homogeneous Media.

(a) Wave types.

There are many interesting physical phenomena which can be described as waves. In each case there is a field component $F(\vec{r}, t)$ which in time-invariant media is usually written as

$$F(\vec{r}, t) = \mathrm{Re}[\underline{F}(\vec{r})\, e^{j\omega t}], \qquad 1.2(1)$$

where ω is the radian frequency of the wave. If waves of more than one frequency are excited, then the total solution for $F(\vec{r}, t)$ must be written as a superposition or integral (inverse Fourier transform) over solutions of the type of the right-hand side of 1.2(1). In uniform media with propagation in only one direction, the z-direction, we can go further and write

$$\underline{F}(\vec{r}) = \underline{a}\ e^{-jkz}, \qquad 1.2(2)$$

where $\underline{a}$ may either be a constant or a function of the transverse dimensions as in a waveguide, and $2\pi/k$ is the wavelength along the z-axis. Superposition of solutions of the form 1.2(2) is also necessary when more than one wave is present.

The field component $F(\vec{r}, t)$ will be the solution of some wave equation. For example, F might be the voltage or current on a transmission line, the x or y component of the electric or magnetic field in a waveguide or in the ionosphere, a strain component of an elastic wave in a solid, a velocity component of electrons in an electron beam or plasma, a component of the magnetization in a magnetic material, and so on. The wave equations for these examples result from Maxwell's curl and divergence equations, and Newton's laws for force and torque. The last three examples usually involve linearization of the equations.

When solutions of the form of 1.2(1) and 1.2(2) are substituted into the wave equations without any external driving terms or loss, we find which values of k for given values of ω represent waves which can exist without any external excitation. The solutions 1.2(2) are then called the normal modes and the relation between k and ω is called the dispersion relation. In a medium which can support only one type of wave, there are normally two solutions for the wavenumber k for each value of frequency. If the system is symmetric about the plane perpendicular to the propagation direction (z-axis), then two solutions for k exist which represent waves traveling in opposite directions but with equal wavelengths ($k_2 = -k_1$). When $F(\vec{r}, t)$ does not represent the total physical field but only the oscillating component as for waves on an electron beam, both solutions for k may have the same sign, one wave moving slower than the beam and the other faster. For spin waves in magnetic materials there are also static fields, but then $F(\vec{r}, t)$ represents an entire field component transverse to such static fields, and there will again be two solutions with $k_2 = -k_1$. In the limiting case of $k \to 0$, the result is ferromagnetic resonance, with the magnetization precessing in time about the static field, but not in space.

(b) Coupled waves: concepts.

Whenever waves of two different transmission lines or physical types or polarizations are coupled together in a one-dimensional system, there are normally four possible traveling waves, one traveling in each direction for each type. For example, simple directional couplers can have a wave going in each direction in each of the two coupled waveguides or transmission lines. In materials such as yttrium iron garnet rods, spin waves and elastic waves can travel in both directions along the rod axis, but can also be coupled strongly due to the effect of the elastic strain on the magnetization and vice versa. In a case such as this, the two coupled "transmission lines" exist at the same location in space but refer to different physical quantities. Radio waves of two different polarizations of electric field can be coupled by the magnetic field in the ionosphere, and each polarization can have waves traveling in either direction. Another example is the interaction and coupling of electron space charge waves with the electromagnetic fields of waves traveling along a surrounding helical conductor in a traveling wave tube. See the bibliography of this work for some representative references.

The characteristic feature of such situations is that the originally uncoupled waves are no longer normal modes. Hence, a given linear combination of the uncoupled waves entering a coupling region will in general emerge as a different linear combination.

A familiar case is Faraday rotation where linearly-polarized electromagnetic waves are coupled due to the anisotropy of a medium with applied static magnetic field, for example. For homogeneous (uniform) media, however, it is still possible to construct uncoupled normal modes, expressible as linear combinations of the original wave types. Basically this property follows from the fact that the corresponding wave equations all have constant coefficients (see below), in contrast with the situation in continuously-varying media. In homogeneous Faraday rotation media, for example, the new uncoupled normal modes are circularly-polarized linear combinations traveling at different velocities.

Apparently another example are the normal modes in rectangular waveguides, which can be considered to be linear combinations of plane waves reflected from the transverse boundaries of the waveguide. These linear combinations do not change with propagation distance, however, and therefore do not represent the same kind of coupled wave situation. Furthermore, from each set of these plane waves there can be constructed only one normal mode traveling in each direction, instead of two. Coupling in the sense to be used in this section occurs, for example, when two waveguides are placed side by side with holes in the wall between them, as in directional couplers. Then certain linear combinations of the fields in the two waveguides will constitute new uncoupled normal modes. The velocities of these new modes are changed by the coupling just as the resonance frequencies of two coupled resonant circuits differ from each other and from the original frequencies, even if these were identical. Compare the splitting of degenerate quantum-mechanical energy levels by a perturbation.

The analysis of coupled waves in homogeneous media is simplified by the fundamental absence of coupling between waves with wavenumbers of opposite sign. Normally, this absence is equivalent to the absence of reflected waves. It is possible, nonetheless, for a wave carrying energy in one direction to be coupled to a "backward wave" carrying energy in the reverse direction. The phase velocity of the two waves, however, is in the same direction; the backward wave has its energy flow vector opposite to its phase velocity, and is of a different physical type than the incident wave. Furthermore, the amplitudes of the normal modes do not change with distance; power conservation is assured by the fact that the wavevectors k_i are complex. (See JOHNSON 1965 and LOUISELL 1960 for examples.) Only in inhomogeneous media or in the presence of boundaries will true reflections be generated. Therefore, in four wave situations with homogeneous media only two waves will be coupled, so that finding the normal mode wavenumbers reduces to a 2 x 2 eigenvalue problem, and the effects of coupling are describable by 2 x 2 matrices (see below). These statements hold for all the relevant physical situations treated in the two references cited above, including directional couplers, contradirectional couplers, backward wave oscillators, and traveling wave tubes.

A difficulty may nevertheless arise if the two coupled waves, carrying different kinds of energy such as electromagnetic and electro-mechanical energy in electron beam tubes, are strongly coupled, since it may then be very difficult to determine the coupled wave equations. This can occur if an electron beam and a guided-wave helix are very close together in a traveling wave tube, for example, so that a simple perturbation analysis of the effect of the fields of one wave on the other breaks down. In such a case, one may use a variational method to approximate the eigenvalues k_i without actually ever explicitly finding the wave equations (HAUS 1958). This method resembles the use of variational methods to determine the propagation constant in waveguides of difficult cross section, such as ridge waveguides. There is also some similarity to the situation in quantum mechanics where one solves for the perturbed eigenvalues using matrix notation when the perturbation ("coupling") is weak, but uses variational methods to approximate the eigenvalues when the perturbation terms (coupling terms) are hard to calculate explicitly.

(c) Coupled waves: Faraday rotation illustration.

Media exhibiting Faraday rotation are special cases of anisotropic media, in which the permittivity ϵ or the permeability μ is a tensor.

Faraday rotation occurs as the result of the application of a static magnetic field H in a plasma (gaseous or solid state), in a magnetic insulator, and near an absorption frequency in many materials. In the former two cases, the effect is normally noticed near microwave frequencies, since the effect is strongest near the electron cyclotron resonance and the ferromagnetic resonance, respectively. In each of those cases, the resonance frequency is approximately 2.8 MHz per oersted of applied field H. The resonance is manifested in a tensor $\underline{\underline{\epsilon}}$ for the plasma, and a tensor $\underline{\underline{\mu}}$ in the magnetic insulator. Chapter 4 treats cases where $\underline{\underline{\mu}}$ is not only dependent on frequency but also on wave-number, which in turn is a function of distance. Optical Faraday rotation occurs in both magnetic and non-magnetic materials due to the Zeeman effect, and is manifested through a tensor $\underline{\underline{\epsilon}}$. Such an effect occurs when application of a magnetic field can split the resonance absorption frequency of an optical transition into an odd number of frequencies, allowing normal modes with circular polarization to propagate at different speeds. (See, for example, MASON 1966 for references.)

Now assume that the medium in question has a tensor $\underline{\underline{\epsilon}}$. The treatment with tensor $\underline{\underline{\mu}}$ is entirely analogous. Assume also that the magnetic field H is applied in the z-direction. For propagation in the z-direction, there will be no z-components of the dynamic fields in infinite media, so that we will be concerned only with the components ϵ_{xx}, ϵ_{xy}, ϵ_{yx}, and ϵ_{yy}. Moreover, for the usual cases of Faraday rotation, $\epsilon_{yx} = -\epsilon_{xy} = \epsilon^*_{xy}$, and $\epsilon_{xx} = \epsilon_{yy}$. Then the components of Maxwell's curl equations can be written as:

$$\underline{E}_x' = -j\omega\mu_o \underline{H}_y ; \qquad \underline{H}_y' = -j\omega(\epsilon_{xx}\underline{E}_x + \epsilon_{xy}\underline{E}_y)$$

1.2(3)

$$\underline{E}_y' = +j\omega\mu_o \underline{H}_x ; \qquad \underline{H}_x = j\omega(-\epsilon_{xy}\underline{E}_x + \epsilon_{xx}\underline{E}_y),$$

where the prime denotes differentiation with respect to the propagation distance z, and the fields are assumed to be written as in 1.2(1).

Equations 1.2(3) constitute a set of four coupled first order differential equations which can be written in matrix form:

$$U'(z) = DU(z) \qquad 1.2(4)$$

where

$$U(z) = \begin{bmatrix} \underline{E}_x(z) \\ \underline{H}_y(z) \\ \underline{E}_y(z) \\ \underline{H}_x(z) \end{bmatrix} \qquad 1.2(5)$$

and

$$D = j\omega \begin{bmatrix} 0 & -\mu_o & 0 & 0 \\ -\epsilon_{xx} & 0 & -\epsilon_{xy} & 0 \\ 0 & 0 & 0 & \mu_o \\ -\epsilon_{xy} & 0 & \epsilon_{xx} & 0 \end{bmatrix} \qquad 1.2(6)$$

To find the normal modes N_i and their wavenumbers k_i, one can diagonalize the matrix D as in Appendix 1. The four eigenvalues of D are then just jk_i, i = 1, 2, 3, 4. Furthermore, the fields in 1.2(5) will be expressible as linear combinations of the normal modes through the matrix L which diagonalizes D:

$$U = LN, \qquad 1.2(7)$$

where $L^{-1}DL$ is diagonal with elements $-jk_i$ (see Appendix 1). From 1.2(7) and 1.2(4) we see that

$$N' = L^{-1}DLN. \qquad 1.2(8)$$

Since $L^{-1}DL$ is diagonal, the solutions for N_i are simply

$$N_i = \underline{a}_i \exp(-jk_i z), \qquad 1.2(9)$$

which is the desired normal mode form (compare 1.2(2)). In continuously-varying inhomogeneous media, the matrix D is a function of position, and then instead of normal modes N_i one finds quasi-normal modes Q_i with $U = LQ$ instead of 1.2(7) (see section 2.2). Equation 1.2(8) then takes on a more complicated form.

Note for isotropic media ϵ_{xy} and ϵ_{yx} are both zero, so that ordinary linearly-polarized plane waves can propagate with either $\underline{E}_y$ and $\underline{H}_x$ both zero or $\underline{E}_x$ and $\underline{H}_y$ both zero (see equations 1.2(3)). These are called the x-polarized and y-polarized waves, respectively. Note in this case that the matrix D in 1.2(6) splits into two 2 x 2 sub-matrices down the main diagonal. The wavenumbers are then

$$k_1 = -k_2 = k_3 = -k_4 = \omega(\mu_o \epsilon_{xx})^{\frac{1}{2}}, \qquad 1.2(10)$$

with corresponding normal mode solutions

$$N_1 = \underline{E}_x + \eta \underline{H}_y; \quad N_2 = \underline{E}_x - \eta \underline{H}_y$$

$$N_3 = \underline{E}_y - \eta \underline{H}_x; \quad N_4 = \underline{E}_y + \eta \underline{H}_x, \qquad 1.2(11)$$

where $\eta \equiv (\mu_o / \epsilon_{xx})^{\frac{1}{2}}$ is an effective impedance, by analogy with similar equations for the voltage and current on a transmission line.

Since we are considering homogeneous media, there are no reflections, as indicated by the fact that N_1 and N_2 are uncoupled, as well as N_3 and N_4. In sections 2.2 and 3.2 we consider how N_1 and N_2 can be coupled by inhomogeneity. Section 4.3 treats a case where all four waves are coupled by inhomogeneity in an anisotropic medium.

For the present case of Faraday rotation the anisotropy couples N_1 and N_3. Since N_2 and N_4 do not enter the analysis, it is simpler to proceed by eliminating $\underline{H}_y$ and $\underline{H}_x$, treating explicitly only the coupling between $\underline{E}_x$ and $\underline{E}_y$. This is done by combining equations in 1.2(3), with the result:

$$\underline{E}_x'' + \omega^2 \mu_o (\epsilon_{xx} \underline{E}_x + \epsilon_{xy} \underline{E}_y) = 0 \qquad 1.2(12a)$$

$$\underline{E}_y'' + \omega^2 \mu_o (-\epsilon_{xy} \underline{E}_x + \epsilon_{xx} \underline{E}_y) = 0. \qquad 1.2(12b)$$

These are now two coupled second-order equations. By adding and subtracting the first equation to j times the second, however, we obtain the uncoupled equations:

$$(\underline{E}_x \pm j\underline{E}_y)'' + \omega^2 \mu_o (\epsilon_{xx} \mp j\epsilon_{xy}) (\underline{E}_x \pm j\underline{E}_y) = 0. \qquad 1.2(13)$$

Assuming normal mode solutions of the form exp $(-jk_{\pm} z)$ for the uncoupled linear combinations

$$\underline{E}_{\pm} = \underline{E}_x \pm j\underline{E}_y \qquad 1.2(14)$$

we find from 1.2(13) that

$$k_{\pm}^2 = \omega^2 \mu_o (\epsilon_{xx} \mp j\epsilon_{xy}) \equiv \omega^2 \mu_o \epsilon_{\pm} , \qquad 1.2(15)$$

as could have been found also by diagonalizing D in 1.2(6). If only $\underline{E}_+$ is propagating ($\underline{E}_- = 0$), then $\underline{E}_y = j\underline{E}_x$ and equation 1.2(1) shows that $\underline{E}_y(z,t)$ lags $\underline{E}_x(z,t)$ by 90^o. Such a wave is called positively circularly-polarized. With the negatively circularly-polarized wave $\underline{E}_-$, $\underline{E}_y(z,t)$ leads $\underline{E}_x(z,t)$ by 90^o.

We could also go further and eliminate $\underline{E}_y$, say, from 1.2(12). Taking the second derivative of 1.2(12a), substituting for $\underline{E}_y''$ from 1.2(12b) and for $\underline{E}_y$ from the original form of 1.2(12a), we obtain

$$E_x''''(z) + 2\omega^2 \mu_o \epsilon_{xx} \underline{E}_x''(z) + (\omega^2 \mu_o)^2 (\epsilon_{xx}^2 - \epsilon_{xy}^2) \underline{E}_x(z) = 0. \qquad 1.2(16)$$

Again, assuming solutions of the form of 1.2(2), we find a quadratic polynomial equation for k^2, whose solutions are again given by 1.2(15). Note that information on polarization has been lost in going to 1.2(16). If, however, we set $V_1 = \underline{E}_x$, $V_2 = \underline{E}_x'$, $V_3 = \underline{E}_x''$, and $V_4 = \underline{E}_x'''$, we find that an equation similar to 1.2(4) results:

$$V'(z) = D_o V(z), \qquad 1.2(17)$$

where D_o is given as

$$D_o = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ d_1 & 0 & d_2 & 0 \end{bmatrix}. \qquad 1.2(18)$$

From 1.2(16), $d_1 = -(\omega^2\mu_o)^2(\epsilon_{xx}^2 - \epsilon_{xy}^2)$, and $d_2 = -2\omega^2\mu_o\epsilon_{xx}$. By diagonalizing D_o just as D was diagonalized, we find again the normal modes N_i, but now written in terms of $\underline{E}_x$ and its derivatives, instead of explicitly in terms of $\underline{E}_x$, $\underline{H}_y$, $\underline{E}_y$, and $\underline{H}_x$ as before. A case in an inhomogeneous medium where it is convenient to transform a fourth-order differential equation to the form of 1.2(17) and 1.2(18) is examined in section 4.3.

To demonstrate rotation of the plane of polarization using the matrix formulation of this section, write

$$U_f = L_f N_f , \qquad 1.2(19)$$

with

$$U_f \equiv \begin{bmatrix} \underline{E}_x \\ \underline{E}_y \end{bmatrix} \quad \text{and} \quad N_f \equiv \begin{bmatrix} \underline{E}_+ \\ \underline{E}_- \end{bmatrix}, \qquad 1.2(20)$$

the subscript f standing for "Faraday." From the definition 1.2(14) we have

$$L_f = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -j & +j \end{bmatrix} . \qquad 1.2(21)$$

Since the components of N_f have the normal mode form of 1.2(9), we can write

$$\underline{E}_\pm (z_b)/\underline{E}_\pm(z_a) = \exp [-jk_\pm (z_b - z_a)] . \qquad 1.2(22)$$

Thus

$$N_f(z_b) = S\, N_f(z_a), \qquad 1.2(23)$$

where

$$S = \begin{bmatrix} \exp(-jk_+\ell) & 0 \\ 0 & \exp(-jk_-\ell) \end{bmatrix}, \qquad 1.2(24)$$

and $\ell \equiv z_b - z_a$. Hence at the output, from 1.2(19) and 1.2(23):

$$U_f(z_b) = L_f \, S \, L_f^{-1} \, U_f(z_a) \, . \qquad 1.2(25)$$

Equation 1.2(25) expresses what happens in coupled mode situations: a wave incident at $z = a_a$ is decomposed into normal modes which propagate uncoupled through to $z = z_b$ where it is desired to determine the effect of the coupling on the original waves.

Now assume that at the input $z = z_a$ to the Faraday rotation medium the wave is linearly polarized in the x-direction, so that

$$U_f(z_a) = \begin{bmatrix} E_o \\ 0 \end{bmatrix} \qquad 1.2(26)$$

Also neglect any reflections at the input and output $z = z_b$ due to any discontinuities. Then explicitly performing the calculation indicated in 1.2(25), we find

$$U_f(z_b) = \frac{E_o}{2} \begin{bmatrix} \exp(-jk_+\ell) + \exp(-jk_-\ell) \\ -j\exp(-jk_+\ell) + j\exp(-jk_-\ell) \end{bmatrix} . \qquad 1.2(27)$$

The angle of rotation is given by

$$\varphi = \arctan\left[E_y(z,t)/E_x(z,t)\right] . \qquad 1.2(28)$$

Since the ratio E_y/E_x does not change with time for a linearly polarized wave when it is again a normal mode outside the Faraday rotation medium, we can evaluate 1.2(28) at some time $\omega t = n2\pi$. Thus from 1.2(1), 1.2(27) and 1.2(28) we find

$$\varphi(z_b) = \arctan\left(\frac{\sin k_-\ell - \sin k_+\ell}{\cos k_-\ell + \cos k_+\ell}\right) = \left(\frac{k_- - k_+}{2}\right)\ell \, . \qquad 1.2(29)$$

Thus the Faraday rotation is due to the different wavenumbers and hence different velocities of the two circularly-polarized waves.

(d) Reflection and transmission coefficients at boundaries.

Consider now a sharp boundary located at $z = z_o$. If the media on both sides allow only two waves, one traveling in each direction, a wave incident upon the boundary will generate a reflected and a transmitted wave. Assume, for example, that the incident wave is an x-polarized electromagnetic wave traveling in the z-direction with amplitude E_o. Then in the region to the left of the boundary,

$$\underline{E}_x = E_o \left[\exp(-jk_1^- z) + \underline{R} \exp(-jk_2^- z) \right], \qquad 1.2(30)$$

while in the region to the right:

$$\underline{E}_x = E_o \, \underline{T} \exp(-jk_1^+ z), \qquad 1.2(31)$$

with the fields written as in 1.2(2). We assume $k_2 = -k_1$ in both regions, as in 1.2(10).

To determine the unknowns $\underline{R}$ and $\underline{T}$, two boundary conditions are necessary. For multi-wave media with n waves allowed on each side of the boundary, there will need to be n boundary conditions, since the amplitude of the single incident wave is known, and only transmitted waves are allowed on the other side of the boundary. For the electromagnetic case considered here, the boundary conditions are continuity of $E_x(z,t)$ and $H_y(z,t)$ at $z = z_o$, or in view of Maxwell's equations and 1.2(1), the continuity of $\underline{E}_x(z)$ and $\frac{1}{\mu} \underline{E}_x'(z)$. Application of the conditions to 1.2(30) and 1.2(31) yields:

$$\underline{R} = \frac{k_1^-/\mu_- - k_1^+/\mu_+}{k_1^-/\mu_- + k_1^+/\mu_+} = \frac{Z_+ - Z_-}{Z_+ + Z_-} \qquad 1.2(32)$$

and

$$\underline{T} = \frac{2\,k_1^-/\mu_-}{k_1^-/\mu_- + k_1^+/\mu_+} = \frac{2\,Z_+}{Z_+ + Z_-} \qquad 1.2(33)$$

where $Z = (\mu/\epsilon)^{\frac{1}{2}} = \omega\mu/k$ is the effective impedance as was η in 1.2(11). The reflection coefficient with the waves written in terms of the magnetic field is the negative of that given in 1.2(32).

When μ is continuous across the boundary, so that $\mu_+ = \mu_-$, the boundary conditions reduce to continuity of $\underline{E}_x$ and its derivative $\underline{E}_x'(z)$. Then 1.2(32) takes the form

$$\underline{R} = (k_1^- - k_1^+)/(k_1^- + k_1^+) \,. \qquad 1.2(34)$$

RYDBECK 1960 showed for 4-wave media in which one field and its first three derivatives are continuous, that the reflection and transmission coefficients have the form of a product of three factors such as the right-hand side 1.2(34). This fact actually follows quickly from a property of the determinants of alternate matrices mentioned in Appendix 1, when Cramer's method is used to solve for the reflection and transmission coefficients.

Now recall that the total power flow for the waves in 1.2(30) and 1.2(31) can be calculated from the Poynting vector. In this case the z-component of electromagnetic power S_z is just $E_x(z,t)\, H_y(z,t)$. With the fields written as in 1.2(1), we find that the time average $\langle S_z \rangle$ is given as (see, for example, RAMO, WHINNERY, and VAN DUZER 1965):

$$\langle S_z \rangle = \frac{1}{2} \mathrm{Re}\, (\underline{E}_x \underline{H}_y^*). \qquad 1.2(35)$$

Substituting from 1.2(30) and 1.2(31), noting that $\underline{H}_y = j\underline{E}_x'/\omega\mu$, we obtain:

$$\langle S_z \rangle = \frac{|E_o|^2}{2 Z_-} (1 - |\underline{R}|^2) = \frac{|E_o|^2}{2 Z_+} |\underline{T}|^2 . \qquad 1.2(36)$$

Note that the ratio of power in the reflected wave to that in the incident wave is simply $|\underline{R}|^2$. If the impedance Z_+ differs from Z_-, however, the ratio of transmitted power to incident power is not $|\underline{T}|^2$, but $Z_- |\underline{T}|^2 / Z_+ = k_1^+ |\underline{T}|^2 / k_1^-$. Similarly, for multi-wave media with different "impedances" for different waves, the ratio of reflected power in a certain wave ℓ to incident power in a wave i will be given directly in terms of the relevant reflection coefficient $\underline{R}_{i\ell}$ only when $|k_\ell| = |k_i|$. Otherwise the ratio may be $|k_\ell| \, |R_{i\ell}|^2 / |k_i|$, or may have even a more complicated factor multiplying $|R_{i\ell}|^2$ when the waves i and ℓ are not of the same physical type (see subsections 4.2(c) and 4.3(b), for example).

Furthermore, for multi-wave media with $|k_\ell| \neq |k_i|$, the total time-averaged power $\langle S_z \rangle$ will not be the simple generalization of 1.2(36). When only propagating waves are present, there will be an interchange of power between the waves of different physical type in a manner analogous to the "beating" of two musical notes of different frequencies. Since this interchange is periodic in space there will be no net interchange after a spatial average is taken. On the other hand, if evanescent or decaying waves are present, with imaginary or complex k, respectively, then power can be permanently transformed from one physical type to another. Thus, even though the power in each physical "channel" may be required to be separately continuous at the boundary, a net energy transformation may take place as a wave is decaying away from the boundary (see MORGENTHALER 1967 for an example of this effect).

CHAPTER 2

SOLUTION OF COUPLED WAVE PROBLEMS IN MEDIA WITH CONTINUOUSLY-VARYING PARAMETERS

2.1 Introduction

In this chapter we review the mathematical concepts behind various methods of solving coupled wave problems where the parameters vary significantly in only one direction, the direction of propagation. The distinction between wave propagation in such media and that in corresponding linear homogeneous media is as big as the distinction between differential equations with variable coefficients and those with constant coefficients only. This chapter is concerned primarily with situations where the coefficients vary monotonically or with one relative maximum or minimum; hence, there is no explicit treatment of wave propagation in periodic or random media. Although almost all of the material of this chapter has existed in scattered places in the literature, the overall picture and certain of the specific results are often not well understood.

When a certain wave can propagate independently of all others in a bulk homogeneous medium, the situation is often describable by a second-order ordinary differential equation, allowing for two solutions, one representing propagation in each direction. If then some physical coupling mechanism is introduced between this wave and another, this second-order differential equation becomes coupled to another second-order equation. The result can be expressed as one fourth-order equation. As long as the coupling and the wave velocities do not change with distance, however, the resulting differential equation has constant coefficients and hence has exponential, wavelike solutions. As was pointed out in section 1.2, these solutions, called normal modes, propagate independently of each other; they are uncoupled. The only effect of the physical coupling mechanism is that a certain normal mode may be expressible as a different linear combination of the original waves at two different positions along the direction of propagation. The original waves may thus be coupled. The analysis of this situation for the case of Faraday rotation was given in subsection 1.2(c).

In continuously-varying media (to be referred to hereafter as "inhomogeneous" media), it is no longer possible to find simple exponential wavelike solutions as in 1.2(9). One can always find such normal mode solutions for differential equations with constant coefficients, but the best that is possible for propagation in inhomogeneous media is to find approximate quasi-normal mode WKB solutions of differential equations with variable coefficients. Such solutions have been derived for certain fourth-order systems by RYDBECK 1960 and 1967 and BUDDEN 1961, for example, as well as by SCHLÖMANN and JOSEPH 1964, although they are not always identified as "WKB" or "quasi-normal mode" approximate solutions. Furthermore, the form of the "WKB solutions" constructed recently by VASILE and LAROSA 1968a was incorrect. Hence in section 2.2 we show how to construct the quasi-normal mode solutions in general, for systems of any order. A rigorous and complete mathematical development of these quasi-normal modes was accomplished by KELLER and KELLER 1962, starting from a system of first-order coupled differential equations. Section 2.2 in connection with Appendix 1 goes further to show how to construct these modes explicitly as approximately linearly-independent "basis" solutions for all fourth-order differential equations and for sets of coupled second-order equations and coupled first-order equations.

Typically, the physics of a situation produces two coupled second-order equations describing coupling between two wave types. In cases for which the variation of the parameters with distance is slow compared to the wavelengths involved, one can with good accuracy usually neglect the two reflected waves and just consider the coupling between two waves traveling in the same direction. This situation can often be modeled by two coupled first-order equations extracted from the original coupled equations. From these first-order equations quasi-normal modes can be constructed. Such modes were called "tapered" or "warped" normal modes by COOK 1955, FOX 1955, and LOUISELL 1955, who found that directional couplers with a wider bandwidth could be constructed by varying the coupling and/or transverse dimensions of the waveguides continuously with distance. The corresponding directional couplers built with waveguides with constant parameters were extremely frequency-sensitive due to "beating" which occurred. When reflections are important, quasi-normal mode solutions may still be useful for the analysis, as in the WKB method (see the discussion below and section 2.4).

Methods of solving coupled differential equations by successive approximations are almost always based on quasi-normal mode solutions as the "zero-order" approximations. For the first approximation it is assumed that the incident wave (quasi-normal mode) has a constant amplitude throughout the inhomogeneous coupling region. The excitation of the other modes is then considered one at a time. (Compare time-dependent perturbation theory in quantum mechanics, with "incident wave" replaced by "initial state.") LOUISELL 1955 applied a successive approximations method to find the coupling between the quasi-normal modes described above in tapered directional couplers. Note that subsections 3.2(b) and 4.2(b) describe an analytical solution of the relevant second-order equation, which has the same form as that of the magnetoelastic coupling problem when reflections are neglected.

In section 2.3 we describe various applications which have been reported in the literature of successive approximations methods using coupled first- and second-order equations. When the complete set of n coupled first-order equations for an n^{th} order system are used, the coupling terms come directly from the analysis which determines the quasi-normal modes. In this case, the coupling terms are proportional to the derivatives of the parameters of the medium, and go to zero when the medium becomes homogeneous. The successive approximations approach then involves matrizants as described by KELLER and KELLER 1962, similar to related approaches in the control theory of time-varying systems. Approaches using coupled second-order equations may or may not be written so that the coupling terms involve only parameter derivatives, since the system may or may not have been transformed first to normal modes.

Note that the discussion so far has been mostly in terms of coupled first- and second-order equations. It is always possible, however, to rewrite four first-order or two second-order coupled equations in terms of one fourth-order equation (see Appendix 2). This procedure can be useful in comparing the various kinds of coupled mode situations, especially since the corresponding dispersion relations for homogeneous media usually involve the solution of similar fourth-order polynomials in the wavenumber. Furthermore, there are important cases where solution of the fourth-order differential equation seems to offer the only hope for obtaining useful analytical results (in distinction to useful numerical results obtainable on a computer). These cases occur, for example, when neither wave type can propagate past a certain point, called the "turning point," but an incoming wave of one type will, in general, couple strongly to both types of reflected wave.

For cases where reflections are substantial, successive approximations are no longer useful, since the amplitude of the incident normal mode is not approximately constant throughout the inhomogeneous region. In other words, there is then strong coupling between the various quasi-normal modes, and the approximate validity of the WKB quasi-normal modes breaks down. This can happen when either the parameters of the medium, or the "local" wavenumbers of the quasi-normal modes vary rapidly. Such a region will be termed a "critical region" in the following. Then we usually look for solutions of the n^{th} order differential equation with variable coefficients describing the system with n waves.

A critical region may be one of rapid variation of the medium parameters, or where two roots of the dispersion relation become equal, or where one root goes to zero. At a large distance from such a region, however, we can usually specify that the fields consist of an incident wave of one type and one or more reflected waves. On the side of the region opposite to where the incident wave is present there will be the physical restriction that only waves carrying energy away from the critical region may be present, as long as no reflections from some distant point are allowed. If certain waves are cut off, so that the wavenumber is imaginary in the lossless case, this restriction is equivalent to requiring that only waves which are evanescent (as opposed to growing) may be present on the side opposite the incident wave. Care is required in certain cases to identify which waves carry energy away from the critical point, since for "backward waves" the energy flow is directed opposite to the phase velocity (see, for example, HEADING 1969a; also STIX 1965b, and KUEHL 1967a). Backward waves occur in the examples of section 4. 3.

There exist basically three methods for finding the reflection and transmission coefficients at critical regions. These coefficients refer to the amplitudes of plane waves representing independent waves in exterior homogeneous regions, or of the quasi-plane wave WKB solutions of section 2. 2 representing almost-independent waves traveling in weakly-inhomogeneous exterior regions, outside the critical region. For very thin (thinner than any wavelength of the waves on either side) critical regions, the ordinary boundary conditions will provide these coefficients, through a set of simultaneous algebraic equations. The number of boundary conditions must be equal to the number of unknown amplitudes, since there must be as many equations as unknowns. As discussed in section 1. 2, if there are more than two waves allowed in each region, there will be more than one reflected wave and one transmitted wave at the boundary, so more than the usual two boundary conditions are required. Boundary conditions involve the requirement of continuity of certain field quantities, arising from physical arguments about forces and torques remaining finite.

The use of a large number of boundaries of small height to model a smooth spatial variation forms the basis of the numerical schemes developed by ALTMAN and CORY 1969a and 1969b. Their method is called the "generalized thin film optical method" and has been applied successfully to coupled wave problems in the ionosphere, even when more than two waves are involved. An iterative procedure is applied after the transfer coefficient matrices for an elementary layer are derived, in terms of a geometric series of matrix products. The method can be adapted to yield transfer matrices of intermediate "layers" as well as the overall reflection and transmission coefficients of the medium. In the present work, numerical schemes such as this will not be discussed further, since it is generally desirable if possible to have an analytic solution for a quick assessment of the various factors. The work of Altman and Cory cited above, however, includes numerical graphical plots which are quite helpful in understanding the coupling and reflections at numerous places in the ionisphere. Their work may also be consulted for a discussion of earlier numerical approaches in the literature.

The other two methods depend upon the properties of the differential equation which describes the fields within the critical region. The distinction between these two methods is based on the location and type of the singular points of the differential equation. Near such points it is impossible to write all the solutions in terms of single-valued convergent power series. One method is the WKB method; the other involves the use of generalized hypergeometric functions. The areas of differences and similarities between these two methods are discussed in section 2.5, together with an explanation of the reasons that generalized hypergeometric functions are so useful. Expositions of some of the mathematical details of the WKB method and of generalized hypergeometric functions appear in sections 2.4 and 3.1, respectively. Applications to physical problems appear in section 4.3, and in sections 3.2, 3.4, 4.2, 4.3, and 4.4.

The WKB method treats cases where the inhomogeneity extends effectively over all space, although the critical region is localized. Outside the critical region, the WKB quasi-normal mode solutions of section 2.2 will be valid in the sense that they are approximately linearly independent basis solutions for the field, since the coupling terms are small there. Some approximating differential equation is used for the fields in the critical region. The basic problem is then to identify the asymptotic expansions of the solutions of this differential equation as linear combinations of the WKB solutions at the "edge" of the critical region. This "matching" technique usually works well only in the limit of the large value of some parameter, such as inverse static field gradient in the critical region, and when the original differential equation used to derive the WKB solutions is itself valid even at large distances from the critical region. If some approximation, such as a "quasistatic" approximation used to derive the original equation, breaks down too close to the coupling region, there may be difficulty in identifying the linear combinations referred to above. Similar trouble can arise from other coupling, such as magnetoelastic coupling, occurring in regions too close to the critical region. See subsection 4.3(d) for an explicit discussion of these phenomena in the case of medium-k to high-k spin wave conversion at a turning point.

The WKB treatment of the second-order Airy equation forms the basis of phase-integral methods applied to a broader class of problems. The Airy equation describes a cutoff situation where the square of the "local" wavenumber passes through zero so that no energy can propagate further; all is reflected. When loss is included, a phase integral may account for the corrections. Multi-wave problems can sometimes also be approximated in certain regions by an equation resembling the Airy equation. Instead of describing a cutoff situation, however, this equation may now describe the confluence of two wave-numbers at a branch point. A phase integral may then be used to obtain an estimate of the reflection coefficients, or sometimes of the transmission coefficients. See BUDDEN 1961, RYDBECK 1967, and the discussion in part 2.4(b)2 below.

In cases where the inhomogeneity is localized, the linearly independent solutions outside the critical region are just the normal modes of a homogeneous medium, with simple exponential variation with distance (see section 1.2). For certain variations of the medium parameters in the inhomogeneous region, analytic solutions of the corresponding differential equations can be obtained in terms of generalized hypergeometric equations. Far away from the inhomogeneous region these solutions can be represented well by the first term of convergent power series expanded about positive or negative infinity. These first terms are easily identified with the exponential normal modes. Note that if the width of the inhomogeneous region is allowed to become infinitesimally thin, the reflection and transmission coefficients obtained from these solutions should agree with those obtained from the boundary conditions.

When the original differential equation has a form away from the critical region which is significantly different from that of the approximating equation solved by one of the above methods, it may be profitable to find the corrections. In this case the approximating equation is called the "comparison equation," and its solutions serve as zero-order solutions in a new successive approximations scheme. The differences between the original and the comparison equations can be considered as driving terms of a differential equation, which is solved by the variation of parameters method using the solutions of the comparison equation as basis functions. See subsection 2.3(b) for cases when the comparison equation is second-order.

Although this discussion has been in terms of spatially inhomogeneous media, very similar situations exist in time-varying media. The differential equations which arise are often almost identical in the two different kinds of media. See, for example, subsection 4.2(b). In parametric interactions and in quantum-mechanical time-dependent perturbation theory, the coupling parameters of the media usually vary sinusoidally with space and/or time, however.

2.2 WKB Quasi-Normal Modes

(a) General formalism.

As in equation 1.2(4), let us again assume that we have a set of n coupled first-order equations which can be written in matrix form

$$U' = DU \qquad 2.2(1)$$

where U is an n x 1 column matrix, and D is an n x n square matrix. If we start originally with coupled second-order equations, or a single fourth-order equation, for example, Appendix 1 shows how to convert them to the form of 2.2(1). Now we would like to diagonalize D to find normal modes as in 1.2(7) and following. Instead of the normal-mode matrix N in 1.2(7), let us introduce Q, in anticipation of the fact that we will have to be content with quasi-normal modes in inhomogeneous media:

$$U = LQ, \qquad 2.2(2)$$

where L is the n x n "linear combination" matrix, and Q is a column matrix. Indeed, instead of an equation of the form 1.2(8), we find (CLEMMOW and HEADING 1954):

$$Q' = L^{-1}DLQ - L^{-1}L'Q. \qquad 2.2(3)$$

The extra term in 2.2(3) arises because of the extra term generated by taking

$$U' = LQ' + L'Q, \qquad 2.2(4)$$

and then substituting in 2.2(1) and 2.2(2). If we neglected this extra term because of slow variations in L, then we would obtain:

$$Q_i \approx -jk_i Q_i, \qquad 2.2(5)$$

where the $-jk_i$ are defined to be the eigenvalues of D and hence the diagonal elements of $L^{-1}DL$ (see Appendix 1):

$$-jk_i \equiv (L^{-1}DL)_{ii} . \qquad 2.2(6)$$

Even so, however, the solutions of 2.2(5) for the Q_i would not have the same simple exponential form of 1.2(9), but rather:

$$Q_i \approx \underline{a}_i \exp\left[-j\int^{z} k_i\,(\xi)\,d\xi\right], \qquad 2.2(7)$$

since the "local wavenumbers" k_i are functions of the distance z. The exponential factor in 2.2(7) is called an "exponential phase integral."

To construct quasi-normal mode solutions which at least resemble 2.2(7), we first must eliminate the "self-coupling" terms along the diagonal of $(-L^{-1}L')$. As was shown in RYDBECK 1960 and 1967, by calculating the relevant transmission coefficient for a boundary of infinitesimal height, these terms represent the amount of "forward scattering" into the same wave by the inhomogeneity. The off-diagonal terms of $(-L^{-1}L')$ represent scattering into the other modes. Proceeding, we now write

$$Q = BX , \qquad 2.2(8)$$

where B is assumed to be an n x n diagonal matrix and will be shown to consist of the WKB amplitudes which must multiply the exponential phase integrals in 2.2(7). The column matrix X will be shown to consist of exponential phase-integral solutions.

Following the same steps as led to 2.2(3) after combining 2.2(8) with 2.2(2), we find:

$$X' = [(LB)^{-1}D(LB)]X - (LB)^{-1}(LB)'X . \qquad 2.2(9)$$

Note that

$$(LB)^{-1}D(LB) = B^{-1}(L^{-1}DL)B = L^{-1}DL , \qquad 2.2(10)$$

since B and $(L^{-1}DL)$ are assumed to be diagonal and hence commute. Furthermore, we define the n x n coupling matrix C as:

$$C \equiv (LB)^{-1}(LB)' = B^{-1}B' + B^{-1}(L^{-1}L')B \qquad 2.2(11)$$

The diagonal elements C_{ii} of C are from 2.2(11)

$$C_{ii} = (B_{ii})^{-1}B'_{ii} + (L^{-1}L')_{ii} , \qquad 2.2(12)$$

since B is diagonal. For C to be diagonal as desired, $C_{ii} = 0$ for $i = 1, 2, \cdots, n$. Then 2.2(12) represents a set of uncoupled differential equations with solutions:

$$B_{ii}(z) = \exp\left[-\int (L^{-1}L')_{ii}\, d\xi\right] ; \quad C_{ii} = 0. \qquad 2.2(13)$$

To the extent that cross-couplings between waves are small so that the off-diagonal elements of C can be neglected, the solutions for X_i in 2.2(9) will be the exponential phase integrals of 2.2(7) with arbitrary constants A_i:

$$X_i = A_i \exp\left[-j \int^{z} k_i(\xi)\, d\xi\right] , \qquad 2.2(14)$$

where we have noted 2.2(10), 2.2(11), and 2.2(13). The quasi-normal mode solutions Q_i from 2.2(8) are then simply

$$Q_i = B_{ii} X_i , \qquad 2.2(15)$$

with B_{ii} and X_i given by 2.2(13) and 2.2(14), respectively.

The solution outlined above for the "WKB amplitudes" B_{ii} was derived by KELLER and KELLER 1962 and equivalently by BUDDEN 1961. RYDBECK 1960 and 1967 also derived the form of these "WKB amplitudes" in general directly from the equivalent of the diagonal terms of $(-L^{-1}L')$ in 2.2(3), by arguing that they represent "forward scattering" terms. In this case, however, his wavenumbers k_i were taken from normal-mode type solutions of the n^{th} order equation including terms in gradients of the medium parameters. These solutions for k_i differ slightly from those of 2.2(6), calculated as if the medium were homogeneous (see Appendix 1).

HEADING 1961 and RYDBECK 1967 have shown that whenever 2.2(1) is obtained from a single n^{th} order differential equation.

$$(L^{-1}L')_{ii} = \sum_{\substack{\ell=1 \\ \ell \neq i}}^{n} \frac{k_i'}{k_i - k_\ell} , \qquad 2.2(16)$$

provided that L is chosen to have the form of A1(11) (see Appendix 1). When $n = 2$ and $k_2 = -k_1$, $(L^{-1}L')_{11} = k_1'/2k_1 = (L^{-1}L')_{22}$, so that the solution in 2.2(13) becomes $B_{11} = B_{22} = (k_1)^{-\frac{1}{2}}$, the familiar result for ordinary WKB solutions. However, when $n = 4$, $B_{ii} \neq (k_i)^{-\frac{1}{2}}$. For example, if $k_2 = -k_1$ and $k_4 = -k_3$, then from 2.2(16) we have

$$(L^{-1}L')_{11} = \left(\frac{k_1'}{2k_1}\right)\left[1 + \frac{(4k_1^2)}{(k_1^2 - k_3^2)}\right]. \qquad 2.2(17)$$

If $k_1^2 k_3^2$ is constant, we can integrate 2.2(17) to obtain, from 2.2(13),

$$B_{11}(z) = \left[k_1^3 (k_1^2 - k_3^2)\right]^{-\frac{1}{2}} ; \quad k_1 k_3 = \text{constant}. \qquad 2.2(18)$$

Similarly, $B_{22} = B_{11}$ and $B_{44} = B_{33}$, with

$$B_{33}(z) = \left[k_3^3 (k_3^2 - k_1^2)\right]^{-\frac{1}{2}} ; \quad k_1 k_3 = \text{constant}. \qquad 2.2(19)$$

The fact that $B_{ii} \neq (k_i)^{-\frac{1}{2}}$ for general n^{th} order systems has not always been appreciated in the literature (for example, VASILE and LAROSA 1968a), although BUDDEN 1961 gives the correct solutions when $n = 4$ for radio waves in the ionosphere, including the effects of the magnetic field. Also, PEARLSTEIN and BHADRA 1969 derive the correct solutions when $n = 4$ for plasma waves at the upper hybrid frequency. Their derivation proceeds in the fashion of geometrical optics by assuming solutions of the form $Q_i = \exp[S(z)]$, and then solving by iteration for $S(z)$ assuming $S''(z) \ll (S')^2$, etc. Such calculations involve approximate solutions of nonlinear n^{th} order Riccati equations, however, without the elegrance and generality of the method leading to 2.2(13). Furthermore, the n^{th} order differential equation is needed at the start with such a method, and none of the coupling terms represented by the off-diagonal elements of C are readily obtainable.

For systems not initially described by a single n^{th} order differential equation, the solutions for the B_{ii} will generally be different from those in 2.2(18) and 2.2(19), since $L^{-1}L'$ will have a different form. As shown in Appendix 2, however, it is always possible to reduce a set of four coupled first-order equations or two coupled second-order equations to one fourth-order equation. It is not necessary, moreover, to know the explicit form for inhomogeneous media of this fourth-order equation to obtain L in the form of A1(11). As remarked in Appendix 1, all that need be known are the eigenvalues $\lambda_i = -jk_i$ of the matrix D in 2.2(1).

These eigenvalues are the solutions of the fourth-order dispersion relation for homogeneous media. Note that whenever the system is reduced to a fourth-order system to obtain the simple form A1(11) for L, the quasi-normal modes using L^{-1} will be expressible in terms of U_1 and its first three derivatives, rather than in terms of the original U_1, U_2, U_3, and U_4. Even if it is not desired to reduce the system in this way, it is still possible directly from coupled first- or second-order equations to choose the first row of L to consist of constants. This is done using equation A1(15), where the cofactors $c_{4\ell}$ there are those of $(D - \lambda_i I)$. For two coupled second-order equations, Appendix 1 shows the relevant steps explicitly. Similar remarks hold for n^{th} order systems.

Now note that 2.2(2) implies

$$U_1 = \sum_{i=1}^{n} L_{1i} Q_i . \qquad 2.2(20)$$

When the L_{1i} are all constants independent of distance, the right-hand side of 2.2(20) shows that the Q_i constitute quasi-independent "basic" solutions for U_1, which is usually some characteristic field. As indicated in the preceding paragraph, the L_{1i} can always be chosen to satisfy this requirement. U_1 is then just a constant linear combination of the "partial waves" Q_i. Furthermore, to satisfy power conservation, we then expect that the power in each Q_i must be approximately constant whenever the quasi-normal mode amplitudes A_i are approximately constant. In fact, the characteristic feature of WKB quasi-normal mode solutions is this constancy of power flow in each mode when cross-coupling between modes is neglected. We can also infer that the power flow in any mode Q_i will be proportional to $|A_i|^2$, multiplied at most by constants independent of distance and of i. This result is proved explicitly for a fourth-order spin wave system in part 4.3(b)3.

When the off-diagonal coupling elements of the matrix C in 2.2(11) must be considered, the quasi-normal mode amplitudes A_i in 2.2(14) are no longer constants. To examine when these variations might be significant, write 2.2(14) as

$$X_i = E_{ii} A_i ; \quad (X = EA), \tag{2.2(21)}$$

with

$$E_{ii} \equiv \exp\left[-j \int^{z} k_i(\xi)\, d\xi\right]. \tag{2.2(22)}$$

The E_{ii} are exponential phase-integral factors, and will be considered to be elements of an n x n diagonal matrix E. From the property 2.2(6) and 2.2(22), we find

$$E' = (L^{-1} DL) E = E(L^{-1} DL), \tag{2.2(23)}$$

since $(L^{-1} DL)$ and E are both diagonal. Furthermore, 2.2(9) with 2.2(10) and 2.2(11) becomes

$$X' = (L^{-1} DL) X - CX. \tag{2.2(24)}$$

Combining 2.2(21), 2.2(23), and 2.2(24), we finally obtain

$$A' = -E^{-1} CEA, \tag{2.2(25)}$$

or, in component notation,

$$A_i'(z) = -\sum_{\ell \neq i} \exp\left[j \int^{z} (k_i - k_\ell)\, d\xi\right] C_{i\ell} A_\ell. \tag{2.2(26)}$$

The equations represented by 2.2(26) are the coupled first-order equations for the quasi-normal mode amplitudes, where the $C_{i\ell}$ depend upon the derivatives of the parameters of the medium through 2.2(11). The original equations represented by 2.2(1), on the other hand, generally contain coupling terms which do not vanish even for homogeneous media. The transformation between 2.2(1) and 2.2(26) is given from 2.2(2), 2.2(15), and 2.2(21) as:

$$U = LBEA, \qquad 2.2(27)$$

where B and E are diagonal with elements given by 2.2(13) and 2.2(22). Once the Q_i are specified by the diagonalization of D, the transformation $Q_i = B_{ii}E_{ii}A_i$ leading to coupled equations of the form 2.2(26) with no self-coupling terms is unique to within a factor independent of z. To prove this fact, first assume the contrary. Then we should be able to write $A_i(z) = g_i(z)\underline{A}_i(z)$, where the $\underline{A}_i(z)$ are new variables which must satisfy an equation of the form of 2.2(26). However, we would have $A_i' = g_i\underline{A}_i' + g_i'\underline{A}_i$. Substituting for A_i' from 2.2(26), we obtain after division by g_i the self-coupling term $-(g_i'/g_i)\underline{A}_i$ in the equation for $\underline{A}_i'$, thus contradicting our assumption.

Note finally that the procedure outlined in this section can be applied to find the quasi-normal modes in a systematic manner for any problem, including those treated by SCHLÖMANN and JOSEPH 1964 ("strong coupling" case), LOUISELL 1955 and KIRCHNER 1966 (see part 4.2(b)1). Each of these papers reduced the coupled equations to the form of 2.2(24) where C has zero diagonal elements, but did not indicate a systematic procedure for doing so. LOUISELL 1955 first simplified the problem by neglecting coupling terms representing reflections before finding the quasi-normal modes. His modes were thus found from a second-order system of two coupled first-order equations.

(b) Example: reflection of electromagnetic waves.

As an example for the techniques of this section, consider the propagation of an x-polarized electromagnetic wave in an isotropic medium with variable dielectric constant $\epsilon(z)$. Since the medium is isotropic, there will be no coupling to y-polarized waves, as in the Faraday rotation example of subsection 1.2(c). However, in an inhomogeneous medium we now may have reflections. Writing $U_1 = \underline{E}_x(z)$ and $U_2 = \underline{H}_y(z)$, we find from Maxwell's equations 1.2(3) that the matrix D in 2.2(1) becomes:

$$D = -j\omega \begin{bmatrix} 0 & \mu_o \\ \epsilon(z) & 0 \end{bmatrix} . \qquad 2.2(28)$$

Diagonalizing, we find the eigenvalues $-jk_i$ satisfy

$$k_1 = -k_2 = \omega(\mu_o\epsilon)^{\frac{1}{2}} , \qquad 2.2(29)$$

and that we can write L in the form

$$L = \begin{bmatrix} 1 & 1 \\ \left(\frac{\epsilon}{\mu_o}\right)^{\frac{1}{2}} & -\left(\frac{\epsilon}{\mu_o}\right)^{\frac{1}{2}} \end{bmatrix} , \qquad 2.2(30)$$

producing after inversion the normal modes N_1 and N_2 of equation 1.2(11). Since $(\epsilon/\mu_o)^{\frac{1}{2}} = k_1/\omega\mu_o$ from 2.2(29), we have

$$L^{-1}L' = \frac{k_1'}{2k_1} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} . \qquad 2.2(31)$$

Now explicitly evaluating the "WKB amplitudes" B_{ii} using 2.2(13), we obtain

$$B_{11} = B_{22} = (k_1)^{-\frac{1}{2}} . \qquad 2.2(32)$$

Combining 2.2(27), 2.2(22), and 2.2(32) yields

$$\underline{E}_x(z) = (k_1)^{-\frac{1}{2}} \left[A_1 \exp\left(-j \int^z k_1 \, d\xi\right) + A_2 \exp\left(+j \int^z k_1 \, d\xi\right)\right], \qquad 2.2(33)$$

which is the ordinary WKB solution, representing waves traveling in the positive and negative z directions. Note from 1.2(35) that the average z component of electromagnetic power $\langle S_z \rangle$ in each wave is $|\underline{E}_x|^2/2Z$ where $Z = (\mu_o/\epsilon)^{\frac{1}{2}} = \omega\mu_o/k$. Thus we see from 2.2(33) that the power in each wave is approximately constant as long as A_1 and A_2 are approximately constant.

The coupling matrix C acting to change the amplitudes A_1 and A_2 according to 2.2(26) is, from 2.2(11), 2.2(31), and 2.2(32):

$$C = \frac{-k_1'}{2k_1} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} . \qquad 2.2(34)$$

Hence whenever $k_1(z)$ (or $\epsilon^{\frac{1}{2}}$, from 2.2(29)) varies rapidly, the reflections will be large. Also, if $\epsilon(z)$ approaches zero (cutoff), substantial reflections will result even in weakly-inhomogeneous media.

Analytical solutions for the reflections in this problem are given in section 3.2 for certain variations in $\epsilon(z)$. Similar solutions for situations where the anisotropy of the medium also couples waves traveling in the same direction are given in section 4.3. (See parts 4.3(b)3 and 4.3(e).)

2.3 Successive Approximation Methods.

(a) From coupled first-order equations.

Recall from section 2.2 that the "WKB amplitudes" 2.2(13) could be derived by considering the "forward-scattering" transmission coefficients at a boundary of infinitesimal height. BREMMER 1951 considered the WKB quasi-normal mode solution obtained in this way for a second-order system to be the first term in a geometrical optical series, higher-order terms resulting from multiple reflections described by the off-diagonal terms of $(-L^{-1}L')$. Reference to treatments on the mathematical properties of such series are contained in BELLMAN 1964, page 102. The series constructed by KELLER and KELLER 1962, however, was shown to converge under more general conditions.

This last series was developed using a matrizant-integral procedure based on first-order coupled quasi-normal mode equations similar to 2.2(26). Such a procedure is now common in the treatment of time-varying control systems with state variables. The first term in the series is the usual first approximation, equivalent to the assumption that the amplitude A_i of the incident quasi-normal mode Q_i is constant throughout the coupling region.

Since the off-diagonal elements $C_{i\ell}$ in 2.2(26) are non-zero only when spatial derivatives of the parameters of the medium are present, it is expected that the coupling will be small when these spatial derivatives are small. If we start with only one of the A_i nonzero, say A_1, and assume small coupling, equations 2.2(26) can be integrated by assuming that A_1 is constant throughout the coupling region. Furthermore, the $C_{i\ell}$ often are significant only in a small region, so that the contribution in this region is the most important, and then the limits on the integrals can be pushed to infinity, with good accuracy. The parameters of the medium are also expanded in a power series. This is one approach used by SCHLÖMANN and JOSEPH 1964 in calculating magnetoelastic coupling efficiency. His WKB solutions were derived in a different manner, since the field elements of his matrix corresponding to our U were arranged in a manner prohibiting the setting of all the components in the first row of L equal to unity. Appendix 1, however, shows how a slight rearrangement allows the systematized method of this paper to be used.

If a better approximation is desired, it is necessary to use the $A_i(z)$, $i \neq 1$, calculated in the first-order approximation, then to calculate the variation of $A_1(z)$ through integration of 2.2(26), and then use this corrected value to correct the values of $A_i(z)$, and so on. This procedure is generally so involved that it is not used in practice. If the coupling is too strong to give good accuracy in the first approximation, another method is usually sought.

The coupling of the first-order equations of 2.2(26) can become strong when the parameters of the medium vary rapidly. In that case, it is possible that the coupling terms of the corresponding coupled second-order equations are then small. Such is the case with magneto-elastic coupling, when the parameters still vary slowly enough that reflections may be neglected. Then one uses the method of successive approximation in the coupled second-order equations, instead (the "weak coupling" case of SCHLÖMANN and JOSEPH 1964). On the other hand, when the coupling between the second-order equations becomes strong ("strong coupling" case), it may happen that the corresponding first-order equations are weakly coupled, so that the method described above is useful.

The usefulness of the successive approximations based on 2.2(26) also breaks down when the waves are close to a singularity in B or C. Equations 2.2(33) and 2.2(18) show that this can happen when one of the wavenumbers k_i approaches zero, or when two wavenumbers approach equality. Usually these situations imply that the waves are cut off beyond the point of singularity so that strong reflections are generated. Such is the case with the generation of high-k spin waves by medium-k spin waves ("magnetostatic waves"), and vice versa, at a "turning point." (See section 4.3.) In these cases the amplitudes A_i of the incident waves cannot be even approximately constant in the coupling region, since beyond that point is the cutoff region where the amplitudes decay rapidly to zero, as the wavenumbers k_i become imaginary. Successive approximations are then useless, whether based on the first- or second-order coupled equations. The techniques of other subsections may then be tried.

Before leaving the subject of coupled first-order equations, note the similarity between equations 2.2(20), 2.2(26), and those of time-dependent perturbation theory in quantum mechanics. Instead of an expansion in terms of WKB quasi-normal modes, there is an expansion in terms of the eigenfunctions of the unperturbed time-independent Hamiltonian. The method of successive approximations is used when the perturbation is small, to calculate the transition probabilities between energy levels. Actually, whenever a successive approximations scheme is applied to coupled first-order equations, the appearance will be similar, as in the solution of state-variable problems in time-varying control systems, for example, mentioned above. BELLMAN 1964, pages 32-51, may be consulted to see the relation to other perturbation techniques, including those used for nonlinear equations.

(b) From coupled second-order equations.

Coupled second-order equations such as in equations 1.2(12) are often the most natural form of description of a system, particularly when the coupling terms are small. The coupling terms in one equation may be functions of the field of the second equation or its derivative. In this form, one quickly sees how the dispersion relations for the two uncoupled systems are affected by these terms.

In attempting to apply successive approximations methods to such equations when the coupling terms are weak, usually both coupling terms are neglected at first. The uncoupled second-order equations are then solved, often in terms of WKB quasi-normal mode solutions, with $n = 2$ instead of $n = 4$. (As discussed in connection with first-order coupled equations, such WKB "solutions" are good approximations only in regions of space where there are negligible reflections of the energy in each solution.) Then one coupling term is calculated from these solutions and is interpreted as a driving term in the other equation. The solution to the coupled equation then follows the method of variation of parameters using the solutions to the corresponding uncoupled equation as basis functions. See SCHLÖMANN and JOSEPH 1964, and BURMAN 1967, for examples. HORTON 1966 takes the Laplace transform of the coupled second-order equations before applying this method. Higher approximations may be obtained by iteration.

In other cases, when the solutions to the uncoupled equations are also difficult to obtain, it may be convenient in a successive approximation approach also to neglect some terms in the uncoupled equations at first. Then one solves the resulting equations exactly in terms of known functions (for example, BURMAN 1967). These resulting equations are of the same kind as the comparison equations mentioned at the end of section 2.1. The solutions to the original equations are then obtained in a manner similar to that described above, using the correction terms as driving terms, and then applying the variation of parameters method. Neglect of the correction terms mentioned may be equivalent to assuming that the wavelengths of the waves in question are much smaller than free-space wavelengths, and this approximation is then called a quasi-static (electrostatic or magnetostatic) approximation. Usually, however, such an approximation is made in the process of deriving the coupled wave equations to begin with (see section 4.1).

When an inhomogeneous region is bounded by homogeneous regions a kind of Born approximation can be used (BURMAN 1967). First the Green's functions are found for the uncoupled wave equations satisfied in the homogeneous regions. Then a particular solution is found in terms of the fields which enter the equations for the inhomogeneous region as driving terms. This has the form of an integral equation. However, if the medium is slowly varying and only one wave type is incident from the homogeneous region, one particular solution gives the field of the other wave type generated in the inhomogeneity.

In a uniform, homogeneous, linear medium, there is no coupling of normal modes, as was mentioned in section 1.2. In the absence of nonlinearities or random irregularities, the normal modes are only coupled in the presence of an inhomogeneity. Thus it is often advantageous to transform variables so that the only coupling terms arise from derivatives of the parameters of the medium, as in equations 2.2(26) for coupled first-order equations. For example, instead of using equations 1.2(12), one first uses the matrix L_f of 1.2(21). The resulting two second-order equations are then coupled by inhomogeneities in the medium. This is the idea behind Försterling's coupled second-order equations which are used in analyzing wave propagation in the ionosphere. The normal modes are the ordinary and extraordinary electric fields with their associated magnetic fields.

Often the two kinds of modes are strongly coupled only in certain isolated regions. Outside those regions, one can just be concerned with how the inhomogeneity affects the quasi-normal ordinary or extraordinary modes separately (see, for example, BUDDEN 1961). If there is another region where strong reflection of a quasi-normal mode occurs, the WKB phase-integral methods of part 2.4(b)2 may be used. Furthermore, when the extraordinary and ordinary waves do couple, it often happens that reflections are weak. Then it may be possible to model the coupling between the two forward-traveling waves by one single second-order equation using new variables (see HOUGARDY and SAXON 1962). The phase-integral method may then also be adapted to handle such coupling (see BUDDEN 1961 and part 2.4(b)2 below).

2.4 The WKB Method.

(a) Matching quasi-normal mode solutions to asymptotic expansions.

In a critical region, where n WKB solutions of the type of section 2.2 break down, one needs some information about the exact solutions to the corresponding n^{th} order differential equation. This may be the original differential equation, valid over the entire one-dimensional space under consideration, or it may be an approximation valid near the critical region only. Such an approximation may be obtained by replacing the coefficients of each derivative in the differential equation by the first term of a power series for the coefficient, expanded about the critical point. This type of approximation will be particularly useful when the critical region is narrow, since outside this region the WKB quasi-normal mode solutions are satisfactory.

The next step is to try to connect the solutions of this n^{th} order equation to the quasi-normal mode solutions, at the edge of the critical region. Since the quasi-normal mode solutions are written in terms of exponential phase integrals, as in 2.2(14) and 2.2(15), it is natural to look for asymptotic solutions to the differential equation which also can be written conveniently in terms of exponential phase integrals. For this reason, a saddle point method (either the steepest descents or the stationary phase methods) is invariably used, as described below. The solutions are called asymptotic because they become increasingly better approximations to exact solutions as some parameter becomes large. This parameter is usually some normalized measure of distance from the critical point, and often is proportional to some characteristic of the medium such as inverse static field gradient. If this parameter is not large, then the asymptotic solutions are reasonably valid only at large distances from the critical point. In such a case, if the differential equation used to derive the asymptotic solutions becomes a bad approximation far from the critical point, then of course this method breaks down.

To obtain asymptotic solutions by a saddle point method, it is first necessary to write the solutions of the differential equation in the form of a contour integral of the following form:

$$I_\ell(z) = \int_{C_\ell} \exp\,[f(z,t)]\, dt \qquad 2.4(1)$$

A common case where this can be done is when the differential equation can be solved by Laplace's method. (See LANDAU and LIFSHITZ, Quantum Mechanics, mathematical appendix, for example.) In this case, the integral in 2.4(1) is basically a Laplace transform. Any extra terms in the integrand outside the exponential can be absorbed into f(z, t) by first taking the logarithm. Furthermore, almost all common second-order differential equations can be written in terms of the hypergeometric differential equation and its generalizations, such as the confluent hypergeometric equation. Certain higher order equations can also be transformed into generalized hypergeometric equations. In such a case it is possible to express the solutions in terms of Barnes contour integral representations, where the integrand is written in terms of the ratio of products of gamma functions. By taking the logarithm, these terms are absorbed into f(z, t). Since the logarithm of a gamma function is conveniently approximated by Stirling's formula for large values of the argument, it is possible to evaluate 2.4(1) by a saddle point method whenever the contour can be deformed to corresponding regions of t. See Chapter 3 for cases when explicit convergent power series can be obtained instead of asymptotic solutions. A recent book (LUKE 1969) summarizes information on generalized hypergeometric functions and their relations to other functions in an organized and accessible form. See also ERDELYI et al. 1953.

To solve 2.4(1), first find at fixed z the points t_o where

$$f'(t_o) = 0 \qquad 2.4(2)$$

where the prime denotes differentiation with respect to t. In general there will be n such "saddle" points, corresponding to the n WKB quasi-normal mode solutions of an n^{th} order differential equation. Then, following BUDDEN 1961, for example, expand the exponent f(t) in a Taylor series about $t = t_o$:

$$\exp [f(t)] = \exp f_o \exp\left[\frac{1}{2!}(t - t_o)^2 f''_o + \frac{1}{3!}(t - t_o)^3 f'''_o + \cdots\right] \qquad 2.4(3)$$

where $f_o = f(t_o)$, $f''_o = f''(t_o)$, etc. For the method of steepest descents, one writes 2.4(3) as

$$\exp [f(t)] = \exp f_o \exp (-\sigma^2) \qquad 2.4(4)$$

with σ real, while for the method of stationary phase write:

$$\exp [f(t)] = \exp f_o \exp (i\tau^2) \qquad 2.4(5)$$

with τ real. Which method is chosen depends upon the relative convenience of the transformation of variables from $(t - t_o)$ to σ or τ. This change of variables generally involves writing $(t - t_o)$ as a series in

σ or τ, respectively. Since σ and τ are defined by 2.4(4) and 2.4(5) as power series in $(t - t_o)$ through 2.4(3), this last operation involves an inversion of power series, which may be very difficult to do. In this connection, Lagrange's expansion or power series "reversion" formulas may be useful (see ABRAMOWITZ and STEGUN 1964).

Usually there are n or more contours C_ℓ for which the integral $I_\ell(z)$ in 2.4(1) represents a solution of the corresponding n^{th} order differential equation in the independent variable z. If one deforms each contour to follow paths of steepest descent or stationary phase, these paths will pass through up to n of the saddle points defined by 2.4(2). Regardless of whether one uses a steepest descent path as in 2.4(4) or a stationary phase path as in 2.4(5), the first term of the contribution $I_{\ell m}$ to I_ℓ from each saddle point m has the same form

$$I_{\ell m} \sim \pm \frac{(2\pi)^{\frac{1}{2}}}{|A^{\frac{1}{2}}|} \exp[f_o] \exp[\tfrac{1}{2} i(\pi - \alpha)] \qquad 2.4(6)$$

Here A and α are defined through $Ae^{i\alpha} \equiv f_o'' \equiv f''(t_o)$, with $0 \le \alpha < 2\pi$.

The sign of $I_{\ell m}$ depends on the direction θ of the contour at the saddle point m and is positive if $-\frac{1}{2}\pi \le \theta \le \frac{1}{2}\pi$. The factor $\pi^{\frac{1}{2}}$ arises from an integration $\int_{-\infty}^{+\infty} e^{-\sigma^2} d\sigma$ or $\int_{-\infty}^{+\infty} e^{i\tau^2} d\tau$. Note that $\exp[f_o]$ factors out of the integration, since $f_o = f(t_o)$ is not a function of σ or τ. It is this factor which must be identified with some WKB integral exponential $\exp[-j\int^z k_i d\xi]$ as in 2.2(14). The other factors in 2.4(6) turn out to be the WKB amplitudes $B_{ii} = \exp\left[-\int^z (L^{-1}L')_{ii}\, d\xi\right]$ of 2.2(13). Note that the saddle points t_o are functions of the distance of travel z, and hence A and α in 2.4(6) will also be functions of z. See subsection 4.3(d) for a demonstration of these phenomena in a fourth-order differential equation treated by the WKB method.

The central issue of the WKB method lies in finding through which saddle points the various contours C_ℓ can be deformed to pass, as a function of z, and then in identifying the terms 2.4(6) with WKB quasi-normal mode solutions. The solutions to the differential equation are defined in terms of fixed undeformed contours. In general, the same contour C_ℓ will pick up different saddle points on different sides of the critical region, and may represent a linear combination of propagating waves on one side and evanescent waves on the other side. This is a manifestation of Stokes' phenomenon which occurs when a differential equation has an irregular singular point at $z = \infty$. The relation between Stokes' phenomenon and the similar functioning circuit relations for equations with regular points at $z = \infty$ is discussed in section 2.5. The desired total solution then is, in turn, a linear combination of the I_ℓ, determined by the requirement that only one incident wave be present, and that the waves on the opposite side of the critical region from the incident wave must be either evanescent or traveling away from the critical region. See, for example, subsection 4.3(d).

Difficulties arise in this method when the steepest descent or stationary phase contours pass through regions where some essential approximation such as Stirling's formula, used to find the saddle points and the steepest descent or stationary phase contours in the first place, is no longer valid. In any case, the contribution to $I_{\ell m}$ represented by 2.4(6) is only the first term in a series. Unless the distance from the critical region or some characteristic parameter is large, the higher order terms may actually be larger than the first-order terms from other saddle points. These difficulties become particularly troublesome with differential equations of higher order than two (see subsection 4.3(d)). Mathematically, the series which arises for $I_{\ell m}$ in the saddle point method is ordinarily an asymptotic expansion in the Poincaré sense as is shown by Watson's lemma. See, for example, CARRIER, KROOK, and PEARSON 1966. Whenever Stokes' phenomenon appears, these series are divergent. For general mathematical treatments of these problems, and references to earlier work, see HEADING 1962, WILCOX 1964, WASOW 1965, for example.

(b) The Airy equation.

1. Basic solutions.

The most common equation used in the WKB method is the Airy equation

$$F'' - zF = 0 \tag{2.4(7)}$$

where the prime denotes differentiation with respect to z. The solutions of 2.4(7) can be written in terms of powers of z multiplying Bessel functions of order one-third. The well-known asymptotic expansions of the Bessel functions in terms of exponentials can then be used to match to WKB quasi-normal mode solutions. Recent treatments, however, almost universally consider directly the Airy integrals which are the solutions of 2.4(7) by Laplace's method. Then the saddle-point method described following 2.4(1) is used to connect to the WKB solutions. The Airy integrals are more natural solutions since they are analytic about $z = 0$, which is an ordinary point of 2.4(7), whereas Bessel functions of order one-third have a singularity at the origin. Since, however, the Airy equation has an irregular singular point at infinity, the Airy integral solutions exhibit Stokes' phenomenon.

The "dispersion relation" calculated from 2.4(7) as if z were constant gives

$$k^2 = -z , \qquad 2.4(8)$$

showing that for z greater than zero k is imaginary and the wave is cut off in that region. For this reason z = 0 is called a "turning point." The WKB solutions obtained in the manner of section 2.2 represent a wave incident from z_s less than zero and a wave reflected back from z = 0. The total solution for F for z < 0 is then the linear combination (see 2.2(33)):

$$k^{-\frac{1}{2}} \left\{ \exp\left[-j \int_{z_s}^{z} k dz\right] + R \exp\left[+ j \int_{z_s}^{z} k dz\right]\right\} . \qquad 2.4(9)$$

R is the reflection coefficient and the ratio of the amplitudes of the two waves at z_s, the "starting point." Now the Airy integral solution which represents an evanescent wave for z > 0 matches onto 2.4(9) for z < 0, if

$$R = j \exp\left[-2j \int_{z_s}^{z_o} k dz\right] \qquad 2.4(10)$$

where z_o is the place where k(z) = 0. (See BUDDEN 1961, Chapter 16.) Clearly, for a lossless medium, z_o = 0 from equation 2.4(10), so that $|R| = 1$, representing complete reflection of energy. The phase integral in 2.4(10) is just a function of the reference point z_s, just as with reflections on lossless transmission lines. To include the effects of loss, one can in practical problems often replace 2.4(7) by

$$F'' - (z + jZ)F = 0 \qquad 2.4(11)$$

in which case 2.4(10) still holds, but now $z_o = -jZ$. For example, for spin waves Z is proportional to the linewidth, while Z is proportional to collision frequency for radio waves.) Thus to find the magnitude $|R|$ of the reflection coefficient, one must evaluate the integral $\text{Im}\left(\int k\, dz\right)$ not just from the limit $z = z_s$ to $z = 0$, as there is an added contribution along the line $z = 0$ to $z = -jZ$. This added contribution, as well as the factor j in 2.4(10), are the corrections to simple ray-theoretical approaches. The corrections are necessary because ray theory breaks down when the WKB solutions break down, near $z = 0$.

2. Phase-integral methods.

The use of 2.4(10) and variations thereof to find R for various lossless and lossy media is called a "phase-integral method" for determining R. Note, however, that the form of 2.4(10) is strictly correct only when the original differential equation has the form of 2.4(7) or 2.4(11). For other differential equations, a different formula would have to be calculated, based on new asymptotic expansions (see HEADING 1962). Nevertheless, 2.4(10) is often approximately valid for other differential equations when the medium parameters are slowly varying and the branch points of the equation are far apart.

Comparisons of exact "full wave" results and those obtained from phase-integral methods based on the Airy equation (Stokes' equation) are contained in Chapter 20 of BUDDEN 1961, for second-order equations similar to 2.4(7), but with variations of $k^2(z)$ other than that of 2.4(8). When $k^2(z)$ is parabolic instead of linear with z, for example, differences in the results become substantial for certain regions of the medium parameters. To quote Budden, "the phase integral formula is unreliable because the branch points ... are so close together that there is no region between them where the WKB solutions are good approximations." When $k^2(z)$ follows a hyperbolic tangent variation, the formula for the reflection coefficient is reproduced in 3.2(a)5, from an exact solution of the differential equation using hypergeometric functions. Budden shows that the reflection coefficient R found from the phase-integral formula 2.4(10) agrees with the expression from the exact solution when $k^2(z)$ has a slow variation. In applying the phase-integral formula, however, it is necessary to use some intuition to determine around which branch point of the integrand the integral should be taken.

As mentioned previously, it is sometimes possible to reduce, for example, a fourth-order system to one second-order approximating equation in certain regions of space. If this second-order equation is the Airy equation, then 2.4(10) can again be used. For the case of the coupling of extraordinary and ordinary waves in the ionosphere, when reflections are neglected, this indeed is possible after transformation of the independent variable. In that case R represents a transmission coefficient with z_s above the coupling region and 2.4(10) is then called a phase integral formula for coupling. (See BUDDEN 1961, Chapter 20.) In such a case, it is generally assumed that the variable parameter involved has a linear variation with distance z. Since the phase integral formula is not rigorous, it would probably be preferable, if possible, to use exact solutions of the coupled wave equations such as described in parts 3.2(c) and 4.2(b)1 below.

For example, RYDBECK 1967 has shown that the coupling between cyclotron waves and "whistler" modes in four-wave ionized media can in certain regions be described approximately by:

$$F'' + \left(\frac{k_3 - k_1}{2}\right)^2 F = 0 . \qquad 2.4(12)$$

Here k_3, representing an incident cyclotron wave, and k_1, representing a reflected "whistler" mode, are both functions of distance z. At some points z_1 and z_2, called the branch points, $k_3 = k_1$. At these two points the group velocities of the waves go to zero and the waves contain equal amounts of the two types of power flow (kinetic and electromagnetic). Between these two points no energy can propagate because k_1 and k_3 become imaginary. This situation is similar to the turning point problem discussed in section 4.3, except that in that case the waves are cutoff beyond the second branch point also. By analogy with 2.4(7) and 2.4(10), Rydbeck from 2.4(12) takes the coefficient R for energy reflected from a wave incident upon a branch point into a wave of the same type traveling in the opposite direction to be (except for a factor of 2 in the exponent):

$$|R| = \exp\left[-2 \left| \int_{z_1}^{z_2} \mathrm{Im}\left(\frac{k_3 - k_1}{2}\right) dz \right| \right], \qquad 2.4(13)$$

as long as $(k_3 - k_1)^2$ varies linearly with z. It is then assumed that $(1 - |R|^2)$ represents the amount of energy converted to the other wave type. There is no rigorous justification of this assumption or of formula 2.4(13), however. Note also that WKB quasi-normal mode solutions are not valid between the branch points. (See the quotation from BUDDEN 1961 given above.) Results obtained using formula 2.4(13) are compared in section 4.3 with those from the WKB method and from generalized hypergeometric functions.

The whole subject of modeling the coupling involved in n^{th} order equations by extracting different single second-order equations in different regions of space is treated by HEADING 1961. See also HEADING 1967 for a specific application to electroacoustic waves. When reflections and coupling occur in the same region, as for examples in sections 4.2, 4.3, and 4.4, the "embedding method" breaks down. Then solutions of the complete n^{th} order differential equation must be sought. Note in this connection that HEADING 1962 (page 23) showed that the attempt of ECKERSLEY 1950 to establish a phase-integral method for fourth-order equations by generalization of 2.4(10) was without mathematical justification.

(c) Other second-order equations.

Matching of WKB quasi-normal mode solutions to solutions derived from saddle-point integrations has been successful for certain other cases of the equation

$$F'' + [C + g(z)]F = 0 \qquad 2.4(14)$$

besides the Airy equation for which g(z) is linear in z. BUDDEN 1961, for instance, gives several examples. When g(z) is parabolic in z or simply proportional to the square of z, the solutions are parabolic cylinder functions, whose asymptotic expansions match onto WKB solutions. With these functions it is possible to consider two isolated critical regions. See HEADING 1962. When g(z) is exponential in z, Hankel functions result.

Equation 2.4(14) is useful in studies of the vertical propagation of waves in the ionosphere or other medium assumed to be vertically stratified, when the medium is assumed to be isotropic so that a maximum of two waves are coupled. In the case of the ionosphere, this means that the effect of the magnetic field is neglected. Then g(z) is proportional to the density of ionized particles. To consider oblique incidence upon the ionosphere in the x-z plane, it is only necessary to take C in 2.4(14) to be a certain function of the angle of incidence, but independent of z, as long as the wave is polarized with its electric field in the y (horizontal) direction. In these cases, g(z) is just proportional to the square of the refractive index of the ionosphere. If, however, the wave has a component of electric field in the z direction (vertical polarization), then-g(z) is proportional to $n^2(z) - n(z)\left[\frac{1}{n(z)}\right]''$, where n(z) is the refractive index (BREMMER 1958). Then when the refractive index is linear in z, for example, the resulting equation is Whittaker's confluent hypergeometric equation instead of the Airy equation 2.4(7). A list of the cases solved in the literature for vertical polarization is contained in HEADING 1969b, while WESTCOTT 1969a gives all the cases whose solutions can be written in terms of hypergeometric functions, by generalizing a method of HEADING 1965. By using other generalizations of this method, WESTCOTT 1968, and WESTCOTT 1969b, have systematically derived solutions for propagation in spherically stratified and cylindrically stratified media, respectively. These latter solutions are for horizontal polarization only, however.

When the medium is anisotropic, certain reflection, transmission, and absorption processes can still be modeled by 2.4(14) as long as only two of the four or more possible waves are interacting strongly in the region of interest. In many regions of the ionosphere, for example, the extraordinary waves are essentially uncoupled from the ordinary waves. To consider the propagation of the former, g(z) can be considered as proportional to the square of the index of refraction for the extraordinary waves. Now, however, new phenomena occur. For example, at a resonance point g(z) may have a pole, in which case C + g(z) may have an isolated infinity, or a zero and an infinity close together. The solutions for F are then in terms of Bessel functions of order unity, or Whittaker's confluent hypergeometric functions, respectively. STIX 1960 and 1962, has given a treatment of the former case (isolated pole) to try to understand the absorption of energy in an inhomogeneous plasma in plasma heating experiments. VASILE and LAROSA 1968, consider an "isolated pole" case when the extraordinary wave is a circularly-polarized medium-k "magnetostatic wave" in an anisotropic magnetic insulator.

(d) Fourth-order equations.

The seeming possibility that the refractive index of the extraordinary wave may approach infinity raises some interesting problems. First of all, the WKB method applied to the case of an isolated pole mentioned in the last paragraph gives the result that all the energy is absorbed near that point, even when the losses in the medium are negligibly small, and the derivatives of the medium parameters arbitrarily large. Attempts to explain this result included a supposition that most of the energy in the wave is diverted sideways instead of being absorbed (or reflected) when the losses are small (see, for example, BUDDEN 1961, section 21.16). Also, if one allows the refractive index to approach infinity, it turns out that the amount of power radiated by an antenna in such a medium approaches infinity ("the infinity catastrophe"). The September 1966 issue of "Radio Science" was largely concerned with this question.

The resolution of the paradoxes just mentioned lies basically in the fact that the refractive index does not actually go to infinity. Instead, the extraordinary wave couples to waves which have finite but large refractive indexes (or wavenumbers) but which were neglected. When the effects of finite temperature are included, it is seen that the extraordinary transverse electromagnetic waves in a plasma are coupled to longitudinal "Bernstein" (electrokinetic) waves. Similarly, including the effects of exchange in a magnetic insulator shows the possibility of high-k exchange spin waves with different dispersion characteristics than the medium-k (extraordinary electromagnetic) magnetostatic spin waves. Both types of high-k, propagating, coherent waves were known earlier (longitudinal plasma waves: BERNSTEIN 1958; exchange spin waves: KITTEL 1958b), and a qualitative indication of possible coupling between these waves and the corresponding extraordinary waves appeared in GINZBURG 1961, and FLETCHER and KITTEL 1960, respectively. However, it was assumed at first that the high-k spin waves, for example, were excited by a process describable by an Airy equation such as 2.4(7) directly from free-space ordinary electromagnetic waves, by a large net dipole moment where the spin-wave refractive index went to zero (at $z = 0$). (See SCHLÖMANN 1961 and 1964a, and the discussion in subsection 4.3(a).)

The introduction of the high-k wave type prevented the refractive index from going to infinity, but it introduced a situation where it was impossible to avoid a fourth-order system arising from the coupling of the two extraordinary waves with the two high-k waves. The ordinary waves, of course, were neglected long since. To include them again would require a sixth-order system. The first to apply the WKB method to the fourth-order system described above was STIX 1965a and 1965b, by using an equation of WASOW 1950:

$$F'''' + p_o^2 (zF'' + F) = 0 \qquad 2.4(15)$$

Wasow found solutions of this equation using the WKB method for application to the Orr-Sommerfeld theory of hydrodynamical stability. A more flexible equation was solved asymptotically for the same purpose by RABENSTEIN 1958:

$$F'''' + p_o^2 (zF'' + p_1 F' + p_2 F) = 0 \qquad 2.4(16)$$

Several workers have recently applied these results of the WKB treatments by Wasow and Rabenstein to various problems in plasma physics. Since the more recent papers do not refer to each other, it may be helpful to present here a summary. KOPECKY, PREINHALTER, and VACLAVIK 1969 use Rabenstein's results to treat more thoroughly the wave transformations described by STIX 1965b, and by GORMAN 1966. These included transformations from transverse cold plasma waves to short-wavelength Bernstein modes in plasmas with radially decreasing charged particle density, at the upper and lower hybrid resonance frequencies. A magnetic field is assumed to be applied axially in the plasma, perpendicular to the propagation direction. Since the short-wavelength waves are readily absorbed, plasma heating can occur, especially when the ions are involved as at the lower hybrid frequency. Certain "Landauer" radiation observed from a hot plasma can also be explained (see also BEKEFI 1966). Using Rabenstein's results, KOPECKY and PREINHALTER 1969 explain another linear process leading to plasma heating, namely the transformation from electrostatic waves excited by an electron beam into short-wavelength Langmuir waves at the plasma frequency. The waves are assumed to be traveling with a component of the wavevector parallel to the beam, which is also parallel to an applied magnetic field and the direction of the variation in electron density. KUEHL 1967a applied Rabenstein's results to the coupling of extraordinary waves to high-k waves at the upper hybrid resonance frequency near the second electron cyclotron harmonic in a plasma. The waves are assumed to propagate perpendicular to a static magnetic field and to be incident on a nonuniform plasma half-space. In such a case, the second-order equation for the extraordinary waves alone corresponds to 2.4(14), where C + g(z) has a zero close to an infinity. PEARLSTEIN and BHADRA 1969 calculated the dispersion of the upper hybrid mode propagating parallel to a uniform magnetic field but perpendicular to the density variation. Finally, TIMOFEEV 1968 used the results of Wasow to show that certain flute oscillations are stable, without employing expansions in the Larmor radius.

The WKB method for these fourth-order systems is not completely satisfactory, however, since at finite distances from the critical point the error terms in certain saddle-point contributions to the asymptotic expansions of solutions of 2.4(15) or 2.4(16) can be larger than the first term contributions from other saddle points. (Recall that 2.4(6) represented only the first terms in these expansions.) Practically, this means that the WKB method sometimes predicts complete conversion of the extraordinary wave energy into the high-k wave regardless of the steepness of the inhomogeneity, whereas in fact there is some energy reflected back into the medium-k extraordinary wave. A general warning that errors in asymptotic expansions might lead to such problems is given by HEADING 1968. Furthermore, for this same reason, the WKB method applied to the second-order equation 2.4(14) when g(z) has a pole predicted that there never would be any reflected energy, regardless of the gradients in the medium parameters. Hence we are led to look for more exact solutions of fourth-order equations. These solutions, and a detailed discussion of when the WKB method applied to fourth-order systems describable by 2.4(15) and 2.4(16) breaks down, are contained in subsection 4.3(e). Subsection 4.3(d) evaluates the first error terms explicitly.

2.5 Mathematical Concepts Involved in Closed Form Solutions.

The WKB method of section 2.4 involved asymptotic series expanded about $z = \infty$, which was always an irregular singular point, meaning that at least some of the solutions had essential singularities there. Each solution was a linear combination of these asymptotic series whose lowest order term has the form of a WKB partial wave (see equation 2.4(6). In those cases, there were no power series representations of the solutions in positive powers of $1/z$ valid for 2π radians around infinity. At certain angles it then is necessary to change the constants multiplying the various asymptotic series, which is equivalent to changing the linear combination. This feature is called Stokes' phenomenon, and the lines radiating from infinity where these changes occur are called Stokes lines.

In the methods of Chapter 3, however, the solutions are expanded about one or more regular singular points of a differential equation which is a transformation of the original one. In these cases, except for a multiplicative factor, the power series are all convergent and single-valued about the singular point. To make the entire power series single-valued, it is necessary to change the phase of the multiplicative factor at branch cuts. In distinction to the Stokes line, however, a branch cut is encountered only once in a revolution of 2π radians about the singular point for these equations. Furthermore, because these equations are transformed from the original ones, the branch cut can be moved out of the way of the relevant range of the new independent variable.

(a) Singular points of linear differential equations.

To amplify on the above statements, it is helpful to discuss singular points of differential equations in more detail. As a result, it will become clear why hypergeometric equations and their generalizations are almost invariably used when one needs a model which has explicit analytical solutions.

First of all, near an ordinary (nonsingular) point of the linear differential equation:

$$u_o(z)F^{(q)}(z) + u_1(z)F^{(q-1)}(z) + \cdots + u_{q-1}(z)F'(z) + u_q(z)F(z) = 0 \qquad 2.5(1)$$

all the coefficients $u_\ell(z)$ of the various derivative terms are analytic (holomorphic). (A linear equation involves only the first power of the dependent variable F and its derivatives.) Normally the coefficients $u_\ell(z)$ are analytic over almost all of the complex domain of the independent variable. A theorem of differential equations shows that in such regions the solutions can all be written in terms of series in integer powers of the independent variable, expanded about some ordinary point and convergent (and hence analytic) in every neighborhood of that point where the coefficients $u_\ell(z)$ are analytic (see, for example, CODDINGTON 1961, theorem 3.12). These series have the form

$$F_g(z) = \sum_{r=0}^{\infty} \alpha_r (z - z_o)^r \qquad g = 1, 2, \ldots, q \quad 2.5(2)$$

convergent for $|z - z_o| < R$, say, where z_o is an ordinary point and R is a function of z_o. For a q^{th} order equation, there will be q linearly independent solutions of the form of 2.5(2), but with different constants α_r.

At a singular point, however, one or more of the coefficients $u_\ell(z)$

of the differential equation 2.5(1) has a pole or worse behavior. If $u_\ell(z)/u_o(z)$, $\ell = 1, 2, \cdots, q$, has at worst an ℓ^{th} order pole at $z = z_o$, then z_o is called a regular singular point, because none of the solutions of 2.5(1) then has an essential singularity at z_o (Fuch's theorem - see KORN and KORN 1961, section 9.3 - 6). Then about this point the solutions have instead of 2.5(2) the form

$$F_g(z) = (z - z_o)^{\lambda_g} \sum_{r=0}^{\infty} \alpha_r (z - z_o)^r \qquad g = 1, 2, \cdots, q \qquad 2.5(3)$$

where λ_g is not in general an integer. (When λ_g is a negative integer, 2.5(3) represents a Laurent series.) Thus z_o is a branch point of the solutions. For a q^{th} order equation, there will be q possibilities for λ_g, found by solving a q^{th} order polynomial equation called the indicial equation. If two of the roots for λ_g coincide or differ by an integer, then in order to obtain a complete linearly independent set of q solutions, some of the q solutions of the form 2.5(3) must be replaced by other solutions found through the differentiation and limiting processes in Frobenius' method. In these latter solutions the multiplicative factor in front of the infinite series may involve logarithmic functions of $(z - z_o)$.

The differential equation (2.5(1) is said to have a regular singular point at infinity if after the transformation of the independent variable $z = 1/t$ the resulting equation has a regular singular point at $t = 0$. A necessary and sufficient condition for this to be true is that the ratios $u_\ell(z)/u_o(z)$, $\ell = 1, 2, \cdots, q$, must be analytic at $z = \infty$ and have a zero of at least order ℓ there (see theorem 6.2 of CODDINGTON and LEVINSON 1955, Chapter 4). If for all $\ell = 1, 2, \cdots, q$, the zeros are of higher order than ℓ, then $z = \infty$ may even be an ordinary point. If any of the zeros has order less than ℓ, however, then $z = \infty$ is an irregular singular point. When infinity is a regular singular point, the solutions will be of the form of 2.5(3) with $(z - z_o)$ replaced by $1/z$, as long as no two of the roots λ_g are equal or differ by an integer. The infinite series corresponding to 2.5(3) are then all convergent for $|z|$ larger than some finite number.

When infinity is an irregular singular point, the only way to write the solutions in the form of 2.5(3) with $1/z$ instead of $z - z_o$ is to include negative powers of $1/z$ as well as positive powers (KORN and KORN 1961, section 9.3 - 6). Such a form would be useless in solving wave propagation problems, however. What is usually done instead is to attempt to find solutions similar to 2.5(3) with only positive powers in the infinite series. Instead of the factor $(1/z)^{\lambda}$, one tries to obtain factors which can be identified for large z with the WKB quasi-normal mode solutions of section 2.2. (For this purpose, $\alpha_o = 1$.) Solutions in this form are most readily found by the saddle-point integration methods described in section 2.4; the series are then asymptotic expansions:

$$F_h \sim B_{hh}(z) \exp\left[-j \int k_h \, dz\right] \sum_{r=0}^{\infty} \beta_r \left(\frac{1}{z}\right)^r \qquad h = 1, 2, \cdots, q \qquad 2.5(4)$$

The $B_{hh}(z)$ are the WKB amplitudes of equations 2.2(13). Each solution of 2.5(1), for a limited region of large $|z|$, is a certain linear combination of the q almost-linearly-independent solutions 2.5(4).

(b) Stokes' phenomenon

When infinity is an irregular singular point of 2.5(1), the power series in 2.5(4) cannot be convergent in a complete neighborhood of infinity.

In fact, they are divergent series, but the first few terms alone give an increasingly better approximation to the actual function represented by 2.5(4) for increasingly large z. (Ordinary convergent series give an increasingly better approximation at fixed z as the number of terms included in the sum is increased. Series which for a fixed number of terms give an increasingly better approximation as z is increased are called asymptotic series in the Poincaré sense; in certain cases such a series may also be convergent, but not in our cases.) To keep the series in 2.5(4) reasonably good approximations to the F_h as one moves round the neighborhood of infinity, it turns out to be necessary to change the linear combination of the various F_h in each solution. As mentioned in the last subsection, this effect can be visualized as the change in the saddle points through which the contours of the contour integrals 2.4(1) representing each solution may be deformed to pass. Then, for example, the same solution will be given by different linear combinations of WKB solutions as $z \to -\infty$ and $z \to +\infty$; in this way, the reflection and transmission coefficients are determined.

Good treatments of this Stokes' phenomenon for various common differential equations are found in BUDDEN 1961, and HEADING 1962. Both these authors also show when it is necessary to have stricter requirements on the asymptotic series than those of the Poincaré definition, when dealing with wave propagation problems where one must be able to identify the waves. See WILOX 1964, WASOW 1965, and MILLINGTON 1969, for recent discussions of other aspects of Stokes' phenomenon.

(c) Circuit relations.

Since the situation for regular singular points is much better defined, we are led to look for solutions similar to 2.5(4) which are of the form

$$G_h^+(\zeta) = (\zeta)^{-\mu_h} \sum_{r=0}^{\infty} \beta_r (\zeta)^{-r} \qquad h = 1, 2, \cdots, q \qquad 2.5(5)$$

corresponding to a regular singular point at $\zeta = \infty$. Here the $G_h^+(\zeta)$ are to be solutions of some new equation obtained from 2.5(1) by transforming the independent variable z to ζ. Now it is more convenient to look for solutions of differential equations representing wave propagation in media which become homogeneous near $z = \infty$, in contrast to the usual situation in the WKB method where WKB quasi-normal mode solutions as in 2.5(4) are still necessary near $z = \infty$. In other words, we wish $(\zeta)^{-\mu_h}$ to represent a simple, normal mode of the form $\exp[-jk_h z]$ near $\zeta = \infty$. This can be done if ζ represents the following transformation of the independent variable z in 2.5(1):

$$\zeta = -Ae^{az}, \qquad a > 0 \qquad 2.5(6)$$

In this case, $jk_h = \mu_h a$, or

$$\mu_h = jk_h/a \qquad 2.5(7)$$

The factor $(-A)^{-\mu_h}$ multiplies the normal mode.

Furthermore, 2.5(6) shows that as z approaches $-\infty$, ζ approaches zero, whereas ζ goes to minus infinity (if A is positive) when z goes to $+\infty$. Assume that the critical region is centered near $z = 0$. Then in order to obtain reflection and transmission coefficients, it is necessary that the actual physical solution should represent a different linear combination of normal modes for large negative z where there are incident and reflected waves than for large positive z where there are only transmitted (or evanescent) waves. This can be done if $\zeta = 0$ as well as $\zeta = \infty$ is also a regular singular point of the transformed differential equation for $G(\zeta) = F[z(\zeta)]$. Then the series in 2.5(5) will not converge near $\zeta = 0$ and new series must be developed which converge near $\zeta = 0$:

$$G_g^-(\zeta) = (\zeta)^{\lambda_g} \sum_{r=0}^{\infty} \alpha_r \zeta^r \qquad g = 1, 2, \cdots, q \qquad 2.5(8)$$

(Here the $\lambda_g = -jk_g/a$ for z near $-\infty$. Compare equation 2.5(7).)

Although none of the series in 2.5(5) converges over the whole ζ plane, it may be possible to define uniquely, through contour integrals, q solutions which coincide near $\zeta = \infty$ with those of 2.5(5), but which turn out to be various linear combinations of the $G_g^-(\zeta)$ of 2.5(8) near $\zeta = 0$. These relations between the series are called circuit relations and represent analytic continuations of the series. The entire set of these linear transformations between the two sets of n linearly-independent solutions can be expressed in terms of a q x q matrix P, the eigenvalues τ_g of which are equal to exp $[2\pi j\lambda_g]$, when the G_h^+ of 2.5(5) are expressed in terms of the G_g^- of 2.5(8) (see for example KORN and KORN 1961, section 9.3 - 6). Once the circuit relations are obtained, it is an easy matter to find the reflection and transmission coefficients.

From the preceding discussion it is clear that the differential equation for $G(\zeta) = F[z(\zeta)]$ must have more than just one regular singular point in order to be useful. In fact, there are no linear differential equations with just one regular singular point and no other singular points (see, for example, FRIEDRICHS 1965, section V.8).

Furthermore, there is basically only one kind of equation with just two regular singular points and no other singular points. These are called the Euler equations of order q. The two singular points can be assumed to be at zero and infinity, since simple transformations of the independent variable will accomplish this. The theorems quoted before and after equation 2.5(3) then require that the ratio of the ℓ^{th} coefficient to the zeroth coefficient in the differential equation must have a zero of at least order ℓ at $\zeta = \infty$, but a pole of no higher than order ℓ at $\zeta = 0$, $\ell = 1, 2, \cdots, q$. Thus the Euler equation must have the form:

$$\zeta^q G^{(q)}(\zeta) + v_1 \zeta^{q-1} G^{(q-1)}(\zeta) + \cdots + v_{q-1} \zeta G'(\zeta) + v_q G(\zeta) = 0 \qquad 2.5(9)$$

where the v_ℓ are constants. After applying the transform inverse to 2.5(6), we find that the field function $F(z) = G[\zeta(z)]$ satisfies a q^{th} order equation of the form of 2.5(1), where the u_ℓ are all constants. In order to model propagation in an inhomogeneous medium, however, some of the coefficients in the original differential equation for F(z) must be functions of z. Thus we conclude that the Euler equations are useless for our needs.

For example, the second-order Euler equation is

$$\zeta^2 \frac{d^2G}{d\zeta^2} + v_1 \zeta \frac{dG}{d\zeta} + v_2 G = 0. \qquad 2.5(10)$$

Defining the operator

$$\theta \equiv \zeta \frac{d}{d\zeta} \qquad 2.5(11)$$

we find that

$$\theta^2 - \theta = \zeta^2 \frac{d^2}{d\zeta} \qquad 2.5(12)$$

and thus that 2.5(10) can be written in the form

$$(\theta + \rho_1)(\theta + \rho_2) G(\zeta) = 0 \qquad 2.5(13)$$

where

$$\rho_1 + \rho_2 + 1 = v_1 \qquad 2.5(14)$$

and

$$\rho_1 \rho_2 = v_2 \qquad 2.5(15)$$

Writing the equation in terms of θ is useful for two reasons. First, by noting that

$$\frac{d}{dz} = \left(\frac{d\zeta}{dz}\right) \frac{d}{d\zeta} \qquad 2.5(16)$$

and by applying 2.5(6), we find easily that

$$\frac{d}{dz} F(z) = \frac{d}{dz} G[\zeta(z)] = a\zeta \frac{d}{d\zeta} G[\zeta(z)] = a\theta G(\zeta) \qquad 2.5(17)$$

Thus it is seen quickly that when $G(\zeta)$ satisfies 2.5(10), F(z) must satisfy

$$\left(\frac{d}{dz} + a\rho_1\right)\left(\frac{d}{dz} + a\rho_2\right) F(z) = 0 \qquad 2.5(18)$$

which clearly has constant coefficients. Secondly, in preparation for future work with the hypergeometric functions, note that writing an equation such as 2.5(13) in terms of θ facilitates finding the power series solutions in the form of 2.5(8). For, assuming $G(\zeta)$ has that form,

$$\theta G(\zeta) = \zeta \frac{d}{d\zeta}\left[\zeta^{\lambda} \sum_{r=0}^{\infty} \alpha_r \zeta^r\right] = \zeta \frac{d}{d\zeta}\left[\sum_{r=0}^{\infty} \alpha_r \zeta^{(r+\lambda)}\right]$$

$$= \sum_{r=0}^{\infty} (r+\lambda)\,\alpha_r \zeta^{(r+\lambda)} \qquad 2.5(19)$$

where the subscript g on the λ was dropped. Similarly,

$$\theta^2 G(\zeta) = \sum_{r=0}^{\infty} (r+\lambda)^2 \alpha_r \zeta^{(r+\lambda)} \qquad 2.5(20)$$

Thus, it is easy to find the recursion relations that the α_r must satisfy. For equation 2.5(13), in fact,

$$\left[(r+\lambda)^2 + (r+\lambda)(\rho_1+\rho_2) + \rho_1\rho_2\right]\alpha_r = 0, \quad r = 0, 1, \cdots \qquad 2.5(21)$$

Either the term in brackets or the α_r must be zero for all r, and there thus is no real "recursion relation" for α_{r+1} in terms of α_r, etc. Equations 2.5(21) can be satisfied by setting all the $\alpha_r = 0$ for $r \neq 0$, $\alpha_o = 1$, and

$$\lambda^2 + \lambda(\rho_1 + \rho_2) + \rho_1\rho_2 = 0 \qquad 2.5(22)$$

Equation 2.5(22) is the indicial equation for the regular singular point at $\zeta = 0$; because of the form of 2.5(13), it has also the appearance of a dispersion relation. Indeed, equation 2.5(6) taken with 2.5(8) and the results of 2.5(21) shows that

$$F(z) = G[\zeta(z)] = (-A)^{-\lambda} \exp[-jkz] \qquad 2.5(23)$$

when k is defined in terms of λ through a relation similar to 2.5(7).

The solutions for F(z) thus represent normal modes which are nowhere interacting, as was expected from 2.5(18) which has constant coefficients (see section 1.2).

Thus we are led to consider equations for $G(\zeta)$ with three regular singular points: at zero, at infinity, and somewhere else. Differential equations with this property are all called hypergeometric equations for q = 2; higher order equations with three regular singular points behave very similarly and are thus called generalized hypergeometric equations. These equations are equivalent to some interesting equations for F(z) using the transformation 2.5(6); no longer do the resulting equations corresponding to 2.5(18) have constant coefficients. The possible variations of coefficients which are useful and can be generated by hypergeometric equations are described in the latter part of this section. No longer are the recursion relations trivial, but they can be explicitly solved.

Differential equations with more than three regular singular prints are hard to use because the recursion relations are difficult to solve, although they do provide somewhat more variety in the choice of profiles. Heun's equation, for example, is the second-order equation with four regular singular points; its recursion relation for the α_r involves three terms at once. Thus it does not seem to be possible to find explicit solutions for the power series which can be analytically continued to produce reflection and transmission coefficients. Recently, however, (SLUIJTER 1965), Heun's equation arose in the study of an extraordinary wave propagating perpendicularly to a static magnetic field, across an inhomogeneous plasma with a certain variation of electron density. In this case it was apparently possible to match by inspection the linear combinations of power series expanded about $\zeta = 0$ and $\zeta = \infty$, in order to find the reflection and transmission coefficients.

Another possibility is to use differential equations for $G(\zeta)$ which have an irregular singular point at $\zeta = \infty$ as well as a regular singular point at $\zeta = 0$. These equations result from hypergeometric equations or their generalizations whenever one of the three singular points is allowed to coalesce with another; such equations are thus called "confluent" hypergeometric equations. Confluent equations seem to be useful only when the waves are strongly attenuated in the region $z \to +\infty$ ($\zeta \to -\infty$) opposite to the region of the incident wave $z \to -\infty$ ($\zeta \to 0$). The reason for this is that because of the irregular singular point at $\zeta = \infty$, the equation for F(z) represents a medium which remains inhomogeneous as z goes to infinity. Thus it is impossible to find solutions of the form of 2.5(5) representing plane waves for $z \to \infty$. A saddle point method such as is used in the WKB method can, however, be used to the form of the solutions near infinity similar to 2.5(4). Since the physical waves are strongly attenuated, it is only necessary to eliminate the mathematical solutions which represent growing waves; there is no worry about error terms in the asymptotic expansions because in any case they do not represent growing waves. Then near the regular singularity at $\zeta = 0$, the solutions can be found as linear combinations of the normal modes represented by 2.5(8). These normal modes are easily identifiable as the incident and reflected modes in the region $z \to -\infty$, which is homogeneous. The example for second-order equations with this behavior is called the exponential profile. See, for example, BUDDEN 1961. There the functions used are Hankel functions which are a special case of confluent hypergeometric functions. Apparently the only application in the literature of hypergeometric-type equations to the solution of fourth-order wave propagation problems involved such a use of a confluent hypergeometric equation (HEADING and WHIPPLE 1952). The physical problem involved an anisotropic medium with exponential variation of charged-particle density. In Chapter 4 of this work, fourth-order generalized hypergeometric equations, not of the confluent type, are applied to certain problems of finding reflection and transmission coefficients.

(d) Summary

Since the discussion in this section has been concerned largely with singular points, it is important to clarify the overall picture in terms of physical applications. In particular, note that whenever a differential equation such as 2.5(1) is used to find reflection and possibly transmission coefficients for a field F(z) far from a critical region, that equation must have an irregular singular point at infinity. Otherwise, a solution of the form of 2.5(2) or 2.5(3), with $(z - z_o)$ replaced by z^{-1}, will be valid in a complete neighborhood of infinity. This would mean, in turn, that the same linear combination of normal modes which is present at $z = -\infty$ (incidence region) will also be present at $z = +\infty$ (transmission region). Hence there would be no reflection at all. The point is that when the transformation 2.5(6) is used, the irregular singular point at $z = \infty$ is split into two singular points, a regular singular point at $\zeta = 0$ $(z = -\infty)$ and another singular point, usually regular, at $\zeta = -\infty$ $(z = +\infty)$. In this case, however, simple normal mode solutions corresponding to homogeneous regions near $z = -\infty$ result from the necessary form 2.5(8) of the solutions near $\zeta = 0$.

It would be very difficult to modify the transformation 2.5(6) to have the solutions 2.5(8) represent WKB-type quasi-normal modes propagating in an inhomogeneous region near $z = -\infty$, especially with equations of higher order than two. This difficulty is caused by the varying WKB amplitudes $B_{ii}(z)$ as in 2.2(13), multiplying the exponential phase integrals. On the other hand, it is difficult to apply the WKB method to media which become homogeneous outside the critical region, as z goes to minus infinity. This difficulty results from the limitations of Laplace's method used to find solutions of a differential equation in terms of contour integrals 2.4(1). Laplace's method generally works only when the coefficients $u_{\ell}(z)$ of the various derivative terms in the equation 2.5(1) are polynomial functions of the independent variable z. All non-trivial polynomial functions represent inhomogeneities extending to infinity, since they never become constant.

Because the differential equation describing wave propagation does have an irregular singular point at infinity, for both the WKB method and the method of Chapter 3 using generalized hypergeometric equations, there is a basic similarity between the two methods. Both involve contour integral solutions representing different linear combinations of WKB quasi-normal modes or real normal modes on different sides of the critical region near $z = 0$. In the WKB method the different linear combinations result from the Stokes phenomenon representing charges in asymptotic expansions; in the method using non-confluent hypergeometric functions they result from circuit relations representing analytical continuations. Solutions with confluent hypergeometric equations may involve a mixture of both methods.

CHAPTER 3

GENERALIZED HYPERGEOMETRIC FUNCTION SOLUTIONS

Hypergeometric functions and their generalizations seem to be the most useful for solving problems of coupled waves in inhomogeneous media because they not only are relatively simple to apply, but also still allow for a substantial diversity of inhomogeneities. In section 3.1, we outline the mathematical properties of these functions which allow them to be used successfully to find reflection and transmission coefficients. These properties basically have been well known in the mathematical literature, but we formulate them in a way here which allows the functions to be used conveniently in coupled wave problems. Following the concepts of section 2.5, the hypergeometric differential equations are written in a way which allows direct correspondence with the dispersion relations in limiting homogeneous regions. Furthermore, the hypergeometric function solutions are constructed in such a way that they reduce to normal mode solutions in those regions. The major results of 4.1 are the circuit relations 4.1(29) and 4.1(31) which relate the various normal mode solutions, and from which the reflection and transmission coefficients can be calculated.

Section 3.2 contains applications of these solutions to various second-order wave equations, especially that of the electric field in an infinite dielectric with varying permittivity. Although most of the results in 3.2(a) for this problem have been obtained long ago in the literature (EPSTEIN 1930b), they are repeated here for comparison with the new solutions in Chapter 4. In part 3.2(a) 5 we also point out how to find solutions for the derivative field, the magnetic field, for which the wave equation is considerably more complicated. An extension of this simple observation allows hypergeometric functions to be applied to the solution of a much broader class of higher-order equations than would otherwise be possible, as is shown in part (f) of section 3.3.

Furthermore, subsection 3.2(b) presents a new application of hypergeometric equations to nonuniform transmission lines, or equivalently, a dielectric with both varying permittivity and permeability. These solutions are constructed by using a different transformation of the independent variable than is customary. The results for the reflection coefficient are valid from the very slowly varying taper limit all the way to the limit of an abrupt transition between two transmission lines. In subsection 3.2(c) we point out how certain problems written in coupled mode form can be solved using the methods of 3.2(a) and 3.2(b). Solutions are developed which apply when either the propagation "constants" (wavenumbers) or coupling parameters or both are varying quantities. The relations of the solutions described in 3.2 to others in the literature are shown in subsection 3.2(d).

Section 3.3 systematically presents a new method of solving differential equations for reflection and transmission coefficients. Such an equation must contain only one monotonically-varying parameter, but linear combinations of this parameter may occur in any of the coefficients of the various derivative terms. Many solutions found using this method are applied in Chapter 4 to microwave magnetoelastic delay lines with a nonuniform static magnetic field. Section 3.4 extends the results in 3.3 for a fourth-order system in which the varying parameter is not monotonic, but possesses a valley or hump. The solutions in subsection 3.2(a) appear as special cases of the method described in 3.3 and 3.4.

3.1 Generalized Hypergeometric Differential Equations: Basic Properties

(a) Power series solutions for use in reflection and transmission problems.

1. General concepts.

To find the most general differential equation with three regular points, which may be assumed to be at $\zeta = 0, 1,$ and ∞, a systematic method known as Fuch's method may be used. This method is based upon the theorems quoted before and after equation 2.5(3) concerning the requirements on the coefficients of the various derivative terms near a regular point. This same method showed that the Euler equation 2.5(9) was the most general equation with just two regular singular points, at $\zeta = 0$ and ∞. The result for second-order equations with three regular singular points is called the Gaussian hypergeometric equation

$$\zeta(1-\zeta)\frac{d^2G}{d\zeta^2} + [c - (a+b+1)\zeta]\frac{dG}{d\zeta} - abG = 0 \qquad 3.1(1)$$

where a, b, and c are any unequal complex parameters.

As an aid in developing power series solutions for 3.1(1) and for future ease in identifying the normal modes which they represent, introduce again the operator $\theta \equiv \zeta\frac{d}{d\zeta}$ as in 2.5(11). Using 2.5(11) and 2.5(12) shows that 3.1(1) can then be written in the form (COPSON 1935):

$$\theta(\theta + c - 1)\, G(\zeta) = \zeta(\theta + a)(\theta + b)\, G(\zeta) \qquad 3.1(2)$$

By assuming for $G(\zeta)$ an expansion about $\zeta = 0$:

$$G^{-}(\zeta) = (-\zeta)^{\lambda} \sum_{r=0}^{\infty} \alpha_r \zeta^r \qquad 3.1(3)$$

we can then use the properties of the operator θ demonstrated in 2.5(19) and 2.5(20) to derive the recursion relation corresponding to 2.5(21). Setting the coefficient of α_o equal to zero provides the indicial equation similar to 2.5(22), with roots $\lambda = 0$ and $\lambda = 1 - c$. The solution of 3.1(3) with $\lambda = 0$ and $\alpha_o = 1$ is then found using the remaining relations between α_r and α_{r-1}:

$$\alpha_r = \frac{(a+r-1)(b+r-1)}{(c+r-1)\,(r)} \alpha_{r-1} \qquad 3.1(4)$$

This solution is called the hypergeometric series, denoted by

$${}_2F_1\left({a, b \atop c}\Big|\zeta\right) = \frac{\Gamma(c)}{\Gamma(a)\,\Gamma(b)} \sum_{r=0}^{\infty} \frac{\Gamma(a+r)\,\Gamma(b+r)}{\Gamma(c+r)} \frac{\zeta^r}{r!} \qquad 3.1(5)$$

where $\Gamma(a)$ is the gamma function. The gamma function arises from the recursion relation 3.1(4) because of the property

$$\Gamma(a+1) = a\,\Gamma(a) \qquad 3.1(6)$$

which implies

$$a(a+1)\cdots(a+r-1) = \frac{\Gamma(a+r)}{\Gamma(a)} \qquad 3.1(7)$$

For integers r,

$$\Gamma(r+1) = r! = r(r-1)\cdots 2\cdot 1 \qquad 3.1(8)$$

We define $0! \equiv 1$, so that the zeroth order term in 3.1(5) is unity, as was specified.

Generalizations of the hypergeometric series simply involve adding or subtracting factors to the numerator and denominator inside the summation in 3.1(5). For example, a ${}_4F_3$ function has four factors in the numerator and three in the denominator, allowing for seven arbitrary constants like a, b, and c. Those parameters such as a and b which appear in the numerator are added to the top in the symbol corresponding to the left of equation 3.1(5), while denominator parameters such as c appear in the bottom part of the symbol. A simple application of the ratio test shows that 3.1(5) converges for $|\zeta| < 1$, as do all series for ${}_pF_{q-1}$ expanded about $\zeta = 0$, when $p = q$. When $p > q$ the series diverge for all $\zeta \neq 0$, whereas when $p < q$ the series converge for all finite ζ. These latter series are called confluent hypergeometric series. In general, there are q different solutions for λ in 3.1(3). If $p = q$ we can also find series solutions of 3.1(2) and its generalizations expanded about $\zeta = \infty$:

$$G^{+}(\zeta) = (-\zeta)^{-\mu} \sum_{r=0}^{\infty} \beta_r \zeta^{-r} \qquad 3.1(9)$$

For equation 3.1(2), the solutions for μ in this expansion are $\mu = a$ and $\mu = b$.

For the solution of wave propagation problems, we use transformation 2.5(6) : $\zeta = -A\, e^{\alpha z}$. In order to identify normal mode solutions at $\zeta = \infty$ $(z = +\infty)$ and $\zeta = 0$ $(z = -\infty)$ more readily, it is advantageous to write equation 3.1(2) and its generalizations directly in terms of the solutions for λ and μ in 3.1(3) and 3.1(9) at $\zeta = 0$ and (when $p = q$) at $\zeta = \infty$. The reason for this advantage is that, from 2.5(6), $\lambda_g = -jk_g/\alpha$ near $z = -\infty$, and $\mu_h = jk_h/\alpha$ near $z = +\infty$. Thus the solutions λ_g and μ_h are directly related to the solutions for the dispersion relations in k in the homogeneous regions near $z = -\infty$, and (when $p = q$) at $z = +\infty$, when it is required that 3.1(3) and 3.1(9) represent normal modes of the form exp [-jkz] at those limits. Accordingly, we look for solutions of the equation

$$(\theta - \rho_1)(\theta - \rho_2)\cdots(\theta - \rho_q)G(\zeta) = \zeta(\theta + \sigma_1)(\theta + \sigma_2)\cdots(\theta + \sigma_p)G(\zeta) \quad 3.1(10)$$

Now we expect that at the incident region $z = -\infty (\zeta = 0)$ the ρ_g will be solutions for λ in 3.1(3) and that thus

$$\rho_g = \frac{-jk_g^-}{\alpha}, \quad g = 1, 2, \cdots, q. \quad 3.1(11)$$

Similarly, we expect that, when $p = q$, the σ_h will be solutions for μ in 3.1(9) and:

$$\sigma_h = \frac{jk_h^+}{\alpha}, \quad h = 1, 2, \cdots, q. \quad 3.1(12)$$

Extrapolating from the discussion concerning equation 3.1(2), when $p < q$ the series 3.1(3) converge for all finite ζ. Then $\zeta = \infty$ is an irregular singular point of 3.1(10) and no solution such as 3.1(9) exists there. The only way then to find the solutions at $z = \infty$ is to perform saddle-point contour integrations to obtain suitable asymptotic expansions. These points are elaborated below.

2. Solutions near the origin.

To find explicit solutions of 3.1(10) near $\zeta = 0$, apply the θ operators to 3.1(3), obtaining after cancellation of $(-\zeta)^\lambda$:

$$\sum_{r=0}^{\infty} (\lambda+r-\rho_1)(\lambda+r-\rho_2)\cdots(\lambda+r-\rho_q)\,\alpha_r\,\zeta^r$$

$$= \sum_{r=0}^{\infty} (\lambda+r+\sigma_1)(\lambda+r+\sigma_2)\cdots(\lambda+r+\sigma_p)\,\alpha_r\,\zeta^{r+1} \qquad 3.1(13)$$

$$= \sum_{r=1}^{\infty} (\lambda+r-1+\sigma_1)(\lambda+r-1+\sigma_2)\cdots(\lambda+r-1+\sigma_p)\,\alpha_{r-1}\,\zeta^r$$

For r = 0 and assuming $\alpha_o = 1$, we must have

$$(\lambda-\rho_1)(\lambda-\rho_2)\cdots(\lambda-\rho_q) = 0 \qquad 3.1(14)$$

showing that indeed the ρ_g are roots of the indicial equation, which because of 3.1(11) is really a dispersion relation. For r greater than zero we have from 3.1(13) α_r in terms of α_{r-1}:

$$\alpha_r = \frac{(\lambda+r-1+\sigma_1)(\lambda+r-1+\sigma_2)\cdots(\lambda+r-1+\sigma_p)}{(\lambda+r-\rho_1)(\lambda+r-\rho_2)\cdots(\lambda+r-\rho_q)}\,\alpha_{r-1} \qquad 3.1(15)$$

Setting λ equal to one of the roots ρ_g, we find (compare equations 3.1(3), 3.1(4) and 3.1(5)):

$$G_g^-\left(\zeta\middle|{}^{\rho_q}_{\sigma_p}\right) = (-\zeta)^{\rho_g}\frac{\Pi'^{q}_{i=1}\,\Gamma(\rho_g-\rho_i+1)}{\Pi^{p}_{i=1}\,\Gamma(\rho_g+\sigma_i)}\sum_{r=0}^{\infty}\frac{\Pi^{p}_{i=1}\,\Gamma(\rho_g+\sigma_i+r)}{\Pi'^{q}_{i=1}\,\Gamma(\rho_g-\rho_i+1+r)}\,\frac{\zeta^r}{r!}$$

3.1(16)

In this notation $\Pi'^{q}_{i=1}$ denotes a product of all the terms from i = 1 to i = q except for i = g. By comparison of 3.1(16) with 3.1(5) and recalling the discussion about generalized ${}_pF_{q-1}$ functions, we conclude:

$$G_g^-\left(\zeta \middle| {\rho_q \atop \sigma_p}\right) = (-\zeta)^{\rho_g}$$

$$\cdot\ {}_pF_{q-1}\left({\rho_g+\sigma_1,\ \rho_g+\sigma_2,\ \cdots,\ \rho_g+\sigma_p \atop \rho_g-\rho_1+1,\ \cdots\ \rho_g-\rho_{g-1}+1,\ \rho_g-\rho_{g+1}+1,\ \cdots,\ \rho_g-\rho_q+1} \middle| \zeta\right)$$

$$\equiv\ (-\zeta)^{\rho_g}\ {}_pF_{q-1}\left({\rho_g+\sigma_p \atop \rho_g-\rho_q'+1} \middle| \zeta\right) \qquad 3.1(17)$$

where ρ_q' denotes all the ρ_i, $i = 1, 2, \cdots, q$ except for $i = g$. The first term of the ${}_pF_{q-1}$ is unity, since it corresponds to α_o.

3. Solutions near infinity.

When $p = q$, we next need to look for solutions to 3.1(10) of the form 3.1(9) near $\zeta = \infty$. Applying the operator $\theta = \zeta \frac{d}{d\zeta}$ to 3.1(9) we get

$$\theta \sum_{r=0}^{\infty} \beta_r\ \zeta^{-r-\mu} = \sum_{r=0}^{\infty} (-r-\mu)\,\beta_r\ \zeta^{-r-\mu} \qquad 3.1(18)$$

and so on for θ^2, etc. Then instead of 3.1(13) we have, after cancellation of $(-\zeta)^{-\mu}$:

$$\sum_{r=0}^{\infty} (-\mu-r-\rho_1)(-\mu-r-\rho_2)\cdots(-\mu-r-\rho_q)\,\beta_r\,\zeta^{-r}$$

$$= \sum_{r=0}^{\infty} (-\mu-r+\sigma_1)(-\mu-r+\sigma_2)\cdots(-\mu-r+\sigma_q)\,\beta_r\ \zeta^{-r+1} \qquad 3.1(19)$$

$$= \sum_{r=-1}^{\infty} (-\mu-r-1+\sigma_1)(-\mu-r-1+\sigma_2)\cdots(-\mu-r-1+\sigma_q)\,\beta_{r+1}\,\zeta^{-r}$$

The recursion relation is found by again equating coefficients of equal powers of ζ. Now, however, the highest power is ζ^1 corresponding to r = - 1. Hence the indicial equation is now

$$(-\mu+\sigma_1)(-\mu+\sigma_2)\cdots(-\mu+\sigma_q) = 0 \qquad 3.1(20)$$

which is the analog of 3.1(14). Next, by multiplying both sides of 3.1(19) by $(-1)^q$, we obtain the recursion relation

$$\beta_{r+1} = \frac{(\mu+r+\rho_1)(\mu+r+\rho_2)\cdots(\mu+r+\rho_q)}{(\mu+r+1-\sigma_1)(\mu+r+1-\sigma_2)\cdots(\mu+r+1-\sigma_q)}\,\beta_r$$

Reducing the order of the dummy index by one,

$$\beta_r = \frac{(\mu+r-1+\rho_1)(\mu+r-1+\rho_2)\cdots(\mu+r-1+\rho_q)}{(\mu+r-\sigma_1)(\mu+r-\sigma_2)\cdots(\mu+r-\sigma_q)}\,\beta_{r-1} \qquad 3.1(21)$$

which facilitates comparison with 3.1(15) and 3.1(4). Setting μ now equal to one of the roots σ_h, we have

$$G_h^+\left(\zeta\,\Big|\,{\sigma_q \atop \rho_q}\right) = (-\zeta)^{-\sigma_h}\,\frac{\Pi'^{q}_{i=1}\,\Gamma(\sigma_h-\sigma_i+1)}{\Pi^{q}_{i=1}\,\Gamma(\sigma_h+\rho_i)}\sum_{r=0}^{\infty}\frac{\Pi^{q}_{i=1}\,\Gamma(\sigma_h+\rho_i+r)}{\Pi'^{q}_{i=1}\,\Gamma(\sigma_h-\sigma_i+1+r)}\,\frac{\zeta^{-r}}{r!}$$

3.1(22)

Comparison of 3.1(22) with 3.1(16) and the definition 3.1(17) shows that

$$G_h^+\left(\zeta\,\Big|\,{\sigma_q \atop \rho_q}\right) = (-\zeta)^{-\sigma_h}\;{}_qF_{q-1}\left({\sigma_h+\rho_q \atop \sigma_h-\sigma'_q+1}\,\Big|\,\zeta^{-1}\right) \qquad 3.1(23)$$

When no confusion can result, we write in shorthand notation $G_g^-(\zeta)$ and $G_h^+(\zeta)$ for the functions in 3.1(16) and 3.1(22), omitting the parameters.

(b) Contour integral solutions: asymptotic expansions and analytical continuations.

The series in 3.1(16) can be clearly identified in terms of normal modes when ζ approaches zero ($z \to -\infty$) because $G_g^-(\zeta)$ approaches $(-\zeta)^{\rho_g}$ in this limit, and $\zeta = -A\, e^{\alpha z}$ according to the transformation 2.5(6). In order to identify the behavior of these solutions near $\zeta = \infty$ ($z \to +\infty$), however, it is necessary to define contour integral solutions valid over the entire ζ plane which reduce to 3.1(16) near $\zeta = 0$. When $p = q$ these contour integrals represent the analytical continuation of the $G_g^-(\zeta)$, since then 3.1(16) converge only for $|\zeta| < 1$. In this case certain linear combinations of the contour integrals will reduce to 3.1(22) near $\zeta = \infty$ and hence they can also be interpreted as the analytical continuation of the series 3.1(22), which converge only for $|\zeta| > 1$. The resulting relations between the $G_h^+(\zeta)$ and the $G_g^-(\zeta)$ are the circuit relations. These are given in equation 3.1(31). When $p < q$, the series 3.1(16) converge for all finite ζ but are not identifiable in terms of normal modes near $\zeta = \infty$. In that case, these same kind of contour integrals, which reduce to 3.1(16) near $\zeta = 0$, can be used to find asymptotic expansions of the solutions near $\zeta = \infty$. These expansions are then identifiable as linear combinations of evanescent and growing waves because of the nature of the saddle point integration methods used to find them (see section 2.4).

The basis for being able to define such contour integral solutions in a simple way lies in the properties of the gamma function $\Gamma(w)$; namely, $\Gamma(w)$ has simple poles at $w = 0, -1, -2, \cdots$, and the residue at $w = -r$ is equal to $(-1)^r/r!$, r a non-negative integer ($0! \equiv 1$). A systematic notation and a catalog of properties for such integrals has been developed in terms of Meijer's G-function, defined as

$$G_{p,q}^{m,n}\left(\zeta \middle| \begin{matrix} a_1, \cdots, a_p \\ b_1, \cdots, b_q \end{matrix}\right)$$

$$= (2\pi j)^{-1} \int \frac{\Pi_{i=1}^{m} \Gamma(b_i - s) \Pi_{i=1}^{n} \Gamma(1 - a_i + s)}{\Pi_{i=m+1}^{q} \Gamma(1 - b_i + s) \Pi_{i=n+1}^{p} \Gamma(a_i - s)} \zeta^s \, ds \qquad 3.1(24)$$

Here $0 \le m \le q$, $0 \le n \le p$, and an empty product is interpreted as unity. It is also assumed that $(a_h - b_g)$ is not a positive integer, $g = 1, 2, \cdots, m$ and $h = 1, 2, \cdots, n$, so that none of the poles of $\Gamma(b_g - s)$ coincides with a pole of $\Gamma(1 - a_h + s)$. Shorthand notation for the function in 3.1(24) is $G_{p,q}^{m,n}\left(\zeta \middle| \begin{matrix} a_p \\ b_q \end{matrix}\right)$ or simply $G_{p,q}^{m,n}(\zeta)$, used when no confusion can result. For our purposes, when $p = q$ we define a contour $L = L_o$ in 3.1(24) to go from $-j\infty$ to $+j\infty$, wiggling if necessary in such a way that all the poles of $\Gamma(b_g - s)$, $g = 1, 2, \cdots, m$ lie to the right of the path and all the poles of $\Gamma(1 - a_h + s)$, $h = 1, 2, \cdots, n$ lie to the left. When $p \le q$, we can define contours L_g to begin and end at $s = +\infty$, encircling all of the poles of $\Gamma(b_g - s)$, where g is any one number from 1 to q, once in the negative direction, but none of the poles of $\Gamma(1 - a_h - s)$, $h = 1, 2, \cdots, n$. The reason for these choices will become apparent shortly. The integrals defined above converge in the cases of interest to us (see LUKE 1969, section 5.2). Note that $\zeta^s = \exp(s \ln \zeta)$.

If we assume $0 < |\zeta| < 1$, the contour L_o can be closed by drawing a semicircle in the right-half s-plane where $Re(s) > 0$; then the contribution from this semicircle vanishes (Jordan's lemma) as in the theory of inverse Laplace transforms. If $m = 1$, the integration over L_o will give the same result as that over L_1, or by interchange of b_1 with b_g, with the integration over L_g. Then the only poles encountered are those of $\Gamma(b_g - s)$, occurring at $b_g - s = -r$, or $s = b_g + r$, $r = 0, 1, 2, \cdots$, each with residue $(-1)^r/r!$. In this way we obtain a power series proportional to the one in 3.1(16), if we identify $b_g = \rho_g$ and $a_h - 1 = \sigma_h$.

When $p < q$, the series for $G_g^-(\zeta)$ in 3.1(16) converges for all finite ζ, and then the function in 3.1(24) with $m = q$, $n = p$ may be used to determine the asymptotic behavior of various linear combinations of the $G_g^-(\zeta)$ for $\zeta \to \infty$. This is done by using a saddle-point integration along a contour, which like L_g, begins and ends at $s = +\infty$, but encircles all of the poles of $\Gamma(b_i - s)$, $i = 1, 2, \cdots, q$. Since it encircles all of the poles, this contour can be distorted to lie far from any pole, and then Stirling's formula for the gamma function in 3.1(24) can be used. By taking $G_{p,q}^{q,p}\left(e^{n2\pi j}\zeta\right)$ one can obtain three other independent linear combinations of the $G_g^-(\zeta)$. The asymptotic expansions of these linear combinations then show which of them represent decaying waves for $\zeta \to \infty$ ($z \to +\infty$). Finally, by choosing the linear combination of these latter solutions which represents only one incident wave of the desired type for $\zeta \to 0$ ($z \to -\infty$), the final solution is obtained, which contains the reflection coefficients in it. This procedure was used by HEADING and WHIPPLE 1952, in terms of ${}_0F_3(\zeta)$ (or $G_{0,4}^{4,0}(\zeta)$) functions for the oblique reflection of long radio waves in the ionosphere in the presence of a vertical magnetic field. The density of ionized particles was assumed to increase exponentially from a constant value as $z \to -\infty$ to an infinitely large value as $z \to +\infty$.

Usually the functions described in this subsection are of most use when $p = q$. Then the objective is to relate the $G_g^-(\zeta)$ through Meijer's $G_{p,q}^{m,n}(\zeta)$ functions to the series 3.1(22) for $G_h^+(\zeta)$, which converge for $|\zeta| > 1$. Since $\zeta^s = \exp[s \ln \zeta]$, we now must close the contour L_o in the left-hand s-plane, where Re s < 0, so that the contribution from the integration along the infinite semicircle vanishes. Now 3.1(24) can be evaluated by taking the sum of residues as before, except that the poles to be counted now are those of $\Gamma(1 - a_i + s)$ instead of those of $\Gamma(b_i - s)$. That is, we have poles at $s = a_i - 1 - r$ with residue $(-1)^r/r!$ for the function $\Gamma(1 - a_i + s)$. In this way, the same function which involved powers such as $\zeta^{b_g} = \zeta^{\rho_g}$ for $|\zeta| < 1$ involves $\zeta^{-\sigma_h}$ for $|\zeta| > 1$. When $n > 1$, moreover, we will have to consider poles arising from n roots σ_h. Thus the series 3.1(16) are analytically continued outside the unit circle when p = q. Note in this connection that definition 3.1(24) in connection with the above discussion shows that

$$G^{m,n}_{p,q}\left(\zeta \middle| {a_p \atop b_q}\right) = G^{n,m}_{q,p}\left(\zeta^{-1} \middle| {1 - b_q \atop 1 - a_p}\right) \qquad (\arg(\zeta^{-1}) = -\arg\zeta) \qquad 3.1(25)$$

c. Circuit relations.

We now wish to find explicitly how the $G^-_g(\zeta)$ can be written as linear combinations of the $G^+_h(\zeta)$, and vice versa, when p = q. Actually, we are interested in the relationship of the normal modes at $z = \pm\infty$. Because of the minus sign in transformation 2.5(6): $\zeta = -e^{\alpha z}$ (assuming A = 1), the normal modes $\exp(-jk^-_g z)$ are represented by $G^-_g(\zeta)$, as $\zeta \to 0$. Similarly as $\zeta \to -\infty$, the normal modes $\exp(-jk^+_h z)$ are represented by $G^+_h(\zeta)$. (See equations 3.1(11) and 3.1(12) for ρ_g and σ_h in terms of k^-_g and k^+_h.)

Recall that the series in 3.1(16) and 3.1(22) could be expressed in terms of the generalized hypergeometric functions ${}_qF_{q-1}$ in 3.1(17) and 3.1(23), so that:

$$G^-_g(\zeta) = (-\zeta)^{\rho_g} \; {}_qF_{q-1}\left({\rho_g + \sigma_q \atop \rho_g - \rho'_q + 1} \middle| \zeta\right) \qquad 3.1(26)$$

$$G_h^+(\zeta) = (-\zeta)^{-\sigma_h} {}_qF_{q-1}\left({\sigma_h + \rho_q \atop \sigma_h - \sigma'_q + 1}\middle|\zeta^{-1}\right) \qquad 3.1(27)$$

We have also just seen how a Meijer's G-function $G_{q,q}^{m,n}(\zeta)$ with $m = 1$ and $n = q$, say, can be proportional to one $G_g^-(\zeta)$ for $|\zeta| < 1$, and to a linear combination of q $G_h^+(\zeta)$ functions for $|\zeta| > 1$. The explicit relationships between the Meijer's G-functions and the ${}_qF_{q-1}$ functions may be found as special cases of the formulas in LUKE 1969, section 5.2. As a result then of the analytical continuation formula 3.1(25) we will then be able to write (LUKE 1969, equation 5.3.3):

$$\begin{aligned} {}_qF_{q-1}\left({a_q \atop b_{q-1}}\middle|\zeta\right) &= \frac{\Gamma(b_{q-1})}{\Gamma(a_q)} \\ &\cdot \left[\sum_{h=1}^{q} \Gamma(a_h) \frac{\Pi'^{q}_{i=1}\,\Gamma(a_i - a_h)\,(\zeta^{-1} e^{i\pi})^{a_h}}{\Pi^{q-1}_{i=1}\,\Gamma(b_i - a_h)} \; {}_qF_{q-1}\left({1 + a_h - b_{q-1},\ a_h \atop 1 + a_h - a'_q}\middle|\zeta^{-1}\right)\right] \end{aligned} \qquad 3.1(28)$$

where $0 < \arg\zeta < 2\pi$. As before, a_q stands for all the a_i, $i = 1, 2, \cdots, q$; and $\Gamma(a_q)$ stands for a q-fold product of the $\Gamma(a_i)$. The prime denotes that $i = h$ is excluded.

To make correspondence between the functions in 3.1(26) and 3.1(27), we first identify $a_i = \rho_g + \sigma_i$, $i = 1, 2, \cdots, q$ and $b_i = \rho_g - \rho_i + 1$, $i = 1, 2, \cdots, q$, $i \neq g$. After cancelling terms, we then can write

$$G_g^-(\zeta) = \sum_{h=1}^{q} M_{gh}\, G_h^+(\zeta) \qquad 3.1(29)$$

where

$$M_{gh} \equiv \prod_{\substack{i=1\\ i\neq g}}^{q} \left(\frac{\Gamma(1 - \rho_i + \rho_g)}{\Gamma(1 - \rho_i - \sigma_h)}\right) \prod_{\substack{i=1\\ i\neq h}}^{q} \left(\frac{\Gamma(\sigma_i - \sigma_h)}{\Gamma(\sigma_i + \rho_g)}\right) \qquad 3.1(30)$$

To find the inverse relationship, identify $a_i = \sigma_h + \rho_i$, $i = 1, 2, \cdots, q$ and $b_i = \sigma_h - \sigma_i + 1$, $i = 1, 2, \cdots, q$, $i \neq h$, and exchange $\zeta \leftrightarrow \zeta^{-1}$, with a change of argument to keep ζ^{-1} in the range $0 < \arg \zeta^{-1} < 2\pi$. The result is very similar to 3.1(29) and 3.1(30), except that $g \leftrightarrow h$ and $\sigma \leftrightarrow \rho$:

$$G_h^+(\zeta) = \sum_{g=1}^{q} P_{hg} \, G_g^-(\zeta) \qquad 3.1(31)$$

where

$$P_{hg} \equiv \Pi_{\substack{i=1 \\ i \neq h}}^{q} \left(\frac{\Gamma(1 - \sigma_i + \sigma_h)}{\Gamma(1 - \sigma_i - \rho_g)} \right) \Pi_{\substack{i=1 \\ i \neq g}}^{q} \left(\frac{\Gamma(\rho_i - \rho_g)}{\Gamma(\rho_i + \sigma_h)} \right) \qquad 3.1(32)$$

These are the desired circuit relations and form the starting point for the calculation of reflection and transmission coefficients from generalized hypergeometric differential equations of order q. Since the set of $G_g^-(\zeta)$ and $G_h^+(\zeta)$ functions are linearly independent, $P = M^{-1}$, where the matrix components are given by 3.1(30) and 3.1(32).

3.2 Examples of Determination of Reflection and Transmission Coefficients from Second-Order Equations

(a) Equations with one varying parameter.

1. Basic equation.

Now that we have the circuit relations for the solutions of 3.1(10) when p = q, we would like to know what kind of inhomogeneous variations can be modeled by 3.1(10). Consider first the second-order case, where p = 2 = q. In particular, we would be interested in solutions to an equation such as 1.2(12) for a component of the electric field when the medium is isotropic ($\epsilon_{xy} = 0$):

$$\underline{E}_x''(z) + k^2(z)\,\underline{E}_x(z) \;=\; 0 \qquad 3.2(1)$$

where

$$k^2(z) \;\equiv\; \omega^2 \mu_0\, \epsilon(z) \qquad 3.2(2)$$

Then the question is, what kinds of variations of the dielectric constant $\epsilon(z)$ give rise to hypergeometric equations with convenient solutions?

EPSTEIN 1930b was the first to realize that the circuit relations of hypergeometric functions could be used to find reflection coefficients for an inhomogeneous $\epsilon(z)$ in an equation such as 3.2(1). His treatment, however, was in terms of the hypergeometric equation 3.1(2) with solutions ${}_2F_1(\zeta)$. Both in order to transform to a useful equation in z and to derive the analytic continuations, equation 3.1(10) is more convenient. As noted before, its parameters are more closely related to the solutions of the dispersion relation on either side of an inhomogeneous region. The second-order equation 3.1(10) is:

$$(\theta - \rho_1)(\theta - \rho_2)\, G(\zeta) \;=\; \zeta(\theta + \sigma_1)(\theta + \sigma_2)\, G(\zeta) \qquad 3.2(3)$$

where $\theta \equiv \zeta \frac{d}{d\zeta}$.

2. <u>Transformation of independent variable: smooth transition of parameter.</u>

Following EPSTEIN 1930b, we use transformation 2.5(6) : $\zeta = -A e^{\alpha z}$, $\alpha > 0$, and assume

$$\underline{E}_x(z) = G[\zeta(z)] \qquad 3.2(4)$$

Then

$$\underline{E}_x'(z) = \alpha\zeta \frac{d}{d\zeta} G(\zeta) \equiv \alpha\theta G(\zeta) \qquad 3.2(5)$$

where the prime denotes differentiation with respect to z. Using these definitions then, 3.2(3) is equivalent to

$$(1-\zeta)\underline{E}_x''(z) - \alpha\left[\rho_1 + \rho_2 + \zeta(\sigma_1 + \sigma_2)\right]\underline{E}_x'(z) + \alpha^2\left(\rho_1\rho_2 - \zeta\sigma_1\sigma_2\right)\underline{E}_x(z) = 0 \qquad 3.2(6)$$

In order for 3.2(6) to correspond to equation 3.2(1), which has no first derivative term, we must have

$$\rho_1 = -\rho_2 \text{ and } \sigma_1 = -\sigma_2 \qquad 3.2(7)$$

Also we have

$$k^2(z) \equiv \omega^2\mu_o\epsilon(z) = \alpha^2(\rho_1\rho_2 - \zeta\sigma_1\sigma_2)/(1-\zeta) \qquad 3.2(8)$$

Clearly, when ζ goes to zero ($z \to -\infty$), the solution for $k^2(z)$ is $\alpha^2\rho_1\rho_2 = -\alpha^2\rho_1^2$. Similarly, when ζ approaches infinity ($z \to +\infty$), $k^2(z)$ approaches $\alpha^2\sigma_1\sigma_2 = -\alpha^2\sigma_1^2$. Thus, as mentioned before, the ρ_i and σ_i are simply related to the limiting wavenumbers (see equations 3.1(11) and 3.1(12). Further, if we assume $A = 1$ in the transformation

2.5(6), we can write

$$\frac{\zeta}{(1-\zeta)} = -\frac{1}{2}\left(1+\tanh\frac{\alpha z}{2}\right) \qquad 3.2(9)$$

so that 3.2(8) becomes

$$k^2(z) \equiv \omega^2 \mu_o \epsilon(z) = \alpha^2\left[\rho_1\rho_2 + \frac{1}{2}\left(\sigma_1\sigma_2 - \rho_1\rho_2\right)\left(1+\tanh\frac{\alpha z}{2}\right)\right] \qquad 3.2(10)$$

representating a smooth transition from one value of dielectric constant to another.

3. Transformation of dependent and independent variables: symmetric hump or valley in the parameter.

To obtain a different kind of transition, it is necessary to transform the dependent variable as well as the independent variable. Instead of 3.2(4), one then writes

$$G[\zeta(z)] = f(z)\, \underline{E}_x(z) \qquad 3.2(11)$$

and equation 3.2(5) is modified to become

$$\alpha\theta G(\zeta) = f'(z)\, \underline{E}_x(z) + f(z)\, \underline{E}'_x(z) \qquad 3.2(12)$$

Instead of 3.2(4) we obtain, after dividing through by $(1-\zeta)$ and f(z),

$$0 = \underline{E}''_x + \underline{E}'_x\left[2\frac{f'}{f} + \alpha\,\frac{(r_1 - \zeta s_1)}{(1-\zeta)}\right]$$

$$+ \underline{E}_x\left[\frac{f''}{f} + \alpha\frac{f'}{f}\,\frac{(r_1-\zeta s_1)}{(1-\zeta)} + \alpha^2\,\frac{(r_2-\zeta s_2)}{(1-\zeta)}\right] \qquad 3.2(13)$$

where the primes all denote differentiation with respect to z, and

$$r_1 \equiv -\rho_1 - \rho_2, \ r_2 \equiv (-\rho_1)(-\rho_2) \text{ and } s_1 \equiv \sigma_1 + \sigma_2, \ s_2 \equiv \sigma_1\sigma_2 \qquad 3.2(14)$$

The condition that the coefficient of $\underline{E}'_x$ vanish is now equivalent to a differential equation for f(z), whose solution is

$$f[z(\zeta)] = (-\zeta)^{-r_1/2} (1-\zeta)^{(r_1 - s_1)/2} \qquad 3.2(15)$$

As before in equation

$$\zeta = -A e^{\alpha z} \qquad 3.2(16)$$

Evaluating the two derivatives of f(z) by the chain rule and substituting into 3.2(13), we find by comparison with 3.2(1) and 3.2(2) that now $k^2(z)$ can be written in the form

$$k^2(z) = \alpha^2 \left[c_1 + \frac{c_2 \zeta}{1-\zeta} + \frac{c_3 \zeta}{(1-\zeta)^2} \right] \qquad 3.2(17)$$

The coefficients in this partial fraction expansion 3.2(17) are found by multiplying through by $(1-\zeta)^2$ and then equating equal powers of ζ, with the result:

$$c_1 = -\frac{1}{4} r_1^2 + r_2; \quad c_2 = \frac{1}{4}(s_1^2 - r_1^2) + (r_2 - s_2);$$

and 3.2(18)

$$c_3 = -\frac{1}{4}(r_1 - s_1)(r_1 - s_1 + 2)$$

When $r_1 = s_1 = 0$, this result reduces to 3.2(10), for which $f(z) \equiv 1$. However, a new kind of profile is now possible, since

$$\frac{\zeta}{(1-\zeta)^2} = -\frac{1}{4} \operatorname{sech}^2\left(\frac{\alpha z}{2}\right) \qquad 3.2(19)$$

when A = 1 in 3.2(16).

For example, taking $c_2 = 0$ shows that $k^2(z)$ of 3.2(17) can take the form:

$$k^2(z) = \alpha^2\left[c_1 - \frac{1}{4} c_3 \operatorname{sech}^2\left(\frac{\alpha z}{2}\right)\right] \qquad 3.2(20)$$

representing a case where $\epsilon(z)$ has a symmetrical hump (or valley) near $z = 0$. Alternatively, if $k^2(z)$ in 3.2(20) goes negative near $z = 0$, we have a model for a quantum-mechanical tunneling problem. Note that now, however, ρ_i and σ_i, $i = 1, 2$, are not so simply related to the wave-numbers at $z = \pm\infty$, because of the transformation 3.2(11) of the dependent variables. In fact, 3.2(15) and 3.2(11), in conjunction with 3.1(16), show that as $\zeta \to 0$, $\underline{E}_x(z) \xrightarrow[(z \to -\infty)]{} (-\zeta)^{+r_1/2} (-\zeta)^{\rho_i}$. Similarly, using 3.1(22) which shows the limiting behavior of $G(\zeta)$ as $\zeta \to \infty$, we see $\underline{E}_x(z) \xrightarrow[(z \to +\infty)]{} (-\zeta)^{s_1/2} (-\zeta)^{-\sigma_i}$. In order for these limiting forms to represent normal modes traveling in opposite directions at $z \to -\infty$, respectively, we must have

$$(\tfrac{1}{2} r_1 + \rho_1) = \frac{-jk_1^-}{\alpha} = \frac{+jk_2^-}{\alpha} = -(\tfrac{1}{2} r_1 + \rho_2) \qquad 3.2(21)$$

and

$$(\tfrac{1}{2} s_1 - \sigma_1) = \frac{-jk_1^+}{\alpha} = \frac{+jk_2^+}{\alpha} = -(\tfrac{1}{2} s_1 - \sigma_2) \qquad 3.2(22)$$

Both of these relations are satisfied identically, by virtue of the definitions 3.2(14). Note that in general r_1, r_2, s_1, and s_2 must be determined from c_1, c_2, and c_3 through 3.2(18). Since this is an overdetermined system, we can arbitrarily choose one of the parameters. For convenience, choose

$$r_1 = 0 \qquad 3.2(23)$$

Then s_1 through 3.2(18) is given directly in terms of c_3, which from

3.2(20) is proportional to the depth of the valley in $k^2(z)$.

4. General expression for the reflection coefficient.

The reflection coefficients for all these cases are easily found using the circuit relations 3.1(29) and 3.1(31). We choose the σ_h which makes $G_h^+(\zeta)$ represent an evanescent or transmitted wave for $z \to +\infty$ of the form exp [-jkz], where $\mathrm{Re}(k) \geq 0$ and $\mathrm{Im}(k) \leq 0$. In view of 3.2(4) and 3.1(23) this is equivalent to requiring that the σ_h which we choose must have $\mathrm{Re}(\sigma_1) \geq 0$ and $\mathrm{Im}(\sigma_h) \geq 0$ for the case where $k^2(z)$ satisfies 3.2(10). If 3.2(20), or more generally 3.2(17, is satisfied, then we must have instead $\mathrm{Re}\,(s_1/2 - \sigma_1) \leq 0$ and $\mathrm{Im}\,(s_1/2 - \sigma_1) \leq 0$. Label as σ_1 the choice of σ_h which satisfies these requirements. Furthermore, we must identify the ρ_g which causes the $G_g^-(\zeta)$ of 3.1(16) to represent an incident wave. The requirement is $\mathrm{Im}\,(\frac{1}{2} r_1 + \rho_g) < 0$. Call the ρ_g which satisfies this requirement ρ_1. Then ρ_2 corresponds to the reflected wave. Then from 3.1(31) we write

$$G_1^+\left(\zeta\left|{}^{\sigma_1,\sigma_2}_{\rho_1,\rho_2}\right.\right) = P_{11}\, G_1^-\left(\zeta\left|{}^{\rho_1,\rho_2}_{\sigma_1,\sigma_2}\right.\right) + P_{12}\, G_2^-\left(\zeta\left|{}^{\rho_1,\rho_2}_{\sigma_1,\sigma_2}\right.\right) \qquad 3.2(24)$$

The reflection coefficient is simply P_{12}/P_{11}, which from 3.1(32) is

$$R_E \equiv \frac{P_{12}}{P_{11}} = \frac{\Gamma(\rho_1-\rho_2)}{\Gamma(\rho_2-\rho_1)}\,\frac{\Gamma(\rho_2+\sigma_1)}{\Gamma(\rho_1+\sigma_1)}\,\frac{\Gamma(1-\sigma_2-\rho_1)}{\Gamma(1-\sigma_2-\rho_2)} \qquad 3.2(25)$$

(Note that the factor f(z) in 3.2(11) cancels out when evaluating R_E.)

5. Magnitude of reflection coefficient for smooth transition in a lossless medium.

Recall the discussion following 3.2(20) and note that for any $k^2(z)$

satisfying 3.2(9) we must have $r_1 = 0$. Then $\rho_1 = -\rho_2$ and hence in a lossless medium they are complex conjugates: $\rho_1 = \rho_2^*$. To evaluate the magnitude of the reflection coefficient, we also note the property

$$\Gamma(w^*) = [\Gamma(w)]^* \qquad 3.2(26)$$

and $|R|^2 = RR^*$. Consider the case 3.2(10). If $k^2(z)$ goes negative as z approaches positive infinity (cutoff situation), σ_1 and σ_2 are real. Then in 3.2(25) the numerator is the complex conjugate of the denominator so that $|P_{12}/P_{11}|^2 = 1$, as expected with a cutoff when $k^2(z)$ goes through zero. In the absence of a cutoff, $\sigma_1 = -\sigma_2 = \sigma_1^*$. Further, note another property of the gamma function:

$$\Gamma(w)\,\Gamma(-w) = -\frac{\pi}{w}\csc w \qquad 3.2(27)$$

Using 3.1(6) and 3.2(24) and 3.2(25) shows that we can then write

$$\left|\frac{P_{12}}{P_{11}}\right|^2 = \frac{(\rho_1+\sigma_1)\sin[\pi(\rho_1+\sigma_1)]}{(-\rho_1+\sigma_1)\sin[\pi(-\rho_1+\sigma_1)]}\ \frac{(-\rho+\sigma_1)^2}{(+\rho_1+\sigma_1)^2}\ \frac{(\rho_1+\sigma_1)\sin[\pi(\rho_1+\sigma_1)]}{(-\rho_1+\sigma_1)\sin[\pi(-\rho_1+\sigma_1)]}$$

$$= \frac{\sin^2[\pi(\rho_1+\sigma_1)]}{\sin^2[\pi(-\rho_1+\sigma_1)]} \qquad 3.2(28)$$

From 3.2(10) it is clear that α is a measure of the steepness of the transition. Furthermore, from 3.1(11) and 3.1(12) we have

$$\rho_1 = -jk_1^-/\alpha \text{ and } \sigma_1 = +jk_1^+/\alpha$$

Thus when $|k_1^+ - k_1^-| >> \alpha$, which corresponds to a transition which is much wider than any local wavelength, the hyperbolic sines arising from 3.2(28) can be replaced by exponentials, showing that

$$\left|\frac{P_{12}}{P_{11}}\right| \rightarrow \exp(-2k_1/\alpha) \qquad 3.2(29)$$

where k_1 is taken to be the smaller of k_1^+ and k_1^-. Hence as the transition becomes more gradual, α gets smaller, and the reflection goes to zero. On the other hand, if the width of the transition becomes much smaller than any local wavelength, α becomes very large. Then we can replace the hyperbolic sines arising from 3.2(28) by the first terms in their power series, producing

$$\left|\frac{P_{12}}{P_{11}}\right| \rightarrow \frac{|k_1^+ - k_1^-|}{|k_1^+ + k_1^-|} \qquad 3.2(30)$$

Notice that this is just the value of the reflection coefficient 1.2(34) calculated from the boundary conditions of continuous $\underline{E}_x(z)$ and its first derivative, which is proportional to $\underline{H}_y(z)$. Thus the results of the hypergeometric equation include the sharp boundary as a special case, as they should. Note that 3.2(30) is independent of α in this limit.

It is possible to include the effects of loss in the treatment simply by assigning an appropriate imaginary part to the dielectric constant $\epsilon(z)$ in 3.2(2). This in turn produces ρ_i and σ_i with real as well as imaginary parts. The reflection coefficient is still given by expression 3.2(25). Now, however, it is more difficult to simplify.

6. Reflection coefficient for the derivative field.

Recall that the above discussion has been in terms of a component of the electric field, $\underline{E}_x(z)$, which satisfies 3.2(1). As indicated in Chapter 1, the reflection coefficient for the corresponding component of magnetic field, $\underline{H}_y(z)$, must have the same magnitude as that for $\underline{E}_x(z)$, since power flow can be expressed in terms of either $|\underline{E}_x(z)|^2$ or $|\underline{H}_y(z)|^2$. In fact for plane waves, which exist in the above examples at large $|z|$,

$$\langle S_z \rangle = \frac{1}{2}\sqrt{\frac{\epsilon}{\mu}}\,|\underline{E}_x(z)|^2 = \frac{1}{2}\sqrt{\frac{\mu}{\epsilon}}\,|\underline{H}_y(z)|^2 \qquad 3.2(31)$$

where $\underline{E}_y = 0 = \underline{H}_x$. $\langle S_z \rangle$ is the time averaged power flow. Thus we would probably expect that $\underline{H}_y(z)$ satisfies equation 3.2(1) as well as $\underline{E}_x(z)$. The relevant components of the two curl Maxwell's equations in an isotropic medium where $\epsilon_{xy} = 0$ are:

$$\underline{E}'_x(z) = -j\omega\mu_o\,\underline{H}_y(z) \qquad 3.2(32)$$

and

$$\underline{H}'_y(z) = -j\omega\epsilon(z)\,\underline{E}_x(z) \qquad 3.2(33)$$

where we have set $\epsilon(z) \equiv \epsilon_{xx}$. Taking the first derivative of 3.2(33), however, produces a term in $\epsilon'(z)$ as well as one in $\underline{E}'_x(z)$. Substituting 3.2(32) and 3.2(33) into the result produces

$$\underline{H}''_y(z) - \frac{\epsilon'(z)}{\epsilon(z)}\underline{H}'_y(z) + \omega^2\mu_o\,\epsilon(z)\,\underline{H}_y(z) = 0 \qquad 3.2(34)$$

Thus instead of 3.2(1) we have an additional first derivative term.

Moreover, if $k^2(z) = \omega^2\mu_o\,\epsilon(z)$ has the form of 3.2(10) or 3.2(20) or some combination in order that 3.2(1) be soluble in terms of hypergeometric functions, then the $\epsilon'(z)/\epsilon(z)$ term in 3.2(34) will be a very complicated hyperbolic function. In that case, it does not appear that 3.2(34) has solutions of the hypergeometric type. Even making the standard transformation to eliminate the first derivative term in 3.2(34):

$$\underline{H}_y(z) \equiv h(z) \exp\left[\frac{1}{2}\int \frac{\epsilon'}{\epsilon}\, dz\right] = \epsilon^{\frac{1}{2}}\, h(z) \qquad 3.2(35)$$

we only get an equation with a horrible term multiplying h(z):

$$h''(z) + \left[\omega^2 \mu_o \epsilon - \frac{1}{4}\left(\frac{\epsilon'}{\epsilon}\right)^2 + \frac{1}{2}\frac{\epsilon''}{\epsilon}\right] h(z) = 0 \qquad 3.2(36)$$

Thus we conclude that equation 3.2(34) cannot be solved in terms of hypergeometric functions in the manner we have been considering in this subsection, when $\epsilon(z)$ has the form of 3.2(17) or its special cases 3.2(10) and 3.2(20).

There is nevertheless a way to solve 3.2(34), by a technique which is particularly important for similar types of fourth-order equations to be encountered later. We have the solutions of 3.2(1) for $\underline{E}_x(z)$ in terms of hypergeometric functions which satisfy relations such as equations 3.1(29) and 3.1(31). By equation 3.2(32), $\underline{H}_y(z)$ is basically just the derivative $\underline{E}'_x(z)$, since we assume μ_o is constant. Hence, to find the relations for $\underline{H}_y(z)$ corresponding to 3.1(29) and 3.1(31) in the limits of $z \to \pm\infty$, simply take the derivative with respect to z in those limits of the leading terms of the $G^-_g(\zeta)$ and $G^+_h(\zeta)$ appearing in 3.1(29) and 3.1(31). As is seen by reference to equations 3.1(17) and 3.1(22), these leading terms involve the factors $(-\zeta)^{\rho_g}$ and $(-\zeta)^{-\sigma_h}$. According to the transformation 3.2(16) with A = 1, and the identification of ρ_g and σ_h in 3.1(11) and 3.1(12), these two factors are equal to $\exp(-jk^-_g z)$ and $\exp(-jk^+_h z)$, respectively. Hence we have the desired relationship for the normal mode components of $\underline{H}_y(z)$:

$$-\alpha\sigma_h \exp(-\alpha\sigma_h z) \leftrightarrow \sum_{g=1}^{q} P_{hg}\, \alpha\rho_g \exp(\alpha\rho_g z) \qquad 3.2(37)$$

If, as before, we assume that ρ_1 represents the wave incident from $z = -\infty (\zeta = 0)$, the reflection coefficient R_H for the magnetic field is simply

$$R_H = \frac{\alpha \rho_2 \; P_{12}}{\alpha \rho_1 \; P_{11}} = - \frac{P_{12}}{P_{11}} = - R_E \qquad 3.2(38)$$

because $\rho_1 = -\rho_2$ for any profile of $\epsilon(z)$ satisfying 3.2(17) (see 3.2(23) and 3.2(14)). The result in 3.2(38) is just as was expected, with $|R_E| = |R_H|$ to satisfy power flow requirements.

7. Transmission coefficients.

Consider first the smooth transition represented by 3.2(10). Then in the limits of $z \rightarrow \pm\infty$, the circuit relation 3.2(24) becomes, because of 3.1(11) and 3.1(12):

$$\exp(-jk_1^+ z) \leftrightarrow P_{11} \exp(-jk_1^- z) + P_{12} \exp(-jk_2^- z) \qquad 3.2(39)$$

Hence the transmission coefficient T_E for the electric field, with k_1^- representing the incident wavenumber, is simply

$$T_E = 1/P_{11} \qquad 3.2(40)$$

When the more general variation of $k^2(z)$ in 3.2(17) is allowed, the expression for T_E is only multiplied by $\exp(-\alpha s_1 z/2)$, which for lossless media is just a phase factor.

Similarly, for a smooth transition relation 3.2(37) for the magnetic field becomes

$$-jk_1^+ \exp(-jk_1^+ z) \leftrightarrow -jk_1^- P_{11} \exp(-jk_1^- z) - jk_2^- P_{12} \exp(-jk_2^- z) \qquad 3.2(41)$$

Hence

$$T_H = \frac{k_1^+}{k_1^-} \frac{1}{P_{11}} = \frac{k_1^+}{k_1^-} T_E \qquad 3.2(42)$$

Note that the magnitude of T_H is not the same as T_E when $\epsilon^+ \neq \epsilon^-$ because the power flow expressions in 3.2(31) involve ϵ in the numerator when using $|\underline{E}_x(z)|^2$, but in the denominator when using $|\underline{H}_y(z)|^2$. Thus the transmission coefficients must in general be different in order to give the same transmission of power for the same physical problem.

(b) Equations with two simultaneously-varying parameters. The telegraphist's equation. Nonuniform transmission lines.

If the permeability μ as well as the permittivity ϵ vary, then neither the equation for $\underline{E}_x(z)$ nor that for $\underline{H}_y(z)$ has the simple form 3.2(1). In fact, we then have

$$\underline{E}_x''(z) - \frac{\mu'(z)}{\mu(z)} \underline{E}_x'(z) + \omega^2 \mu(z)\, \epsilon(z)\, \underline{E}_x(z) = 0 \qquad 3.2(43)$$

which is the dual of 3.2(24), with the interchange of $\underline{E}_x$ and $\underline{H}_y$ and ϵ and μ. An equation of this type is called a telegraphist's equation, since it is satisfied by the voltage and current on a nonuniform transmission line, with the voltage V replacing $\underline{E}_x$, the impedance Z replacing μ, and the admittance Y replacing ϵ.

Equation 3.2(43) can be solved, however, by using a more general transformation of the independent variable than 3.2(16). Choose

$$\zeta = -\exp\left(\alpha \int_0^z \mu \, d\xi\right), \quad \alpha > 0 \qquad 3.2(44)$$

Then, as long as μ remains positive, this transformation still exhibits the property that as $z \to -\infty$, $\zeta \to 0^-$ and as $z \to +\infty$, $\zeta \to -\infty$. Now, since $d/dz\,(\alpha \int_0^z \mu\, d\xi) = \alpha\mu(z)$, however, the hypergeometric equation 3.2(3) is transformed to the following, instead of 3.2(6):

$$0 = \underline{E}_x'' + \underline{E}_x' \left\{ -\frac{\mu'}{\mu} + \alpha\mu \left[r_1 + (r_1 - s_1) \frac{\zeta}{1-\zeta} \right] \right\}$$

$$+ \underline{E}_x \left\{ \alpha^2 \mu^2 \left[r_2 + (r_2 - s_2) \frac{\zeta}{1-\zeta} \right] \right\} \qquad 3.2(45)$$

where the definitions 3.2(14) are used. If we now set $r_1 = s_1 = 0$, and

$$r_2 = \frac{\omega^2 \epsilon_-}{\alpha^2 \mu_-}, \qquad s_2 = \frac{\omega^2 \epsilon_+}{\alpha^2 \mu_+}, \qquad 3.2(46)$$

3.2(45) can be identified with 3.2(43).

Furthermore,

$$\rho_1 = -\rho_2 = -j \frac{\omega}{\alpha} \sqrt{\frac{\epsilon_-}{\mu_-}}$$

and 3.2(47)

$$\sigma_1 = -\sigma_2 = +j \frac{\omega}{\alpha} \sqrt{\frac{\epsilon_+}{\mu_+}}$$

To evaluate the parameter α in terms of physical quantities, note that in general from 3.2(44) and 3.2(45):

$$\frac{\omega^2}{\alpha^2} \frac{\epsilon(z)}{\mu(z)} = r_2 - \frac{1}{2}(r_2 - s_2)\left[1 + \tanh\left(\frac{\alpha}{2} \int_0^z \mu\, d\xi\right)\right] \qquad 3.2(48)$$

(Compare 3.2(10).) It might be possible to treat variations more general than 3.2(48) if the dependent variable is also transformed (compare 3.2(11) and 3.2(17)). Taking the derivative of 3.2(48) with respect to z at z = 0, and using 3.2(46) we find

$$\alpha = \frac{4\left(\frac{\epsilon}{\mu}\right)'_o}{\mu(o)\left(\frac{\epsilon_+}{\mu_+} - \frac{\epsilon_-}{\mu_-}\right)} . \qquad 3.2(49)$$

Thus we have found solutions of 3.2(43) in terms of r_2, s_2, and α, all of which depend only upon the ratio μ/ϵ which is just the square of the effective impedance Z of the medium. The permeability and permittivity can have any arbitrary variation with $\mu > 0$, as long as their ratio satisfies 3.2(48). In the limit of a sharp boundary, the reflection coefficient reduces to

$$|R_E| = \left|\frac{Z_- - Z_+}{Z_- + Z_+}\right| , \qquad 3.2(50)$$

as can be seen by following the steps from 3.2(28) to 3.2(30). Equation 3.2(50) is of course the same as that obtainable from the boundary conditions, with $Z = \sqrt{\mu/\epsilon}$. This treatment also includes the one already treated with μ a constant, as a special case.

WESTCOTT 1969c has obtained many other solutions of the telegraphist's equation 3.2(43) for special variations using hypergeometric functions, but does not appear to have derived this solution, or to have shown that 3.2(50) holds in the limit of a sharp boundary. His treatment, however, contains other analytical solutions previously derived in the literature as special cases. Note that the treatment of this section and those of WESTCOTT 1969c do not involve any assumption that the parameters of the transmission line are slowly varying. Most treatments of nonuniform transmission lines in the literature do in practice make this assumption. Hence they involve some kind of successive approximations scheme and cannot be used anywhere near the region where 3.2(50) begins to be a good approximation.

(c) Coupled mode equations for inhomogeneous media.

The most general form of the coupled mode equations for two interacting modes a_1 and a_2 is:

$$\begin{aligned} a_1' + j\beta_1 a_1 &= c_{12} a_2 \\ a_2' + j\beta_2 a_2 &= c_{21} a_1 , \end{aligned} \qquad 3.2(51)$$

where the β_i are the propagation constants of the uncoupled modes. The coupling coefficients are related for power conservation by

$$c_{12} = \mp c_{21}^{*} , \qquad 3.2(52)$$

For waves traveling in the same (opposite) direction, use the upper (lower) sign. Usually equations 3.2(51) result from a four-wave problem in which it is assumed that only two waves interact strongly at one time. For active coupling or when energy is transferred to other modes, or is lost, the β_i will be complex (see WALDRON 1967, JOHNSON 1965).

In an inhomogeneous medium, the β's and c's may be functions of distance z. By combining equations 3.2(51) we obtain

$$a_1'' + \left[j(\beta_1 + \beta_2) - \frac{c_{12}'}{c_{12}} \right] a_1' + \left(j\beta_1' - j\beta_1 \frac{c_{12}'}{c_{12}} - \beta_1\beta_2 - c_{12}c_{21} \right) a_1 = 0. \qquad 3.2(53)$$

In Chapter 4 on magnetoelastic delay lines, an equation of this form is solved in terms of hypergeometric functions for coupled magnetoelastic modes. In that case, c_{12} and c_{21} are constants and only one of the β's is varying. For that solution, this β has a variation as in 3.2(10). Note also that the telegraphist's equation 3.2(24) has the same form as 3.2(53), with $\beta_1 = 0 = \beta_2$ and $c_{12} = -j\omega\epsilon(z)$, $c_{21} = -j\omega\mu(z)$. In this connection, compare Maxwell's curl equation 3.2(32) and 3.2(33) with 3.2(51).

Sometimes it may be helpful to simplify the form of 3.2(53). Introduce the new dependent variable $\underline{a}_i$:

$$a_i = \underline{a}_i \exp\left(-j\int_o^z \beta_i \, d\xi\right), \quad i = 1, 2. \qquad 3.2(54)$$

Substituting into 3.2(51) we find

$$\begin{aligned} \underline{a}_1' &= c_{12}\, h(z)\, \underline{a}_2 \\ \underline{a}_2' &= c_{21}\, h^{-1}(z)\, \underline{a}_1 , \end{aligned} \qquad 3.2(55)$$

where

$$h(z) \equiv \exp\left(j\int_o^z (\beta_1 - \beta_2)\, d\xi\right). \qquad 3.2(56)$$

Combining the coupled equations 3.2(55) gives now instead of 3.2(53):

$$\underline{a}_1'' - \left[\frac{c_{12}'}{c_{12}} + j(\beta_1 - \beta_2)\right]\underline{a}_1' - c_{12}c_{21}\,\underline{a}_1 = 0 \qquad 3.2(57)$$

If now we apply a transformation of the independent variable of the form of 3.2(44):

$$\zeta = -\exp\left[\alpha\int_o^z c_{12}\, d\xi\right] \qquad 3.2(58)$$

we obtain from the hypergeometric equation 3.2(3) an equation similar to 3.2(45), except that now we have $\underline{a}_1$ instead of $\underline{E}_x(z)$ and $c_{12}(z)$ instead of $\mu(z)$. This equation may thus be useful in solving coupled wave problems by hypergeometric functions. The problem must be stated, however, in such a way that the modes a_1 and a_2 are coupled in homogeneous regions as well as in the inhomogeneous ones. Otherwise c_{12} will go to zero outside of the transition region, and the transformation 3.2(57) will not have the desired property that $\zeta \to 0^-$ as $z \to -\infty$ and $\zeta \to -\infty$ as $z \to +\infty$. Recall from Chapter 1 that it is always possible to find normal modes which are uncoupled in a homogeneous medium. In the same way, by taking linear combinations of these uncoupled modes we can obtain coupled ones.

See LOUISELL 1955 for an analysis of tapered mode directional couplers, where the propagation constants β_1, β_2 and the coupling coefficients jc are both functions of distance along the coupled transmission lines. His original wave amplitudes are such that they are coupled even in homogeneous regions, provided that the transmission lines are physically coupled there also. He then transformed to normal modes which are coupled only when the propagation and coupling parameters are varying. Then he obtained solutions using a successive approximation method when the parameters are slowly varying. Using transformation 3.2(58), however, we also have a solution for 3.2(57) in terms of hypergeometric functions, provided we take $c_{12} = jc = c_{21}$, $r_2 = -1/\alpha^2 = s_2$, and

$$-\frac{\beta_1 - \beta_2}{j\alpha c} = r_1 - \frac{1}{2}(r_1 - s_1)\left[1 + \tanh\left(\frac{j\alpha}{2}\int_0^z c\, d\xi\right)\right] \qquad 3.2(59)$$

(Compare 3.2(44), 3.2(45), and 3.2(48) with 3.2(58), 3.2(57), and

3.2(59).) Energy conservation requires that c be real (see 3.2(52)), so α is then required to be imaginary to allow 3.2(58) to be a useful transformation. Note that it may be possible to use variations more general than 3.2(59) if the dependent variable is transformed further (compare 3.2(11) and 3.2(17).

In these coupled mode problems, the circuit relations 3.1(29) in terms of the matrix elements M_{gh} are more useful than those in terms of P_{hg}. Now we start with an incident wave at $z = -\infty$ and wish to find how much energy is transferred into the two transmitted waves. Note that reflections are ignored in the second-order treatment. When the parameters change very quickly, reflections will become important, and it is then necessary to use a complete four-wave treatment to obtain the reflection coefficients as well as the transmission coefficients. An example of such a treatment in connection with coupled magnetoelastic waves is given in Chapter 4.

(d) Summary of related solutions in the literature.

As mentioned earlier, the application of the circuit relations for second-order hypergeometric functions to the solution of the wave propagation equation 3.2(1) was first done by EPSTEIN 1930b. His treatment used the transformation of the independent variable 3.2(16), $\zeta = -A\,e^{\alpha z}$, and also included the transformation of the independent variable corresponding to equation 3.2(15). Thus the most general profile for $k^2(z)$ handled in his analysis was that of our equation 3.2(17). A summary of his treatment, including graphical and numerical results for the special cases 3.2(1) and 3.2(20) is contained in BREKHOVSKIKH 1960 for example. This latter treatment also compares the results for reflection coefficients obtained from the hypergeometric equation with those obtained from the WKB method, using a parabolic variation of $k^2(z)$ sandwiched between two homogeneous regions. HEADING 1965 gives a systematic survey of the applications of hypergeometric equations to problems of the propagation of horizontally polarized waves in a vertically stratified ionosphere, including problems involving various confluent hypergeometric functions. Because they have an irregular singular point at $\zeta = \infty$, however, such confluent hypergeometric equations must be solved by the WKB method of section 2.3 rather than by the methods of this section.

WESTCOTT 1969a has given a similar treatment of the use of hypergeometric equations in problems involving vertically polarized waves propagating in an inhomogeneous but isotropic ionosphere. As mentioned under part (c) of section 2.4, the equation for such waves is more complicated than that for a horizontally polarized wave in the presence of the same variation in ion density (see BREMMER 1958). HEADING 1969b uses these solutions together with those derived for horizontally polarized waves to find the polarization of reflected waves when the incident waves are elliptically polarized (a combination of horizontal and vertical polarization).

Finally, the papers WESTCOTT 1968 and WESTCOTT 1969b apply the hypergeometric equation to find solutions for electromagnetic wave propagation in spherically stratified iostropic media, and for transverse (ordinary) electromagnetic wave propagation in a cylindrically stratified axially magnetized plasma, respectively. The propagation of extraordinary waves in the latter situation has been discussed in terms of the WKB method, section 2.4(c).

PEARLSTEIN 1965 solves a problem in plasma instability using an equation such as 3.2(1) with $k^2(z)$ having the general variation 3.2(17). His results are then compared with those derived using the exponential phase integral of WKB solutions in a manner such as is commonly used to find ground states to Schrodinger's equation with slowly varying potentials. Solutions corresponding to eigenstates of the potential well are stable; others are unstable since they grow exponentially in the "cutoff region" (they do not have "energy" equal to an "eigenvalue"). Under certain assumptions, Pearlstein shows that the two methods yield approximately the same requirement for stability. Clearly, several different kinds of potential wells can be modeled by 3.2(17) since that form allows for a symmetrical well ($-\operatorname{sech}^2(\alpha z/2)$ - equation 3.2(20)) with an added variation which can make one side higher than the other (the $\tanh(\alpha z/2)$ term of equation 3.2(10). No applications to quantum mechanical problems of the hypergeometric equation seem to be generally known, however, since usually the WKB treatments are good approximations, unless the potential is quickly varying.

Until now, all the discussion has been in terms of treatments which have used the transformation of the independent variable of equation 3.2(16): $\zeta = -Ae^{\alpha z}$. More general transformations are possible which can give useful results without too much labor for profiles of $k^2(z)$ more general than 3.2(17). RAWER 1939, for example, considers transformations which reduce to 3.2(16) in the limits $\zeta \to 0$ ($z \to -\infty$) and $\zeta \to \infty$ ($z \to +\infty$). BURMAN and GOULD 1965 have introduced the transformation $\zeta = -\frac{1}{2}(e^{az} + e^{bz})$. Using this transformation, they were able to treat a case where $k^2(z)$ tends to the same limit as z goes to plus or minus infinity, but has an asymmetrical profile near $z = 0$, in distinction to the $\operatorname{sech}^2(\alpha z/2)$ profile of 3.2(20), which is symmetric. HEADING 1967a has found a solution for a similar profile in $k^2(z)$ by systematically deriving the profiles which result from more general transformations than 3.2(16). His treatment is similar to RAWER 1939, but he gives more explicit solutions, particularly for a special case where $k^2(z)$ has both a barrier and a well. Also, many of the solutions derived by WESTCOTT 1969c for nonuniform transmission lines using hypergeometric functions involved more general transformations.

3.3 Differential Equations of Arbitrary Order Having One Varying Parameter with a Smooth Transition: A "Standard Procedure" for Finding Reflection and Transmission Coefficients

In this section, a new method is described for solving in an orderly manner certain wave equations for multi-wave inhomogeneous media, using generalized hypergeometric functions. All of the solutions described in section 3.2 which involved transformation of the independent variable only can be treated by special cases of this step-by-step method. Applications of this method to second-, fourth-, and sixth-order equations are given in sections 4.2, 4.3, and 4.4.

(a) Identification of the related differential equation for a homogeneous medium.

Many equations, especially those of higher order than second, are too complicated in form to be transformed conveniently to a generalized hypergeometric equation. Not only do they contain a varying parameter p(z) such as permittivity, but they also contain terms with the derivative of this parameter. See, for example, equations 3.2(34) and 3.2(43) which are of the telegraphist's type.

In a homogeneous region where the parameter p(z) is constant, however, these derivative terms drop out and an equation of much simpler form such as 3.2(1) results, with constant coefficients. Such an equation can then easily be solved by assuming normal mode solutions of the form exp (-jkz), where k is a constant, inversely proportional to the wavelength. In this manner, the equation is transformed to a polynomial equation in k, which is equivalent to the dispersion relation of the system. When this dispersion relation is biquadratic in k, it expresses the property that the system allows two normal modes traveling in opposite directions with the same wavelength for each solution for k^2.

When the parameter p(z) in the original differential equation has a "smooth transition" variation of the hyperbolic tangent form of 3.2(10), then as z approaches plus or minus infinity, the variation ends and any terms involving derivatives of the parameter become negligible. Hence in these infinite limits the original equation approaches the form of the one with constant coefficients described above. What is done in the method described in this section, however, is to look for solutions of the original differential equation without the terms involving derivatives of the parameter p(z), but still allowing the parameter itself to vary as in 3.2(10). This new equation will be labeled the "source equation."

The "source equation" has the same form as the one satisfied by the real physical system in homogeneous regions, and hence resembles the dispersion relation in form, but it does not have constant coefficients. Thus its solutions may not correspond to any physical fields. Often, however, solutions of the original differential equation can be expressed as linear combinations of various derivatives of the solutions of this "source equation." Thus solutions of 3.2(34) for $\underline{H}_y(z)$ could be expressed as the first derivative of solutions of 3.2(1), for example. See part (f) below for an outline of this technique in more general cases.

(b) <u>Transformation of the "source equation" into a generalized hypergeometric differential equation.</u>

The "source equation" has the form

$$\sum_{m=0}^{q} w_m(z)\, F^{(q-m)}(z) = 0 , \qquad 3.3(1)$$

where q is the order of the differential equation, $w_o(z)$ is chosen to be unity, and $F^{(n)}(z)$ denotes the n^{th} derivative of F(z) with respect to z. If we call the varying parameter p(z), then any of the $w_m(z)$, $m > 0$, may have any linear combination of p(z), but none of its derivatives. The dispersion relation corresponding to 3.3(1) is

$$\sum_{m=0}^{q} w_m \, k^{q-m} = 0 , \qquad 3.3(2)$$

where the w_m are evaluated in homogeneous regions. For systems with dispersion relations which are biquadratic in k, all the $w_m(z)$ with odd m must be zero.

A generalized hypergeometric equation of order q can be written in the form

$$(\theta - \rho_1)(\theta - \rho_2)\cdots(\theta - \rho_q)\,G(\zeta) = \zeta(\theta + \sigma_1)(\theta + \sigma_2)\cdots(\theta + \sigma_n)\,G(\zeta), \qquad 3.3(3)$$

where as before in 3.2(3):

$$\theta \equiv \zeta \frac{d}{d\zeta} . \qquad 3.3(4)$$

An equation of the form of 3.3(3) with $q = 4$, $n = 0$ was used by HEADING and WHIPPLE 1952 in the solution of a coupled wave problem in the ionosphere, with p(z) representing electron density increasing exponentially with altitude, in the presence of a magnetic field (see also HEADING 1955). The right-hand side of 3.3(3) was then just $\zeta G(\zeta)$. No other solution of a coupled wave problem for reflection and transmission coefficients seems to have been given in the literature for $q > 2$. For media which become homogeneous as z approaches positive or negative infinity, so that p(z) becomes constant in those limits, we must choose $n = q$. This must be true to apply the method of this section.

The simplest transformation of the independent variable which will transform 3.3(3) into the form 3.3(1) is as in 3.2(16):

$$\zeta = -e^{\alpha z} \qquad 3.3(5)$$

As was done in section 3.2 in equation 3.2(4), we consider G to be a composite function of z and write

$$F(z) \equiv G[\zeta(z)]. \qquad 3.3(6)$$

Then as in 3.2(5) we find that

$$F^{(1)}(z) = \alpha\zeta \frac{d}{d\zeta} G(\zeta) = \alpha\theta\, G(\zeta), \qquad 3.3(7)$$

using the definition 3.3(4). By repeated differentiation of 3.3(7) we can write in general, for any m:

$$F^{m}(z) = \alpha^{m}\, \theta^{m}\, G(\zeta), \qquad 3.3(8)$$

since θ commutes with itself.

Proceeding then by collecting terms in equal powers of θ, 3.3(3) can be rewritten in the form, when n = q:

$$\sum_{m=0}^{q} (r_m - \zeta s_m)\, \theta^{q-m}\, G(\zeta) = 0, \qquad 3.3(9)$$

where r_m and s_m are defined as the sum of all the distinct products of the quantities $(-\rho_i)$ and σ_i, respectively, taken m at a time. Thus

$$r_1 = \sum_{i=1}^{q} (-\rho_i), \quad r_2 = \sum_{h=1}^{q} \sum_{i=h+1}^{q} (-\rho_h)(-\rho_i),$$

etc. We define $r_o = 1 = s_o$. Note now that $(1 - \zeta)$ never goes to zero for any real value of z under the transformation 3.3(5) with real α. Thus, dividing 3.3(9) through by $(1-\zeta)$ and also multiplying by α^q, we get

$$\alpha^q \sum_{m=0}^{q} \left[r_m + (r_m - s_m)\frac{\zeta}{1-\zeta}\right] \theta^{q-m}\, G(\zeta) = 0. \qquad 3.3(10)$$

Using 3.3(8), this may be rewritten in terms of an equation for F(z):

$$\sum_{m=0}^{q} \alpha^m \left[r_m + (r_m - s_m)\frac{\zeta}{1-\zeta}\right] F^{(q-m)}(z) = 0. \qquad 3.3(11)$$

Now in view of 3.3(5)

$$\zeta/(1-\zeta) = -\frac{1}{2}\left(1 + \tanh\frac{\alpha z}{2}\right), \qquad 3.3(12)$$

so that for 3.3(11) to correspond to 3.3(1) we must have for each $w_m(z)$, $m > 0$:

$$w_m(z) = \alpha^m \left[r_m + \frac{1}{2}\left(s_m - r_m\right)\left(1 + \tanh\frac{\alpha z}{2}\right)\right]. \qquad 3.3(13)$$

Thereby the transformation from 3.3(3) to 3.3(1) is complete.

Since there is only one available hyperbolic tangent function with one α, we see that we can handle only one varying parameter p(z), but that it can appear in as many of the coefficients $w_m(z)$ as desired, combined with various constant parameters r_m and s_m. Note that the hyperbolic tangent represents a "smooth transition." By transforming the dependent variable also as in 3.2(11) it is possible to obtain in certain cases a valley or hump in a $w_m(z)$ and hence in p(z). An example for q = 4 is treated below in section 3.4. The circumstances under which this can be done are more restricted, however, than for the second-order case q = 2 described in part 3.2(a) 3.

(c) Identification of parameters in the hypergeometric equation from the dispersion relation in the limiting homogeneous regions of the medium.

Note from 3.3(13) that in the limits $z \to \pm\infty$, the coefficients w_m become

$$w_m^- = \alpha^m r_m , \qquad w_m^+ = \alpha^m s_m , \qquad 3.3(14)$$

where the superscripts ± refer to $z = \pm\infty$. Thus the r_m and s_m are related to the coefficients in the limiting dispersion relations of the form 3.3(2). Recall that the r_m and s_m are defined in terms of the ρ_i and σ_i, following 3.3(9).

Now label the solutions of the dispersion relations k_i^- and k_i^+, $i = 1, 2, \cdots, q$. Then 3.3(2) can be rewritten in the form at $z = -\infty$

$$(k^- - k_1^-)(k^- - k_2^-)\cdots(k^- - k_q^-) = 0 , \qquad 3.3(15)$$

and similarly at $z = +\infty$. Recall from 3.3(5) that $z \to -\infty$ and $z \to +\infty$ correspond to $\zeta \to 0$ and $\zeta \to -\infty$, respectively. Observing that the hypergeometric differential equation 3.3(3) looks very much like 3.3(15) in those limits, we see that the ρ's and σ's must be related to the k_i^- and k_i^+, respectively Since in the limits we are assuming normal mode solutions of the form

$$F^-(z) = \exp(-jk^- z) , \qquad 3.3(16)$$

we see from 3.3(7) that

$$\theta G^-(\zeta) = \frac{jk^-}{\alpha} F^-(z) , \qquad 3.3(17)$$

and similarly for the + superscripts. Hence comparison of 3.3(3) and 3.3(15) now gives the explicit relations:

$$\rho_i = -jk_i^-/\alpha, \quad \sigma_i = +jk_i^+/\alpha; \qquad i = 1, 2, \cdots, q. \qquad 3.3(18)$$

These relations are the same as in 3.1(11) and 3.1(12), where it was shown how the leading terms in the power series solutions of 3.3(3) in the limits $\zeta = 0$ and $\zeta = -\infty$, $G_g^-(\zeta)$ and $G_h^+(\zeta)$ respectively, could represent normal modes.

(d) Identification of the transformation parameter in terms of the "width of the transition."

There remains one more parameter to be determined in the transformation from 3.3(3) to 3.3(1). This is α, arising in the transformation 3.3(5) of the independent variable. To accomplish this identification, take the derivative of 3.3(13):

$$w_m'(z) = \frac{1}{4}\alpha^m (s_m - r_m)\, \alpha\, \mathrm{sech}^2 (\alpha z/2) \qquad 3.3(19)$$

where the prime denotes differentiation with respect to z. Using relations 3.3(14) and evaluating 3.3(19) at z = 0, we obtain:

$$\alpha = 4 w_m'(0)/(w_m^+ - w_m^-) \qquad 3.3(20)$$

for every $w_m(z)$ which is not identically constant. Relation 3.3(20) emphasizes the fact that if the parameter p(z) occurs within more than one coefficient $w_m(z)$ of the original differential equation, it can at most be multiplied by a different constant and/or added to a different constant each time. Reference to figure 1 in connection with 3.3(20) shows that α is inversely proportional to the effective length L of the hyperbolic tangent transition. Thus it is an effective "transition wavenumber."

Sometimes it is convenient to express α in terms of the gradient of $w_m(z)$ at a critical point which is not located at the center (half-way point) of the transition, $z = 0$. From 3.3(19) we have, dropping the subscript m:

$$w'(0) = \cosh^2(\alpha z/2)\, w'(z). \qquad 3.3(21)$$

Furthermore, using the relations 3.3(14) write 3.3(13) in the form

$$\tanh\left(\frac{\alpha z}{2}\right) = 2\left(\frac{w(z) - w^-}{w^+ - w^-}\right) - 1. \qquad 3.3(22)$$

Now using the identity $\cosh^2 x = (1 - \tanh^2 x)^{-1}$, substituting 3.3(22) into 3.3(21) and that in turn into 3.3(20):

$$\alpha = w'(z)/\left[(w(z) - w^-)\left(1 - \frac{w(z) - w^-}{w^+ - w^-}\right)\right]. \qquad 3.3(23)$$

If $w(z)$ is close to w^- then this can be approximated by

$$\alpha \approx w'(z)/(w(z) - w^-), \qquad 3.3(24)$$

regardless of the value of w^+, as long as $w(z) - w^- \ll w^+ - w^-$. This feature is a result of the fact that the hyperbolic tangent behaves for large arguments like the difference between unity and an exponential. By writing 3.3(22) in a slightly different form, α can be written as in 3.3(23) with $w^+ - w(z)$ replacing $w(z) - w^-$. Then for z such that $w^+ - w(z) \ll w^+ - w^-$ we have α expressed in the same form as 3.3(24), but with $w^+ - w(z)$ replacing $w(z) - w^-$.

(e) Solution of the "source equation" in terms of the circuit relations of hypergeometric functions.

Through equations 3.3(18) and 3.3(20) we have now related the parameters ρ_i and σ_i of the hypergeometric equation 3.3(3) to the characteristics of the medium corresponding to the "source equation" 3.3(1). Furthermore, the solutions $G[\zeta(z)]$ of 3.3(3) are identical to the solutions F(z) of 3.3(1) by virtue of 3.3(6). Consequently, the solutions can be conveniently expressed by the $G_g^-(\zeta)$, which are related to the $G_h^+(\zeta)$ solutions by the circuit relations 3.1(29). The inverse relationships are given in 3.1(31). Recall from expressions 3.1(16) and 3.1(22) together with 3.3(18) and 3.3(5) that the leading term in each $G_g^-(\zeta)$ and $G_h^+(\zeta)$ is equal to a normal mode of the form exp (-jkz) for z approaching negative and positive infinity, respectively.

(f) Extraction of the solutions of the original equation from those of the "source equation."

When the original differential equation contains derivatives of a parameter p(z), its solutions H(z) can often be written in terms of a linear combination of derivatives of the solutions F(z) of the "source equation" (3.3(1). By construction, this equation contains only p(z) but no terms involving p'(z) or higher derivatives. Assume that the original equation contains derivatives of p(z) up to a maximum of order ℓ. Then it is necessary that the solutions H(z) include at least a term in $F^{(\ell)}(z)$. For example, consider the second-order equation

$$H_1^{(2)}(z) - \frac{p^{(1)}(z)}{p(z)} H_1^{(1)}(z) + p(z) H_1(z) = 0 \qquad 3.3(25)$$

and the corresponding source equation

$$F_1^{(2)}(z) + p(z)\, F_1(z) = 0 . \qquad 3.3(26)$$

(Compare equations 3.2(34) and 3.2(1) for magnetic and electric fields.) Now take the derivative of 3.3(26):

$$F_1^{(3)}(z) + p(z)\, F_1^{(1)}(z) + p^{(1)}(z)\, F_1(z) = 0. \qquad 3.3(27)$$

Substituting for $F_1(z)$ from 3.3(26) and 3.3(27) we find:

$$F_1^{(3)}(z) - \frac{p^{(1)}(z)}{p(z)}\, F_1^{(2)}(z) + p(z)\, F_1^{(1)}(z) = 0 . \qquad 3.3(28)$$

This now has the same form as the original equation 3.3(25). Clearly we can take

$$H_1(z) = F_1^{(1)}(z). \qquad 3.3(29)$$

The same procedure applied to the fourth-order equation:

$$H_2^{(4)} - \frac{p^{(1)}}{p}\, H_2^{(3)} + (a_2 + p)\, H_2^{(2)} - \frac{p^{(1)}}{p}\, a_2\, H_2^{(1)} + b_2 p H_2 = 0 \qquad 3.3(30)$$

with the source equation

$$F_2^{(4)} + (a_2 + p)\, F_2^{(2)} + b_2\, p F_2 = 0 \qquad 3.3(31)$$

again gives the result that

$$H_2(z) = F_2^{(1)}(z) \qquad 3.3(32)$$

where a_2 and c_4 are some constants.

Consider now an equation containing the second derivative of p(z):

$$H_3^{(4)} + (a_2 + p) H_3^{(2)} + 2 p^{(1)} H_3^{(1)} + (p^{(2)} + b_2 p) H_3 = 0 . \qquad 3.3(33)$$

Eliminating the terms in $p^{(1)}$ and $p^{(2)}$, we see that $H_3(z)$ has the same source equation 3.3(31) as $H_2(z)$. Taking the second derivative of 3.3(31):

$$F_2^{(6)} + (a_2 + p) F_2^{(4)} + 2 p^{(1)} F_2^{(3)} + (p^{(2)} + b_2 p) F_2^{(2)} + 2 b_2 p^{(1)} F_2^{(1)} + b_2 p^{(2)} F_2 = 0, \qquad 3.3(34)$$

we see that the first four terms look like 3.3(33). Assume now that the solutions of 3.3(33) are related to $F_2(z)$ by

$$H_3(z) = F_2^{(2)}(z) + b_2 F_2(z). \qquad 3.3(35)$$

In order to prove this result it is simplest to proceed first by applying the differential operator of 3.3(31) to 3.3(35). Then substitute for the $F_2^{(6)}(z)$ and $b_2 F_2^{(4)}(z)$ terms appearing in $H_3^{(4)}$ by using equations 3.3(34) and 3.3(31), respectively:

$$\begin{aligned} H_3^{(4)} + (a_2 + p) H_3^{(2)} + b_2 p H_3 = \\ -(a_2 + p) F_2^{(4)} - 2 p^{(1)} F_2^{(3)} - (p^{(2)} + b_2 p) F_2^{(2)} - 2 b_2 p^{(1)} F_2^{(1)} - b_2 p^{(2)} F_2 \\ + b_2 [-(a_2 + p) F_2^{(2)} - b_2 p F_2] \\ + (a_2 + p) F_2^{(4)} + b_2 (a_2 + p) F_2^{(2)} + b_2 p F_2^{(2)} + b_2^2 p F_2 . \end{aligned} \qquad 3.3(36)$$

Clearly all the terms cancel except for those involving $p^{(1)}$ and $p^{(2)}$. The remaining terms on the right-hand side of 3.3(36), however, are equal to $2 p^{(1)} H_3^{(1)} + p^{(2)} H_3$, with H_3 expressed by 3.3(35). Hence we have shown that 3.3(35) gives the solutions of 3.3(33) in terms of the solutions $F_2(z)$ of its source equation 3.3(31).

Simply by taking the second derivative of the source equation

$$F_4^{(4)} + p\,F_4^{(2)} + a_4\,F_4 \;=\; 0 \qquad 3.3(37)$$

for the system

$$H_4^{(4)} + p\,H_4^{(2)} + 2\,p^{(1)}\,H_4^{(1)} + (p^{(2)} + a_{4)}\,H_4 \;=\; 0\,, \qquad 3.3(38)$$

we see that

$$H_4 \;=\; F_4^{(2)} \qquad 3.3(39)$$

Finally, in Chapter 4 there is some interest in a sixth-order differential equation modeling simultaneous interactions between "magnetostatic" spin waves, exchange-dominated spin waves, and elastic waves in microwave magnetoelastic delay lines. This equation has the form:

$$\begin{aligned} &H_5^{(6)} + (a_5 + p)\,H_5^{(4)} + 4\,p^{(1)}\,H_5^{(3)} + (6\,p^{(2)} + b_5 + c_5\,p)\,H_5^{(2)} \\ &\quad + (4\,p^{(3)} + 2\,c_5\,p^{(1)})\,H_5^{(1)} + (p^{(4)} + c_5\,p^{(2)} + b_5\,c_5)\,H_5 \;=\; 0. \end{aligned} \qquad 3.3(40)$$

Equations 3.3(33) and 3.3(38) are actually special cases of 3.3(40) in various limits, when $a_2 = a_5$, $a_4 = b_5$, and $b_2 = c_5$. The source equation for 3.3(40) is

$$F_5^{(6)} + (a_5 + p)\,F_5^{(4)} + (b_5 + c_5\,p)\,F_5^{(2)} + b_5\,c_5\,F_5 \;=\; 0. \qquad 3.3(41)$$

By using the same technique which followed 3.3(35), we can similarly prove that

$$H_5 = F_5^{(4)} + c_5 F_5^{(2)} . \qquad 3.3(42)$$

(g) <u>Construction of proper linear combinations of solutions to give the reflection and transmission coefficients.</u>

In section 3.2 (a) 4 it was shown that for second-order equations it is a simple matter to determine the reflection coefficients from the circuit relations for the hypergeometric functions, equation 3.1(31). All that is necessary is to label the wavenumber for the transmitted wave k_1^+ and then to write G_1^+ in terms of G_1^- and G_2^- from 3.1(31), where G_1^- represents the incident wave. For a wave incident from $z = -\infty$, there can be only one transmitted wave. On the other hand, for second-order coupled mode problems it was mentioned in section 3.2(c) that there is one incident wave and two transmitted waves. It was inherent in the approximations leading to such a second-order system, however, that reflected waves were neglected. Thus for those coupled mode problems one uses the circuit relation 3.1(29) for the single incident wave G_1^- in terms of the two transmitted waves G_1^+ and G_2^+.

For fourth- and higher-order systems, however, there can in general be two or more transmitted waves as well as two or more reflected waves, for each incident wave. Then the reflection and transmission coefficients are no longer given directly by 3.1(31) or 3.1(29). It is necessary to take the linear combination of the transmitted waves which makes the coefficients of all the incident waves zero except for that of the desired incident wave, which should be unity.

Consider fourth-order systems, and label the solution of the "source equation" F. Label the wavenumbers which correspond to transmitted (or evanescent) waves as $z \to +\infty$ as k_1^+ and k_3^+. Then write

F as the following linear combination of these two waves:

$$F = G_1^+ + J\,G_3^+ \qquad 3.3(43)$$

By 3.1(31) which expresses the G_h^+ in terms of G_g^-, we can also write 3.3(43) as

$$F = \sum_{g=1}^{4} (P_{1g} + J\,P_{3g})\,G_g^- \,. \qquad 3.3(44)$$

Now consider G_1^- and G_3^- to be the possible incident waves, and assume that in fact only G_1^- is incident, with unit amplitude. Thus the coefficient of G_3^- must vanish:

$$J_1 = -P_{13}/P_{33} \,. \qquad 3.3(45)$$

Also, in order for the incident wave to have unit amplitude we must now divide 3.3(43) and 3.3(44) through by $P_{11} + J_1 P_{31}$, with the result:

$$F_1 = G_1^- + R_{12}\,G_2^- + R_{14}\,G_4^- = T_{11}\,G_1^+ + T_{13}\,G_3^+ , \qquad 3.3(46)$$

where the R's and the T's are called the reflection and transmission factors for the "source equation" 3.3(1).

If the original differential equation for the system is written in terms of H(z) which can be expressed in terms of F(z) through the procedures of 5.3(f), we will have to take derivatives of 3.3(46) to get the physical reflection and transmission coefficients. To do this, recall that the G's represent normal modes of the form exp (-jkz) as $z \to \pm\infty$. Thus, as in the second-order example in section 3.2(a) 7,

$$F_1^{(1)}(z) \leftrightarrow -jk_1^-\,G_1^- - jk_2^-\,R_{12}\,G_2^- - jk_4^-\,R_{14}\,G_4^- \leftrightarrow -jk_1^+\,T_{11}\,G_1^+ - jk_3^+\,T_{13}\,G_3^+$$

3.3(47)

and

$$F_1^{(2)}(z) \leftrightarrow -(k_1^-)^2\, G_1^- - (k_2^-)^2 R_{12}\, G_2^- - (k_4^-)^2 R_{14}\, G_4^-$$

$$\leftrightarrow -(k_1^+)^2\, T_{11}\, G_1^+ - (k_3^+)^2\, T_{13}\, G_3^+ \qquad 3.3(48)$$

For example, if the physical field H(z) is given as $F_1^{(1)}(z)$, then the reflection and transmission coefficients for an incident wave of type G_1^- are found by dividing 3.3(47) through by $-jk_1^-$.

Now we must evaluate the R's and T's. From equations 3.3(43) through 3.3(46) we see that

$$T_{11} = \left(P_{11} - \frac{P_{13}P_{31}}{P_{33}}\right)^{-1} \qquad 3.3(49)$$

$$T_{13} = \left(P_{31} - \frac{P_{13}P_{11}}{P_{33}}\right)^{-1} \qquad 3.3(50)$$

$$R_{12} = T_{11}\left(P_{12} - \frac{P_{13}P_{32}}{P_{33}}\right) \qquad 3.3(51)$$

$$R_{14} = T_{11}\left(P_{14} - \frac{P_{13}P_{34}}{P_{33}}\right) \qquad 3.3(52)$$

Note from 3.1(31) that the P_{hg} for fourth-order systems (q = 4) involve the product and quotient of 12 gamma functions. Thus we would like to simplify the expressions in 3.3(49) through 3.3(52).

First, note from 3.1(31) that P_{hg} is invariant under any interchange $\rho_\ell \leftrightarrow \rho_m$ or $\sigma_\ell \leftrightarrow \sigma_m$ as long as $m \neq g \neq \ell$ or $m \neq h \neq \ell$, respectively. Further, observe that the operation $\rho_g \leftrightarrow \rho_\ell$ transforms P_{hg} to $P_{h\ell}$ and $\sigma_h \leftrightarrow \sigma_\ell$ transforms P_{hg} to $P_{\ell g}$. Hence the combined operation $\rho_g \leftrightarrow \rho_h$,

$\sigma_h \leftrightarrow \sigma_g$ transforms P_{hg} to P_{gh}. As a result, from 3.3(49) through 3.3(52):

$$(\sigma_1 \leftrightarrow \sigma_3)_{op} T_{11} = T_{13} \qquad 3.3(53)$$

and

$$(\rho_2 \leftrightarrow \rho_4)_{op} R_{12} = R_{14} \qquad 3.3(54)$$

Also note that operating on T_{11}^{-1} by $(\rho_1 \leftrightarrow \rho_2)_{op}$ produces the term $P_{12} - P_{13}P_{32}/P_{33}$ which occurs in 3.3(51). Thus

$$T_{11}[(\rho_1 \leftrightarrow \rho_2)_{op} T_{11}^{-1}] = R_{12} \qquad 3.3(55)$$

Furthermore, if we choose the incident wave to be of type 3 instead of type 1, we will find instead of 3.3(45) that J must satisfy

$$J_3 = P_{11}/P_{31} \qquad 3.3(56)$$

with the result, after dividing 3.3(43) and 3.3(44) by $(P_{13} + J_3 P_{33})$:

$$F_3 = G_3^- + R_{32} G_2^- + R_{34} G_4^- = T_{31} G_1^+ + T_{33} G_3^+ , \qquad 3.3(57)$$

Note that the interchange $\rho_1 \leftrightarrow \rho_3$ produces J_3 from J_1 and $(P_{13} + J_3 P_{33})$ from the factor $(P_{11} + J_1 P_{13})$ used to obtain 3.3(46), but leaves invariant the P_{12}, P_{14}, P_{32}, and P_{34} terms multiplying G_2^- and G_4^- in 3.3(44). Thus $(\rho_1 \leftrightarrow \rho_3)$ operating on T_{11}, T_{13}, R_{12}, R_{14} produces T_{31}, T_{33}, R_{32}, and R_{34}, respectively. In summary, we conclude that the only quantity which we must calculate in detail is T_{11}.

Writing out T_{11}^{-1} in 3.3(49) in terms of the gamma functions in the expression 3.1(31) for P_{hg}, we have:

$$T_{11}^{-1} = \frac{\Pi_{i=2}^{4} \Gamma(\rho_i - \rho_1)\,\Gamma(1 - \sigma_i + \sigma_1)}{\Pi_{i=1}^{4} \Gamma(\rho_i + \sigma_1)\,\Gamma(1 - \sigma_i - \rho_1)}\, B, \qquad 3.3(58)$$

where

$$B = \Gamma(\rho_1 + \sigma_1)\,\Gamma(1 - \sigma_1 - \rho_1) - \frac{\Gamma(\rho_3 + \sigma_1)\,\Gamma(1 - \sigma_3 - \rho_1)\,\Gamma(\rho_1 + \sigma_3)\,\Gamma(1 - \sigma_1 - \rho_3)}{\Gamma(\rho_3 + \sigma_3)\,\Gamma(1 - \sigma_3 - \rho_3)} \qquad 3.3(59)$$

Now use the property that

$$\Gamma(x)\,\Gamma(1 - x) = \pi / \sin(\pi x), \qquad 3.3(60)$$

and note from application of ordinary trigonometric identities that

$$\sin\pi(\rho_3 + \sigma_1)\sin\pi(\rho_1 + \sigma_3) - \sin\pi(\rho_1 + \sigma_1)\sin\pi(\rho_3 + \sigma_3) = \sin\pi(\rho_3 - \rho_1)\sin\pi(\sigma_3 - \sigma_1). \qquad 3.3(61)$$

In this way the factor B is evaluated. After transforming back to gamma functions using 3.3(60) again we finally get:

$$T_{11} = \frac{\Gamma(\rho_2+\sigma_1)\,\Gamma(\rho_4+\sigma_1)\,\Gamma(1-\sigma_2-\rho_1)\,\Gamma(1-\sigma_4-\rho_1)\,\Gamma(1-\rho_3+\rho_1)\,\Gamma(\sigma_3-\sigma_1)}{\Gamma(\rho_2-\rho_1)\,\Gamma(\rho_4-\rho_1)\,\Gamma(1-\sigma_2+\sigma_1)\,\Gamma(1-\sigma_4+\sigma_1)\,\Gamma(1-\rho_3-\sigma_1)\,\Gamma(\sigma_3+\rho_1)} \qquad 3.3(62)$$

Interchanging parameters according to 3.3(55) gives the explicit expression for R_{12}:

$$R_{12} = \frac{\Gamma(\rho_1-\rho_2)\,\Gamma(\rho_4-\rho_2)\,\Gamma(\sigma_1+\rho_2)\,\Gamma(\sigma_3+\rho_2)\,\Gamma(1-\sigma_2-\rho_1)\,\Gamma(1-\sigma_4-\rho_1)\,\Gamma(1-\rho_3+\rho_1)}{\Gamma(\rho_2-\rho_1)\,\Gamma(\rho_4-\rho_1)\,\Gamma(\sigma_1+\rho_1)\,\Gamma(\sigma_3+\rho_1)\,\Gamma(1-\sigma_2-\rho_2)\,\Gamma(1-\sigma_4-\rho_2)\,\Gamma(1-\rho_3+\rho_2)} \qquad 3.3(63)$$

All the remaining reflection and transmission factors in 3.3(46) and 3.3(57) can be written down immediately from these expressions by interchanging parameters according to 3.3(53), 3.3(54), and interchanging $\rho_1 \leftrightarrow \rho_3$ to obtain the R_{3i} and T_{3i} from R_{1i} and T_{1i}, respectively.

For systems satisfying a biquadratic dispersion relation in homogeneous regions of space, the first and third order terms in 3.3(1), 3.3(2) and hence 3.3(3) will be absent. It is then possible to simplify equations 3.3(62) and 3.3(63) further by noting that

$$\rho_2 = -\rho_1, \ \rho_4 = -\rho_3, \ \sigma_2 = -\sigma_1, \ \sigma_4 = -\sigma_3. \qquad 3.3(64)$$

This is the usual situation with the waves traveling in one direction having the same wavelength as their counterparts traveling in the other direction (see 3.3(18)).

(h) Analytical expressions for the coefficients in the lossless case in terms of elementary transcendental functions.

For lossless media, it is possible to simplify expressions for the transmission and reflection factors such as 3.3(62) and 3.3(63), in the manner indicated in section 3.2(a) 5. If all the waves are propagating, then in a lossless system the wavenumbers k will all be purely real, and in view of 3.3(18) the ρ's and σ's will all be purely imaginary. Hence they are the negatives of their complex conjugates. Then it is possible to evaluate the magnitudes $|T_{11}|^2$, $|R_{12}|^2$, etc. in terms of hyperbolic sines (sines with imaginary argument) by using properties 3.2(26), 3.2(27), and 3.3(60) of the gamma function.

When the waves for $z \to +\infty$ are all evanescent, the σ's will all be real. In that case, it may not be possible to express the magnitudes of all the reflection factors in terms of hyperbolic sines. On the other hand, because the ρ's are imaginary and the σ's are real, factors in the numerator of R_{12}, for example, may turn out to be complex conjugates of those in the denominator and will hence simplify to one when the absolute magnitude is taken. This happens when the system is biquadratic so that 3.3(64) holds. See Chapter 4 for specific examples.

Again as in section 3.2(a) 5, when the transformation parameter α is small (see 3.3(20)) the ρ's and σ's from 3.3(18) will tend to be large, and it may be possible to simplify the expressions by noting that hyperbolic sines of large arguments reduce to exponentials. This happens when the transition width is much larger than any local wavelengths or differences in local wavelengths of the system. In the opposite limit of a very sharp transition, the hyperbolic sines may be replaced by the first terms in their power series. Then the parameter α drops out so that the coefficients no longer depend on the transition width, and the results should be the same as those derivable by using the boundary conditions if no assumptions are violated in the meantime (see, for example, 3.2(30)).

(i) Evaluation of the coefficients when loss is included.

In the same limit where the hyperbolic sines may be replaced by exponentials in the lossless case, namely when α is small and the transition width large, it is sometimes possible to simplify expressions such as 3.3(62) and 3.3(63) by using Stirling's approximation for gamma functions of large argument:

$$\Gamma(x) \sim (2\pi)^{\frac{1}{2}} \exp\left[x\,(\ell n\, x - 1) - \frac{\ell n\, x}{2}\right], \quad |\arg x| < \pi. \qquad 3.3(65)$$

In this way only exponentials result.

Alternatively, the gamma functions of complex arguments arising from 3.3(62) and 3.3(63) may be evaluated by means of a computer subroutine. One such subroutine is LOGGAM, written in Fortran by Max Goldstein at New York University, and available as a library subroutine at the M.I.T. computer center.

(j) Discussion of limitations.

1. First of all, the physical problem must be stated in such a way that the independent variable z can range from $-\infty$ to $+\infty$. This is necessary in order that the transformation 3.3(5) from z to ζ can allow for an interesting variation of the parameter p(z), as in 3.3(13), and also in order that the solutions of the equation in ζ represent normal modes in the extreme limits. Hence radial variables ranging from 0 to ∞ are not allowed. Nevertheless, it may be possible to approximate the radial problem in certain regions by one with cartesian coordinates (PEARLSTEIN 1969, for example) and thus apply this method.

2. Secondly, the parameter p(z) is constrained to vary like a hyperbolic tangent, as in 3.3(13). This is often a useful variation, but one might like to be able to treat other variations, such as a hump or valley in p(z). Such a variation can be treated fairly easily when it occurs in a second-order equation, as in section 3.2(a) 3. For higher-order equations this becomes more difficult since the dependent as well as the independent variables are being transformed. See section 3.4 below, however, for one case where a valley can be treated in a fourth-order equation.

3. The procedure of this section is valid only if there is only one varying parameter in the differential equation. This restriction results

since there is only one transformation parameter α (see 3.3(13)). Near certain critical or coupling points, however, other parameters in an equation may be effectively constant to a good approximation. Then one can apply the method of this section if the parameter p(z) is constrained to vary only in a limited range about such a point. See Chapter 4 for examples.

In sections 3.2(b) and 3.2(c), a slightly modified transformation of the independent variable was noted which allowed a second-order differential equation with two varying parameters to be modeled by a hypergeometric equation. Such a transformation is not expected to be too useful for higher-order equations, however, because it results in the appearance of the $(n-1)^{th}$ and all lower derivatives of one of the parameters, in an equation of order n.

4. Finally, no nonlinear combinations of the parameter p(z) are allowed in the original differential equation. This includes rational fractions of linear combinations of p(z). One such combination occurs in the fourth-order equation for "magnetostatic" to exchange-dominated spin wave conversion at a turning point (see section 4.3). Near that point, however, the combination is constant to a very good approximation and hence the problem can be treated within restricted limits on p(z). Note also that SLUIJTER 1967 has solved a second-order equation with a similar rational fraction combination of a parameter p(z) which satisfies a hyperbolic tangent variation as in 3.3(13). There the physical problem was propagation of an extraordinary wave in a plasma with varying electron density perpendicular to a static magnetic field. The solution, however, required transformation to the differential equation with four regular singular points, Heun's equation.

3.4 Fourth-order equations with a turning point having a parameter with a valley or well.

It may be of interest to exhibit a solution for the above-mentioned problem in terms of hypergeometric equations, by extending the method of section 3.3. As with similar second-order equations, it is necessary to transform the dependent variable as well as the independent variable (see subsection 3.2(a)3). First transform the independent variable ζ according to 3.3(5) and thus bring the hypergeometric differential equation 3.3(3) into the form 3.3(11). Now proceed further by transforming the dependent variable as in 3.2(11):

$$F(z) \quad = \quad f(z)\, D(z), \qquad 3.4(1)$$

where D(z) is to be the field in terms of which the differential equation corresponding to the physical problem is written. Now apply Leibniz's theorem for multiple differentiation of a product to 3.4(1):

$$F^{(q-m)}(z) = \sum_{n=0}^{q-m} \frac{(q-m)!}{n!\,(q-m-n)!} f^{(n)}(z)\, D^{(q-m-n)}(z). \qquad 3.4(2)$$

Substituting this into 3.3(11) gives

$$\sum_{m=0}^{q} \sum_{n=0}^{q-m} u_{m+n,\,m}\, D^{(q-m-n)}(z) \quad = \quad 0 \qquad 3.4(3)$$

where

$$u_{\ell,\,m} \equiv \alpha^m \frac{(q-m)!}{(\ell-m)!\,(q-\ell)!} \left[r_m + (r_m - s_m) \frac{\zeta}{1-\zeta} \right] \frac{f^{(\ell-m)}}{f(z)} \qquad 3.4(4)$$

(Recall that $r_o = 1 = s_o$, so that $u_{o,\,o} = 1$.) Changing the dummy index in 3.4(3) according to $\ell = m + n$ gives

$$\sum_{m=0}^{q} \sum_{\ell=m}^{q} u_{\ell,\,m}\, D^{(q-\ell)}(z) \quad = \quad 0 \qquad 3.4(5)$$

Finally, after interchanging the order of summation, we have

$$\sum_{\ell=0}^{q} D^{(q-\ell)}(z) \sum_{m=0}^{\ell} u_{\ell,m}(z) = 0 \qquad 3.4(6)$$

in which we can now identify the coefficients of each derivative of D(z). Note that 3.4(6) is entirely equivalent to the generalized hypergeometric equation 3.3(3), and the solutions of one are related to solutions of the other by 3.4(1) and 3.3(6).

Now specialize to fourth-order equations, with $q = 4$. Further specialize to a biquadratic system satisfying a differential equation of the form

$$D^{(4)}(z) + p(z)\, D^{(2)}(z) + K^4\, D(z) = 0 \qquad 3.4(7)$$

Such an equation is said to have a turning point where $p(z) = 0$, since there the two roots for k^2 of the corresponding biquadratic dispersion relation in k coalesce. Furthermore, none of the solutions for k are real for $p(z) < 0$, corresponding to a cutoff situation (see section 4.3 for more details).

In order for 3.4(6) to have the form 3.4(7), we must first of all have the coefficient of $D^{(3)}(z)$ vanish:

$$\sum_{m=0}^{1} u_{1,m}(z) = 0. \qquad 3.4(8)$$

This is actually a differential equation for f(z), whose solution is

$$f(z) = (-\zeta)^{-r_1/4}\, (1-\zeta)^{(r_1-s_1)/4} \qquad 3.4(9)$$

which is very similar to 3.2(15). Proceeding with the coefficients of

$D^{(1)}(z)$ and $D(z)$, we require:

$$\sum_{m=0}^{3} u_{3,m}(z) = 0 \qquad 3.4(10)$$

and

$$\sum_{m=0}^{4} u_{4,m}(z) = K^4 , \qquad 3.4(11)$$

where the derivatives $f^{(\ell - m)}(z)$ must be calculated from 3.4(9). After considerable algebra, we find that 3.4(10) and 3.4(11) can be satisfied if any one of the following three sets of conditions is satisfied:

$$r_1 = s_1 \qquad 3.4(12a)$$

or

$$r_1 = s_1 + 4,\ r_2 - s_2 = \frac{3}{2}(r_1 + s_1),\ 4(r_2 + s_2) = -3r_1^2 - 12r_1 + 16 \qquad 3.4(12b)$$

or

$$r_1 = s_1 - 4,\ r_2 - s_2 = -\frac{3}{2}(r_1 + s_1) , \qquad 3.4(12c)$$

together with certain other relations

$$t_3 = -\frac{1}{8} t_1^3 + \frac{1}{2} t_1 t_2 \qquad 3.4(13)$$

and

$$t_4 = -\frac{5}{256} t_1^4 + \frac{1}{16} t_1^2 t_2^2 + \left(\frac{K}{\alpha}\right)^4 , \qquad 3.4(14)$$

where t stands for either r or s.

Furthermore, the coefficient $p(z)$ of $D^{(2)}(z)$ then becomes

$$p(z) = -\frac{3}{8}\alpha^2 \Big\{(r_1^2 - \frac{8}{3}r_2) + [r_1^2 - s_1^2 - \frac{8}{3}(r_2 - s_2)] \frac{\zeta}{1-\zeta}$$
$$+ (r_1 - s_1)(r_1 - s_1 + 4) \frac{\zeta}{(1-\zeta)^2}\Big\} . \qquad 3.4(15)$$

Note that if we set $r_1 = s_1 = 0$, $f(z) = 0$ from 3.4(9), and 3.4(15) reduces to the expression for $w_2(z)$ in 3.3(13): a hyperbolic tangent variation. Thus 3.4(15) contains 3.3(13) as a special case. More generally, if we assume 3.4(12a) is satisfied with arbitrary r_1 we see that p(z) still has only a hyperbolic tangent variation. If 3.4(12c) is satisfied, p(z) is identically constant, which is uninteresting.

The only new situation occurs when 3.4(12b) is satisfied. Substituting all of those relations into 3.4(15), and using 3.2(19), we find

$$p(z) = \alpha^2 \left[2 r_2 + 1 + 3 \operatorname{sech}^2\left(\frac{\alpha z}{2}\right)\right], \qquad 3.4(16)$$

which no longer contains the hyperbolic tangent variation. Relation 3.4(16) is similar to the corresponding situation for second-order equations, 3.2(20), except that there is one less available parameter in 3.4(16). All the other parameters are constrained through 3.4(12b), 3.4(13), and 3.4(14). Hence, if the peak in p(z) is made higher by increasing α^2, 3.4(16) requires that the width of the peak must become narrower. For very large α, the result would be something like a delta function. See Fig. 2.

Note from 3.4(16) that if r_2 is chosen to be less than $-\frac{1}{2}$, there are two turning points at some finite $z = \pm z_o$. For $|z| > z_o$, $p(z) < 0$ and the waves are all cutoff. Thus we have the analogy with a potential well. The methods of this chapter work best for determining reflection and transmission coefficients near $z = \pm\infty$. If $r_2 < -\frac{1}{2}$, there can be no propagating incident wave, and there is no physical significance to "reflection and transmission coefficients." For $r_2 > -\frac{1}{2}$, we can calculate the reflection and transmission coefficients, following the method outlined in 3.3(g). In this latter case, however, it would probably be simpler to approximate the problem with two approximately uncoupled second-order equations. On the other hand, with $r_2 < -\frac{1}{2}$, it may be of interest to calculate the solutions $D(z) = f^{-1}(z)\, G[\zeta(z)]$ at finite $|z| < z_o$ (ζ finite and non-zero) where the four possible waves can bounce back and forth. It may be possible in that region to find suitable approximations for $G(\zeta)$ in terms of polynomial and rational approximations to generalized hypergeometric functions (see, for example, LUKE 1969, volume 2, chapters 11 and 12). Note that the parameters r_i and s_i determined by 3.4(16), 3.4(14), 3.4(13), and 3.4(12a) are defined in terms of the ρ's and σ's appearing in the generalized hypergeometric equation 3.3(3) (see after 3.3(9)).

CHAPTER 4

COUPLED WAVES IN MICROWAVE MAGNETOELASTIC DELAY LINES WITH A NONUNIFORM MAGNETIC FIELD

As indicated schematically in Figs. 3 through 5, the waves in the interior of typical microwave magnetoelastic delay lines couple at two points called the crossover and turning points. In section 4.1, we derive the three coupled second-order differential equations which model these physical situations. Two of these equations are combined in section 4.2 to treat the exchange-dominated spin wave to elastic wave conversion at the crossover point. By combining one of these latter equations with the third, we treat in section 4.3 the "magnetostatic" to exchange-dominated spin wave conversion at the turning point. Section 4.4 shows how all three equations may be combined into one sixth-order equation to determine how the coupling at each of the two points affects that at the other.

As indicated in section 1.1, there is no treatment here of nonlinear effects or the propagation of pulses or transients. We also do not consider explicitly the effects of nonuniformities in the directions transverse to the propagation direction. ADDISON, AULD, and COLLINS 1968 may be consulted for a recent application of ray theory to investigate the focusing or defocusing effects of such transverse variations. Modifications in previous theories to account for these variations have also been made by LEWIS and SCOTTER 1969. Finally, this chapter does not contain any detailed discussion of the coupling of "magnetostatic" waves to fine-wire antennas. See, for example, AULD 1963, BURKE and BHAGAT 1967, and DESORMIERE and LEGALL 1969. Nevertheless, an understanding of the solutions of the basic linear one-dimensional differential equations treated in this work is fundamental for better understanding of the effects mentioned above.

4.1 The Basic Differential Equations for One-Dimensional Propagation

(a) Physical introduction.

There are basically three wave types of interest in the interior of microwave magnetoelastic delay lines: elastic waves, whose wavelengths are typically in the micro-meter range; exchange-dominated spin waves, whose wavelengths can be either smaller or larger than those of elastic waves; and the so-called "magnetostatic" spin waves. The first two have such short wavelengths that they are generally much smaller than any of the dimensions of the delay line and hence behave like plane waves. The "magnetostatic" spin wave, however, has longer wavelengths and thus has properties which depend strongly on the transverse dimensions of the sample. In practice, at least one of these transverse dimensions is much smaller than electromagnetic wavelengths at the frequency of operation, so that the "magnetostatic" waves have wavelengths which are still considerably shorter than the wavelengths of ordinary electromagnetic waves. These "magnetostatic" waves are actually the "extraordinary," slow waves resulting from the interaction of electromagnetic fields with the spins in a magnetized material. Note that both types of spin waves represent precession in time and space of the magnetization about its static position.

The predominant field for each wave type satisfies a second-order wave equation. The lattice displacement $\vec{R}$ of a point in the crystalline delay line from its equilibrium position satisfies a well-known equation which is derived from Newton's laws and the relation between stress (restoring force) and strain (resulting from non-zero R) in an elastic medium. The predominant field for exchange-dominated spin waves is the dynamic magnetization $\vec{m}$ which precesses in space and time about the static saturation magnetization M_s, according to the "torque equation." This torque equation results from the relation between torque and the time rate of change of the spatially-averaged spin angular momentum, which is in turn related to the magnetization by the gyromagnetic ratio γ. When the spins on neighboring atoms are appreciably misaligned, the quantum-mechanical exchange interaction produces a restoring force whose contribution to the torque on the magnetization is describable by an effective exchange magnetic field. The existence of this force allows the spin wave to carry "exchange" power as well as electromagnetic power. Often only the exchange power is significant. The exchange force and effective field become important, however, only for short wavelengths since only then are neighboring spins in the crystal appreciably misaligned.

Finally, for the longer-wavelength "magnetostatic" spin waves, the dynamic magnetic field $\vec{h}$ is the predominant field. Only the electromagnetic power flow is usually significant for these spin waves. To derive their characteristic wave equation, it is also necessary to use the torque equation, but without the exchange field. The torque equation must now also be combined with Maxwell's equations. However, these waves get their name from the fact that their characteristics are usually derived to first order by neglecting the dynamic electric field $\vec{e}$ in Maxwell's equations. (The characteristics of the exchange-dominated spin waves are also derived ignoring $\vec{e}$, of course. Hence in this sense they could also be called magnetostatic waves.) It then may be difficult to get a quantitative understanding of how the "magnetostatic" waves couple to the fields of ordinary electromagnetic waves which might arise from a fine-wire antenna located outside the delay line (see Fig. 3). Nevertheless, the sizes of the fine-wire antenna and the delay line are usually much smaller than electromagnetic wavelengths at the frequency of operation, so that only the near-field component of the magnetic field set up by the antenna interacts with the fields inside the delay line. As is well known, this near-field component can also be calculated from Maxwell's equations by using the magnetostatic approximation which ignores $\vec{e}$. Hence it is possible to get a qualitative feel for the coupling of the antenna fields to the "magnetostatic" waves by considering only the magnetic field $\vec{h}$. See the bibliography by subject for a listing of relevant papers in the literature. The work of the present section treats only interactions between the three kinds of waves within the magnetic material.

(b) Derivation of the coupled second-order equations for spatially-varying fields.

1. Summary of results.

Equations corresponding to most of those derived in this section have been obtained elsewhere. SCHLÖMANN 1960 derived the equivalent of the elastic equations of motion 4.1(2) and the torque equation components 4.1(5), without the transverse wavenumber which we neglect later anyway. Equation 4.1(14a) was derived from Maxwell's equations by VASILE and LAROSA 1968a for thin magnetic slabs. Using the technique of VASILE 1967, it is shown here how to derive the similar equation 4.1(14b) for modes with no azimuthal variation in cylindrically symmetric samples. AULD, COLLINS, and WEBB 1968 observed experimentally that, in rods, such a mode was the easiest to excite by fine wire antennas. See their Fig. 9 for the antenna configurations they used, including the full loop which excites no azimuthal magnetic field.

Further, using the methods of VASILE and LAROSE 1968a, we give in 4.1(19) the explicit form of requations 4.1(5) when only one component of transverse magnetic field is present. The elastic equations 4.1(2) in this case become 4.1(20) directly. Equations 4.1(19) and 4.1(20) are useful particularly for the analysis of section 4.4, when the interactions between elastic, exchange-dominated spin, and "magnetostatic" spin waves are all considered simultaneously. In that case, 4.1(14a) is also added, to complete the set of three coupled second-order equations. Finally equation 4.1(22) is derived, apparently for the first time, to show explicitly in a second-order coupled equation how "magnetostatic" waves can couple to the exchange-dominated spin waves.

Since the various assumptions in these derivations are used in several places, and since it is helpful to have all of them listed in one location with a discussion of their physical implications, Appendix 3 is provided for this purpose. The numbers which label each assumption thus refer to this appendix.

2. The equation of motion for the lattice displacement (elastic waves).

To find the equations of motion for the lattice displacement $\vec{R}$, we consult SCHLÖMANN 1960, equation (1). There it is assumed that the static magnetization is parallel to a cubic crystalline axis, which will be labeled the z-axis. As a result, only shear elastic waves interact with a spin-wave traveling in the z-direction. That is, $R_z = 0$, which is assumption 17. SCHLÖMANN 1960 also assumes that the material is elastically isotropic, which is a good approximation for yttrium iron garnet, which is most commonly used. Other assumptions in that paper are taken from KITTEL 1958a. Ignoring the equation for R_z, we thus have

$$D\ddot{R}_x = c_{44}\nabla^2 R_x + (c_{44} + c_{12})\frac{\partial}{\partial x}(\nabla\cdot\vec{R}) + \frac{b_2}{M}m'_x$$

$$D\ddot{R}_y = c_{44}\nabla^2 R_y + (c_{44} + c_{12})\frac{\partial}{\partial y}(\nabla\cdot\vec{R}) + \frac{b_2}{M}m'_y \qquad 4.1(1)$$

where certain changes in notation have been made to avoid confusion with other symbols in this work. Dots and primes denote differentiation with respect to time and space, respectively. The mass density is D, c_{12} and c_{44} are the elastic constants relating stress to strain, b_2 is magnetoelastic coupling coefficient, and M is the saturation magnetization (see assumption 9).

As mentioned previously, it is expected that the elastic waves by themselves will act like plane waves since their wavelengths are so short. If we thus neglect transverse variations and the coupling to the magnetization (the last two terms on the right of equation 4.1(1)), the result is the ordinary wave equation for shear waves. The velocity of sound v_p is then identifiable as $v_p^2 = c_{44}/D$ (p is for "phonon").

More explicitly, if the parameters of the medium do not vary with time, we can Fourier analyze 4.1(1) and consider each harmonic component individually. The fields are then written as in assumption 1. Furthermore, the normal modes for the magnetization are generally almost entirely circularly polarized when they interact with the elastic waves. Thus it is convenient to take linear combinations of 4.1(1) corresponding to the circularly polarized components $\underline{R}_\pm = \underline{R}_x \pm j\underline{R}_y$.

Thus we obtain:

$$-\omega^2 D\underline{R}_\pm = c_{44}\underline{R}''_\pm + \frac{b_2}{M} m'_\pm , \qquad 4.1(2)$$

where all transverse derivatives are neglected. This approximation is roughly equivalent to assumption 12a, if the transverse variations of the elastic displacement $\vec{R}$ have characteristic wavenumbers equal to some k_t. Here we identify the elastic ("phonon") wavenumber k_p as $k_p^2 = \omega^2/v_p^2 = \omega^2 D/c_{44}$. Hence

$$\underline{R}''_\pm + k_p^2 \underline{R}_\pm = -\frac{b_2}{c_{44}M}\underline{m}'_\pm . \qquad 4.1(3)$$

The form of this equation does not change with the system of electromagnetic units, since b_2 and c_{44} both have units of energy density.

3. <u>The torque equation for the magnetization (spin waves).</u>

The torque equation for the magnetization is $\dot{\vec{M}} = \gamma\mu_o \vec{M} \times \vec{H}_{eff}$, where the effective magnetic field $\vec{H}_{eff}$ is composed of the following: H(z), the net internal static magnetic field in the crystal, assumed to be approximately parallel to the magnetization M; the magnetic dipolar field $\vec{h}$ of the precessing magnetization, calculated from Maxwell's equations; the effective exchange field $\lambda\nabla^2\vec{m}$ arising from the quantum-mechanical exchange force tending to align the magnetic moments on neighboring atoms; and the effective field arising from the magnetoelastic energy, calculated from the gradient of the magnetoelastic energy with respect to the magnetization. Assuming that only small signal (linear) excitation is involved, using assumptions 7, 8, and 9, and again considering only shear elastic waves with $R_z = 0$, we obtain for the x- and y- components of the torque equation:

$$\dot{m}_x = -\gamma\mu_o M(h_y + \lambda\nabla^2 m_y) + \gamma\mu_o m_y H + \gamma b_2 R'_y$$
$$\dot{m}_y = \gamma\mu_o M(h_x + \lambda\nabla^2 m_x) - \gamma\mu_o m_x H - \gamma b_2 R'_x \qquad 4.1(4)$$

These equations again can be taken from SCHLÖMANN 1960, equation (1), except that 4.1(4) are written in mks units. See JACKSON 1962, Table 3, for a conversion table for symbols and formulas in going from Gaussian to mks units. Now again consider harmonic components using assumption 1. Also assume that the transverse variation of $\vec{m}$ is as in assumption 4, so that $\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)\vec{m} \equiv \nabla_t^2 \vec{m} \approx -k_t^2 \vec{m}$. (It will turn out that transverse variations in $\vec{m}$ actually can be ignored for our purposes.) Finally, take linear combinations of 4.1(4) corresponding to the circularly polarized components $\underline{m}_\pm = \underline{m}_x \pm j\underline{m}_y$:

$$\underline{m}''_\pm - \left(\frac{H \mp \omega/|\gamma\mu_o|}{M\lambda} + k_t^2\right)\underline{m}_\pm = -\frac{1}{\lambda}\underline{h}_\pm + \frac{b_2}{\mu_o M\lambda}\underline{R}'_\pm \qquad 4.1(5)$$

4. Maxwell's equations for the magnetic field (transverse electric waves).

The third coupled equation is obtained from the Maxwell curl equations, writing the fields as in assumption 1:

$$\nabla \times \vec{\underline{e}} = -j\omega\mu_o(\vec{\underline{m}} + \vec{\underline{h}}) \qquad 4.1(6a)$$

$$\nabla \times \vec{\underline{h}} = j\omega\epsilon\vec{\underline{e}} \qquad 4.1(6b)$$

By taking the curl of 4.1(6b) and substituting into 4.1(6a) we obtain the wave equation for $\underline{\vec{h}}$:

$$\nabla(\nabla \cdot \underline{\vec{h}}) - \nabla^2 \underline{\vec{h}} = k_o^2(\underline{\vec{m}} + \underline{\vec{h}}) , \qquad 4.1(7)$$

where $k_o^2 \equiv \omega^2 \mu_o \epsilon$.

To have only the transverse components of $\underline{\vec{h}}$ appear, as in equation 4.1(5), we must find $\underline{h}_z$ in terms of $\underline{h}_\pm$. This is done through the z-component of 4.1(7):

$$\frac{\partial}{\partial z}(\nabla_t \cdot \underline{\vec{h}}_t) - \nabla_t^2 \underline{h}_z = k_o^2 \underline{h}_z , \qquad 4.1(8)$$

where the definitions $\nabla_t^2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ and $\nabla_t \cdot \underline{\vec{h}}_t \equiv \partial \underline{h}_x / \partial x + \partial \underline{h}_y / \partial y$ are used, and $\underline{m}_z$ is neglected according to assumption 2. Now evaluate $\nabla_t \cdot \underline{\vec{h}}_t$ in terms of $\underline{h}_\pm = h_x \pm jh_y$ by defining the operators $\nabla_\pm \equiv \frac{\partial}{\partial x} \pm j \frac{\partial}{\partial y}$, introduced by VASILE 1967. Then by explicit evaluation, and by using the z-component of 4.1(6b), we find

$$\nabla_+ \underline{h}_- = \nabla_t \cdot \underline{\vec{h}}_t + \omega \epsilon \underline{e}_z \qquad 4.1(9a)$$

$$\nabla_- \underline{h}_+ = \nabla_t \cdot \underline{\vec{h}}_t - \omega \epsilon \underline{e}_z \qquad 4.1(9b)$$

To proceed further, assume $\underline{e}_z \approx 0$ (see assumption 3), corresponding to propagation in a transverse electric mode. Also assume that the variables in equation 4.1(7) can approximately be separated according to assumption 4, so that $\nabla_t^2 \underline{\vec{h}} = - k_t^2 \underline{\vec{h}}$. Thus 4.1(8) becomes

$$\underline{h}_z = \frac{1}{k_o^2 - k_t^2} \frac{\partial}{\partial z}(\nabla_\mp \underline{h}_\pm) , \qquad 4.1(10)$$

where either the top or the bottom signs may be taken.

With the same assumptions used to obtain 4.1(10), and by taking appropriate linear combinations of the transverse components of 4.1(7), we find

$$\nabla_{\pm}(\nabla_{\mp}\underline{h}_{\pm}) + \nabla_{\pm}(\partial \underline{h}_z/\partial z) + k_t^2\underline{h}_{\pm} - \frac{\partial^2}{\partial z^2}\underline{h}_{\pm} = k_o^2(\underline{m}_{\pm} + \underline{h}_{\pm}) \qquad 4.1(11)$$

Next note that $\nabla_{\pm}\nabla_{\mp}\underline{h}_{\pm} = \nabla_t^2\underline{h}_{\pm} = -k_t^2\underline{h}_{\pm}$.

Finally, calculate $\nabla_{\pm}(\partial\underline{h}_z/\partial z)$ in 4.1(11) from equation 4.1(10), noting that $k_t^2 - k_o^2$ is a constant by assumption 4:

$$\nabla_{\pm}(\partial \underline{h}_z/\partial z) = \frac{k_t^2}{k_t^2 - k_o^2}\frac{\partial^2}{\partial z^2}\underline{h}_{\pm}\,. \qquad 4.1(12)$$

Substituting 4.1(12) into 4.1(11) and combining terms gives

$$\underline{h}_{\pm}'' = (k_t^2 - k_o^2)(\underline{m}_{\pm} + \underline{h}_{\pm}), \qquad 4.1(13)$$

where the prime denotes differentiation with respect to z, as before. As was pointed out by VASILE 1967, it is rigorously possible with a magnetic material to neglect $\underline{e}_z$, as we have done only when $k_t^2 >> k_o^2$ (assumption 5). Hence k_o^2 should be neglected in 4.1(13).

In practice, it is more convenient to combine the two equations represented by 4.1(13) into one. For rectangular coordinates, recall that $\underline{h}_{\pm} = \underline{h}_x \pm j\underline{h}_y$ and similarly for $\underline{m}_{\pm}$. Then we obtain by adding the two equations in 4.1(13), and neglecting k_o^2:

$$\underline{h}_x'' - k_t^2\underline{h}_x = k_t^2\underline{m}_x\,. \qquad 4.1(14a)$$

Equation 4.1(14a) was derived for slabs thin in the x-direction by VASILE and LAROSA 1968a, using the "magnetostatic" approximation in which all electric fields are neglected and only the curl $\vec{h}$ and divergence $\vec{h}$ Maxwell equations are used. For cylindrical coordinates r and φ as in a rod, $\underline{h}_\pm = e^{\pm j\varphi}(\underline{h}_r \pm j\underline{h}_\varphi)$ and similarly for $\underline{m}_\pm$. Multiplying the $\underline{h}''_+$ equation by $e^{-j\varphi}$ and adding it to $e^{+j\varphi}$ times the $\underline{h}''_-$ equation in 4.1(13) gives, neglecting k_o^2:

$$\underline{h}''_r - k_t^2 \underline{h}_r = k_t^2 \underline{m}_r . \qquad 4.1(14b)$$

5. <u>The coupled equations with only one transverse component of magnetic field present.</u>

It is common that only one transverse component of $\underline{\vec{h}}$ may be present (see assumptions 6a and 6b), either $\underline{h}_x$ or $\underline{h}_r$, in which case the relevant one of equations 4.1(14a) and 4.1(14b) is used. To determine how this equation then couples to 4.1(5), it is necessary to eliminate one of the circularly-polarized components of $\underline{\vec{m}}$. For example, if $\underline{h}_y = 0$ (assumption 6a), we first find $\underline{m}_-$ in terms of $\underline{h}_x$ from 4.1(5). To do this, first consider 4.1(5) with the right-hand side terms, involving coupling to the magnetic and elastic fields, set to zero. Then the left-hand side is a wave equation for exchange-dominated spin waves with normal-mode solutions $\underline{m}_\pm \approx \exp(-jk_\pm z)$, where

$$k_\pm^2 = -\left(\frac{H \mp \omega/|\gamma\mu_o|}{M\lambda}\right) \qquad 4.1(15)$$

(Typically, k_t^2 is completely negligible compared even to k_+^2, unless H is within about 10^{-6} oersted of $\omega/|\gamma\mu_o|$.)

For propagating waves, the k^2 must be positive. This is only possible for the positive circularly-polarized waves $\underline{m}_+$ (with $\underline{m}_- = 0$). Furthermore, these waves interact with the elastic waves only when $k_+ \approx k_p$, which occurs in practice near where $H \approx \omega/|\gamma\mu_o|$, and with the magnetic field only where k_+ is even smaller (see the dispersion diagram in Fig. 4). Thus from 4.1(5), we see that $|k_+| << |k_-|$ in the regions of interest. The negatively-circularly polarized waves are very rapidly evanescent (k_- very large and imaginary). Consequently, only the positively-circularly polarized waves interact with the elastic waves, so we can set $\underline{R}_- = 0$ (assumption 18), ignoring negatively-circularly polarized elastic waves.

On the other hand, when coupling to the $\vec{\underline{h}}$ field in equation 4.1(5) again becomes important, for smaller $|k_+|$, the magnetization will no longer be exactly positively-circularly polarized because $\vec{\underline{h}}$ is not circularly polarized. However, we can now use assumption 11 to neglect $|\underline{m}''_-| \approx k_+^2 \underline{m}_-$, since as was just mentioned, in this region $|k_+^2| << |k_-^2|$, as given by 4.1(15). Thus, from equation 4.1(5), with assumptions 18, 11, and 6a:

$$\underline{m}_x - j\underline{m}_y \equiv \underline{m}_- \cong \frac{M}{H + \omega/|\gamma\mu_o|}\underline{h}_x , \qquad 4.1(16)$$

where k_t^2 has also been neglected in accordance with assumption 12c. Equation 4.1(16) gives $\underline{m}_y$ in terms of $\underline{m}_x$ and $\underline{h}_x$ so that

$$\underline{m}_+ \equiv \underline{m}_x + j\underline{m}_y = 2\underline{m}_x - \frac{M}{H + \omega/|\gamma\mu_o|}\underline{h}_x . \qquad 4.1(17)$$

Here it is seen that the term proportional to $\underline{h}_x$ gives the deviation from circular polarization in $\vec{\underline{m}}$. With $\underline{R}_- = 0$, however, we can write

$$\underline{R}_+ = 2\underline{R}_x. \qquad 4.1(18)$$

Now substituting 4.1(17) and 4.1(18) into equation 4.1(5) for $\underline{m}_+$, and substituting for $\underline{h}_x''$ in terms of $\underline{m}_x$ and $\underline{h}_x$ through 4.1(14a), we obtain:

$$\underline{m}_x'' + k_{Ex}^2 \underline{m}_x = \frac{-H}{(H + \omega/|\gamma\mu_o|)} \frac{1}{\lambda} \underline{h}_x + \frac{b_2}{\mu_o M \lambda} \underline{R}_x', \qquad 4.1(19)$$

with $k_{Ex}^2 \equiv k_+^2$ as given by 4.1(15). A term approximately equal to k_t^2 has been neglected compared with k_{Ex}^2, in accordance with assumption 13. Since $\omega/|\gamma\mu_o|$ is close to H in the regions of interest, the only change in form from 4.1(5) is an extra factor of approximately $\frac{1}{2}$ multiplying the $\vec{\underline{h}}$ term, arising because $\vec{\underline{m}}$ is almost circularly polarized but $\vec{\underline{h}}$ is not. The two equations represented by 4.1(3) can be combined simply by adding them together to obtain:

$$\underline{R}_x'' + k_p^2 \underline{R}_x = -\frac{b_2}{c_{44} M} \underline{m}_x' \qquad 4.1(20)$$

Finally, equation 4.1(14a) is repeated here:

$$\underline{h}_x'' - k_t^2 \underline{h}_x = k_t^2 \underline{m}_x \qquad 4.1(21)$$

Equations 4.1(19) through 4.1(21) form the starting point for consideration of the complete wave coupling problem, where interactions between elastic displacement, magnetization, and magnetic field are simultaneously treated (see section 4.4). Note that if instead of $\underline{h}_y = 0$ we had $\underline{h}_\varphi = 0$ in a cylindrical rod (assumption 6b), then equations 4.1(19) through 4.1(21) would have exactly the same form, with r replacing x everywhere.

6. <u>The wave equation for "magnetostatic" spin waves.</u>

When the right-hand sides of equations 4.1(19) and 4.1(20) are neglected, we obtain the wave equations for the exchange-dominated spin waves and the elastic waves, respectively. However, the wave equation for the "magnetostatic" waves is not obtained by neglecting the right-hand side of 4.1(21). The reason for this is that "magnetostatic" waves intrinsically involve both magnetization and magnetic field. They do not involve exchange field terms, represented by λ, however. To obtain the characteristic features of the "magnetostatic" waves, ignore $\underline{R}'_x$ and find $\underline{m}_x$ in terms of $\underline{m}''_x$ and $\underline{h}_x$ from 4.1(19). Then substitute into 4.1(21):

$$\underline{h}''_x + k^2_{MS}\underline{h}_x = +\frac{M\lambda k_t^2}{H - \omega/|\gamma\mu_o|}\underline{m}''_x , \qquad 4.1(22)$$

where

$$k^2_{MS} \equiv k_t^2\left[\frac{H^2 + HM - (\omega/|\gamma\mu_o|)^2}{(\omega/|\gamma\mu_o|)^2 - H^2}\right] \qquad 4.1(23)$$

Notice that the wavenumber of the "magnetostatic" waves becomes very small near where the field satisfies $H^2 + HM = (\omega/|\gamma\mu_o|)$. The long-wavelength electromagnetic fields will couple to these waves in this region. When H approaches resonance at $H = \omega/|\gamma\mu_o|$, however, k^2_{MS} becomes very large (see the dispersion diagram in Fig. 4). The coupling term on the right-hand side of 4.1(22) also becomes large near resonance, showing that coupling to the exchange-dominated spin waves (whose existence depends upon the exchange parameter λ) occurs there also. An equation of the form 4.1(22) does not appear to have been exhibited previously in the literature.

4.2 High-k spin-elastic wave conversion at magnetoelastic crossover point

(a) Introduction.

1. Summary of results in the literature.

At the place in a magnetized sample where the wavenumber k_m for exchange-dominated spin-waves becomes equal to that (k_p) of the elastic waves, conversion from one type to another can occur. (See the dispersion relation in Figs. 4 and 5.) When the field is nonuniform, as in Fig. 3, this conversion can only occur at one point, the "crossover point," since only k_m varies with the magnetic field. Since the wavelengths involved are so short compared with typical transverse dimensions of the samples, transverse variations in 4.1(5) can be neglected, and consequently the waves can be assumed to be circularly polarized (assumptions 13, 16, and 18). If the high-frequency magnetic field is similarly neglected (assumption 19a), equations 4.1(3) and 4.1(5) become

$$v'' + k_p^2 v = -au' \qquad 4.2(1a)$$

$$u'' + k_m^2 u = av' \qquad 4.2(1b)$$

where

$$k_m^2(z) = \frac{\omega/|\gamma\mu_o| - H(z)}{M\lambda} \qquad 4.2(2)$$

and

$$v = (c_{44})^{\frac{1}{2}} \underline{R}_+$$

$$u = (\mu_o \lambda)^{\frac{1}{2}} \underline{m}_+ \qquad 4.2(3)$$

$$a = \frac{b_2}{M} (\lambda \mu_o c_{44})^{-\frac{1}{2}}$$

Equations 4.2(1) are identical to those derived by SCHLÖMANN and JOSEPH 1964, who used definitions and normalizations corresponding to 4.2(2) and 4.2(3), except with some different symbols and Gaussian units.

As was mentioned in Chapter 2, the first treatment of spin/elastic wave conversion by SCHLÖMANN and JOSEPH 1964 used successive approximations methods to derive conversion efficiencies in limits of small and large gradients of magnetic field at the crossover point. For large gradients, the "weak coupling" approximation was used involving the two coupled second order equations 4.1(1), since little power is then converted from one wave type to the other. When the gradients are small, on the other hand, much power is converted (the "strong coupling limit"). In this case it was more convenient first to write 4.1(1) in the form of four coupled first-order equations and then transform to hybrid magnetoelastic modes, which would be the normal modes in a homogeneous medium. In a weak inhomogeneity, these modes are coupled weakly (see equation 2.2(26) and remarks following), and thus another successive approximation scheme was applied. The results are reproduced in Fig. 6.

Later, KIRCHNER 1966 showed that if reflected waves could be neglected, 4.1(1) could be reduced to a set of only two coupled first order equations, of the same form as the equations in LOUISELL 1955 for tapered mode directional couplers. These type of equations are usually called the "coupled mode equations" for the inhomogeneous system. By numerically integrating these equations, Kirchner obtained a smooth curve for the conversion efficiency, reducing to the results of Schlömann and Joseph in the appropriate limits (see again Fig. 6). Finally, REZENDE and MORGENTHALER 1969a derived coupled mode equations of the same form for the case when spin/elastic-wave conversion occurs in time-varying, rather than spatially-varying, magnetic fields. These equations were also integrated numerically, and a plot of the conversion versus time was obtained.

2. Summary of the results of this section.

In this section, we show how solutions of generalized hypergeometric differential equations can be used to obtain the reflection and transmission (conversion) coefficients.

A simple transformation allows the results to be obtained analytically for arbitrary gradients of the magnetic field, having a variation of the form shown in Fig. 1. In part (b), it is shown how an equation of second order gives directly the conversion efficiency, without the need for numerical integration. This is done by applying the simple transformation to the coupled mode equations derived by KIRCHNER 1966 and REZENDE and MORGENTHALER 1969a.

By using a fourth-order equation derived from 4.1(1), the analysis in part (c) is able to include the effects of reflections. These are shown to be significant when the field gradients become large or the wavenumber k_m becomes small so that the WKB solutions break down strongly (see Fig. 8). It is also demonstrated that the magnitude of these reflections depends upon the nature of the overall variation of k_m, whereas the basic spin-elastic conversion (transmission) depends only upon the local magnetic field gradient at the crossover. The results for all the reflection and transmission power fractions are summarized schematically in Fig. 7. Note that these results are analytical expressions derived without a computer. It should also be noted here that information on the phases of the reflection and transmission coefficients can also be obtained from the treatment of this section. Then, however, a computer evaluation of gamma functions of complex arguments is necessary (see part 3.3(i)).

Note that there is a negligible amount of reflected elastic (phonon) energy with an incident spin wave (magnon), in contradiction to the result of SCHLÖMANN and JOSEPH 1964 obtained by successive approximation. This reflection becomes significant only when the magnetic field transition is so abrupt that it is thinner than the wavelengths at the crossover point. Then of course the boundary conditions would apply. Thus we conclude that successive approximations are not useful in calculating reflected waves. A similar conclusion holds for the waves in section 4.3. In that case, moreover, there are no transmitted waves.

(b) <u>Solution of the second order "coupled mode" equations with hypergeometric functions</u>.

1. Spatially-varying fields.

By neglecting reflections, KIRCHNER 1966 derived from equations 4.2(1) the following coupled equations:

$$
\begin{aligned}
A_m' + jk_m(z)\, A_m &= j\tfrac{1}{2} a A_p \\
A_p' + jk_p\, A_p &= j\tfrac{1}{2} a A_m\,,
\end{aligned}
\qquad 4.2(4)
$$

where the A_m and A_p are normalized linear combinations of u, u', v, and v', such that the quantities $|A_m|^2$ and $|A_p|^2$ represent power flow in the spin and elastic waves, respectively. The prime denotes differentiation with respect to z. As shown in deriving equation 3.2(53), equations 4.2(4) can be combined into one of the form:

$$A_p'' + j(k_m(z) + k_p)A_p' + \left(\frac{a^2}{4} - k_m(z)k_p\right)A_p = 0 \qquad 4.2(5)$$

(The equation for A_m is identical.)

Using the transformation in 3.3(5) from z to ζ, we find that 4.2(5) becomes a hypergeometric equation, if $k_m(z)$ has a tanh $(\frac{\alpha z}{2})$ variation (see 3.3(13)). This variation is shown schematically in Fig. 1. Then the parameter α is given by 3.3(20) or 3.3(23), and the quantities ρ_i and σ_i are related through 3.3(17) and 3.3(18) to the solutions k_i^- and k_i^+ of the dispersion relation at $z = \pm\infty$. The solutions of 4.2(5) can be written in normal mode form in these limits as in 3.3(16), with $A_p(z)$ replacing F(z). The dispersion relation for the homogeneous regions, from 4.2(5) and 3.3(16) is:

$$-k^2 + (k_m + k_p)k + (\frac{a^2}{4} - k_m k_p) = 0, \qquad 4.2(6)$$

with solutions

$$k = \frac{k_m + k_p}{2} \pm \left[\left(\frac{k_m - k_p}{2}\right)^2 + \frac{a^2}{4}\right]^{\frac{1}{2}} \qquad 4.2(7)$$

Label the solutions of 4.2(7) corresponding to elastic waves as k_1 and those for spin waves as k_2.

Now consider an incident phonon (elastic wave) as in Fig. 6. The solution of the hypergeometric equation corresponding to 4.2(5) is then

$$A_p(z) = G_1^-[\zeta(z)] = M_{11}G_1^+[\zeta(z)] + M_{12}G_2^+[\zeta(z)] \qquad 4.2(8)$$

where G_1^-, G_1^+, and G_2^+ represent the incident phonon, transmitted phonon, and transmitted magnon (spin wave), respectively. The transmission coefficient for the transmitted magnon is M_{12}, where from 3.1(30)

$$M_{12} = \frac{\Gamma(1-\rho_2+\rho_1)\,\Gamma(\sigma_1-\sigma_2)}{\Gamma(1-\rho_2-\sigma_2)\,\Gamma(\sigma_1+\rho_1)} \qquad 4.2(9)$$

Due to the normalizations included in the definitions of A_m and A_p, the power conversion efficiency η_{pm} is

$$\eta_{pm} = \left| \frac{A^+_{m2}}{A^-_{p1}} \right|^2 \qquad 4.2(10)$$

From the second of equations 4.2(4) we find the relation between A^+_{m2} and A^+_{p2}. Moreover, the ratio A^+_{p2} to A^-_{p1} is from 4.2(8) simply M_{12}. Hence

$$\eta_{pm} = \left| \frac{2}{a}(k_p - k_2^+) \right|^2 \left| M_{12} \right|^2 \qquad 4.2(11)$$

To evaluate $|M_{12}|^2$ under the assumption of a lossless medium, proceed as in subsection 3.2(a)5. Using the properties 3.2(26), 3.1(6) and 3.3(60) of the gamma function, and identifying the σ's and ρ's through 3.3(17) and 3.3(18) we obtain

$$|M_{12}|^2 = \frac{(k_1^+ - k_1^-)(k_2^- - k_1^-)\sinh\left[\frac{\pi}{\alpha}(k_1^+ - k_1^-)\right]\sinh\left[\frac{\pi}{\alpha}(k_2^- - k_2^+)\right]}{(k_2^+ - k_2^-)(k_1^+ - k_2^+)\sinh\left[\frac{\pi}{\alpha}(k_2^- - k_1^-)\right]\sinh\left[\frac{\pi}{\alpha}(k_1^+ - k_2^+)\right]} \qquad 4.2(12)$$

For the conversion from an elastic wave to a spin wave, the wavenumbers are as shown schematically in the top half of Fig. 7 with k_3 replaced by k_2. Hence, when $\pi k_p >> \alpha$, all of the sinh factors in 4.2(12) can be approximated by exponentials, except for the one involving $\pi(k_1^+ - k_1^-)/\alpha$. In fact

$$\sinh\beta \approx \tfrac{1}{2} e^{\beta}, \qquad \beta \gg 1, \qquad 4.2(13a)$$

and

$$2\sinh\beta\, e^{-\beta} = 1 - e^{-2\beta}. \qquad 4.2(13b)$$

In this way, $|M_{12}|^2$ becomes, approximately:

$$|M_{12}|^2 = \frac{(k_1^+ - k_1^-)(k_2^- - k_1^-)}{(k_2^+ - k_2^-)(k_1^+ - k_2^+)}(1 - x); \qquad \pi k_p >> \alpha \qquad 4.2(14)$$

where

$$x \equiv \exp\left[-\frac{2\pi}{\alpha}(k_1^+ - k_1^-)\right] \qquad 4.2(15)$$

Next note from 4.2(7) that

$$k - k_p = \frac{k_m - k_p}{2}\left\{1 \pm \left[1 + \left(\frac{a}{k_m - k_p}\right)^2\right]^{\frac{1}{2}}\right\} \qquad 4.2(16)$$

If $(k_m - k_p)^2 >> a^2$ both at $z = -\infty$ and $z = +\infty$, then the waves at those limits are almost purely spin or elastic waves. From 4.2(16), we then also have

$$k_1^+ - k_1^- \cong \frac{a^2}{4} \frac{(k_m^- - k_m^+)}{(k_p - k_m^+)(k_m^- - k_p)} \qquad 4.2(17)$$

Furthermore, from 3.3(23) we can write

$$\alpha = \frac{k_m'\big|_{cr}(k_m^+ - k_m^-)}{(k_p - k_m^-)(k_m^+ - k_p)} \qquad 4.2(18)$$

Thus

$$x = \exp(-\pi a^2/2|k'_m|_{cr}); \quad (k_m^{\pm} - k_p)^2 \gg a^2 \qquad 4.2(19)$$

regardless of whether the crossover point occurs at the center of the hyperbolic tangent variation or not. If $M\lambda$ is constant, from 4.2(2) we see that

$$x = \exp(-H'_{crit}/|H'|) \qquad 4.2(20)$$

where

$$H'_{crit} \equiv \pi a^2 k_p M\lambda , \qquad 4.2(21)$$

and H' is to be evaluated at the magnetoelastic crossover.

To obtain η_{pm}, combine 4.2(11), 4.2(14), and 4.2(16) under the assumption $(k_m^{\pm} - k_p^2)^2 \gg a^2$. In this case, using 4.2(17), it is seen that the multiplicative factors in 4.2(14) are all cancelled out, leaving

$$\eta_{pm} = (1 - x) \qquad 4.2(22)$$

With 4.2(20), this is the same as the result obtained numerically by KIRCHNER 1966 (see Fig. 6). Note that it involves only the assumptions listed in 4.2(14) and 4.2(19), plus the neglect of reflected waves. It is not necessary that $(k_m^+ + k_m^-)/2 = k_p$ (symmetrical transition). Clearly the result 4.2(22) depends only on the field gradient at the crossover, and not on the shape of the curve.

2. Time-varying fields.

REZENDE and MORGENTHALER 1969a have derived the "coupled mode" equations for spin and elastic waves in a time-varying internal magnetic field H(t), assuming negligible spatial variation. Such a field can be produced by pulsing a Helmholtz coil wrapped around a cylindrical rod of magnetic material, magnetized initially to saturation along the cylinder axis. Their equation (65) has the same form as 4.2(4):

$$\dot{P} - j\omega_p P = j\left(\frac{b}{2}\right) Q$$
$$\dot{Q} - j\omega_k(t) Q = j\left(\frac{b}{2}\right) P , \qquad 4.2(23)$$

where the dot denotes differentiation with respect to time. Combining equations 4.2(23) results in

$$\ddot{P} - j(\omega_k(t) + \omega_p)\dot{P} + \left(\frac{b^2}{4} - \omega_k(t)\omega_p\right) P = 0 , \qquad 4.2(24)$$

which has the same form as 4.2(5). Here $\omega_k(t)$ and ω_p are the magnon and phonon frequencies, respectively, and b is related to the magneto-elastic splitting parameter a. The magnon frequency is given by 4.2(2), where it is assumed k_m^2, M, and λ are constants. Thus $\omega_k(t)$ is simply proportional to H(t). The variables P and Q are normalized such that their absolute magnitudes are proportional to the phonon and magnon momenta, respectively.

The analysis of 4.2(a)1 can be used for this case with the following changes. Instead of transformation 3.3(5) from z to ζ, we write

$$\zeta = -e^{\alpha t} \qquad 4.2(25)$$

in which case 4.2(24) is transformed to the hypergeometric equation 3.3(3) provided that $\omega_k(t)$ has a tanh $(\alpha t/2)$ variation (compare 3.3(13)). Instead of 3.3(16), write

$$P^{\pm}(t) = \exp(+ j\omega^{\pm} t) \qquad 4.2(26)$$

and then 3.3(17) and 3.3(18) become

$$\rho_i = + j\omega_i^{-}/\alpha , \qquad \sigma_i = - j\omega_i^{+}/\alpha , \qquad 4.2(27)$$

with

$$\alpha = \frac{4\dot{H}(t = 0)}{(H_+ - H_-)} . \qquad 4.2(28)$$

The plus and minus signs refer to $t = \pm\infty$, respectively. It is assumed that there is an elastic wave (phonon) incident from $t = -\infty$. Finally, the dispersion relation is

$$\omega = \frac{\omega_k + \omega_p}{2} \pm \left[\left(\frac{\omega_k - \omega_p}{2} \right)^2 + \frac{b^2}{4} \right]^{\frac{1}{2}} . \qquad 4.2(29)$$

For an incident phonon, we assume that the frequency $\omega_k(t)$ increases with time through the crossover point $\omega_k = \omega_p$ so that $\omega_2^+ > \omega_1^- \gtrsim \omega_1^+ > \omega_2^-$ (see Fig. 4, noting that k is constant). Now the analysis can be carried out exactly as in 4.2(a)1, with the result that the elastic-to-spin momentum conversion efficiency is

$$\eta_{pm} = (1 - x); \quad \pi\omega_p \gg \alpha, \ (\omega_k^{\pm} - \omega_p)^2 >> b^2 , \qquad 4.2(30)$$

with

$$x = \exp(- \dot{H}_{crit} / |\dot{H}|) , \qquad 4.2(31)$$

and

$$\dot{H}_{crit} = \pi b^2/2 |\gamma\mu_o| . \qquad 4.2(32)$$

The same critical gradient was obtained by REZENDE and MORGENTHALER

1969a.

c. Solution of fourth-order equation including the effects of reflections.

1. Derivation of results.

The two equations in 4.2(1) can be combined into one fourth-order equation in the following way. First take the second derivative of 4.2(1a). Then to find u''' in terms of v and its derivatives, take the first derivative of 4.2(1b). Substitute for u' in this new equation, using 4.2(1a). Note that since k_m^2 is a function of z, there will also be a term involving u in the first derivative of 4.2(1b). However, use 4.2(1b) itself to find u in terms of v' and u''. Finally, find u'' in terms of v''' and v' by taking the first derivative of 4.2(1a). The result of all these substitutions is:

$$v'''' - \frac{(k_m^2)'}{(k_m^2)} v''' + (k_m^2 + k_p^2 + a^2) v'' - \frac{(k_m^2)'}{(k_m^2)} (k_p^2 + a^2) v' + k_m^2 k_p^2 v = 0 \qquad 4.2(33)$$

Looking back to subsection 3.3(f), we see that this equation is identical in form to equation 3.3(30). In a similar manner, an equation for u having the same form as 3.3(33) may be derived from equations 4.2(1). Thus the equations for v and u both have the same "source equation," 3.3(31). Identifying there $p(z) = k_m^2(z)$, $a_2 = k_p^2 + a^2$, $b_2 = k_p^2$, and dropping the subscript of F, we find

$$v(z) = F'(z) \qquad 4.2(34a)$$

$$u(z) = F''(z) + k_p^2 F(z) \, , \qquad 4.2(34b)$$

with

$$F''''(z) + (k_m^2(z) + k_p^2 + a^2)\,F''(z) + k_m^2(z)\,k_p^2\,F(z) = 0 \qquad 4.2(35)$$

In homogeneous regions where k_m^2 is constant, the solutions for F(z) can be written as normal modes exp (- jkz), producing from 4.2(35) the biquadratic dispersion relation:

$$k^4 - (k_m^2 + k_p^2 + a^2)\,k^2 + k_m^2 k_p^2 = 0 \qquad 4.2(36)$$

The solutions of 4.2(36) can be written in the form:

$$k^2 - k_p^2 = \frac{(k_m^2 - k_p^2 + a^2)}{2}\left\{1 \pm \left[1 + \left(\frac{2a\,k_p}{k_m^2 - k_p^2 + a^2}\right)^2\right]^{\frac{1}{2}}\right\} \qquad 4.2(37)$$

Thus for each wave type represented by a solution of 4.2(37) for k^2, there are two waves having the same magnitude of k, but traveling in opposite directions. Near the crossover point $k_m = k_p$, the dispersion relation can be approximated by 4.2(6), with solutions of the form of 4.2(16), as was shown by SCHLÖMANN and JOSEPH 1964. Then, however, the reflected waves are neglected.

Since 4.2(35) has only one varying parameter, $k_m^2(z)$, it can be transformed into a generalized hypergeometric equation, using the method of subsection 3.3(b), as long as $k_m^2(z)$ has a hyperbolic tangent variation with z (compare the solutions in 4.2(b)1). The solutions of this hypergeometric equation can be written in the form of 3.1(31). By taking linear combinations of these solutions according to the prescription in subsection 3.3(g), the reflection and transmission factors $R_{i\ell}$ and $T_{i\ell}$ for the two types of incident waves are derived. Two of these are given in 3.3(62) and 3.3(63) after which it is remarked how to obtain the others by interchange of symbols.

The limiting forms of $v(z) = F'(z)$ at $z = \pm\infty$ are then given by 3.3(47), where the various G's are hypergeometric functions which reduce to normal mode form in those limits. Now label the incident and transmitted elastic waves by the subscript 1 and the reflected wave by 2. Similarly, subscripts 3 and 4 refer to spin waves traveling in the plus and minus z-directions, respectively. Note since the dispersion relation 4.2(36) is biquadratic, that $k_2 = -k_1$, $k_4 = -k_3$. Thus we find from 3.3(64) that

$$v_2^-/v_1^- = -R_{12}, \qquad v_4^-/v_1^- = -R_{14}k_3^-/k_1^-$$
$$v_1^+/v_1^- = T_{11}k_1^+/k_1^-, \qquad v_3^+/v_1^- = T_{13}k_3^+/k_1^- \qquad 4.2(38)$$

For an incident spin wave having a small but finite v_3^-, four ratios similar to 4.2(38) are found, with the first subscript in the R's and T's changed to 3, and k_1^- changed to k_3^-. The ratio v_2^-/v_3^- is $-R_{32}k_1^-/k_3^-$.

From equation (15) of SCHLÖMANN and JOSEPH 1964, the time averaged total power flow can be written

$$\langle S \rangle = -\frac{1}{4}\,\omega\,\mathrm{Im}\,(u^* u' + v^* v' + av^* u) \qquad 4.2(39)$$

The first term represents exchange (magnetic) power $\langle S_m \rangle$, the second term elastic power $\langle S_p \rangle$, and the last term magnetoelastic power. Note that each type of wave is identified by its wavenumber, and each wave will have some u and v, related by equations 4.1(1). In the limiting homogeneous regions, from 4.1(1a) for example:

$$u = \frac{j(k^2 - k_p^2)}{ak} v . \qquad 4.2(40)$$

Except very near the magnetoelastic crossover, each wave will be either predominantly magnetic or elastic in character, with $|u| >> |v|$ or $|v| >> |u|$. Even at the crossover where $u \approx v$ and $k_m = k_p$, however, the magnetoelastic power term in 4.2(39) is the smallest of the three power terms if $a < k_p$, which is always true in practice. Thus the magnetoelastic power is neglected.

In order to calculate the exchange power in a spin wave from solutions such as 4.2(38) for v, it is necessary to use 4.2(40). For example, to calculate the conversion efficiency η_{pm} from incident phonon to transmitted magnon (spin wave), write using 4.2(39) and 4.2(38):

$$\eta_{pm} \equiv \frac{\langle S_m \rangle_3^+}{\langle S_p \rangle_1^-} = \left|\frac{u_3^+}{v_1^-}\right|^2 \frac{\mathrm{Re}\, k_3^+}{\mathrm{Re}\, k_1^-} = \left|\frac{(k_3^+)^2 - k_p^2}{a k_3^+}\right|^2 \left|\frac{v_3^+}{v_1^-}\right|^2 \frac{\mathrm{Re}\, k_3^+}{\mathrm{Re}\, k_1^-}$$

$$= \left|\frac{(k_3^+)^2 - k_p^2}{a k_1^-}\right|^2 |T_{13}|^2 \frac{\mathrm{Re}\, k_3^+}{\mathrm{Re}\, k_1^-} . \qquad 4.2(41)$$

The other power ratios for the various reflected and transmitted waves are found similarly. Even the reflection factors $|R_{i\ell}|^2$ will need to be multiplied by factors if the reflected wave is not of the same type as the incident wave. In a lossless medium the wavenumbers k_1^-, k_3^+, etc., will all be real, as long as $(k_m^{\pm})^2 > 0$.

Using 3.3(53) to obtain T_{13} from T_{11}, and noting 3.3(64), we can obtain η_{pm} in 4.2(41) in the same manner as in 4.2(b)1, with the result

$$\eta_{pm} = \left| \frac{k_p^2 - (k_3^+)^2}{a k_1^-} \right|^2 \frac{[(k_3^-)^2 - (k_1^-)^2][(k_1^+)^2 - (k_1^-)^2]}{[(k_3^-)^2 - (k_3^+)^2][(k_1^+)^2 - (k_3^+)^2]} (1 - x)(1 - y), \quad 4.2(42)$$

provided that $\pi k_p >> \alpha$, with x as defined in 4.2(15) and

$$y \equiv \exp(-4\pi k_3^+/\alpha) \quad 4.2(43)$$

Now α is defined from 3.3(20) in terms of $k_m^2(z)$, rather than in terms of $k_m(z)$ as in 4.2(b)1. Note that in the opposite limit $\pi k_p << \alpha$, instead of 4.2(13a) we have $\sinh \beta \approx \beta$, $\beta << 1$. Then the form 4.2(42) is no longer valid, but the answer no longer depends on α and hence not on the field gradient. In that case, the thickness of the hyperbolic tangent transition is less than the wavelengths at the crossover, so that the answers should then be those obtainable from boundary conditions.

To simplify 4.2(42), note from the solutions 4.2(37) to the dispersion relation when $(k_m^{\pm})^2 - k_p^2 + a^2 >> 2a\, k_p$ that

$$k_1^2 - k_p^2 \cong -a^2 k_p^2/(k_m^2 - k_p^2 + a^2). \quad 4.2(44)$$

Then following steps similar to those involved in 4.2(17) and 4.2(18), and by neglecting a^2 compared with $(k_m^{\pm})^2 - k_p^2$, we find

$$\frac{2\pi}{\alpha}(k_1^+ - k_1^-) \cong -\frac{\pi a^2 k_p}{(k_m^2)'_{cr}}, \quad 4.2(45)$$

whether or not the crossover point occurs at z = 0 in the tanh $(\alpha z/2)$ variation for $k_m^2(z)$. If $M\lambda$ is constant, 4.2(45) and 4.2(2) combined with the definition 4.2(15) for x show that again x has the form of 4.2(20), with the same critical gradient H'_{crit} as in 4.2(21). Furthermore, the multiplicative terms in 4.2(42) cancel out when it is observed from 4.2(45) that

$$(k_1^+)^2 - (k_1^-)^2 \cong a^2(k_m^-)^2/[(k_m^-)^2 - k_p^2], \qquad 4.2(46)$$

when $k_m^+ << k_p$ and $a << k_p$. We also use the facts that $k_1^+ \approx k_p$, $k_3^+ \approx k_m^{\pm}$, since we are assuming $k_m^2 - k_p^2 >> 2ak_p$. Thus we finally obtain

$$\eta_{pm} \cong (1 - x)(1 - y) \qquad 4.2(47)$$

The results of a representative computer plot of η_{pm} from 4.2(41) without any assumptions being made are shown in Fig. 8. (The evaluation by computer involved the use of the subroutine LOGGAM mentioned in subsection 3.3(i).)

2. Discussion of results.

Figure 8 shows that the result $\eta_{pm} \cong 1 - x$, derived in 4.2(b)1, is valid until quite large field gradients are reached in this example. The numbers for k_p and $H'_c \equiv H'_{crit}$ are taken from typical experimental conditions for coupling of spin waves and longitudinal elastic waves in HU, REZENDE, and MORGENTHALER 1970. Reflections are important only when y becomes significant, which occurs when the appropriate WKB solutions break down. The criterion for this breakdown for propagating spin waves describable by the second-order equation 4.1(1b), without the magnetoelastic coupling term, is $(k_m)'/k_m^2 \gtrsim 1$. (See section 2.2.) Thus reflections are generated for large gradients or small k_m. The minimum value of k_m ($k_m^{Min} \approx k_3^+$ for an incident elastic wave - see the top half of Fig. 7) may be taken as the value at the turning point. See section 4.3 below for a discussion of how exchange spin-wave power is reflected from this turning point and largely converted into electromagnetic power in a reflected "magnetostatic" spin wave. There it is shown that the wavenumber k_{TP} at the turning point is $(k_t^2/2\lambda)^{\frac{1}{4}}$, where k_t is the transverse wavenumber of the sample. For Fig. 8 it is assumed that $k_t \cong 10\ \mathrm{cm}^{-1}$, a reasonable value for a cylindrical YIG rod of diameter 0.3 cm.

It is interesting that the factor y defined by 4.2(43) is identical in form to the expression for the fraction of electromagnetic power reflected from an incident plane wave traveling in a dielectric whose permittivity varies in a hyperbolic tangent manner. This power fraction is given as the square of the reflection coefficient R_E in equation 3.2(29), an expression first derived by EPSTEIN 1930b. The reflection in that case is also proportional to the severity of the breakdown in the WKB solutions. Furthermore, for the same shape curve and same gradients, these reflections become larger as either the original or final wavenumbers, k^- or k^+, approach zero. Physically, this corresponds to approaching a cutoff situation where the wavenumber becomes negative. In this latter case, of course, all of the energy is reflected. We conclude that the reflections represented by the factor y in 4.2(47) depend most strongly on the shape of the variation in k_m near its minimum value, and upon that minimum itself. The relative distribution of the energy in the transmitted waves is determined only by the gradient $(k_m^2)'$ or H' at the crossover point, however, and is insensitive to the variation in k_m^2 away from that point.

With these interpretations in mind, it is instructive to look at

Fig. 7, where are summarized all the reflected and transmitted power fractions for an incident phonon or magnon. These results are obtained in the same manner as 4.2(47), starting from expressions corresponding to 4.2(41). Figure 7 indicates schematically how these results may be interpreted. For the incident phonon (elastic wave), a fraction x of the energy is transmitted as phonon power. The remainder (1 - x) is transmitted into the spin wave. Before this energy reaches the homogeneous region, however, a fraction y is reflected from near the minimum in k_m due to WKB breakdown. This amount y(1 - x) is in turn split into two parts at the crossover point, a fraction (1 - x) of it being reconverted to a reflected phonon. Note that the fraction x represents the amount of power which is not converted at the crossover point, regardless of whether phonon or magnon power is incident.

The results for an incident magnon are shown in the bottom half of Fig. 7. The starting value of k_m is chosen to be less than k_p, because this is the way the situation occurs in practice. Once energy gets into spin waves with wavenumbers greater than k_p, it is usually lost, unless it can again be converted into another kind of elastic wave, such as a shear wave with higher k_p. Figure 7 shows that again a fraction y of the magnon power is reflected. The remainder is split at the crossover according to the usual factor x. Now, however, the wavenumbers are so large that there is negligible power reflected into the oppositely-traveling phonon, contrary to the result in SCHLÖMANN and JOSEPH 1964.

The factor ϵ in the expression for the reflected phonon power in Fig. 7 becomes noticeable only when α approaches πk_p (L in Fig. 7 is the "length of the transition," defined in Fig. 1 as $4/\alpha$). In this case, however, the approximation leading to 4.2(42) breaks down, corresponding to the fact that L is then smaller than the wavelengths at the crossover point. Then the amount of reflected and transmitted power should become insensitive to L because there is effectively a sharp boundary. This result is confirmed in Fig. 8 in the region of very high gradients. In that region it is possible to use the boundary conditions to obtain the reflections and transmissions. Results obtained in that manner will be the same as those obtained in this section only if $k_m^2(z)$ is constrained to keep the tanh (z) variation. In order physically to obtain such an abrupt transition in the magnetic field H(z) it would undoubtedly be necessary to have a non-constant magnetization in order to satisfy $\nabla \cdot \vec{B} = 0$.

Note now, however, that SCHLÖMANN and JOSEPH 1964 obtained a significant amount of power in the reflected phonon, even with relatively small field gradients (see their Fig. 3). Their analysis involved a successive approximations approach with $k_m(z)$ expanded in a power series about the crossover point. It is concluded that such an analysis for reflected waves is not dependable. Reflections generally depend upon breakdown of WKB solutions, but successive approximations schemes, as mentioned in section 2.2, depend for their success upon the approximate validity of WKB quasi-normal mode solutions. (Equations (51) in SCHLÖMANN and JOSEPH 1964 are in fact WKB quasi-normal mode solutions of the type of equation 2.2(14.)

In the turning point problems of section 4.3, the WKB quasi-normal mode solutions break down drastically due to the presence of a cutoff point for all propagating waves. Then successive approximations are of no help. Finally, in section 4.4 an analysis is presented which treats the simultaneous presence of crossover and turning points.

4.3 Medium-k to high-k spin wave conversion at a fourth-order turning point.

(a) Introduction.

1. Physical and historical background.

The essential feature of conversion from medium-k ("magnetostatic") to high-k (exchange-dominated) spin waves is that it occurs where the energy velocity in a lossless medium goes to zero. Such a point is called a turning point by analogy with quantum mechanical potential wells, since beyond it energy cannot propagate. Turning points arise in magnetoelastic delay lines because the wavenumbers k_{Ex}^2 (the same as k_m^2 in 4.2(2) and k_+^2 in 4.1(15)) and k_{MS}^2 (given by 4.1(23)) depend on the static internal magnetic field H(z), which is a function of distance in nonellipsoidal samples such as rods (see Fig. 3, based on the formula of SOMMERFELD 1964). For reference we reproduce here equations 4.1(15) and 4.1(23):

$$k_{EX}^2 = \frac{H_{TP} - H}{M\lambda} \qquad 4.3(1)$$

$$k_{MS}^2 = k_t^2 \left(\frac{H^2 + HM - H_{TP}^2}{H_{TP}^2 - H^2}\right) \qquad 4.3(2)$$

where

$$H_{TP} \equiv \omega / |\gamma \mu_o| \qquad 4.3(3)$$

Equation 4.3(1) shows that the wavenumber k_{EX}, usually very large because of the small magnitude of the exchange parameter λ, becomes very small near $H(z) = H_{TP}$. In fact, for $H(z) > H_{TP}$, k_{EX}^2 is negative, indicating that exchange-dominated spin waves are cutoff, or evanescent, in such regions. Since a small wavenumber corresponds to a large wavelength, it was originally expected that exchange-dominated spin waves could couple directly to long-wavelength (low-k) electromagnetic radiation near where $H(z) = H_{TP}$ (see SCHLÖMANN 1961 and 1964).

More recent evidence, however, has indicated that low-k electromagnetic radiation couples to exchange-dominated spin waves only through the intermediary of the "magnetostatic" waves. First of all, it was noticed that echoes having the characteristics of "magnetostatic" wave echoes were observed even when the static applied field was large enough that a turning point existed with the sample (see KEDZIE 1966 and 1968, LEWIS and LACKLISON 1966). Previously, "magnetostatic" wave echoes were observed in rods only when H(z) was somewhat less than H_{TP} throughout the length of the sample, so that the wave could propagate ($k_{MS}^2 > 0$) practically from one end of the rod to the other and back. Later, by Bragg diffraction of infrared laser light, COLLINS and WILSON 1968 confirmed the evidence that "magnetostatic" waves could propagate when a turning point was present. Furthermore AULD, COLLINS, and WEBB 1968 showed in a series of experiments with fine wire antenna excitation of spin waves in rods, that the best efficiency of excitation occurred when the antenna was near the end of the rod, not near the turning point. Near the end of the rod where H(z) is smaller, k_{MS} in 4.3(2) would tend to be small, approximately equal to the wave-number $k_o = \omega(\mu_o \epsilon)^{\frac{1}{2}}$ of electromagnetic radiation from the antenna. In that region, however, k_{EX}^2 would be extremely large, precluding any coupling to electromagnetic radiation.

To understand in more detail how a "magnetostatic" wave can serve as an intermediary between electromagnetic radiation and exchange-dominated spin waves, recall the discussion following 4.1(23). Where the numerator of 4.3(2) becomes small, the wavenumber k_{MS} becomes comparable to electromagnetic wavenumbers. Then the "magnetostatic" wave should be able to couple to electromagnetic fields in a manner similar to that postulated by SCHLÖMANN 1964 for electromagnetic coupling to exchange-dominated spin waves, as mentioned above. For higher magnetic field H(z) (corresponding to positions further toward the interior of the rod in Fig. 3), k_{MS} in 4.3(2) becomes larger. Because $\lambda k_t^2 <<< 1$ for the usual practical situations, k_{MS} is usually much smaller than k_{EX}, as is seen by reference to 4.3(1) and 4.3(2). However, for H(z) approaching H_{TP}, the denominator in 4.3(2) becomes very small, and k_{MS} rapidly approaches infinity. In plasma terminology this phenomenon represents a resonance of the extraordinary electromagnetic wave (the "magnetostatic" wave in our case), and occurs in our case at the field required for ferromagnetic resonance in an infinite sample. Recall that at the same field H_{TP} the exchange-dominated spin wave experiences a cutoff. Neither type of spin wave can propagate for higher internal fields H(z), since then k_{MS}^2 and k_{EX}^2 both become negative.

As will be shown below, the complete dispersion relation for the spin system results from the confluence of the two second order dispersion relations in 4.3(1) and 4.3(2) (see also Fig. 4). As a consequence, the wavenumber for the combined system still changes rapidly with magnetic field near the turning point (see Fig. 5), but does not go either to zero or to infinity. This wavenumber k_{TP} is nevertheless much greater than electromagnetic wavenumbers, so it is quite understandable that AULD, COLLINS, and WEBB 1968 found that coupling from electromagnetic radiation to the spins does not occur at the turning point. Basically, the wavenumber does not go to zero because of the necessity of including the effects of the finite transverse dimensions of the material when the wavelengths become comparable in size to those dimensions. On the other hand, the wavenumber does not go to infinity because the exchange forces tend to keep spins on neighboring atoms closely aligned, and hence tend to preclude variations with extremely short wavelengths.

A dispersion relation of the form of Fig. 4 showing the confluence of the exchange-dominated and "magnetostatic" spin waves first appeared in FLETCHER and KITTEL 1960. COLLINS 1967 was the first, however, to use this dispersion relation to conclude that "magnetostatic" spin waves could act as an intermediary between the electromagnetic and exchange-dominated spin waves, as described above. This conclusion followed the experimental evidence that waves of the "magnetostatic type" could exist simultaneously with the high wavenumber magnetoelastic waves (KEDZIE 1966 and LEWIS and LACKLISON 1966). AULD, COLLINS, and WEBB 1968 explained the conversion further using a wavevector diagram noting the fact that "magnetostatic" waves act like guided waves formed from the resolution of two plane waves traveling at an angle with respect to the static magnetic field (z-direction).

2. Summary of the results of this section.

The combined dispersion relation for the spin system was first derived rigorously by VASILE and LAROSA 1968a by first deriving the relevant fourth order differential equation. In part 4.3(b)1 below we rederive this equation by combining the two coupled wave equations 4.1(19) and 4.1(23), or equivalently 4.1(19) and 4.1(21). We also indicate how a similar equation can be derived for modes with no azimuthal variation in cylindrically-symmetric rods.

Next we find the WKB quasi-normal mode solutions for the fourth order spin system in a nonuniform static field H(z) by following the procedure outlined in section 2.2. After expressions for electromagnetic and "exchange" power flow are developed, it is shown how the total power in each of these quasi-normal modes is conserved. As was mentioned in Chapter 2, this power conservation is the essential characteristic of the WKB solutions, regardless of the order of the differential equation. Consequently, the "amplitude" factors multiplying the exponential phase integrals in these solutions do not have the same form for fourth order systems as for second order systems. This latter fact was not realized by VASILE and LAROSA 1968a, who then concluded that there was no difficulty in matching two WKB solutions representing a "magnetostatic" and an exchange-dominated spin wave at the turning point. They took the second-order form $(k)^{-\frac{1}{2}}$, where k is the "local" wavenumber of the quasi-normal mode, as the "amplitude" factor. As mentioned above, the wavenumber at the turning point does not go either to zero or to infinity. Rather, the wavenumbers for the "magnetostatic" and exchange-dominated spin waves simply become equal there. The "amplitude" factors in the properly-constructed WKB solutions, however, are inversely proportional to the difference in these wavenumbers, and hence become very large near the turning point. For this reason, it is impossible to know how to match the WKB solutions at that point. Furthermore, for the same reason it is impossible to apply a successive approximations scheme based on equation 2.2(26), as is shown below in part 4.3(b)3.

When a successive approximations scheme fails, it is customary to turn to the WKB method, as was mentioned in Chapter 2. Recall that the WKB quasi-normal mode solutions can be found as approximate solutions of coupled first order equations, when the coupling is weak. For second order systems, this weak coupling is equivalent to very little power being reflected from an incident wave. Successive approximations schemes are based on sets of coupled first or second order equations. The WKB method, however, uses asymptotic expansions of solutions of the one higher-order differential equation which completely describes the system. Such a treatment is necessary for a second order system, for example, when a wave in an inhomogeneous medium reaches a cutoff. The asymptotic solutions of the complete equation are then matched to the quasi-normal modes in regions where the latter are approximately valid.

In subsection 4.3(d) below, we apply the WKB method to a fourth order equation of the type derived by VASILE and LAROSA 1968a and again in part 4.3(b)1. This treatment is based on asymptotic expansions of a similar equation developed in the Orr-Sommerfeld theory of hydrodynamical stability by WASOW 1950 and RABENSTEIN 1958. Their work has been applied to at least three distinct situations in plasma physics, as was indicated in section 2.4. However, this present work appears to contain the first application to the problem of "magnetostatic" to exchange-dominated spin wave conversion.

The result of the WKB method is that "magnetostatic" waves are completely converted into exchange-dominated spin waves at the turning point and vice versa. As is shown in subsection 4.3(d), however, this result is strictly valid only when either the magnetic field gradient at the turning point is very small, or when the waves can be observed at reasonably large distances from the turning point. Otherwise error terms in the asymptotic expansions can be so large as to mask other reflected waves.

In microwave magnetoelastic delay lines, the typical existence of reasonably high gradients of the internal static magnetic field H(z) and the close proximity of the magnetoelastic coupling point to the turning point thus introduces some doubt about the reliability of the WKB result. Furthermore, KEDZIE 1966 and 1968 and LEWIS and LACKLISON 1966 in pulse-echo experiments have observed varying amounts of "magnetostatic" waves reflected directly from the turning point when a "magnetostatic" wave was apparently incident. LEWIS and SCOTTER 1969 concluded from continuous wave absorption resonances that there was some reflection of this type, but most of the energy was converted to the exchange-dominated spin wave. Finally, absorption experiments by KOHANE, SCHLÖMANN, and JOSEPH 1965 may indicate that exchange-dominated spin waves are reflected from the turning point when waves of the same type are incident. This conclusion may follow because the high Q of the observed magnetoelastic resonances implied the existence of standing waves, while the energy reflected near the magnetoelastic cross-over as calculated by the method in subsection 4.2(d) may not be sufficient to account for such standing waves. See section 4.4 for a treatment by generalized hypergeometric functions of the combined (sixth-order) turning point-crossover point problem.

In subsection 4.3(e) below we improve on the WKB method for magnetostatic to exchange spin wave conversion by solving the fourth order equation in the manner of section 3.3 using generalized hypergeometric functions. In this way, reflections into both types of spin waves are treated naturally, regardless of which type is incident. We show how the results reduce to those of the WKB method when either the magnetic field gradient at the turning point is quite small, or when the wavenumbers of the incident and reflected waves are far from the wavenumber k_{TP} at the turning point. Recall that these conditions are exactly those under which the error terms of the asymptotic expansions in the WKB method are negligible.

A difficulty in interpretation which remains is how to determine what the wavenumbers of the incident waves should be in this model. The "magnetostatic" wave for low enough $H(z) < H_{TP}$ eventually couples to the ordinary electromagnetic wave, but long before that the exchange-dominated spin wave will couple to an elastic wave. Another difficulty may result from the fact that the variation of H(z) in a rod, for example, does not correspond to the hyperbolic tangent variation assumed for the generalized hypergeometric function treatment (compare H(z) in Fig. 3 and Fig. 1). In certain configurations with reasonably thin films, however, a hyperbolic tangent variation may be accurate.

The analysis of subsection 4. 3(e) is also extended to include the effects of loss. Note in Fig. 9 that loss tends to split the real part of the wavenumber dispersion relation at the turning point, in a manner which seems similar to the magnetoelastic splitting at the crossover point (see Fig. 4). One would expect that such a splitting might generate increased reflections, as well as absorption. The results, however, are somewhat difficult to interpret for the reasons given in the preceding paragraph.

Finally, the expressions for the reflection coefficients found from generalized hypergeometric functions are compared near the end of 4. 3(e) with the results of a phase-integral analysis given in subsection 4. 3(c). This analysis is based on the phase-integral method of RYDBECK 1967, described in part 2. 4(b)2. A conclusion from these analyses is that power reflected from a wave incident upon the turning point into the oppositely-traveling wave of the same type becomes substantial only when the magnetic field gradient H'(z) at the turning point exceeds a critical value. For magnetic fields H(z) with variations more like Fig. 3 than Fig. 1, this critical gradient is probably proportional to $M\lambda k_{TP}^3$, where $k_{TP} = (k_t^2/2\lambda)^{\frac{1}{4}}$ is the wavenumber for both types of spin waves at the turning point (see equation 4.3(7) and following).

4.3(b) Basic properties.

1. Derivation of basic fourth-order differential equation and dispersion relation.

When there is only one transverse component of magnetic field present, the coupled wave equations for the "magnetostatic" and exchange-dominated spin waves take the form of 4.1(22) and 4.1(19). In the latter equation, we now neglect the coupling to elastic waves represented by the derivative of the lattice displacement (assumption 19b). Recall that while 4.1(19) and 4.1(22) are written assuming $\underline{h}_y = 0$, if the subscript x is replaced everywhere by r, these equations also hold for waves with no azimuthal (φ) component of magnetic field in cylindrically symmetric rods (see assumptions 6). Under certain circumstances, "magnetostatic" waves with more than one transverse magnetic field component can be excited, but it is expected that they will couple to exchange-dominated spin waves in a manner similar to that described in this section.

Rather than using 4.1(22), it is simpler to use 4.1(21), which with 4.1(19) is equivalent to 4.1(22) (see the derivation immediately preceding 4.1(22)). Thus we have:

$$\underline{h}_x'' - k_t^2 \underline{h}_x = k_t^2 \underline{m}_x \qquad 4.3(4)$$

and

$$\underline{m}_x'' + k_{EX}^2 \underline{m}_x = \frac{-H}{(H+H_{TP})\lambda} \underline{h}_x , \qquad 4.3(5)$$

with k_{EX}^2 and H_{TP} defined as in 4.3(1) and 4.3(3), respectively. Now simply take the second derivative of 4.3(4), substitute for $\underline{m}_x''$ using 4.3(5) and for $\underline{m}_x$ using 4.3(4) itself. The result is:

$$\underline{h}_x'''' + k_{EX}^2 \underline{h}_x + k_{EX}^2 k_{MS}^2 \underline{h}_x = 0 \qquad 4.3(6)$$

where k_t^2 is again ignored compared with k_{EX}^2 in accordance with assumption 13. Recall from 4.3(1) and 4.3(2) that k_{EX}^2 and k_{MS}^2 go to zero and infinity at the turning point, respectively. For fields H(z) reasonably close to H_{TP}, however, the product $k_{EX}^2 k_{MS}^2$ is approximately a constant:

$$k_{EX}^2 k_{MS}^2 = \frac{k_t^2}{\lambda} \frac{(H^2 + HM - H_{TP}^2)}{(H_{TP} + H)M} \cong \frac{k_t^2}{2\lambda} \equiv k_{TP}^4 ; \; H \cong H_{TP} \qquad 4.3(7)$$

Here k_{TP} is the wavenumber at the turning point, as is shown below.

The dispersion relation corresponding to 4.3(6) is

$$k^4 - k_{EX}^2 k^2 + k_{EX}^2 k_{MS}^2 = 0, \qquad 4.3(8)$$

valid for homogeneous regions where k_{EX} and k_{MS} are not functions of position, so that the solutions of 4.3(6) can be written as normal modes of the form exp (-jkz). The solutions of 4.3(8) are

$$k^2 = \frac{k_{EX}^2}{2}\left[1 \pm \left(1 - \frac{4k_{MS}^2}{k_{EX}^2}\right)^{\frac{1}{2}}\right]. \qquad 4.3(9)$$

Note that when $k_{EX} >> k_{MS}$ (far from the turning point), the two solutions in 4.3(9) reduce to approximately the unperturbed values $k^2 = k_{MS}^2$ and $k^2 = k_{EX}^2$.

On the other hand, when $k_{EX} = 2k_{MS}$, the two solutions in 4.3(9) coalesce. Since $\lambda k_t^2 << 1$, 4.3(1) and 4.3(2) show that this coalescence can only occur when $H \cong H_{TP}$. Thus $k_{EX}^2 k_{MS}^2 \cong k_{TP}^4$, and we see from 4.3(9) that the common value of k^2 to which the two solutions coalesce is k_{TP}^2. (See the dispersion relations in Figs. 4 and 5.) From the condition $k_{EX} = 2k_{MS}$, we find from 4.3(1) and 4.3(2) that this coalescence occurs for H(z) slightly displaced from what has been defined in H_{TP}. Explicitly:

$$k = k_{TP} \text{ when } H_{TP} - H = M(2\lambda k_t^2)^{\frac{1}{2}} \qquad 4.3(10)$$

Typically the difference from H_{TP} is less than 0.1 oersted. The reason for the difference is that H_{TP} was defined as the value of H(z) which would cause k_{EX} to go to zero (cutoff) when the "magnetostatic" waves are neglected.

VASILE and LAROSA 1968a have labeled the place where 4.3(10) is satisfied as the branch point, since even in lossless media the solutions for k^2 in 4.3(9) become complex conjugates when $|H - H_{TP}| \le M(2\lambda k_t^2)^{\frac{1}{2}}$. These authors used this fact to show qualitatively how one type of spin wave couples to the other near this point. Finally, for $H - H_{TP} > M(2\lambda k_t^2)^{\frac{1}{2}}$, no energy propagates in lossless media, since then both solutions for k^2 become negative. The corresponding solutions for k are hence both imaginary, representing evanescent waves. (See Fig. 6 of VASILE and LAROSA 1968a for a plot of the wavenumber loci.)

To treat the effects of losses, it is customary to replace H in

4.3(1) and 4.3(2) by $H + j\frac{\Delta H}{2}$, where ΔH is the full spin-wave linewidth. Reasonably near the turning point, 4.3(7) shows that $4k_{MS}^2/k_{EX}^2 = 4(k_{TP}/k_{EX})^4 = 2M^2(\lambda k_t^2)/(H_{TP}^2 - H^2)$. Then even at $H = H_{TP}$, we see that the two solutions for k^2 from 4.3(9) are basically the unperturbed values of 4.3(1) and 4.3(2), as long as $(\Delta H)^2 >> (\Delta H_{crit})^2$, where

$$\Delta H_{crit} = M(8\lambda k_t^2)^{\frac{1}{2}} . \qquad 4.3(11)$$

For rods of yttrium iron garnet (YIG) such as those used in practice, $\Delta H_{crit} \approx 0.1$ oersted. This is approximately the observed value for the linewidth at room temperature in the purest YIG now commercially available (see VASILE and LAROSA 1968c, and also COMSTOCK 1965). Note from Fig. 9 how the dispersion relation does split for linewidths greater than ΔH_{crit}. In that figure, $k_t = 20\ cm^{-1}$, $M\lambda = 5 \times 10^{-9}\ Oe\ cm^2$, $\Delta H = 0.36$ Oe, and $\Delta H_{crit} = 0.17$ Oe. (See also KEDZIE 1968.)

2. <u>Derivations of expressions for power flow.</u>

As indicated earlier, the "magnetoelastic" spin wave carries mostly electromagnetic power, since it is really an extraordinary electromagnetic wave. The exchange-dominated spin wave, on the other hand, carries "exchange" power resulting from the exchange forces acting to align neighboring spins. Now we wish to find explicit expressions for these power flows and write them in terms of the field $\underline{h}_x$ satisfying the differential equation 4.3(6).

Electromagnetic power flow.

The z-component of the time-averaged electromagnetic Poynting vector is, in complex notation:

$$\langle(\vec{S}_{EM})_z\rangle = \frac{1}{2}\,\mathrm{Re}\,(\vec{\underline{e}} \times \vec{\underline{h}}^*)_z\,, \qquad 4.3(12)$$

where the brackets ⟨ ⟩ represent a time average, and fields are assumed to be written in terms of time-harmonic components as in assumption 1. Then Maxwell's curl equations take the form of 4.1(6), of which 4.1(6b) is

$$\nabla \times \vec{\underline{h}} = j\omega\epsilon\,\vec{\underline{e}} \qquad 4.3(13)$$

Recall that in section 4.1 we neglected $\underline{e}_z$ (assumption 3). If we also again assume that $\frac{\partial}{\partial y} \approx 0$ so that $\underline{h}_y = 0$ (assumption 6a), we find $\underline{e}_x = 0$ also. Using the y-component of 4.3(13), however, we find

$$-\,\underline{e}_y\,\underline{h}_x^* = \frac{1}{j\omega\epsilon}\left(\frac{\partial \underline{h}_z}{\partial x} - \frac{\partial \underline{h}_x}{\partial z}\right)\underline{h}_x^*\,. \qquad 4.3(14)$$

All that remains to evaluate 4.3(12) is now to find $\partial\underline{h}_z/\partial x$ in terms of $\underline{h}_x$. This can be done by applying the operators $\nabla_\pm$ to the two equations represented by 4.1(10), then noting that $\nabla_\pm\nabla_\mp = \nabla_t^2 = -k_t^2$ by assumption 4, and finally adding the resulting two equations:

$$\partial\underline{h}_z/\partial x = \frac{k_t^2}{(k_t^2 - k_o^2)}\,\partial\underline{h}_x/\partial z \qquad 4.3(15)$$

Combining 4.3(14) and 4.3(15) along with 4.3(11) produces:

$$\langle S_{EM}\rangle = \frac{1}{2}\,\frac{\omega\mu_o}{(k_t^2 - k_o^2)}\,\mathrm{Im}\,(\underline{h}_x'\,\underline{h}_x^*)\,, \qquad 4.3(16)$$

where we again introduce the prime to denote differentiation with respect to z, and drop the subscript z on S_{EM}.

Note that if we had neglected k_o^2 in 4.3(15) with respect to k_t^2

(assumption 5), the combination of 4.3(14) and 4.3(15) would give zero for the power flow. This situation shows again that assuming $\underline{e}_z = 0$ and $k_t^2 >> k_o^2$ is equivalent to the "magnetostatic" approximation in which the right-hand side of 4.3(13) is zero. $\langle (S_{EM})_z \rangle$ can be calculated in this latter case only if the $\nabla \times \vec{\underline{e}}$ Maxwell equation is used to calculate $\vec{\underline{e}}$ from known $\vec{\underline{h}}$ and $\vec{\underline{m}}$. This is the case when plane wave solutions are assumed. Now, however, if we assume normal mode solutions for $\underline{h}_x$ of the form exp (-jkz), we obtain from 4.3(16):

$$\langle S_{EM} \rangle = -\frac{1}{2} \frac{\omega\mu_o}{(k_t^2 - k_o^2)} \mathrm{Re}(k) \, |\underline{h}_x|^2 , \qquad 4.3(17)$$

where it is now usually permissible to neglect k_o^2 compared with k_t^2 (assumption 5). Also recall that normal mode solutions are not the same as plane waves, since the transverse variations of the fields are not assumed to have exponential form.

The power flow expression 4.3(17) shows that "magnetostatic" waves with $k_t^2 > k_o^2$ are backward waves: the power flows in a direction opposite to that of the phase velocity ω/k. This fact could also have been deduced from the dispersion relation in Fig. 4 where it is clear that the group velocity $\partial\omega/\partial k$ is negative for "magnetostatic" waves. The reason these waves can propagate in closed structures with $k_t^2 >> k_o^2$ is that the dynamic permeability is large and negative in the relevant regions of frequency and wavenumber, allowing an additional surface (evanescent) wave component to match the boundary conditions at the transverse walls. (See, for example, the discussions in VASILE and LAROSA 1968b and also THOMPSON 1962.) In ordinary waveguides, on the other hand, k_t^2 must be less than k_o^2 for propagating modes to exist, in which case 4.3(17) shows that they are ordinary forward waves: $\langle (S_{EM})_z \rangle$ is positive if Re(k) is positive.

For modes in cylindrically symmetric rods satisfying $\frac{\partial}{\partial \varphi} = 0 = \underline{h}_\varphi$ (assumption 6b), an expression of exactly the same form as 4.3(14) is obtained from 4.3(13), with y and x replaced by φ and r, respectively. To go from 4.1(10) to the equation corresponding to 4.3(15), note that $\nabla_\pm \equiv \frac{\partial}{\partial x} \pm j\frac{\partial}{\partial y} = e^{\pm j\varphi}\left(\frac{\partial}{\partial r} \pm j\frac{1}{r}\frac{\partial}{\partial \varphi}\right)$ and $\underline{h}_\pm \equiv \underline{h}_x \pm j\underline{h}_y = e^{\pm j\varphi}(\underline{h}_r \pm j\underline{h}_\varphi)$. By writing all the terms out explicitly, we find

$\nabla_\pm \nabla_\mp \underline{h}_\pm = \partial^2 \underline{h}_r/\partial r^2 + \frac{1}{r}\partial \underline{h}_r/\partial r + \frac{1}{r^2}\partial^2 \underline{h}_\varphi/\partial \varphi^2 = \nabla_t^2 \underline{h}_\pm$, just as before, where the last part of this equation follows from the formula for the Laplacian in cylindrical coordinates. Using assumption 4 again, $\nabla_t^2 = -k_t^2$. Finally 4.3(15) with r replacing x is obtained by combining the two equations resulting from 4.1(10) in the same manner that 4.1(14b) was obtained from 4.1(13). Thus the power flow expressions for modes with no azimuthal variation become 4.3(16) and 4.3(17), if x is replaced by r.

Exchange power flow.

The expression for exchange power flow was obtained by SCHLÖMANN and JOSEPH 1964, assuming that the magnetization is circularly polarized (our assumption 16). This assumption is valid for both kinds of spin waves as long as the wavenumber k is much greater than k_t, as can be seen from 4.1(17) and 4.1(21). Thus we can take the z-component of the time-averaged exchange power flow $\langle S_m \rangle$ from 4.2(41) as $-\frac{1}{2}\omega \,\mathrm{Im}\,(u^* u')$, which in light of the normalization 4.2(3) becomes

$$\langle S_m \rangle = -\tfrac{1}{4}\mu_o \lambda \omega \,\mathrm{Im}\,(\underline{m}_+^* \underline{m}_+'). \qquad 4.3(18)$$

For circularly polarized waves with $\underline{m}_- = 0$, we have $\underline{m}_+ = 2\,\underline{m}_x$ (or $2\,e^{+j\varphi}\,\underline{m}_r$). Now $\underline{m}_x$ can be found in terms of $\underline{h}''_x$ and $\underline{h}_x$ from 4.1(21). When $|\underline{h}''_x| >> k_t^2\,|\underline{h}_x|$, we can neglect the term in $\underline{h}_x$. Hence in that limit:

$$\langle S_m \rangle \;=\; -\frac{\omega\mu_o}{2\,k_t^2\,k_{TP}^4}\;\mathrm{Im}\left[\left(\underline{h}''_x\right)^* \underline{h}'''_x\right], \qquad 4.3(19)$$

where we have used the definition for k_{TP}^4 in 4.3(7). Finally, assuming normal mode solutions of the form exp (-jkz) for $\underline{h}_x$ we obtain:

$$\langle S_m \rangle \;=\; +\frac{\omega\mu_o}{2\,k_t^2 k_{TP}^4}\;\mathrm{Re}\,(k^5)\,|\underline{h}_x|^2\,. \qquad 4.3(20)$$

Note that $\langle S_m \rangle$ is positive when $\mathrm{Re}(k^5)$ is positive, so that waves such as exchange-dominated spin waves with $|\langle S_m \rangle| > |\langle S_{EM} \rangle|$ are forward waves.

Total power flow.

Combining 4.3(16) and 4.3(19) or 4.3(17) and 4.3(20), we find that the z-component of the total time-averaged power flow is

$$\begin{aligned}\langle S_{tot} \rangle \;&=\; \beta\,\mathrm{Im}\left[k_{TP}^4\,\underline{h}'_x\,\underline{h}_x^* - \left(\underline{h}''_x\right)^* \underline{h}'''_x\right] \\ &=\; \beta\left(\mathrm{Re}\,k^5 - k_{TP}^4\,\mathrm{Re}\,k\right)|\underline{h}_x|^2,\end{aligned} \qquad 4.3(21)$$

where

$$\beta \;=\; \omega\mu_o/2\,k_t^2\,k_{TP}^4\,. \qquad 4.3(22)$$

For lossless media, $\langle S_{tot} \rangle$ will be positive, zero, or negative if the wavenumber k is real and greater than, equal to, or less than k_{TP}, respectively. Note also that beyond the turning point there is no power flow because the wavenumbers are imaginary in that region.

3. Construction of WKB quasi-normal mode solutions.

To use the method outlined in section 2.2, we first convert the fourth-order differential equation 4.3(6) into four coupled first-order equations by setting

$$U_1 = \underline{h}_x, \; U_2 = \underline{h}_x', \; U_3 = \underline{h}_x'', \; U_4 = \underline{h}_x'''. \qquad 4.3(23)$$

(Recall that the prime denotes differentiation with respect to the distance z.) As a result, we have the matrix equation

$$U' = DU , \qquad 4.3(24)$$

where U is a 4-element column matrix, and

$$D = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -k_{TP}^4 & 0 & -k_{EX}^2 & 0 \end{bmatrix} \qquad 4.3(25)$$

Here $k_{EX}^2 k_{MS}^2$ in 4.3(6) is approximated by the constant k_{TP}^4 as in 4.3(7), but k_{EX}^2 varies with distance z according to 4.3(1).

Next write U as the linear combination

$$U = LQ , \qquad 4.3(26)$$

where Q_i are to be the quasi-normal mode solutions. Appendix 1 shows that when the solutions for the wavenumbers in 4.3(9) are all distinct (that is, $k^2 \neq k_{TP}^2$), it is possible to construct the matrix L such that not only is $L^{-1}DL$ diagonal, but the first row of L is composed of constants (see equation A1(11)). The diagonal elements in $L^{-1}DL$ are the eigenvalues $\lambda_i = -jk_i$ of D. (Note that these eigenvalues are functions of distance.) Explicitly then we can write formal solutions for 4.3(6) from 4.3(26) in the form

$$\underline{h}_x = U_1 = \sum_{i=1}^{4} L_{1i} Q_i , \qquad 4.3(27)$$

where the L_{1i} are independent of distance. Furthermore, 4.3(26) with the form of L obtained in Appendix 1 shows that

$$\underline{h}_x^{(m)} = U_{m+1} = \sum_{i=1}^{4} L_{m+1,i} Q_i = \sum_{i=1}^{4} (-jk_i)^m L_{1i} Q_i, \quad m = 0, 1, 2, 3, \qquad 4.3(28)$$

with $\underline{h}_x^m$ being the m^{th} derivative of $\underline{h}_x$ with respect to z.

As in section 2.2, we now write

$$Q = BEA, \qquad 4.3(29)$$

where B is a diagonal matrix such that the $B_{ii} = b_i$ are the "WKB amplitudes" for the quasi-normal modes Q_i, E is a diagonal matrix whose elements E_{ii} are the exponential phase integral factors, and the A_i are the quasi-normal mode amplitudes. Hence

$$Q_i = b_i \exp\left[-j \int^{z} k_i(\xi)\, d\xi\right] A_i \qquad 4.3(30)$$

as in 2.2(14) and 2.2(15). Since the dispersion relation 4.3(8) is biquadratic, choose k_1 and k_3 as positive when real solutions to 4.3(8) exist, and then set

$$k_2 = -k_1 \text{ and } k_4 = -k_3. \qquad 4.3(31)$$

Furthermore, from 2.2(18) and 2.2(19):

$$\begin{aligned} b_2(z) &= b_1(z) = [k_1^3(k_1^2 - k_3^2)]^{-\frac{1}{2}} \\ b_4(z) &= b_3(z) = [k_3^3(k_3^2 - k_1^2)]^{-\frac{1}{2}} \end{aligned} \qquad 4.3(32)$$

Note that these WKB amplitudes b_i get very large as k_3 and k_1 approach the common value k_{TP}. In the following, k_1 will represent a "magnetostatic" spin wave, and k_3 an exchange-dominated spin wave. For real k, this convention implies that $k_1 < k_3$. Then k_2 and k_4 from 4.3(31) represent the corresponding waves traveling in the opposite direction.

To calculate the power flow in each quasi-normal mode, use 4.3(16) and 4.3(19), where the various derivatives of $\underline{h}_x$ are calculated using 4.3(28). Consider the power in mode Q_i, and choose the multiplicative constant L_{1i} in 4.3(27) to be unity. Then the average electromagnetic power flow in a lossless medium in regions where k is purely real is:

$$\langle S_{EM} \rangle = -\beta k_{TP}^4 k_i |b_i|^2 |A_i|^2 , \qquad 4.3(33)$$

where β is defined in 4.3(22). Similarly from 4.3(19):

$$\langle S_m \rangle = \beta k_i^5 |b_i|^2 |A_i|^2 , \qquad 4.3(34)$$

where in deriving 4.3(19) it was assumed that $k_{MS}^2 >> k_t^2$, or equivalently $k_{MS}^2 k_{EX}^2 = k_{TP}^4$. As a consequence, from the dispersion relation 4.3(8) we have

$$k_1 k_3 = k_{TP}^2 . \qquad 4.3(35)$$

Now evaluating the total power flow $\langle S_{tot} \rangle = \langle S_{EM} \rangle + \langle S_m \rangle$, substituting for the b_i from 4.3(32) and using 4.3(35) we find:

$$\langle S_{tot} \rangle = \beta \, |A_i|^2 \, (-k_{TP}^4 \, k_i + k_i^5) |b_i|^2 = \pm \beta |A_i|^2 . \qquad 4.3(36)$$

Here the minus sign applies for i = 1, 4 and the plus sign for i = 2, 3, corresponding to the fact noted earlier that the "magnetostatic" wave **is** a backward wave, carrying energy in the positive z direction when k is negative, and vice versa. We have chosen k_1 to be positive, so k_2 must then represent a "magnetostatic" wave traveling from $z = -\infty$ to $z = +\infty$. Note particularly in 4.3(36) that the total power in a quasi-normal mode is approximately constant as long as its amplitude $|A_i|$ is approximately constant even though the k_i are functions of distance. Hence we have shown that the Q_i have the desired characteristic of WKB quasi-normal mode solutions for this system.

Recall now from section 2.2, that the amplitudes A_i of the quasi-normal modes Q_i are coupled according to

$$A_i' = - \sum_{\ell=1}^{4} \exp\left[j \int^{z} (k_i - k_\ell) \, d\xi \right] C_{i\ell} \, A_\ell \qquad 4.3(37)$$

where C is a 4 x 4 coupling matrix whose diagonal elements C_{ii} are all zero. The off diagonal elements $C_{i\ell}$ become very small in small field gradients and far from the turning point where $k_1 = k_3 = k_{TP}$. In those circumstances, the A_i are practically constant. In view of 4.3(36) and the requirement of power conservation, the formal solution 4.3(27) with constant L_{1i} will then be approximately valid.

Explicitly evaluating the expression $C = B^{-1}B' + B^{-1}(L^{-1}L')B'$ from 2.2(11), using the form of the matrix L given in Appendix 1 and the expressions for $B_{ii} = b_i$ in 4.3(32) for the diagonal matrix B, we obtain

$$C = \begin{bmatrix} 0 & C_{12} & C_{13} & C_{14} \\ C_{12} & 0 & C_{14} & C_{13} \\ C_{13} & -C_{14} & 0 & -C_{12} \\ -C_{14} & C_{13} & -C_{12} & 0 \end{bmatrix}, \qquad 4.3(38)$$

where

$$\begin{aligned} C_{12} &= -k_1'/2k_1 = +k_3'/2k_3 \\ C_{13} &= \frac{k_1' k_3}{k_{TP}(k_3 - k_1)} = -\frac{k_3' k_1}{k_{TP}(k_3 - k_1)} \\ C_{14} &= -\frac{k_1' k_3}{k_{TP}(k_3 + k_1)} = \frac{k_3' k_1}{k_{TP}(k_3 + k_1)} \end{aligned} \qquad 4.3(39)$$

The second equality in each of the expressions in 4.3(39) follows from 4.3(35), which requires that

$$k_1' k_3 + k_1 k_3' = 0. \qquad 4.3(40)$$

Note that the element C_{12} represents reflections from a spin wave of either type traveling in one direction into the oppositely-traveling spin wave of the same type. The form of C_{12} is the same as that which would be derived for these reflections using a simple second-order analysis. Observe that even for small gradients in H(z), C_{12} will become very large at the turning point since k_1 and k_3 change rapidly in that region even for small changes in H(z) (see the dispersion relation in Fig. 4). Since A_1 represents the amplitude of a "magnetostatic" spin wave carrying energy in the negative z direction while A_3 represents that of an exchange-dominated spin wave carrying energy from $z = -\infty$ to the turning point at some finite $z = z_{TP}$, we see that C_{13} represents the coupling between these two waves (or alternatively, between the corresponding pair carrying energy in the opposite directions). This coupling element becomes infinite at the turning point $k_1 = k_3 = k_{TP}$. Finally, C_{14} represents the coupling between one spin wave of each type carrying energy in the same direction. C_{14} becomes large at $H - H_{TP} = M(2\lambda k_t^2)^{\frac{1}{2}}$, where $k_1 = -k_3 = jk_{TP}$ (see the discussion following 4.3(10)). Since the wavenumbers become purely imaginary at this point, however, C_{14} represents primarily the coupling between decaying (evanescent) waves and thus is not usually important.

Since C_{12} and C_{13} become very large near the turning point, the quasi-normal mode amplitudes A_i will change rapidly according to 4.3(37), and the formal solutions for $\underline{h}_x$ is 4.3(6) given by 4.3(27) and 4.3(30) will not be very useful. Ordinarily, one would try a successive approximation (perturbation) approach based on 4.3(37) assuming that the amplitude A_2 or A_3 representing an incident wave is unity, and then integrating to find the waves of the "scattered waves" A_1 and A_4 (see subsection 2.3(a)). However, the amplitude of such an incident wave must actually go rapidly almost to zero near the turning point since it is evanescent beyond that point. Furthermore, energy incident upon

$z = z_{TP}$ from $z = -\infty$ couples not only to the reflected waves Q_1 and Q_4 but also to the evanescent waves beyond the turning point Q_2 and Q_3. Thus it also does not seem feasible to approximate the problem by any second order system, such as is done for certain coupling points in the ionosphere. (See, for example, BUDDEN 1961 sections 20.6 and 20.7, HOUGARDY and SAXON 1963, and part 2.4(b)2 of the present work.) It is intriguing, nevertheless, to apply in the next subsection the phase integral formula 2.4(13) developed by RYDBECK 1967 for such a second-order approximating system, extracted from 4.3(37), 4.3(38), and 4.3(39).

The remaining two subsections indicate how to find the distribution in energy reflected into the two types of spin waves at the turning point. The WKB method uses asymptotic expansions of solutions of an equation modeling 4.3(6) near the turning point to find the approximate multiplicative factors (reflection coefficients) L_{1i} for the formal solution of 4.3(6), in terms of quasi-normal modes valid far from the turning point. The method using generalized hypergeometric functions assumes that the medium becomes homogeneous far from the turning point, so that the solutions can be identified in terms of ordinary normal modes there. These solutions are then compared with the results of the phase-integral method and the WKB method.

4.3(c) Application of the phase-integral method.

In this section we apply formula 2.4(13) developed by RYDBECK 1967 and quoted in part 2.4(b)2. Consider the amplitude A_3 of an exchange-dominated spin wave incident upon the turning point, and that of the reflected "magnetostatic" spin wave, A_1. If we neglect the amplitudes A_2 and A_4 of the other quasi-normal modes, then from

4.3(37) and 4.3(38) we obtain:

$$A_1' = -C_{13} \exp\left[j \int^{z} (k_1 - k_3)\, d\xi \right] A_3 \qquad 4.3(41)$$

$$A_3' = -C_{13} \exp\left[j \int^{z} (k_3 - k_1)\, d\xi \right] A_1$$

Defining

$$\Delta k \equiv (k_3 - k_1)/2 , \qquad 4.3(42)$$

and using 4.3(40), we find for the factor $-C_{13}$ defined by 4.3(39):

$$-C_{13} \equiv \frac{k_3' \; k_1}{k_{TP}(k_3 - k_1)} = \frac{(\Delta k)'}{\Delta k} \left(\frac{k_{TP}}{k_3 + k_1} \right) \qquad 4.3(43)$$

From 4.3(9) we can also write

$$k_{3,1} = \left[\left(\frac{k_{EX}}{2} \right)^2 + k_{TP}^2 \right]^{\frac{1}{2}} \pm \left[\left(\frac{k_{EX}}{2} \right) - k_{TP}^2 \right]^{\frac{1}{2}}, \qquad 4.3(44)$$

so that

$$k_3 + k_1 = 2\left[\left(\frac{k_{EX}}{2} \right)^2 + k_{TP}^2 \right]^{\frac{1}{2}} . \qquad 4.3(45)$$

and

$$2\Delta k = k_3 - k_1 = 2\left[\left(\frac{k_{EX}}{2} \right)^2 - k_{TP}^2 \right]^{\frac{1}{2}} . \qquad 4.3(46)$$

Now the branch points are defined as the places where $k_3 = k_1 = k_{TP}$. From 4.3(36) these occur where $k_{EX}^2 = 4k_{TP}^2$. Using the definition 4.3(1) for k_{EX}^2, and labeling the branch points z_1 and z_2, we thus have:

$$H(z_{1,2}) = H_{TP} \mp 4M\lambda k_{TP}^2 , \qquad 4.3(47)$$

with $H_{TP} = \omega/|\gamma\mu_o|$ as in 4.3(3) (compare 4.3(10) and following). Between the branch points z_1 and z_2, moreover, we can approximate 4.3(45) by $k_3 + k_1 \approx 2\, k_{TP}$, with an error of no more than a factor of $\sqrt{2}$. With the convention $z_2 > z_1$ 4.3(43) then becomes

$$-C_{13} \approx \frac{(\Delta k)'}{2\Delta k}\,; \qquad z_1 \leq z \leq z_2\,. \tag{4.3(48)}$$

Consequently, equations 4.3(41) now take on the form:

$$\begin{aligned} A_1' &= \frac{(\Delta k)'}{2\Delta k} \exp\left[-j\int^z 2\,(\Delta k)\, d\xi\right] A_3 \\ A_3' &= \frac{(\Delta k)'}{2\Delta k} \exp\left[+j\int^z 2\,(\Delta k)\, d\xi\right] A_1 \end{aligned} \tag{4.3(49)}$$

With A_1 replaced by A_2, A_3 replaced by A_1, and Δk replaced by k_1, equations 4.3(49) become identical to those resulting from the combination of 2.2(34) and 2.2(26), in the example of reflections of an x-polarized electromagnetic wave. The equations in that example resulted from Maxwell's equations 1.2(3) with $\epsilon_{xx} = \epsilon(z)$ and $\epsilon_{xy} = 0$. Furthermore, the two coupled first-order equations were combined into one second-order equation, 3.2(1), which has the form $E_x'' + k_1^2\, \underline{E}_x = 0$, with $k_1^2 = \omega^2 \mu_o \epsilon$. By analogy, then, we infer that

$$F'' + (\Delta k)^2 F = 0, \tag{4.3(50)}$$

where

$$F \equiv (\Delta k)^{-\frac{1}{2}} \left[A_1 \exp\left(j\int k_1\, d\xi\right) + A_3 \exp\left(-j\int k_3\, d\xi\right)\right] \tag{4.3(51)}$$

(Compare equation 2.2(33) for $\underline{E}_x(z)$ in terms of quasi-normal mode amplitudes.) Equation 4.3(50) can be proved explicitly by differentiating

4.3(51), followed by substitution of 4.3(49). Note that it has the same form as 2.4(12).

If $(\Delta k)^2$ varies linearly with z, then 4.3(50) has the form of the Airy equation 2.4(7). Hence we postulate that we can apply a phase-integral formula similar to 2.4(10) or 2.4(13):

$$|R| = \exp\left[-2 \operatorname{Im} \int_{z_1}^{z_2} (\Delta k)\, dz\right]. \qquad 4.3(52)$$

RYDBECK 1967 integrated only between the two branch points z_1 and z_2 in 2.4(13) because Δk was real elsewhere. In our case, Δk from 4.3(46) is real for $z < z_1$ $(H < H(z_1))$, but imaginary for all $z > z_1$. However, we may still take the upper limit of integration in 4.3(52) to be z_2, if we assume most of the coupling occurs between z_1 and z_2. Beyond z_2 the energy attenuates rapidly. We also postulate that $|R|$ represents the magnitude of reflection from the incident exchange wave (A_3) into the oppositely-traveling wave of the same type.

If $(\Delta k)^2$ varies linearly with z, then 4.3(46) combined with the definition 4.3(1) for k_{EX}^2 shows that H(z) varies linearly with z. Let H'_{TP} be the constant value of H'(z) in the region of the turning point. Changing the variable of integration in 4.3(52) from z to H(z), we obtain

$$|R| = \exp\left[-\frac{2}{H'_{TP}} \operatorname{Im} \int_{H(z_1)}^{H(z_2)} (\Delta k)\, dH\right]. \qquad 4.3(53)$$

Now let $x = (H - H_{TP} + 4M\lambda k_{TP}^2)/4M\lambda$. Using 4.3(46) and 4.3(47) we then find that 4.3(53) becomes

$$|R| = \exp\left(-\frac{8M\lambda}{H'_{TP}} \int_0^{2k_{TP}^2} x^{\frac{1}{2}}\, dx\right). \qquad 4.3(54)$$

Consequently, we have finally

$$|R|^2 \cong \exp(-30\, M\lambda k^3_{TP}/H'_{TP}) . \qquad 4.3(55)$$

If we interpret 30 $M\lambda k^3_{TP}$ as a critical gradient H'_{crit} in the manner of section 4.2, then for yttrium iron garnet with $M\lambda \approx 5 \times 10^{-9}$ Oe cm^2, and for $k_t = 10$ cm^{-1}, $H'_{crit} \approx 1200$ oersted/cm. Recall that $|R|^2$ is supposed to denote the fraction of power in the reflected "exchange" wave when a wave of the same type is incident. The corresponding result for power reflected into a "magnetostatic" wave when a "magnetostatic" wave is incident is the same as 4.3(55), since in 4.3(38) $C_{24} \equiv C_{13}$.

The validity of the entire procedure leading to 4.3(55) is, of course, open to serious question. First of all, we have only postulated without proof that $|R|$ represents a reflection coefficient of the type indicated. The reflected exchange-dominated spin wave is actually the quasi-normal mode with amplitude A_4, which we neglected in going from 4.3(37) to 4.3(41). There is also no formula to check that $1 - |R|^2$ represents the power converted to the other wave type, as it should for power conservation. Furthermore, the function F which satisfies 4.3(50) has the form of a linear combination of modes, rather than a single mode. Next, the formula 4.3(52) was the result of a plausibility argument only. Finally, the use of z_2 as the upper limit in 4.3(52) is questionable. If a higher limit is used, the formula changes drastically, as 4.3(54) shows. Nevertheless, comparison of the result 4.3(55) with a rigorous solution using generalized hypergeometric functions shows that 4.3(55) may be approximately valid for magnetic field variations as in Fig. 3 (see the discussion near the end of subsection 4.3(e)). In the meantime, however, we apply the WKB method in the next subsection.

4.3(d) Application of the WKB method.

In this subsection, we model the equation 4.3(6) for $\underline{h}_x(z)$ by a simpler one which is approximately valid near the turning point. We then find solutions of this new equation in terms of contour integrals, as indicated in section 2.4. When the magnetic field gradient is small, and far enough from the turning point, asymptotic expansions of these solutions may be identified with various linear combinations of the WKB quasi-normal modes constructed above from 4.3(6).

First of all, near the turning point 4.3(7) and 4.3(35) are valid, so that $k_{EX} k_{MS} = k_1 k_3 = k_{TP}^2$, where k_1 and k_3 are solutions of the dispersion relation 4.3(8). Secondly, we assume that the variation of k_{EX}^2 and H(z) can be approximated by a straight line near the turning point:

$$H_{TP} - H(z) = -H'_{TP} z \qquad 4.3(56)$$

where we assume that $H_{TP} = H(z)$ at $z = 0$. For further work, it is convenient to transform 4.3(6) into an equation with dimensionless quantities. Setting

$$s = -cz, \qquad 4.3(57)$$

where c has the dimensions of a wavenumber, we find that 4.3(6) has the simple form

$$F^{(4)}(s) + 2b^2 s\, F^{(2)}(s) + b^4 F(s) = 0; \;\; F(s) \equiv \underline{h}_x[z(s)] \qquad 4.3(58)$$

as long as

$$b = 2M\lambda\, k_{TP}^3 / H'_{TP}$$

and $$c = k_{TP}/b\ . \qquad 4.3(59)$$

In this notation, H'_{TP} denotes the gradient of H(z) with respect to z evaluated at the turning point z = 0, while $F^{(m)}(s)$ denotes the m^{th} derivative of F(s) with respect to s. Note that b and s are dimensionless, with s proportional to the change in H(z) from z = 0. The minus sign in 4.3(57) allows the propagating side of the turning point to be at positive s when the waves are propagating for negative z, so that energy incident upon the turning point comes from $z = -\infty$, in accordance with the other treatments in this work.

Equation 4.3(58) has the same form as the one studied by RABENSTEIN 1958, 2.4(16). Equivalence is established if we set $p_o^2 = 2b^2$, $p_1 = 0$, and $p_2 = b^2/2$. In the solutions developed by Rabenstein, p_o was assumed to be a large parameter, and thus his solutions can be applied here when b is large or, from 4.3(59), when the gradient H'_{TP} is small. In the equation 2.4(15) used by WASOW 1950, however, there is no independent parameter p_2. To transform 4.3(6) to the form of 2.4(15) then requires that p_o be proportional to H'_{TP}, which means that Wasow's solutions cannot be used for cases with reasonably small gradients H'_{TP}. For large gradients his solutions may also not be applicable, because the constant c in 4.3(57) again will be proportional to H'_{TP} and thus for finite distance z, s may be too small for the asymptotic expansions to be valid.

Rather than try to apply the solutions of RABENSTEIN 1958 which were developed for hydrodynamical instability problems, it is more convenient to find solutions of 4.3(58) in a manner which allows easier identification with traveling waves. Although very similar to Rabenstein's, the treatment outlined below seems more systematic and elegant for our application, and also produces explicit expressions for the error terms in the asymptotic expansions.

The difference lies in the way the steepest descent integrals are done. To begin, however, we assume the same Laplace integral form for the solutions of 4.3(58):

$$F(s) = \int_C e^{st} f(t)\, dt, \qquad 4.3(60)$$

where t is a complex variable. Substitution of 4.3(60) into 4.3(58) produces a first order differential equation for f(t), whose solutions are readily found. Consequently, F(s) has the form

$$F(s) = \int_C t^{-2} \exp\,[g(s, t)]\, dt\,, \qquad 4.3(61)$$

where

$$g(s, t) = \frac{t^3}{6\,b^2} - \frac{b^2}{2t} + st\,. \qquad 4.3(62)$$

The contour C in 4.3(61) can be any for which exp [g(s, t)] goes to zero at the endpoints. Some such contours are shown in Fig. 10.

To find the saddle points t_o in the complex t plane for g(s, t) at constant s, set $g'(t) = 0$, giving

$$t_o^4 + 2b^2\, s t_o^2 + b^4 = 0\,. \qquad 4.3(63)$$

Notice the similarity between this equation and the "local" dispersion relation for k(s) which would be obtained from 4. 3(58) assuming constant s (equivalent to constant k_{EX}^2 or H). In fact, we can identify

$$t_o(s) = -jk(s), \qquad 4.3(64)$$

where k(s) is one of the four "local" wavenumbers. The solutions of 4. 3(63) are (compare 4. 3(44)):

$$t_o = \pm \frac{jb}{\sqrt{2}} \left[(s+1)^{\frac{1}{2}} \mp (s-1)^{\frac{1}{2}} \right]. \qquad 4.3(65)$$

For large s (large $|H(z) - H(o)|$), we have, approximately:

$$t_{o_{1,2}} \cong \pm jb(2s)^{-\frac{1}{2}}, \qquad t_{o_{3,4}} \cong \pm jb(2s)^{\frac{1}{2}}, \qquad 4.3(66)$$

with the labels 1 and 2 denoting "magnetostatic" wave solutions, and 3 and 4 denoting solutions for the exchange spin waves. Clearly $|t_{o_1}| \ll b$ while $|t_{o_3}| \gg b$, for large s. Note from 4. 3(65) for arbitrary s, that

$$-t_{o_1} t_{o_3} = +b^2 = k_1(s)\, k_3(s) \qquad 4.3(67)$$

where we have used 4. 3(64) also. Clearly the second equation in 4. 3(67) is equivalent to 4. 3(35), since from 4. 3(57) and 4. 3(59), $b = k_{TP}/c$ and

$$k(s) = -k(z)/c. \qquad 4.3(68)$$

The next step is to evaluate exp $[g(t_o)]$, which in a steepest descent integration is pulled outside of the integral in 4.3(61). Combining 4.3(62) with s in terms of t_o from 4.3(63), we find

$$g(t_o) = -\frac{1}{3b^2}\left(t_o^3 + 3b^4 t_o^{-1}\right). \qquad 4.3(69)$$

Now evaluate t_o^3 and t_o^{-1} in terms of s from 4.3(65), and compare with the integral of t_o from the same equation to find:

$$g(t_o) = \int^s t_o(s)\, ds = -j\int^s k(s)\, ds\,, \qquad 4.3(70)$$

where the second relation follows from 4.3(64). Observe that exp $[g(t_o)]$ thus represents the exponential phase integral which occurs in the WKB quasi-normal modes. Thus we have finished the first step in identifying the solutions of 4.3(58) in terms of these modes. Equation 4.3(70) follows easily from 4.3(66) in the limit of large s, since then only one of the two terms in 4.3(69) is important, but with some manipulation it can also be derived from 4.3(65) for arbitrary s, as mentioned above.

To apply the method of steepest descents, it is necessary to write $g(t_o) - g(t, s)$ in terms of some parameter σ^2, which will be real and positive on the steepest descent paths and negative on the steepest ascent paths leading away from the saddle point t_o at $\frac{\partial g}{\partial t}(t, s) = 0$. Again substituting for s in g(t, s) in terms of t_o from 4.3(63), we find that

$$\frac{b}{2}\sigma^2 \equiv g(t_o) - g(t) = \frac{b^2}{2}(t - t_o)^2\left(\frac{1}{t_o^2 t} - \frac{t + 2t_o}{3b^4}\right). \qquad 4.3(71)$$

In this case it is not necessary to expand g(t) in a power series about

$t = t_o$, as is often necessary in steepest-descent integrations. Note that σ^2 as defined in 4.3(71) is insensitive to b for t near t_o, since t_o is proportional to b from 4.3(65). Thus we have factored out the dependence on H'_{TP} through the parameter b/2, which becomes large for small H'_{TP}. Now we can write the contribution from the saddle point t_{o_i} to the integral in 4.3(61) for F(s) as:

$$I_i(s) \cong \exp\left[g(t_{o_i})\right] \int_{-\infty}^{+\infty} \exp\left(-\frac{b}{2}\sigma^2\right) t^{-2}(\sigma) \frac{dt}{d\sigma} d\sigma. \qquad 4.3(72)$$

(A minus sign must be added if the direction of positive σ makes an angle of magnitude more than $\pi/2$ with the positive real t direction; see section 2.4.)

In general, a solution for F(s) along some contour C_ℓ will involve more than one contribution $I_m(s)$. We suppose nevertheless that the limits on σ can be taken as infinite when the contour C_ℓ has passed by the region of significant contribution from t_o. For large s, inspection of 4.3(71) in conjunction with 4.3(66) shows that when $|(t - t_o)/t_o| = 1$, σ^2 is approximately b/t_o for $t_{o_{1,2}}$ with $|t_o| \ll b$, and approximately t_o^3/b^3 for $t_{o_{3,4}}$ with $|t_o| >> b$. Thus σ^2 is likely to be large, and $\exp\left[-\frac{b}{2}\sigma^2\right]$ quite small, before the contour C_ℓ begins to encounter contributions from other saddle points. These saddle points for large s are well separated. Near s = 0 (near the turning point), however, t_{o_1} and t_{o_3} coalesce and the treatment described above clearly becomes invalid.

To evaluate the integral in 4.3(72), it is desirable to express $t^{-2}(\sigma)\frac{dt}{d\sigma} = -\frac{d}{d\sigma}\left(\frac{1}{t}\right)$ as a power series in σ. This is evidently possible if $\left(\frac{1}{t} - \frac{1}{t_o}\right)$ can be found as a power series in σ:

$$\frac{1}{t} - \frac{1}{t_o} = c_1\sigma + c_2\sigma^2 + c_3\sigma^3 + \cdots , \qquad 4.3(73)$$

in which case

$$t^{-2}(\sigma)\frac{dt}{d\sigma} = -c_1 - 2c_2\sigma - 3c_3\sigma^2 - \cdots \qquad 4.3(74)$$

Since

$$\int_{-\infty}^{+\infty} \exp\left(-\frac{b}{2}\sigma^2\right)\sigma^{2n}\, d\sigma = \left(\frac{2\pi}{b}\right)^{\frac{1}{2}} \frac{1\cdot 3\cdot 5\cdots(2n-1)}{b^n}, \qquad 4.3(75)$$

we find finally from 4.3(72) and 4.3(74) that

$$I_i(s) \cong -\exp\left[g(t_{o_i})\right]\left(\frac{2\pi}{b}\right)^{\frac{1}{2}} \sum_{n=0}^{\infty} (2n+1)c_{2n+1} \frac{1\cdot 3\cdot 5\cdots(2n-1)}{b^n} \qquad 4.3(76)$$

Note that only the c_m with odd subscripts m contribute.

The coefficients c_n in 4.3(73) can be found if σ in 4.3(71) is written as a function h(r) of $t^{-1} \equiv r$ and $t_o^{-1} \equiv r_o$:

$$h(r) \equiv \sigma = b^{\frac{1}{2}} \frac{(r - r_o)}{rr_o} \left[r_o^2 r - \frac{(r^{-1} + 2r_o^{-1})}{3b^4}\right]^{-\frac{1}{2}}. \qquad 4.3(77)$$

Then we apply Lagrange's formula for series inversion:

$$n!\, c_n = \left\{\frac{d^{n-1}}{dr^{n-1}}\left[\frac{r - r_o}{h(r)}\right]^n\right\}_{r = r_o}, \qquad 4.3(78)$$

from which it follows easily that for $r_o = t_{o_i}^{-1}$,

$$c_1 = b^{3/2}\, t_{o_i}^{-\frac{1}{2}} (b^4 - t_{o_i}^4)^{-\frac{1}{2}}. \qquad 4.3(79)$$

Hence the leading term in 4. 3(76) becomes

$$I_i(s) \cong \mp \exp\left[-j \int^{s} k_i(s)\, ds\right] \left[\frac{2\pi}{k_i^3(k_3^2 - k_1^2)}\right]^{-\frac{1}{2}} b\,, \qquad 4.3(80)$$

where 4. 3(64), 4. 3(67), and 4. 3(70) have been used, and the wavenumbers k are all taken to be functions of s. Since the dimensionless k(s) are related simply to the corresponding k(z) through 4. 3(68), however, the form of the multiplicative factor in 4. 3(80) remains unchanged, except for constants, when written in terms of wavenumbers k(z). Furthermore, the integral $\int^{s} k_i(s)ds$ is similarly just $\int^{z} k_i(z)\, dz$. Thus the $I_i(s)$ are identified exactly with the WKB quasi-normal modes $Q_i(z)$ defined through 4. 3(30) and 4. 3(32).

To determine which saddle points t_{o_i} contribute to the various solutions for F(s) appearing in 4. 3(61), through deformations of the contours C_ℓ shown in Fig. 10, we must at least know where the t_{o_i} lie and in what direction the steepest descent contours pass through those points. The latter information can be found in the limit of large s by noting then from 4. 3(71) that near $t = t_o$,

$$t - t_o \cong \pm b^{-\frac{1}{2}} t_o^{3/2} \sigma\,; \quad |t_o| << b \qquad (\text{"magnetostatic" waves})$$

and 4. 3(81)

$$t - t_o \cong \pm j b^{3/2} t_o^{-\frac{1}{2}} \sigma\,; \quad |t_o| >> b \qquad (\text{"exchange" waves})$$

Then for real σ, the argument ψ of $(t - t_o)$ is given as

$$\psi = \frac{3}{2}\varphi \qquad (\text{"magnetostatic" waves})$$

or 4. 3(82)

$$\psi = \frac{\pi}{2} - \frac{\varphi}{2} \qquad (\text{"exchange" waves})$$

where

$$\psi = \arg (t - t_o), \ \sigma \text{ real}; \ \varphi = \arg t_o . \qquad 4.3(83)$$

Using these equations and 4.3(66), the locations of the saddle points and the directions of the steepest descent contours through those points are shown schematically in Fig. 10 for large positive and negative s.

Denoting the solution for contour C_ℓ as $F_\ell(s)$, we find from inspection of Fig. 10 (compare similar figures in BUDDEN 1962, Chapter 15) that

$$\begin{aligned}
F_1(+\infty) &= I_4 \\
F_2(+\infty) &= -I_4 + I_2 - I_1 + I_3 \\
F_3(+\infty) &= -I_3 \\
F_4(+\infty) &= -I_1 + I_2 \\
F_5(+\infty) &= -I_1 + I_3 \\
F_6(+\infty) &= -I_1 \\
F_7(+\infty) &= -I_2 + I_4
\end{aligned} \qquad 4.3(84)$$

where the minus signs arise for the reason mentioned following 4.3(72). Only four of these solutions are independent since relations can be found between then through Cauchy's formula, as was done by RABENSTEIN 1958. For $s \to -\infty$, where the wavenumbers $k_i(s)$ are purely imaginary, we have:

$$
\begin{aligned}
F_1(-\infty) &= I_4 \\
F_2(-\infty) &= I_3 \\
F_3(-\infty) &= -I_1 - I_3 - I_4 \\
F_4(-\infty) &= -I_1 \\
F_5(-\infty) &= I_2 + I_3 \\
F_6(-\infty) &= I_2 - I_1 - I_4 \\
F_7(-\infty) &= I_2
\end{aligned}
\qquad 4.3(85)
$$

The only allowed waves in the cutoff region for $s \rightarrow -\infty$ are those which are evanescent; all others are exponentially growing. From 4.3(66) and 4.3(70), we see that the evanescent waves are denoted by the subscripts 2 and 3. These same subscripts 2 and 3 represent waves incident from $s = +\infty$ to the turning point, incident "magnetostatic" and "exchange" spin waves, respectively. (Recall that the "magnetostatic" wave is a backward wave carrying energy opposite to the direction of phase increase. See the discussion following 4.3(36).) Thus for $s \rightarrow -\infty$, the only possible physical solutions for F(s) in 4.3(85) are F_2, F_5, and F_7. Looking back to 4.3(84), we see that of these three solutions, only F_5 and F_7 represent situations where only one type of wave is incident from $s = +\infty$. Furthermore, since the I_i are identifiable in terms of the WKB quasi-normal modes, equation 4.3(36) shows that the total power flow in one of these waves I_i is just proportional to the square of its amplitude. Thus 4.3(80) in convection with solution F_5 implies that energy incident in the form of an "exchange" spin wave (I_3) is entirely converted into energy in a reflected "magnetostatic" wave (I_1). Solution F_7 simply implies the reverse conversion. No energy reflected into the wave of the same type as the incident wave is evident.

It is now necessary to examine the higher-order correction terms in 4.3(76). The error ratio E of the first such term to the leading term 4.3(80) is found from 4.3(76), 4.3(79), and the expression for c_3 in 4.3(78) as

$$E = \frac{1}{6} b^2 t_o (b^4 - t_o^4)^{-3} (b^8 - \frac{10}{3} b^4 t_o^4 + \frac{101}{9} t_o^8) . \qquad 4.3(86)$$

Note that E becomes infinite near the turning point where $t_o^4 = b^4$. Away from the turning point,

$$E_{MS} \cong \frac{1}{6} b^{-1} (2s)^{-\frac{1}{2}}; \quad |t_o| << b \qquad (\text{"magnetostatic" waves})$$

$$E_{EX} \cong \frac{101}{54} b^{-1} (2s)^{-3/2}; \quad |t_o| \gg b \qquad (\text{"exchange" waves}) , \qquad 4.3(87)$$

where 4.3(66) has been used.

Substituting 4.3(57) and 4.3(59) into 4.3(87), we find:

$$E_{MS} \cong H'_{TP} \left[72\ Mk_t^2 |H - H(o)|\right]^{-\frac{1}{2}}$$

$$E_{EX} \cong H'_{TP} (M\lambda)^{\frac{1}{2}} |H - H(o)|^{-3/2} . \qquad 4.3(88)$$

When these error ratios become comparable to unity, we can conclude that there may be substantial energy reflected into the same type of wave as is incident upon the turning point. Since both ratios in 4.3(88) are proportional to H'_{TP}, we infer that these latter reflections are stronger for stronger field gradients, as might be expected. Note that the errors also become larger closer to the turning point where H approaches H(o).

One reason we cannot consider $|H - H(o)|$ to be arbitrarily large is that physically this does not occur. Secondly, for rather small $|H - H(o)|$, the exchange-dominated spin waves may couple to elastic waves at the magnetoelastic crossover point, where $k_p = k_{EX}$. Typically, $|H - H(o)|$ is only about 1 oersted or less at this crossover point, if H(o) is the value at the turning point. Then E_{EX} is equal to unity when H'_{TP} is about 10^4 oersted per centimeter. In typical rods of yttrium iron garnet used in practice, H'_{TP} may reach several thousand Oe cm^{-1}. Hence a substantial fraction of exchange power incident upon the turning point may be reflected as exchange power, rather than being converted to electromagnetic power. This would agree with the high_Q magneto-elastic resonances observed by KOHANE, SCHLÖMANN, and JOSEPH 1965.

Furthermore, for somewhat larger $|H - H(o)|$, the "magnetostatic" wave will couple to the electromagnetic fields of a fine wire antenna, or be reflected from the end of the delay line. Assuming $|H - H(o)| \approx \frac{M}{2}$ at that point, and assuming $k_t = 10$ cm^{-1}, we find E_{MS} is unity if H'_{TP} is about 10^5 Oe cm^{-1}. Hence electromagnetic power incident upon the turning point will probably be almost entirely converted to reflected exchange power. A noticeable fraction may well be reflected into electromagnetic power, however, in agreement with experimental results cited in 4. 3(a).

It is interesting to note that E_{MS} calculated here is essentially the same as would be obtained from solution of the second-order "magneto-static" wave equation 4. 1(23) without coupling to the "exchange" waves. Then the asymptotic expansions involved are those for Hankel functions (see RABENSTEIN 1958). See STIX 1962, Chapter 10, for a treatment by the WKB method of the corresponding extraordinary (medium-k) wave having a resonance (wavenumber approaching infinity) in a plasma. Since the high-k waves were not included and the error terms were not considered, it appeared in that analysis that all the incident energy was absorbed by the plasma, regardless of how large the gradients or how small the losses.

When losses are included, the real part of the fourth-order dispersion relation splits, in proportion to the amount of loss, as shown in Fig. 9. It then could be conjectured that fairly substantial gradients might be necessary at the turning point in order to have energy converted from one type of spin wave to the other. In the magnetoelastic coupling case, higher gradients were necessary for energy to jump across the split in the dispersion relation when the splitting parameter a was increased. In medium-k to high-k spin wave conversion, on the other hand, it is clear that large losses will cause much of the energy to be absorbed before it gets to the turning point. Furthermore, the smaller the field gradients, the longer the time it takes a wave to travel through this lossy region. It is not clear from a study of the relative size of the error ratios E whether reflections into the same type of wave as is incident will be increased by excessively small field gradients.

4.3(e) Solution by hypergeometric functions

As an alternative to the WKB method of the last subsection, we now apply the method described in section 3.3 to the solution of the fourth-order equation 4.3(6) for $\underline{h}_x(z)$. Instead of assuming a linear variation of magnetic field H(z) near the turning point, as in the WKB method, we choose H(z) to have a hyperbolic tangent variation such as depicted in Fig. 1. (Solutions of 4.3(6) with H(z) having a symmetrical hump are indicated in section 3.4.) Thus in this method we match the solutions of the fourth-order equation to the normal modes in the homogeneous regions away from the turning point, whereas in the WKB method we had to match to the WKB quasi-normal modes, since the inhomogeneity continued indefinitely even for large distances from the turning point. The advantage of the hypergeometric function solutions is the possibility of studying analytically the effects on reflections of high magnetic field gradients and nearness to the turning point. Recall that in the WKB method we had to be satisfied with conjectures based upon error ratios such as in 4.3(88).

Mathematically, both methods involve contour integral solutions. The WKB contour integrals resemble Laplace transforms and are given in 4.3(60) and 4.3(61), while those for hypergeometric functions resemble inverse Laplace transforms and are given by 3.1(24). In both cases, a given contour integral turns out to be expressible as a different linear combination of normal or quasi-normal modes for different sides of the turning point, allowing reflection coefficients to be determined. In the WKB method, this phenomenon is caused by the motion of saddle points in the complex plane of integration as the physical distance from the turning point is changed, so that a given contour will begin to pick up contributions from new saddle points. With the hypergeometric function method, however, the contour must be closed on different sides of the complex plane depending on which side of the turning point one is located. In that way, residues from different poles are picked up. Solutions using hypergeometric functions have the advantage that the basic mathematical results are given by 3.1(29) through 3.1(32) independent of the exact differential equation involved, whereas in the

WKB method the mathematics has to be redone for each new equation.

To solve 4.3(6), we again make the assumption that the magnetic field H(z) does not deviate too substantially from the turning point value H_{TP}, so that $k_{EX}^2 k_{MS}^2$ can be assumed to have the constant value k_{TP}^4. Otherwise, 4.3(7) shows that the coefficient of $\underline{h}_x(g)$ in 4.3(6) involves a nonlinear combination of H(z), whereas k_{EX}^2 involves H(z) linearly in view of 4.3(1). Recall that to apply the method of 3.3, only linear combinations of one varying parameter are allowed. With this restriction, we have

$$F''''(z) + k_{EX}^2(z)\,F''(z) + k_{TP}^4\,F(z) = 0;\; F(z) \equiv \underline{h}_x(z). \qquad 4.3(89)$$

Since this equation already has the same form as the dispersion relation 4.3(8), we can transform it directly to a generalized hypergeometric equation as in part 3.3(b). That is, in this case the original equation is already the "source equation." Note, however, that $\underline{m}_x(z)$ is given in terms of $\underline{h}_x$ and its second derivative through 4.3(4). Since $k_{MS}^2 >> k_t^2$ not too far from the turning point, we can neglect $\underline{h}_x$ compared with $\underline{h}_x''$ there. Then from 3.3(39) we see that the equation for $\underline{m}_x$ has the form as 3.3(38), since 4.3(89) has the same form as the "source equation" 3.3(37).

By convention, we have been assuming in this section that the subscripts 2 and 3 denote waves incident from z = -∞ on the turning point, the "magnetostatic" and "exchange" spin waves respectively. The solutions of the biquadratic dispersion relation 4.3(8) for k occur in pairs as in 4.3(31): $k_2 = -k_1$, $k_4 = -k_3$, where in this convention, with normal mode solutions of the form exp (-jkz), k_1 and k_3 are the solutions with non-negative real parts. The wavenumber k_2 for the "magnetostatic" wave carrying electromagnetic power in the positive z-direction must have a negative real part since equation 4.3(17) showed that "magnetostatic" waves are backward waves. In the cutoff region for $H(z) > H_{TP}$, the waves must decay with increasing z, which imposes the requirement that both k_2 and k_3 have non-positive imaginary parts. Note that all these requirements are compatible with 4.3(35) and 4.3(31).

In deriving the reflection factor R_{12} in 3.3(63) from the hypergeometric function solution for an equation such as 4.3(89), it was assumed that subscript 1 indicated an incident wave, and 2 a reflected wave of the same type. Since the reverse is true for "magnetostatic" waves in our case, we must interchange the subscripts 1 and 2. Then from 4.3(21) the ratio of time-averaged total power in the reflected and incident "magnetostatic" waves is

$$\frac{\langle S_{tot}\rangle_1}{\langle S_{tot}\rangle_2} = \frac{-|\underline{h}_x|_1^2}{|\underline{h}_x|_2^2} = -|R_{21}|^2, \qquad 4.3(90)$$

since $k_2 = -k_1$. Note that there are no extra factors multiplying $|R_{21}|^2$ in 4.3(90), since we are considering a reflected wave of the same physical type as the incident wave. See the discussion in subsection 1.2(d), and the examples of subsection 4.2(c). Now the ρ_i and σ_i in R_{21} satisfy $\rho_2 = -\rho_1$, $\rho_4 = -\rho_3$, $\sigma_2 = -\sigma_1$, and $\sigma_4 = -\sigma_3$ because of relations 3.3(18) between the ρ_i, σ_i and the k_i. Hence from 3.3(63) we can write

$$R_{21} = \frac{\Gamma(-2\rho_1)\,\Gamma(-\rho_3-\rho_1)\,\Gamma(-\sigma_1+\rho_1)\,\Gamma(\sigma_3+\rho_1)\,\Gamma(1-\sigma_1+\rho_1)\,\Gamma(1+\sigma_3+\rho_1)\,\Gamma(1-\rho_3-\rho_1)}{\Gamma(2\rho_1)\,\Gamma(-\rho_3+\rho_1)\,\Gamma(-\sigma_1-\rho_1)\,\Gamma(\sigma_3-\rho_1)\,\Gamma(1-\sigma_1-\rho_1)\,\Gamma(1+\sigma_3-\rho_1)\,\Gamma(1-\rho_3+\rho_1)} \qquad 4.3(91)$$

In lossless media, the σ_i are all real because the wavenumbers k_i^+

to the right of the turning point are all imaginary (cutoff). On the other hand, the ρ_i are all imaginary because the k_i^- are real (see 3.3(18)). Since $\Gamma(w^*) = \Gamma^*(w)$, we then observe that most of the terms in the numerator of 4.3(91) are the complex conjugates of those in the denominator, and hence drop out when the absolute magnitude $|R_{21}|$ is taken. In fact, we have:

$$|R_{21}|^2 = \left| \frac{\Gamma(-\rho_3 - \rho_1)\,\Gamma(1 - \rho_3 - \rho_1)}{\Gamma(-\rho_3 + \rho_1)\,\Gamma(1 - \rho_3 + \rho_1)} \right|^2 . \qquad 4.3(92)$$

Writing $|\Gamma(w)|^2 = \Gamma(w)\,\Gamma(w^*)$ as in part 3.2(a)5, and again noting the property $\Gamma(w)\,\Gamma(1-w) = \pi/\sin(\pi w)$ as in 3.3(60), we find

$$|R_{21}|^2 = \left| \frac{\sin(\rho_1 - \rho_3)}{\sin(\rho_1 + \rho_3)} \right|^3 = \frac{\sinh\left[\frac{\pi}{\alpha}(k_3^- - k_1^-)\right]}{\sinh\left[\frac{\pi}{\alpha}(k_3^- + k_1^-)\right]} , \qquad 4.3(93)$$

where the second half of this equation follows from 3.3(18).

Since the coefficient of $F''(z)$ in 4.3(89) is $k_{EX}^2(z)$, we find from 3.3(20) that

$$\alpha = 4\,(k_{EX}^2)'_{z=0} / [(k_{EX}^2)^+ - (k_{EX}^2)^-] \qquad 4.3(94)$$

When $M\lambda$ is constant, expression 4.3(1) for k_{EX}^2 shows that

$$\alpha = 4\,H'(o)/(H_+ - H_-) \qquad 4.3(95)$$

The reflection coefficient R_{21} should not depend on the value of H_+ as long as it is greater than H_{TP}, since then the waves are all cutoff for $z = +\infty$. Thus we let $H_+ - H_{TP} \gg H_{TP} - H_-$ to allow us to write α in the form of 3.3(24):

$$\alpha \cong H'_{TP}/(H_{TP} - H_{-}) , \qquad 4.3(96)$$

where H'_{TP} denotes the gradient $H'(z)$ of the magnetic field at the turning point $H = H_{TP}$. Note that we are now assuming that the turning point occurs not at $z = 0$ in the tanh $(\alpha z/2)$ curve for $H(z)$ as in 4.3(95), but considerably to the left of $z = 0$. See in this connection the various tanh $(\alpha z/2)$ curves shown in Fig. 1.

In the limit $\frac{\pi}{\alpha}(k_3^- - k_1^-) >> 1$, the hyperbolic sines in 4.3(93) may be replaced by exponentials, leaving

$$|R_{21}|^2 \cong \exp[- 4\pi k_1^-/\alpha] . \qquad 4.3(97)$$

Thus for decreasing α, the amount of energy reflected into a "magnetostatic" wave when a "magnetostatic" wave is incident upon the turning point decreases to zero. Note that this is in complete agreement with the result of the WKB method in part 4.3(c), since the error ratio E_{MS} in 4.3(88) has generally the same form as α in 4.3(96).

More precisely, from 4.3(9) we see that k_1^- approaches k_{MS} far from the turning point, and that for not too large $H_{TP} - H_{-}$, k_{MS}^- from 4.3(2) is approximately $[\frac{M}{2}k_t^2/(H_{TP} - H_{-})]^{\frac{1}{2}}$. Combining 4.3(96) and 4.3(97) with this latter fact shows that

$$|R_{21}|^2 \cong \exp\left\{- 2\sqrt{2}\,\pi[Mk_t^2 (H_{TP} - H_{-})]^{\frac{1}{2}}/H'_{TP}\right\} . \qquad 4.3(98)$$

Thus $|R_{21}|^2 \cong \exp[- E_{MS}^{-1}]$, showing that this reflection only becomes important when E_{MS} becomes comparable to unity, exactly as was conjectured from the WKB treatment. Note again that for small reflections of this type it is not sufficient just to have a small gradient H'_{TP} at the turning point, but that the incident wave must also start at a value H_- of magnetic field not too close to the value at the turning point. If $H_{TP} - H_-$ is too small, $k_3^- - k_1^-$ will also become small since $k_3 \approx k_1$ at $H = H_{TP}$, and then the assumption leading from 4.3(93) to 4.3(97) also breaks down. Finally, note the similarity of expression 4.3(97) to the reflection factor 4.2(45) for magnetoelastic conversion and that for plane electromagnetic waves traveling in a varying dielectric, equation 3.2(29). All of these factors become large when the WKB quasi-normal mode solutions become poor approximations, due to steep gradients or nearness to a critical point where a wavenumber gets small or, in the present case, coalesces with another wavenumber.

The other reflection coefficients may be obtained explicitly in the manner indicated in part 3.3(g). The fraction of total power reflected from an incident "magnetostatic" wave energy into an "exchange" wave will involve the reflection factor R_{24} multiplied by factors from the total power flow expression 4.3(21). Note that these factors canceled in obtaining $\langle S_{tot} \rangle_2 / \langle S_{tot} \rangle_1$ in 4.3(90). R_{24} may be calculated as indicated by 3.3(54) from R_{21} given by 3.3(63), with the subscripts 1 and 2 interchanged everywhere as discussed above. In the lossless case, the result is just that which conserves power with $|R_{21}|^2$ as in 4.3(93), namely $1 - |R_{21}|^2$. The reflection coefficients R_{31} and R_{34} for an incident exchange wave are obtained from R_{21} and R_{24} by interchanging ρ_1 and ρ_3 before 3.3(64) is used. Again for the lossless case, the fraction of total power reflected from an incident "exchange" wave into the oppositely-traveling wave of the same type is identical to the expression in 4.3(93), as might have been expected from the results of the treatment by the WKB method.

Compare now 4.3(97) with the corresponding result 4.3(65) obtained from the phase integral method of 4.3(c). For this purpose, assume that the hyperbolic tangent transition is symmetric, with $H(o) = H_{TP}$, and $H_{\pm} = H_{TP} \pm 4M\lambda k_{TP}^2$. In that way, H_+ and H_- correspond to the values of H(z) at the branch points z_2 and z_1, respectively, which were the limits of integration on the phase integral. Then α from 4.3(95) is $4H'_{TP}/8M\lambda k_{TP}^2$. Substituting this value into 4.3(97) gives $|R_{21}|^2$ the form $\exp(-H'_{crit}/H'_{TP})$ as in 4.3(55). This last step is not actually valid, since $k_3^- = k_1^-$ at H_- as given above, and thus $\frac{\pi}{\alpha}(k_3^- - k_1^-)$ is not much greater than unity, as is required for 4.3(97). If we assume, nevertheless, that 4.3(97) is approximately valid, we get $H'_{crit} = 8\pi M\lambda k_{TP}^2 k_1^-$. This value is within a factor of two of the H'_{crit} of 4.3(55), since $k_1^- = k_{TP}$ here. Choosing H_- to be somewhat smaller so that k_1^- will be in the region where the WKB solutions would be approximately valid, we find that the factors in the two expressions can become identical. (δH increases faster than k_1^- decreases.) Hence the phase-integral result 4.3(55) may well be approximately valid for situations, as in Fig. 3, where the magnetic field H(z) varies approximately linearly rather than as in Fig. 1. Note finally that H'(z) is a function of the transverse dimensions of the sample, just as is k_{TP}, through 4.3(7). See JOSEPH and SCHLÖMANN 1965 for plots of the internal demagnetizing factor N_{zz} for various sample shapes. The internal magnetic field H(z) is just $H - N_{zz}M$, where H is the external applied field.

When losses are included, the gamma functions in the reflection coefficients may be reduced using Stirling's formula, 3.3(65), under the same conditions that 4.3(97) was valid. The result for the fraction of total power reflected from an incident "exchange" wave into the wave of the same type is:

$$|R_{34}|^2 \cong \exp\left[\frac{8+2\,\ln 2}{\alpha}\,\text{Im}\,k_3^-\right]\exp\left[-\frac{2\pi}{\alpha}\left(\text{Re}\,k_1^- - \text{Im}\,k_1^-\right)\right]. \qquad 4.3(99)$$

This formula was confirmed by a computer evaluation using the subroutine LOGGAM mentioned in part 3.3(i). Note that usually $\text{Im}\,k_1 \ll \text{Re}\,k_1$, but that $|\text{Im}\,k_3|$ can be approximately equal to $\text{Re}\,k_1$ (see Fig. 9). In fact, $|\text{Im}\,k_3|$ becomes greater than $\text{Re}\,k_1$ when the spin wave linewidth ΔH_k exceeds the critical value given in 4.3(11). Since $\text{Im}\,k_3^-$ is negative, the reflection factor $|R_{34}|^2$ is then greatly decreased, corresponding probably to absorption in the lossy medium. There is no indication in 4.3(99) that there would be increased reflections caused by the splitting in the real part of the dispersion relation when ΔH_k exceeds ΔH_{crit}. The value of $|R_{21}|^2$, on the other hand, was seen by computer evaluation consistently to increase slightly when losses were included. The net power reflected does not increase, however, because of the attenuation factor $\exp(-2\,\text{Im}\,k_1^-\,z)$ which must now be included in the expression for $\langle S_{tot}\rangle_1/\langle S_{tot}\rangle_2$.

Note finally that equation 3.4(7) treated in section 3.4 has the same form as 4.3(89) with $k_{EX}^2 = p(z)$ and $k_{TP}^4 = K^4$. It is shown there that hypergeometric function solutions may be obtained when k_{EX}^2 has the $\alpha^2\,\text{sech}^2(\alpha z/2)$ variation of 3.4(16). In that case, 4.3(1) shows that H(z) would have a symmetrical valley as in Fig. 2. In rods of yttrium iron garnet, k_{EX} can typically change from 10^2 to 10^4 cm^{-1} in distances of 10^{-2} to 10^{-4} centimeters. Hence a variation such as that of 3.4(16) might be useful in certain studies of spin waves near the center portion of the rod where $H(z)$ has a symmetrical variation (see Fig. 1).

4.4 Solutions for wave conversion with simultaneous presence of spin-wave turning point and magnetoelastic crossover point.

Since the magnetoelastic crossover point of section 4.2 and the fourth-order turning point of section 4.3 often lie very close together in practical situations, we would like to know if wave conversion at one of these points is affected by the proximity of the other. Typically, the difference between these points amounts to about 1 oersted of internal magnetic field H(z), or equivalently a few micro-meters of distance z, or a factor of about 10 in the wavenumber k. In particular, it would be desirable to know whether the reflections represented by the factor y in equation 4.2(43) and in Fig. 7 still persist when the region to the right of the crossover contains a turning point for the spin waves rather than a homogeneous propagation region (compare Figs. 11 and 7). In this section we find expressions for the overall efficiency for conversion from elastic waves to the medium-k "magnetostatic" spin waves and vice versa. The solutions are found by applying the method described in section 3.3 to the relevant sixth-order differential equation.

(a) Derivation of the sixth-order differential equation.

To treat the simultaneous interaction of "magnetostatic" waves, "exchange" waves, and elastic waves, it is necessary to combine the three coupled second-order differential equations 4.1(19), 4.1(20), and 4.1(21). For convenience, we repeat these equations here:

$$\underline{R}_x'' + k_p^2 \underline{R}_x = -(b_2/c_{44}M)\, \underline{m}_x' , \qquad 4.4(1)$$

$$\underline{m}_x'' + k_{EX}^2 \underline{m}_x = (b_2/\mu_o M\lambda)\, \underline{R}_x' - [H/(H + H_{TP})\lambda]\, \underline{h}_x , \qquad 4.4(2)$$

and

$$\underline{h}_x'' - k_t^2 \underline{h}_x = k_t^2 \underline{m}_x . \qquad 4.4(3)$$

Recall that these equations were written assuming that $\underline{h}_y \approx 0$ (assumption 6(a) - see Appendix 3), but that similar equations could be written with r replacing x everywhere if $\underline{h}_\varphi \approx 0$ instead (assumption 6(b)). Equations 4.4(1) through 4.4(3) are generally valid whenever at least one transverse dimension of the sample and the wavelengths of the waves in question are much smaller than the wavelength of the ordinary electromagnetic wave at the frequency of interest. Specifically, the major applicable assumptions include numbers 1, 2, 3, 4, 5, 6a, 7, 8, 9, 11, 12a, 12c, 13, 17, and 18. (See Appendix 3 for the basic physical implications of these assumptions.)

In obtaining one single sixth-order equation from 4.4(1) through 4.4(3), it is most convenient to choose $\underline{m}_x(z)$ as the dependent variable. Equations 4.4(1) through 4.4(3) show clearly that the magnetic field $\underline{h}_x(z)$ and the lattice displacement $\underline{R}_x$ are coupled only through the magnetization $\underline{m}_x$. Furthermore, in homogeneous regions where the fields have the form exp (-jkz), $\underline{R}_x$ and $\underline{h}_x$ can be obtained directly in terms of $\underline{m}_x$ from 4.4(1) and 4.4(3), respectively. This feature is necessary in order to calculate the dominant power in the various waves.

Specifically, the forms for the time-averaged elastic, exchange, and electromagnetic power flows in homogeneous regions are, respectively:

$$\langle S_p \rangle = \omega c_{44} \, |\underline{R}_x|^2 \, \mathrm{Re}(k) \, , \qquad 4.4(4)$$

$$\langle S_m \rangle = \omega \mu_o \lambda \, |\underline{m}_x|^2 \, \mathrm{Re}(k) \, , \qquad 4.4(5)$$

and

$$\langle S_{EM} \rangle = -\frac{1}{2} (\omega \mu_o / k_t^2) \, |\underline{h}_x|^2 \, \mathrm{Re}(k) \, . \qquad 4.4(6)$$

The above expression for $\langle S_p \rangle$ results from 4.2(39) when the definition in 4.2(3) for v in terms of $\underline{R}_+$ is used. Note also that $\underline{R}_+ = 2\underline{R}_x$ follows from assumption 18 (see 4.1(18)). Equation 4.4(5) follows from 4.3(18) when it is assumed (assumption 16) that the magnetization is also circularly polarized, so that $\underline{m}_+ = 2\underline{m}_x$. This latter assumption is valid for "exchange" spin waves reasonably far from the turning point. Equation 4.4(6) is just 4.3(17), assuming $k_t^2 \gg k_o^2$ (assumption 5). From 4.4(1) and 4.4(3) we now write $|\underline{R}_x|^2$ and $|\underline{h}_x|^2$ in terms of $|\underline{m}_x|^2$, again for homogeneous regions:

$$|\underline{R}_x|^2 = \left(\frac{b_2}{c_{44} M}\right)^2 \left(\frac{k}{k_p^2 - k^2}\right)^2 |\underline{m}_x|^2 \qquad 4.4(7)$$

and

$$|\underline{h}_x|^2 \cong \left(k_t^2 / k^2\right)^2 |\underline{m}_x|^2 , \qquad 4.4(8)$$

where we assume that the wavenumber k is real and considerably larger than k_t. (The latter assumption is necessary in ay case to apply the method of section 3.3, as noted below.) Finally, using the definitions for $a = \frac{b_2}{M} (\lambda \mu_o c_{44})^{-\frac{1}{2}}$ from 4.2(3) and for $k_{TP}^4 = k_t^2/2\lambda$ from 4.3(7), we obtain after combining 4.4(4) and 4.4(7), and 4.4(6) and 4.4(8):

$$\langle S_p \rangle = \omega\mu_o \lambda k \left[ak/(k_p^2 - k^2)\right]^2 |\underline{m}_x|^2 \qquad 4.4(9)$$

and

$$\langle S_{EM} \rangle = -\omega\mu_o \lambda \, k \, (k_{TP}^4/k^4) \, |\underline{m}_x|^2 . \qquad 4.4(10)$$

Together with 4.4(5), these expressions allow all three kinds of power flow to be calculated from $\underline{m}_x$.

Now we outline the derivation of the sixth-order differential equation for $\underline{m}_x$ from 4.4(1) through 4.4(3). The only new assumption is that H(z) varies much slower than $\underline{h}_x(z)$, in accordance with assumption 10, which is certainly completely valid in practical cases. Thus in 4.4(1) through 4.4(3) there is effectively only one varying parameter, k^2_{EX}. In the following steps, there will be occasion to differentiate with respect to z equations 4.4(1) through 4.4(3). For simplicity, label the n^{th} derivative of 4.4(x) as $(x)^{(n)}$. For example, $(2)^{(4)}$ denotes the fourth derivative of the entire equation 4.4(2) with respect to z. Similarly, $\underline{m}_x^{(n)}$ denotes the n^{th} derivative of $\underline{m}_x$.

Step 1. Evaluate $(2)^{(4)}$, yielding an equation having $\underline{m}_x^{(6)}$ and lower derivatives on the left-hand side and $\underline{R}_x^{(5)}$ and $\underline{h}_x^{(4)}$ on the right-hand side.

Step 2. Find $\underline{R}_x^{(5)}$ in terms of $\underline{R}_x^{(3)}$ and $\underline{m}_x^{(4)}$ from $(1)^{(3)}$, and $\underline{h}_x^{(4)}$ in terms of $\underline{h}_x^{(2)}$ and $\underline{m}_x^{(2)}$ from $(3)^{(2)}$. After we have evaluated $\underline{R}_x^{(3)}$ and $\underline{h}_x^{(2)}$, these expressions will be substituted into the right-hand side of $(2)^{(4)}$ obtained in step 1.

Step 3. To find $\underline{R}_x^{(3)}$ in terms of $\underline{m}_x$ and its derivatives, first use $(1)^{(1)}$ to find $\underline{R}_x^{(3)}$ in terms of $\underline{R}_x^{(1)}$ and $\underline{m}_x^{(2)}$. Into this expression substitute for $\underline{R}_x^{(1)}$ in terms of $\underline{m}_x^{(2)}$, $\underline{m}_x$, and $\underline{h}_x$ from (2) itself. Substitute in turn for $\underline{h}_x$ in terms of $\underline{h}_x^{(2)}$ and $\underline{m}_x$ from (3). Finally, substitute for $\underline{h}_x^{(2)}$ in terms of $\underline{R}_x^{(3)}$, $\underline{m}_x^{(4)}$, and lower derivatives of $\underline{m}_x$ using $(2)^{(2)}$. By collecting terms, the equation resulting from all these substitutions can be solved for $\underline{R}_x^{(3)}$ in terms of only $\underline{m}_x^{(4)}$ and lower derivatives of $\underline{m}_x$.

Step 4. We now must find $\underline{h}_x^{(2)}$ in terms of $\underline{m}_x$ and its derivatives. First, we again use $(2)^{(2)}$ to express $\underline{h}_x^{(2)}$ in terms of $\underline{R}_x^{(3)}$, $\underline{m}_x^{(4)}$, and lower derivatives of $\underline{m}_x$. Substituting for $\underline{R}_x^{(3)}$ from the result of step 3 completes the present step.

Step 5. Now $\underline{R}_x^{(5)}$ and $\underline{h}_x^{(4)}$ in step 2 are known in terms of $\underline{m}_x$ and its first four derivatives, since we know $\underline{R}_x^{(3)}$ and $\underline{h}_x^{(2)}$ from steps 3 and 4. Finally, these expressions for $\underline{R}_x^{(5)}$ and $\underline{h}_x^{(4)}$ may be substituted into the right-hand side of $(2)^{(4)}$ from step 1. The result is the following equation for $\underline{m}_x^{(6)}$ in terms of $\underline{m}_x$ and its first four derivatives:

$$\underline{m}_x^{(6)} + \underline{m}_x^{(4)} (k_{EX}^2 + k_p^2 + a^2 - k_t^2) + \underline{m}_x^{(3)} \left[4 (k_{EX}^2)^{(1)}\right]$$

$$+ \underline{m}_x^{(2)} \left\{6 (k_{EX}^2)^{(2)} + k_{EX}^2 k_p^2 + k_t^2 \left[\left(\frac{H}{H + H_{TP}}\right)\frac{1}{\lambda} - \left(k_{EX}^2 + k_p^2 + a^2\right)\right]\right\}$$

$$+ \underline{m}_x^{(1)} \left[4 (k_{EX}^2)^{(3)} + 2 (k_{EX}^2)^{(1)} (k_p^2 - k_t^2)\right] \qquad 4.4(11)$$

$$+ \underline{m}_x \left\{(k_{EX}^2)^{(4)} + (k_{EX}^2)^{(2)} (k_p^2 - k_t^2) + k_p^2 k_t^2 \left[\left(\frac{H}{H + H_{TP}}\right)\frac{1}{\lambda} - k_{EX}^2\right]\right\} = 0.$$

The "source equation" for 4.4(11) is found by eliminating all terms in the derivatives of k_{EX}^2. If we label the dependent variable in this "source equation" as F_o, we find using the method of subsection 3.3(f) that $\underline{m}_x = F_o^{(4)} + (k_p^2 - k_t^2) F_o^{(2)} - k_p^2 k_t^2 F_o$.

To be consistent with the assumptions inherent in 4.4(1) through 4.4(3), however, we must eliminate some terms in 4.4(11). First of all, recall that $k_t^2 << k_p^2$ (assumption 12(a)) was assumed in deriving 4.4(1), or 4.1(20). If assumption 12(a) were ever invalid, as perhaps in a thin film, it would still be possible to derive an equation corresponding to 4.1(20). Assuming $\underline{R}_x$ had a transverse variation describable by k_t^2, this would involve retaining terms in k_t^2. Otherwise, apply assumption 12(a) to 4.4(11). Also, ignore k_t^2 compared with k_{EX}^2

(assumption 13) as was done in deriving 4. 4(2) (4. 1(19)) and in section 4. 3. If this assumption were not valid, 4. 1(19) could also be modified. Next note that in practical situations, $2\lambda k_p^2 << 1$ (assumption 12(b)). To be consistent with assumption 11 used in deriving 4. 1(19), we must also assume $2\lambda k_{EX}^2 << 1$ (see assumption 12(d)). Finally, as in section 4. 3, we must approximate $H/(H + H_{TP})$ by $\frac{1}{2}$ (assumption 14), if we hope to solve 4. 4(11) using the method of section 3. 3. With these observations, 4. 4(11) becomes the following consistent and useful equation:

$$\underline{m}_x^{(6)} + \underline{m}_x^{(4)} (k_{EX}^2 + k_p^2 + a^2) + \underline{m}_x^{(3)} \left[4\ (k_x^2)^{(1)}\right] + \underline{m}_x^{(2)} \left[6(k_{EX}^2)^{(2)} + k_{EX}^2 k_p^2 + k_{TP}^4\right]$$

$$+ \underline{m}_x^{(1)} \left[4(k_{EX}^2)^{(3)} + 2(k_{EX}^2)^{(1)} k_p^2\right] + \underline{m}_x \left[(k_{EX}^2)^{(4)} + (k_{EX}^2)^{(2)} k_p^2 + k_{TP}^4 k_p^2\right] = 0$$

4. 4(12)

The "source equation" for 4. 4(12) is

$$F^{(6)} + F^{(4)}(k_{EX}^2 + k_p^2 + a^2) + F^{(2)} (k_{EX}^2 k_p^2 + k_{TP}^4) + F(k_{TP}^4 k_p^2) = 0. \qquad 4.4(13)$$

After comparing 4. 4(12) and 4. 4(13) with 3. 3(40) and 3. 3(41), we see from 3. 3(42) that

$$\underline{m}_x = F^{(4)} + k_p^2 F^{(2)}. \qquad 4.4(14)$$

When it is not necessary to solve 4. 4(12) using generalized hypergeometric functions, note that 4. 4(12) would also be consistent if $k_t^2\ H/(H + H_{TP})\ \lambda$ in 4. 4(11) were not approximated by $k_{TP}^4 \equiv k_t^2/2\lambda$ as in 4. 4(12) and 4. 4(13). Equation 4. 4(14) would also still be valid. Note, however, that assumption 10 must still be satisfied.

(b) Solution using generalized hypergeometric functions.

We now proceed to use the method of section 3.3 to solve 4.4(13). As before in subsections 4.2(b)2 and 4.3(e), we assume that k_{EX}^2 has a hyperbolic tangent variation. The definition for k_{EX}^2 in 4.3(1) shows that if $M\lambda$ is constant H(z) will also have a hyperbolic tangent variation such as those pictured in Fig. 1 or that at the bottom of Fig. 11. The corresponding variation of the wavenumbers is shown schematically in the top of Fig. 11 (compare Fig. 1 in COLLINS and WILSON 1968). Note from 4.4(13) that the dispersion relation for normal mode solutions exp (-jkz) is cubic in k^2. Thus the solutions for k come in pairs of equal magnetude but opposite sign. Far from the crossover and turning points, the solutions for k can be obtained approximately from the unperturbed second-order dispersion relations.

We assume by convention for regions of propagation that the positive roots of the dispersion relation are k_1, k_3, and k_5, corresponding to elastic waves, "exchange" waves, and "magnetostatic" waves respectively. Also define

$$k_2 = -k_1, \; k_4 = -k_3, \; k_6 = -k_5. \qquad 4.4(15)$$

Since the former two wave types are forward waves, k_1 and k_3 represent waves traveling from $z = -\infty$ to $z = +\infty$. When loss is included, k_1 and k_3 must have negative imaginary parts in order to represent waves which decay rather than grow with increasing z. Even in lossless media, $\text{Im}\,(k_3) < 0$, whenever the "exchange" wave reaches a cutoff region $(H > H_{TP})$ where $k_3^2 < 0$. The "magnetostatic" wave is, however, a backward wave. Thus k_6 with a negative real part will denote a "magnetostatic" wave carrying power in the positive z-direction (see

4.4(10)). Since k_6 must then have a non-positive imaginary part in order that exp $(-jk_6 z)$ = exp $(+ jk_5 z)$ not grow exponentially with increasing z, we conclude that Im $(k_5) \geq 0$. Notice that both types of spin waves are evanescent in the region to the right of the crossover and turning points in Fig. 11. Only the elastic wave is propagating there.

The reflection and transmission factors for the "source equation" 4.4(13) can now be found by generalization of the techniques of subsection 3.3(g), which treated fourth-order systems explicitly. Instead of 3.3(43) and 3.3(44) we write

$$F_1 = G_1^+ + J_{13} G_3^+ + J_{16} G_6^+ = \sum_{g=1}^{6} (P_{1g} + J_{13} P_{3g} + J_{16} P_{6g}) G_g^-, \quad 4.4(16)$$

where the subscript on F and the first subscript on the J's indicate that the elastic wave G_1^- is considered to be the only wave incident from $z = -\infty$. Thus we must determine J_{13} and J_{16} such that the coefficients of the other incident waves G_3^- and G_6^- vanish. This is done by solution of two simultaneous equations for J_{13} and J_{16} with the result

$$J_{13} = (P_{16}P_{63} - P_{13}P_{66})/(P_{33}P_{66} - P_{36}P_{63}), \quad 4.4(17)$$

where the elements P_{hg} are given in 3.1(32). Application of the interchange operator $(\sigma_3 \leftrightarrow \sigma_6)_{op}$ to J_{13} produces J_{16} (see the discussion of the properties of such operators preceding 3.3(53)).

In order for the incident wave to have unit amplitude, we must also divide 4.4(16) by the coefficient of G_1^-. The transmission factor T_{11} for G_1^- to G_1^+ then becomes:

$$T_{11} = (P_{11} + J_{13} P_{31} + J_{16} P_{61})^{-1} \quad 4.4(18)$$

Using the same arguments as in 3.3(g), we find that the reflection factor R_{12} for G_1^- to G_2^- is again given in terms of T_{11} by equation 3.3(55). Furthermore, relations such as 3.3(53) and 3.3(54) also hold, so that all the other transmission and reflection factors may be obtained from T_{11} simply by interchange of symbols. For example, $(\sigma_1 \leftrightarrow \sigma_6)_{op} T_{11} = T_{16}$, and $(\rho_2 \leftrightarrow \rho_5)_{op} R_{12} = R_{15}$. Finally, interchange of ρ_1 and ρ_2 yields the corresponding factors T_{6i} and R_{6i} for the case of an incident "magnetostatic" wave. We thus conclude that only the form of T_{11} need be calculated explicitly, just as for the fourth order case.

To simplify the expression 4.4(18) for T_{11}, we must write J_{13} and J_{16} in terms of the P_{hg} as in 4.4(17) and then cancel out common factors in the products of gamma functions arising from the definition 3.1(32) for the P_{hg}. In the process of combining terms we also use 3.3(60) and relations similar to 3.3(61). As a result, we obtain:

$$T_{11} = \frac{\Gamma(\sigma_3-\sigma_6)\Gamma(1-\sigma_3+\sigma_6)}{C(\rho_1)} \frac{\Pi_{i=1}^{6} \Gamma(\rho_i+\sigma_1)\Gamma(1-\sigma_i-\rho_1)}{\Pi_{i=2}^{6} \Gamma(\rho_i-\rho_1)\Gamma(1-\sigma_i+\sigma_1)}, \qquad 4.4(19)$$

where

$$\begin{aligned} C(\rho_i) \equiv\ & \Gamma(\sigma_3-\sigma_6)\Gamma(1-\sigma_3+\sigma_6)\Gamma(\sigma_1+\rho_i)\Gamma(1-\sigma_1-\rho_i) \\ & + \Gamma(\sigma_6-\sigma_1)\Gamma(1-\sigma_6+\sigma_1)\Gamma(\sigma_3+\rho_i)\Gamma(1-\sigma_3-\rho_i) \qquad 4.4(20) \\ & + \Gamma(\sigma_1-\sigma_3)\Gamma(1-\sigma_1+\sigma_3)\Gamma(\sigma_6+\rho_i)\Gamma(1-\sigma_6-\rho_i) \end{aligned}$$

Note that the last two terms in $C(\rho_i)$ simply involve a cyclic permutation of σ_1, σ_3, and σ_6 in the arguments of the gamma functions. To calculate the conversion from an elastic wave to a reflected "magnetostatic" wave, we first need the reflection factor R_{15}. Using 3.3(55) and interchanging ρ_2 and ρ_5 in the result, we find

$$R_{15} = \frac{C(\rho_5)\ \Pi_{i=1,\ i\neq 5}^{6}\ \Gamma(\rho_i - \rho_5)}{C(\rho_1)\ \Pi_{i=2}^{6}\ \Gamma(\rho_i - \rho_1)}\ \Pi_{i=1}^{6} \left(\frac{\Gamma(1-\sigma_i-\rho_1)}{\Gamma(1-\sigma_i-\rho_5)}\right) \qquad 4.4(21)$$

(To obtain the reflection factor R_{62} for the reverse conversion, simply interchange ρ_1 with ρ_6 and ρ_2 with ρ_5 in 4.4(21).)

The desired conversion efficiency is the ratio $\langle S_{EM}\rangle_5^- / \langle S_p\rangle_1^-$. To write this ratio in terms of 4.4(21), first use the power flow expression 4.4(9) and 4.4(10) to write the ratio in terms of $|\underline{m}_x|^2$:

$$\frac{\langle S_{EM}\rangle_5^-}{\langle S_p\rangle_1^-} = -\frac{k_5^-}{k_1^-}\left(\frac{k_{TP}}{k_5^-}\right)^4 \left[\frac{k_p^2 - (k_1^-)^2}{a\,k_1^-}\right]^{-2} \frac{|\underline{m}_x|_5^2}{|\underline{m}_x|_1^2} \qquad 4.4(22)$$

From 4.4(14) which gives $\underline{m}_x$ in terms of the source function F we can evaluate the ratio of the $|\underline{m}_x|^2$ factors:

$$\frac{|\underline{m}_x|_5^2}{|\underline{m}_x|_1^2} = \left|\frac{(k_5^-)^4 - k_p^2\,(k_5^-)^2}{(k_1^-)^4 - k_p^2\,(k_1^-)^2}\right|^2 \left|\frac{G_5^-}{G_1^-}\right|^2, \qquad 4.4(23)$$

where G_5^- and G_1^- are the waves of interest in 4.4(16). (See the discussion surrounding 3.3(47) and 3.3(48).) Recall from section 3.3(a) that these waves reduce as $z \to -\infty$ to the form $\exp(-jk_5^- z)$ and $\exp(-jk_1^- z)$, respectively. Finally, note that

$$|G_5^-|^2 / |G_1^-|^2 \equiv |R_{15}|^2 \qquad 4.4(24)$$

Combining equations 4.4(22) through 4.4(24) gives the desired result:

$$\frac{\langle S_{EM}\rangle_5^-}{\langle S_p\rangle_1^-} = -\frac{k_{TP}^4}{a^2}\frac{(k_5^-)}{(k_1^-)^7} \left| (k_5^-)^2 - k_p^2 \right|^2 |R_{15}|^2. \qquad 4.4(25)$$

As was the case for similar expressions in part 4.2(c)1, the conversion efficiency is not given by $|R_{15}|^2$ alone, since the elastic and "magneto-elastic" waves represent two different physical wave types. Furthermore, the reflection factors for F(z) must be different from those for $\underline{m}_x(z)$ because of 4.4(14).

In lossless media, the expression for $|R_{15}|^2$ from 4.4(21) can be written in terms of hyperbolic sines only, following the method described in subsections 3.3(h) and illustrated in parts 3.2(a)5 and 4.2(c)1. First of all, note that $C(\rho_i)$ in 4.4(20) can be written in terms of hyperbolic sines (or sines) even in lossy media, in view of the property 3.3(60). Recall from 3.31(18) that ρ_i and σ_i are imaginary when k_i^- and k_i^+ are real, respectively. Hence in our case, all the ρ_i and σ_i are imaginary, except for σ_3, σ_4, σ_5, and σ_6. Furthermore, all of the ρ_i and σ_i occur in pairs with opposite sign, as in 3.3(64), since the dispersion relation is written in terms of k^2 only (see 4.4(15)).

With this information, the remaining factors in $|R_{15}|^2$ from 4.4(21) can be reduced to products and quotients of sines, sometimes with imaginary or complex arguments:

$$|R_{15}|^2 = C(\rho_5)\, r(\rho_1)/C(\rho_1)\, r(\rho_5) \qquad 4.4(26)$$

where

$$C(\rho_i) = \pi^2 \Big\{ [\sin\pi(\sigma_3 + \sigma_5) \sin\pi(\rho_i + \sigma_1)]^{-1} + [\sin\pi(-\sigma_5 - \sigma_1) \sin\pi(\rho_i + \sigma_3)]^{-1} + [\sin\pi(\sigma_1 - \sigma_3) \sin\pi(\rho_i - \sigma_5)]^{-1} \Big\} \qquad 4.4(27)$$

and

$$r(\rho_i) = \rho_i(\rho_i^2 - \rho_3^2) \sin 2\pi\rho_i \sin\pi(\rho_3 - \rho_i) \sin\pi(\rho_3 + \rho_i)$$

$$\cdot \prod_{\ell=1,3,5} (\rho_i^2 - \sigma_\ell^2) [\sin\pi(\rho_i + \sigma_\ell) \sin\pi(\rho_i - \sigma_\ell)]^{-1} . \quad 4.4(28)$$

From 3.3(18), the ρ_i and σ_i in these expressions may be written in terms of the wavenumbers in the limiting homogeneous regions. We can also choose the magnetic field for these limiting regions to be far enough from the turning point and crossover point that the wavenumbers can be found approximately from the unperturbed second-order dispersion relations. More effort, however, is needed to interpret the results.

APPENDIX 1

Calculation of the linear combination matrix L for use in finding the WKB quasi-normal modes

Section 2.2 showed how to find n quasi-normal mode solutions Q_i for a set of n coupled first-order equations of the form of 2.2(1): $U' = DU$. Furthermore, every solution for the field component U_1 could be expressed formally as a linear combination of these Q_i through 2.2(20), which is a component of the transformation $U = LQ$ defined in 2.2(2). For this linear combination to be independent of distance, the first row of L must consist of constants L_{1i}, $i = 1, 2, \cdots, n$. In this case the Q_i then appear as pseudo-basis functions for every solution of 2.2(1) for U_1. Since the amplitudes A_i of these quasi-normal modes Q_i, however, are free to change through coupling to other waves as described by 2.2(26), we expect that it will always be possible to choose the L_{1i} to be constants (see 2.2(27)). In this appendix we show how this can be done systematically for systems initially describable by one fourth-order equation, by two sets of coupled second-order equations, or by four coupled first-order equations. The importance of having constant L_{1i} lies in the fact that the total power in the i^{th} mode Q_i will then normally be proportional simply to $|A_i|^2$. (See the discussion following 2.2(20), and the example leading to 4.3(36).)

Consider first the transformation to four coupled first-order equations. For a fourth order differential equation of the form

$$F''''(z) + v_1 F'''(z) + v_2 F''(z) + v_3 F'(z) + v_4 F(z) = 0, \qquad \text{A1(1)}$$

where the v_i may be functions of z, we obtain 2.2(1) by defining $U_1 = F(z)$,

$U_2 = F'(z)$, $U_3 = F''(z)$, and $U_4(z) = F'''(z)$. Then $U' = D_1 U$ with

$$D_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -v_4 & -v_3 & -v_2 & -v_1 \end{bmatrix} \qquad \text{A1(2)}$$

(Compare 4.3(23) and 4.3(25).)

A set of two coupled second-order equations can be written most generally as

$$\begin{aligned} F_1'' + k_1^2 F_1 &= a_{12} F_2 + b_{12} F_2' \\ F_2'' + k_2^2 F_2 &= a_{21} F_1 + b_{21} F_1' . \end{aligned} \qquad \text{A1(3)}$$

If any first derivative terms appeared on the left-hand sides of A1(3), they could be eliminated by the standard transformation for such a purpose for second order differential equations (see, for example, KORN and KORN 1961, section 9.3-8(c)). Any second derivative terms on the right-hand sides of A1(3) could be eliminated by substituting from the other equation in the pair. If we define $U_1 = F_1(z)$, $U_2 = F_1'(z)$, $U_3 = F_2(z)$, and $U_4 = F_2'(z)$, then equations A1(3) also reduce to the form $U' = D_2 U$, with

$$D_2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -k_1^2 & 0 & a_{12} & b_{12} \\ 0 & 0 & 0 & 1 \\ a_{21} & b_{21} & -k_2^2 & 0 \end{bmatrix} . \qquad \text{A1(4)}$$

Now we have to find the matrix L which diagonalizes D, with eigenvalues λ_i. These eigenvalues are found as solutions of the equation $|D - \lambda I| = 0$, where I is the unit matrix. Explicitly evaluating this determinant, we obtain from D_1:

$$\lambda^4 + \underline{v}_1 \lambda^3 + \underline{v}_2 \lambda^2 + \underline{v}_3 \lambda + \underline{v}_4 = 0, \qquad \text{A1(5)}$$

and from D_2:

$$\lambda^4 + \lambda^2 (\underline{k}_1^2 + \underline{k}_2^2 - \underline{b}_{12}\underline{b}_{21}) + \lambda(-\underline{a}_{12}\,\underline{b}_{21} - \underline{a}_{21}\underline{b}_{12}) + (\underline{k}_1^2\underline{k}_2^2 - \underline{a}_{12}\underline{a}_{21}) = 0 \qquad \text{A1(6)}$$

Clearly, these have the form of the dispersion relations for the respective systems, with solutions $\lambda_i = -jk_i$, as long as all the coefficients $\underline{v}_\ell$, $\underline{k}_1^2$, $\underline{a}_{12}$, $\underline{b}_{12}$, etc., are taken to have the form their counterparts in A1(1) and A1(3) would have in a homogeneous medium. In other words, we drop all terms in A1(1) and A1(2) involving derivatives of the parameters of the system. An analysis can be performed without dropping these terms, but there is commonly no advantage to such a procedure (see RYDBECK 1967).

Next, we wish to observe that an eigenvector for any λ_i is given by any column of the matrix adj $(D - \lambda_i I)$, where the adjoint matrix is the transpose of the original matrix replaced by its cofactors. The possibility of choosing the eigenvectors in this way follows from the fact that $W(\text{adj } W) = |W|\, I$ for any n x n matrix W (see, for example, HILDEBRAND 1965). Indeed, for any eigenvector c_i for λ_i, we must have $(D - \lambda_i I)c_i = 0$, which can be satisfied by non-zero c_i since $|D - \lambda_i I| = 0$. Letting $W = D - \lambda_i I$ in the formula cited just above, therefore,

$$(D - \lambda_i I) \text{ adj } (D - \lambda_i I) = 0, \qquad A1(7)$$

showing that any column of adj $(D - \lambda_i I)$ will serve as an eigenvector c_i of λ_i.

Choose now the i^{th} column vector forming the n x n matrix L to consist of one eigenvector found in the above manner for each eigenvector λ_i. Defining $\boldsymbol{\Lambda}$ as the diagonal matrix formed from these eigenvalues, we find from A1(7):

$$DL = L\boldsymbol{\Lambda}. \qquad A1(8)$$

According to a standard theorem in matrix algebra, the eigenvectors corresponding to unequal eigenvalues are linearly independent. Consequently, if no two λ_i are equal, L has an inverse and A1(8) becomes

$$L^{-1}DL = \boldsymbol{\Lambda}. \qquad A1(9)$$

Thus the diagonalization of D is completed. Since the diagonal elements $\Lambda_{ii} = \lambda_i = -jk_i(z)$, we also obtain 2.2(6).

We still have some latitude in choosing L, however, since we can choose any column of adj $(D - \lambda_i I)$ to be column i of L, and each column can be multiplied by any non-zero constant. Consider, for example,

$$-(D_1 - \lambda_i I) = \begin{bmatrix} \lambda_i & -1 & 0 & 0 \\ 0 & \lambda_i & -1 & 0 \\ 0 & 0 & \lambda_i & -1 \\ \underline{v}_4 & \underline{v}_3 & \underline{v}_2 & \underline{v}_1 + \lambda_i \end{bmatrix} \qquad A1(10)$$

The cofactor c_{41} of this matrix is easily seen to be +1. Thus, the first element in column 4 of -adj $(D_1 - \lambda_i I)$ is also +1, independent of $\lambda_i(z)$. Therefore, the first row of L will consist of ones if we choose column i of L to be column 4 of - adj $(D_1 - \lambda_i I)$. Constants other than unity can be obtained for L_{1i} by multiplying column 4 of -adj$(D_1 - \lambda_i I)$ by the desired L_{1i} before placing it into the matrix L.

When the L_{1i} are all chosen to be unity, the rest of the matrix L is easily found from A1(10). For example, the second element in column 4 of -adj $(D_1 - \lambda_i I)$ is simply the cofactor $c_{42} = \lambda_i$. Similarly, the third and fourth elements are given by $c_{43} = \lambda_i^2$ and $c_{44} = \lambda_i^3$, respectively. In summary, then:

$$L = \begin{bmatrix} 1 & 1 & 1 & 1 \\ \lambda_1 & \lambda_2 & \lambda_3 & \lambda_4 \\ \lambda_1^2 & \lambda_2^2 & \lambda_3^2 & \lambda_4^2 \\ \lambda_1^3 & \lambda_2^3 & \lambda_3^3 & \lambda_4^3 \end{bmatrix} \qquad \text{A1(11)}$$

A matrix with the form of A1(11) is called an alternant matrix. It is also the transpose of a Vandermode matrix. In calculating $(-L^{-1}L')$ for use in section 2.2, and for the result quoted at the end of section 1.2, it is helpful to be aware of the simple formula for the determinant of such matrices, namely:

$$|L| = \prod_{i<\ell} (\lambda_i - \lambda_\ell), \qquad \text{A1(12)}$$

where the product is over all distinct factors with i from 1 to n - 1 and ℓ from 2 to n. When n = 4 as in A1(11):

$$|L| = (\lambda_1 - \lambda_2)(\lambda_1 - \lambda_3)(\lambda_1 - \lambda_4)(\lambda_2 - \lambda_3)(\lambda_2 - \lambda_4)(\lambda_3 - \lambda_4). \qquad \text{A1(13)}$$

The calculation of the matrix L for D_2 is more complicated. First of all, we have:

$$-(D_2 - \lambda_i I) = \begin{bmatrix} \lambda_i & -1 & 0 & 0 \\ \underline{k}_1^2 & \lambda_i & -\underline{a}_{12} & -\underline{b}_{12} \\ 0 & 0 & \lambda_i & -1 \\ -\underline{a}_{21} & -\underline{b}_{21} & \underline{k}_2^2 & \lambda_i \end{bmatrix} \qquad \text{A1(14)}$$

For this matrix, the cofactor c_{41} is $(\underline{a}_{12} + \underline{b}_{12}\,\lambda_i)$. To make the first row of L consist of ones, we can thus choose column i of L to be the fourth column of $-\text{adj}\,(D_2 - \lambda_i I)$, divided by $(\underline{a}_{12} + \underline{b}_{12}\,\lambda_i)$. Note that it is permissible to divide by such a factor even though it varies with z, since we diagonalize D only at fixed z. The diagonalization of D must be different at each point since D varies with z. The other elements $L_{\ell i}$ are given in terms of the cofactors $c_{4\ell}$ of A1(14) as

$$L_{\ell i} = c_{4\ell}(\lambda_i)/c_{41}(\lambda_i); \qquad i, \ell = 1, 2, 3, 4. \qquad \text{A1(15)}$$

where the cofactors are considered as functions of λ_i.

In Appendix 2 it is shown that any system of two coupled second-order equations such as A1(3) can be reduced to one fourth-order equation such as A1(1), with $F = F_1$. The eigenvalue equation A1(6) must, however, remain unchanged, since it corresponds to the dispersion relation of the physical system in homogeneous regions. Thus from comparison of A1(5) and A1(6) we can conclude that, for the equivalent fourth-order equation, $\underline{v}_1 = 0$, $\underline{v}_2 = \underline{k}_1^2 + \underline{k}_2^2 - \underline{b}_{12}\underline{b}_{21}$, etc. The actual coefficients v_ℓ in A1(1) may include derivatives of system parameters and thus cannot be obtained so simply. They can be calculated explicitly by the technique of Appendix 2, or by a different method in RYDBECK 1960. Without such calculation, however, the "WKB amplitudes" B_{ii} can still be derived for the new fourth-order system by using A1(11), 2.2(16), and 2.2(13), since L in Al(11) depends only on the λ_i calculated from A1(6). The quasi-normal modes Q_i will now be expressed in terms of F_1 and its first three derivatives, rather than in terms of F_1, F_1', F_2, and F_2'.

Finally, note that similar statements must hold if the system is initially describable by four coupled first-order equations. Although the form of $-(D - \lambda_i I)$ will generally be more complicated than A.1(10) or A.1(14) for such systems, A1(15) can still be used to obtain L with ones in the first row. Alternatively, the system could first be transformed to one fourth-order equation using the technique of Appendix 2, or A1(11) could be used directly, once the eigenvalues of D are known.

APPENDIX 2

Transformation from four coupled first-order equations to one fourth-order differential equation

In this appendix, we assume that we have a system initially described by four coupled first-order equations. If there are initially two coupled second-order equations, such as A1(3), these can easily be converted to four coupled first-order equations as indicated in connection with equation A1(4). Alternatively, RYDBECK 1960 may be consulted for a closed form expression for the fourth-order equation resulting from A1(3). Much further calculation is involved in that expression, however, to calculate explicitly the coefficients of each derivative term. In the following we show how to find the coefficients of each derivative term in the fourth-order equation resulting from any set of four coupled first-order equations. Because of the nature of the determinants involved, this appendix consists more of an existence theorem than a practical method. In order to apply a solution method such as that of section 3.3 for the fourth-order equation, it will probably be necessary to have a fairly simple set of coupled first-order equations. In such cases, the relevant fourth-order equation can probably be calculated conveniently by multiple substitutions, without using the present method.

The following procedure originated from discussions with A. Platzker. Consider the coupled equations in the form of 2.2(1):

$$U' = DU, \qquad \text{A2(1)}$$

where now D is a 4 x 4 matrix, and U is a 4 x 1 column matrix. In component notation, A2(1) becomes

$$U_i' = \sum_{\ell=1}^{4} D_{i\ell} U_\ell \qquad \text{A2(2)}$$

Now write the next three derivatives of A2(2):

$$U_i'' = \sum_{\ell=1}^{4} (D_{i\ell}' U_\ell + D_{i\ell} U_\ell') \qquad \text{A2(3)}$$

$$U_i''' = \sum_{\ell=1}^{4} (D_{i\ell}'' U_\ell + 2D_{i\ell}' U_\ell' + D_{i\ell} U_\ell'') \qquad \text{A2(4)}$$

$$U_i'''' = \sum_{\ell=1}^{4} (D_{i\ell}''' U_\ell + 3D_{i\ell}'' U_\ell' + 3D_{i\ell}' U_\ell'' + D_{i\ell} U_\ell'''), \qquad \text{A2(5)}$$

where Leibniz' theorem for differentiation of a product has been used to collect terms. Equations A2(2) through A2(5) constitute 16 equations in 20 variables, since i can range from one to four and we have for each i the variables U_i, U_i', U_i'', U_i''', and U_i''''. The desired fourth-order differential equation has 5 variables, namely U_m and its four derivatives for some m. Hence we must eliminate 15 variables using 15 equations.

To effect the desired elimination of the variables, write down the 16 equations in A2(2) through A2(5), placing all the terms in U_m, U_m', U_m'', and U_m''' on the right-hand sides. Consider these as the known or "driving" terms in the system of linear equations. The remaining 16 variables are to be treated as "unknowns," on the left-hand sides of the equations. One of these "unknowns" is U_m''''. Now we can solve for U_m'''' in terms of the lower derivatives of U_m by using Cramer's rule for the solution of simultaneous linear equations. The formal expression is clearly a ratio of two 16 x 16 determinants. Calculation of these determinants is simplified if they appear in block form. This can be accomplished by writing the "unknowns" in the following order, assuming for simplicity that m = 1: U_1'''', U_2'''', U_3'''', U_4''''; U_2, U_2', U_2'', U_2'''; U_3, U_3', U_3'', U_3'''; U_4, U_4', U_4'', U_4'''. The 16 x 16 matrix of coefficients is thus partitioned into 16 4 x 4 submatrix blocks, several of which consist entirely of zeros. Alternatively, the matrix can be considered as consisting of 4 8 x 8 blocks. The evaluation of the two determinants in Cramer's rule may then be simplified by use of formulas given by GANTMACHER 1959, volume I, page 46, as was suggested to the present author by A. Platzker. For example, if an n x n matrix is written in block form

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix} ,$$

where A, B, C, and D are square matrices, the determinant $|M|$ is

$$|M| = |AD - BD^{-1} CD| ; \qquad |D| \neq 0 \qquad \text{A2(6)}$$

or

$$|M| = |AD - ACA^{-1} B| ; \qquad |A| \neq 0 . \qquad \text{A2(7)}$$

(The matrix D here is not to be confused with that of equations A2(1) through A2(5).) If C and D or A and C commute, these formulas may be further simplified in the obvious way. The restrictions $|D| \neq 0$ or $|A| \neq 0$ are then not necessary.

APPENDIX 3

The major assumptions used in deriving the differential equations in section 4.3, and their physical implications

(Not all of these assumptions are used in any given derivation.)

1. Fields can all be written in the form of 1.2(1): $\vec{F}(\vec{r}, t) = \mathrm{Re}\,[\underline{\vec{F}}(\vec{r})\, c^{j\omega t}]$; that is, single frequency excitation is assumed. Only spatial variations are considered in this case; no time variation of the material parameters is assumed. Also pulse excitation is not explicitly handled; a Fourier integral treatment on the results would be required to handle pulses.

2. $m_z \ll h_z$. This assumption may break down only for wavelengths close to those of slowly-varying electromagnetic waves on the one hand or those of the very-rapidly evanescent left circularly polarized waves on the other, as long as only small signal linear excitations are involved (see assumption 7). In the former region, the modes may approach transverse magnetic configurations with $h_z \approx 0$, but then m_z will also become very small. In the latter regions, h_z may also become very small, but then Maxwell's equations may be neglected.

3. $e_z \approx 0$. This implies that a transverse-electric ("TE") type of mode is excited. Figure 3 shows that fine wire antenna configurations normally used are likely to excite TE modes since they provide strong z and transverse components of dynamic magnetic field, but only a transverse component of electric field. This assumption also implies that transverse-electric modes do not couple to transverse-magnetic (TM) modes, which is not true for most modes even in open (unshielded) simple

dielectric waveguides, for example. However, VASILE 1967 has shown for axially magnetized magnetic samples that the axial (z) component of electric field is small compared with all other field quantities when at least one of the transverse dimensions is much smaller than the dielectric wavelength, so that our assumption 5 holds. If such a sample is enclosed by a metallic conductor, the "TE" mode with negligible e_z fills the interior of the waveguide. A rapidly decaying "TM" surface mode of small amplitude is present at the walls, however, to satisfy the boundary condition by cancelling the small e_z of the "TE" volume mode (see also VASILE and LAROSA 1968b). Note that assumption 3 is less restrictive than the ordinary "magnetostatic" assumption that $\nabla \times \vec{\underline{h}} \approx 0$. However, the same equations are obtained for both cases after assumption 5 is also included. The above argument, moreover, shows that assumption 3 is likely to be dependent upon 5.

4. Separation of variables in $\underline{h}_z$. The complex field components $\underline{h}_i(\vec{r})$ satisfying partial differential wave equations can be written in the form $\underline{h}_i(\vec{r}) = \underline{T}_i(x, y)\, Z_i(z)$. This assumption is dependent upon the satisfaction of two conditions:

(a) $\nabla_t^2 \underline{h}_i = -k_t^2 \underline{h}_i$, where $\nabla_t^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ and k_t^2 is a constant independent of z;

(b) The permeability is independent of x and y so that the longitudinal wave equation is also independent of transverse coordinates.

The equation in (a) is the ordinary transverse wave equation for transverse electric modes in waveguides with conducting walls and homogeneous cross sections. In magnetized non-ellipsoidal samples such as rods the magnetic field and hence the permeability is a function of position. Then, even with a conducting wall, the boundary conditions and hence k_t^2 will be dependent on z (see BURKE and BHAGAT 1967). The variation of the magnetic field in the transverse plain invalidates condition b), and results from the necessity of the net internal static magnetic field to satisfy Laplace's equation. Near the ends and center of an axially-magnetized rod, however, the z-variation of this field is almost linear, and the radial variation for long rods is then small (see JOSEPH and SCHLÖMANN 1965, Figs. 12 and 13). Only when there is substantial curvature in the z direction does the radial variation become important. In that case, it may be possible to obtain the transverse wavenumber k_t by finding an effective radius for the rod as in LEWIS and SCOTTER 1969.

The radial field variation can also cause the turning point to be a function of radial distance from the rod axis. The result is a turning point "surface" as drawn for one special case in Fig. 9 of ADDISON and AULD 1968.

In conclusion, separation of variables is most valid near the center of an axially magnetized rod, where the static magnetic field varies little in either the radial or longitudinal directions. The rapid longitudinal variation near the ends of the rod may affect which mode is excited by affecting the effective k_t for each mode. See BURKE and BHAGAT 1967, page 18, for a discussion of this effect and critique of results based on the assumption of excitation of the mode treated by FLETCHER and KITTEL 1960.

5. $k_t^2 >> k_0^2 \equiv \omega^2 \mu \epsilon$. This assumption implies that at least one of the transverse dimensions is much smaller than the dielectric wavelength. and that the fields do not extend too far outside the sample. This assumption is likely to be necessary for assumption 3 to be valid in magnetic samples. The combination of 3 and 5 is equivalent to the "magnetostatic approximation."

6(a) $h_y \approx 0$. This assumption follows from $\nabla \times \vec{\underline{h}} = j\omega\epsilon \vec{\underline{e}}$ and $\underline{e}_z \approx 0$ (assumption 3), if the y-variations in the field are small, since then $0 \approx (\nabla \times \vec{\underline{h}})_z = \frac{\partial \underline{h}_y}{\partial x} - \frac{\partial \underline{h}_x}{\partial y} \cong \frac{\partial \underline{h}_y}{\partial x}$. This situation can arise from two causes. First of all, the sample may be much longer in the y than in the x direction, so that y-variations are relatively negligible. Secondly, a mode may be excited which has no y-variation. For example, the magnetic field produced by a fine wire antenna placed parallel to the y direction at the end of a slab which is thin in the x-direction will have only x and z components.

6(b) $h_\varphi \approx 0$. In a rod of cylindrical symmetry, this situation can occur as long as a mode with no φ-variation is excited, as can be shown again from the z-component of the curl-h equation, assuming $\underline{e}_z \approx 0$. It is easy to construct a fine-wire "loop" antenna which will produce only r and z components of magnetic field at the end of a rod, for example. Such an antenna configuration is pictured in Fig. 9b of AULD, COLLINS, and WEBB 1968. Of three configurations they tried, they found experimentally that this configuration had the lowest optimum insertion loss for the excitation of magnetoelastic waves in rods.

7. $\underline{m}_z << \underline{m}_x, \underline{m}_y$, where it is assumed that the saturation magnetization M and the static magnetic field H(z) lie approximately in the z-direction. Note that $\underline{m}_z$ represents only the dynamic magnetization component: $\underline{m}_z = M(1 - \cos\theta) \approx -\frac{1}{2}M\theta^2$, for $\theta << 1$, where θ is the angle of precession of the total magnetization about the z-axis. On the other hand, the transverse components $\underline{m}_x$ and $\underline{m}_y$ are proportional to $M \sin\theta \approx M\theta$, for $\theta << 1$. This assumption is justified when only small-signal, linear excitations are considered. From the torque equation, assuming time-harmonic fields (assumption 1): $j\omega\underline{m}_z = \gamma\mu_o(\vec{\underline{m}}_t \times \vec{\underline{h}}_t)$, where the subscript t denotes a transverse component. This equation can normally be neglected because the right-hand side contains products of two quantities small compared with M and H(z). To analyze the instabilities, harmonic generation, etc., caused by large-signal excitation, this non-linear equation must be included.

8. $h_z << H(z)$, where H(z) is the net static magnetic field, which in nonellipsoidal samples is a function of position. In every case of interest the dynamic field $\underline{h}_z$ is certainly small compared to the static field. Otherwise fantastic power would have to be in the system, or the demagnetizing field must be so large as almost to cancel the applied field. This latter event may possibly be approximated near the surface of barely-magnetized samples.

9. $\left|\frac{\nabla^2 M}{M}\right| << \left|\frac{\nabla^2 \vec{\underline{m}}}{\vec{\underline{m}}}\right|$, where $\vec{\underline{m}}$ is the dynamic magnetization. In the interior of most samples magnetized to "saturation," the magnetization M varies little or not at all (see JOSEPH and SCHLÖMANN 1965). Certainly the effective "wavelength" of its variation is much longer in any

case than the wavelengths of the waves of interest. This assumption may become invalid only near the boundary between two materials of different magnetization. If it is invalid, then one of the boundary conditions may also be affected. From equation 14 of RADO and WEERTMAN 1959, we can infer that one boundary condition is the continuity of $\lambda\vec{M} \times \frac{d\vec{M}}{dn}$, where $\frac{d}{dn}$ is the derivative taken normal to the plane of the boundary. When assumption 9 is valid, this condition reduces to the continuity of $\lambda\vec{M} \times \frac{d\vec{m}}{dn}$ (see MORGENTHALER 1967). Otherwise there is an additional term $\lambda\vec{m} \times \frac{d\vec{M}}{dn}$.

10. $\left|\frac{d}{dz} \log\left(\frac{H}{H(z) + H_{TP}}\right)\right| << \left|\frac{1}{\underline{\vec{h}}} \frac{d\underline{\vec{h}}}{dz}\right|$, where $H_{TP} = \omega/|\gamma\mu_o|$. This assumption is very similar to 9 and holds generally where 9 does, since the net static field H(z) in the interior of magnetized samples varies slowly compared to wavelengths of interest.

11. $\left|\frac{H + M\lambda k_t^2 + \omega/|\gamma\mu_o|}{M\lambda} \underline{m}_-\right| \gg \left|\frac{d^2\underline{m}^-}{dz^2}\right|$, where $\underline{m}^- \equiv \underline{m}_x - j\underline{m}_y$ is the negatively circularly polarized component of $\underline{m}$. The essence of this assumption is that the wavelengths of the waves of interest are much greater than those of the very-rapidly evanescent negatively circularly polarized waves. When the wavelengths of the positively circularly polarized waves are near those of the elastic waves, this assumption is valid.

12(a) $k_t^2 << k_p^2$. This is true whenever all effective transverse dimensions of the sample are much longer than elastic wavelengths, which are on the order of micro-meters at gigahertz frequencies.

12(b) $2\lambda k_p^2 << 1$. This assumption is normally well satisfied at the low microwave frequencies. Note that assumption 12(a) implies $\lambda k_t^2 <<< 1$ whenever 12(b) is valid.

12(c) $\lambda k_t^2 <<< \frac{H(z) + H_{TP}}{M}$, where $H_{TP} \equiv \omega/|\gamma\mu_o|$. This assumption is valid for usual operating parameters whenever 12(b) is satisfied.

12(d) $2\lambda k_{EX}^2 << 1$. In practice, this assumption breaks down only when the wavenumber of the positively circularly polarized spin wave approaches that of the rapidly-evanescent negatively circularly polarized wave (see assumption 11).

13. k_t^2 is ignored compared with $k_{EX}^2 \equiv (H_{TP} - H(z))/M\lambda$. This assumption is not valid near the turning point where $H(z) \approx H_{TP} = \omega/|\gamma\mu_o|$, and the coefficient of the second derivative term in 4.3(6) gets very small. Including k_t^2 in this term, however, only has the effect of changing very slightly the physical location of the turning point defined as the point where the second derivative term goes to zero. The difference in magnetic field H(z) between the real "turning point" and the shafted one (neglecting k_t^2) is approximately $M\lambda k_t^2 \approx 10^{-6}$ oersted for $k_t \approx 10$ cm^{-1}, which is a reasonable value for k_t in practice. Alternatively, one can view this assumption as being equivalent to shifting the excitation frequency in the definition of k_{EX}^2 by about 2.8 Hertz ($|\gamma\mu_o| \leftrightarrow 2.8$ MHz per oersted).

14. $\frac{H(z) - H_{TP}}{H(z) + H_{TP}} << 1$. This assumption is valid only reasonably near the turning point where $H(z) \approx H_{TP}$. For longer wavelengths

approaching the transverse dimensions of the sample, this approximation no longer gives the correct "magnetostatic" wave dispersion relation. When the wavelengths become comparable to electromagnetic wavelengths, of course, then assumptions 5 and possibly 3 also break down, and the "magnetostatic" approximation is also not valid. Assumption 14 must be made in order to model conveniently wave-conversion at a turning point by generalized hypergeometric differential equations, by the WKB method, or by the phase-integral method.

15. $k_t << \left| \frac{1}{\underline{\vec{m}}} \frac{d\underline{\vec{m}}}{dz} \right|$. This assumption simplifies the expressions for power flow. It is implied by assumption 14, but is also true for the short-wavelength exchange waves when 14 is not valid. It is not valid, however, for "magnetostatic" waves with wavelengths large enough to be comparable with transverse dimensions.

16. $\underline{m}_- \cong 0$, where $\underline{m}_- = \underline{m}_x - j\underline{m}_y$, so that the magnetization is positively circularly polarized. This assumption is equivalent to assumptions 13 and 14 together, as far as determining the differential equation is concerned. It is also a good approximation when assumption 15 is valid. Note that it is <u>not</u> assumed that $\underline{\vec{h}}$ is circularly polarized, so that $\underline{h}_- = 0$. Normally in fact one transverse component of the magnetic field may be zero (assumptions 6(a) or 6(b)), so that $\underline{h}_- = \underline{h}_+$.

17. $\underline{R}_z = 0$. This assumption implies that longitudinal elastic waves are absent, that only shear elastic waves interact with the spin waves. This is true when the magnetization is parallel to a crystal axis in a cubic crystal, for example. In certain other practical cases this may

not be true (see MORGENTHALER, REZENDE, HU, and PLATZKER 1970), but then similar equations can be developed for the conversion to longitudinal waves from spin waves. It is then assumed that the longitudinal elastic waves interact with spin waves at a point distinguishable from where the shear waves interact.

Note: Assumption 17 also implies that the term

$$(c_{44} + c_{12}) \frac{\partial}{\partial z} (\nabla \cdot \vec{\underline{R}}) + \frac{b_2}{M} (\partial \underline{m}_x / \partial x + \partial \underline{m}_y / \partial y)$$

coupling $\underline{R}_z$ and $\vec{\underline{R}}_t$ in the ignored equation for R_z is negligible (see 4.1(1)).

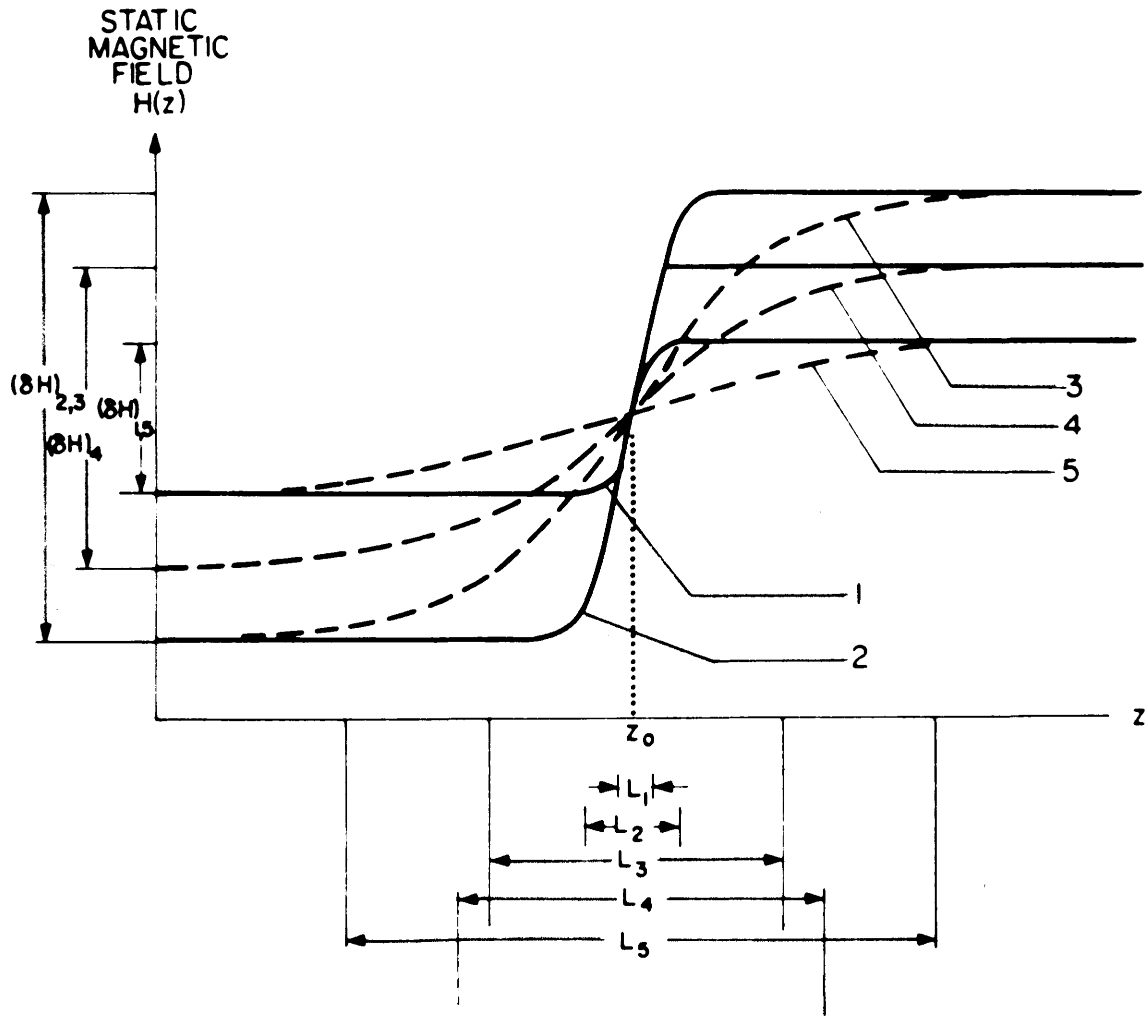
18. $\underline{R}_- = 0$, where $\underline{R}_- = \underline{R}_x - j\underline{R}_y$. This assumption means that the elastic shear system which interacts with the spin system is positively circularly polarized. It is basically equivalent to assumptions 12(a) and 15, and those mentioned preceding 4.1(1).

19(a) $|\vec{\underline{h}}| << \left| \frac{b_2}{\mu_0 M} \frac{d}{dz} \vec{\underline{R}}_t \right|$, where $\vec{\underline{h}}_t$ and $\vec{\underline{R}}_t$ are the transverse components of the dynamic magnetic field and the elastic displacement, respectively. With assumption 6(a), $|\vec{\underline{h}}_t| = |\underline{h}_x|$, or with 6(b), $|\vec{\underline{h}}_t| = |\vec{\underline{h}}_r|$. With assumption 18, $|\vec{\underline{R}}_t| = 2\,|\underline{R}_x|$ or $2\,|\underline{R}_r|$. When 19(a) is valid, the sixth order system decouples so that "magnetostatic" waves do not enter into the fourth-order magnetoelastic coupling equation of the form 4.2(33). Effectively this means that the waves act like plane waves and do not notice the transverse boundaries of the sample.

19(b) $|\vec{\underline{h}}_t| >> \left| \frac{b_2}{\mu_0 M} \frac{d}{dz} \vec{\underline{R}}_t \right|$. When 19(b) is valid, the elastic waves do not influence the conversion of "magnetostatic" to exchange waves, described by the fourth-order turning point equation 4.3(6).

The validity of 19(a) and 19(b) for the respective wave-conversion problems will depend upon how close are the turning point and magneto-elastic crossover point. The further apart these points are, the better the approximation is that the sixth order equation can be treated by the two separate fourth-order equations for which 19(a) is valid in one case and 19(b) in the other. Spreading the points apart is accomplished with larger k_p (slower elastic waves) and smaller k_t (larger transverse dimensions to the sample. Longitudinal elastic waves usually have faster velocities than shear waves.

VARIOUS FIELD VARIATIONS SATISFYING A HYPERBOLIC TANGENT - TANH (z) - FORM



$\frac{4}{a} \equiv \frac{\delta H}{H'(z_0)} \equiv L$: EFFECTIVE TRANSITION WIDTH

NOTE: THE COUPLING POINT NEED NOT BE AT z_0

FIGURE 1

FIELD VARIATIONS HAVING A SYMMETRIC VALLEY
DESCRIBED BY EQUATIONS 3.4 (16) AND 3.2 (20)

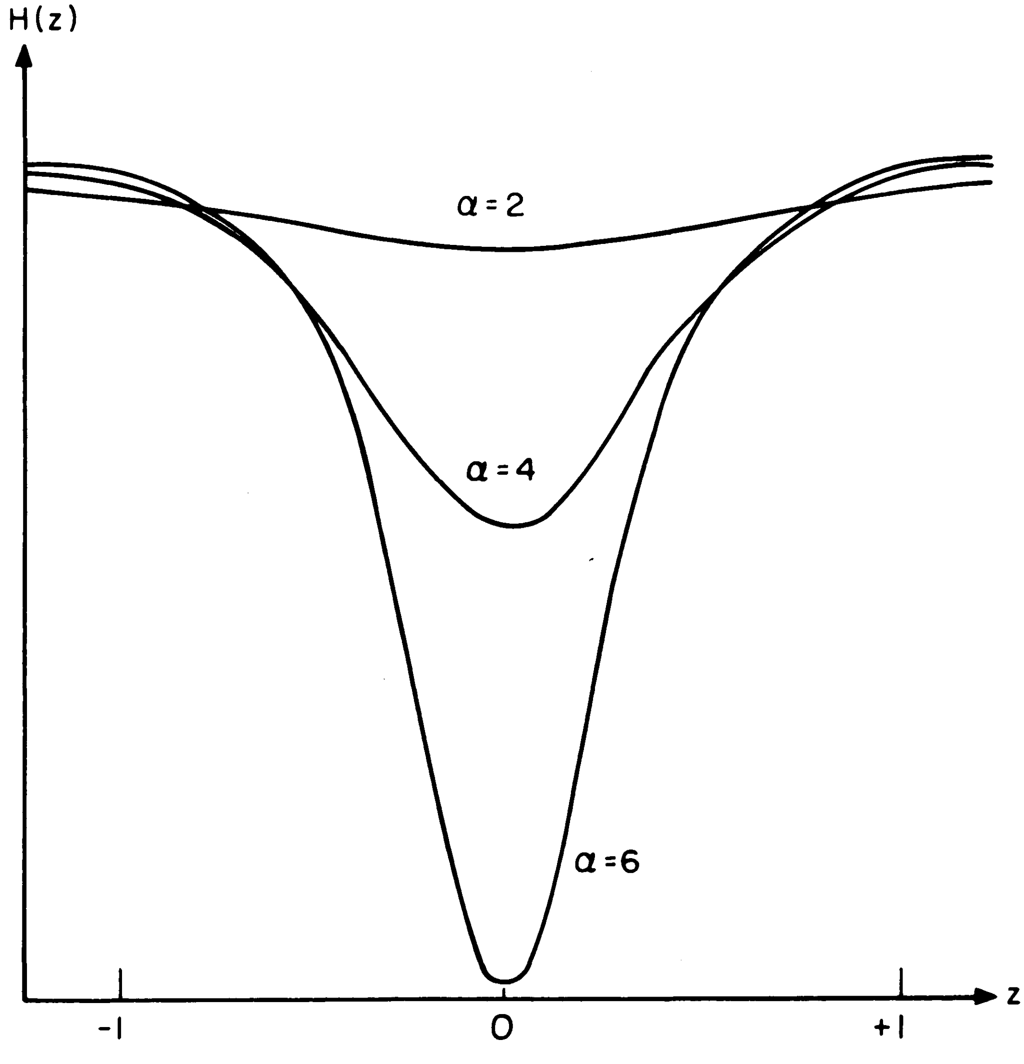


FIGURE 2

EXCITATION OF WAVES IN A MAGNETOELASTIC DELAY LINE

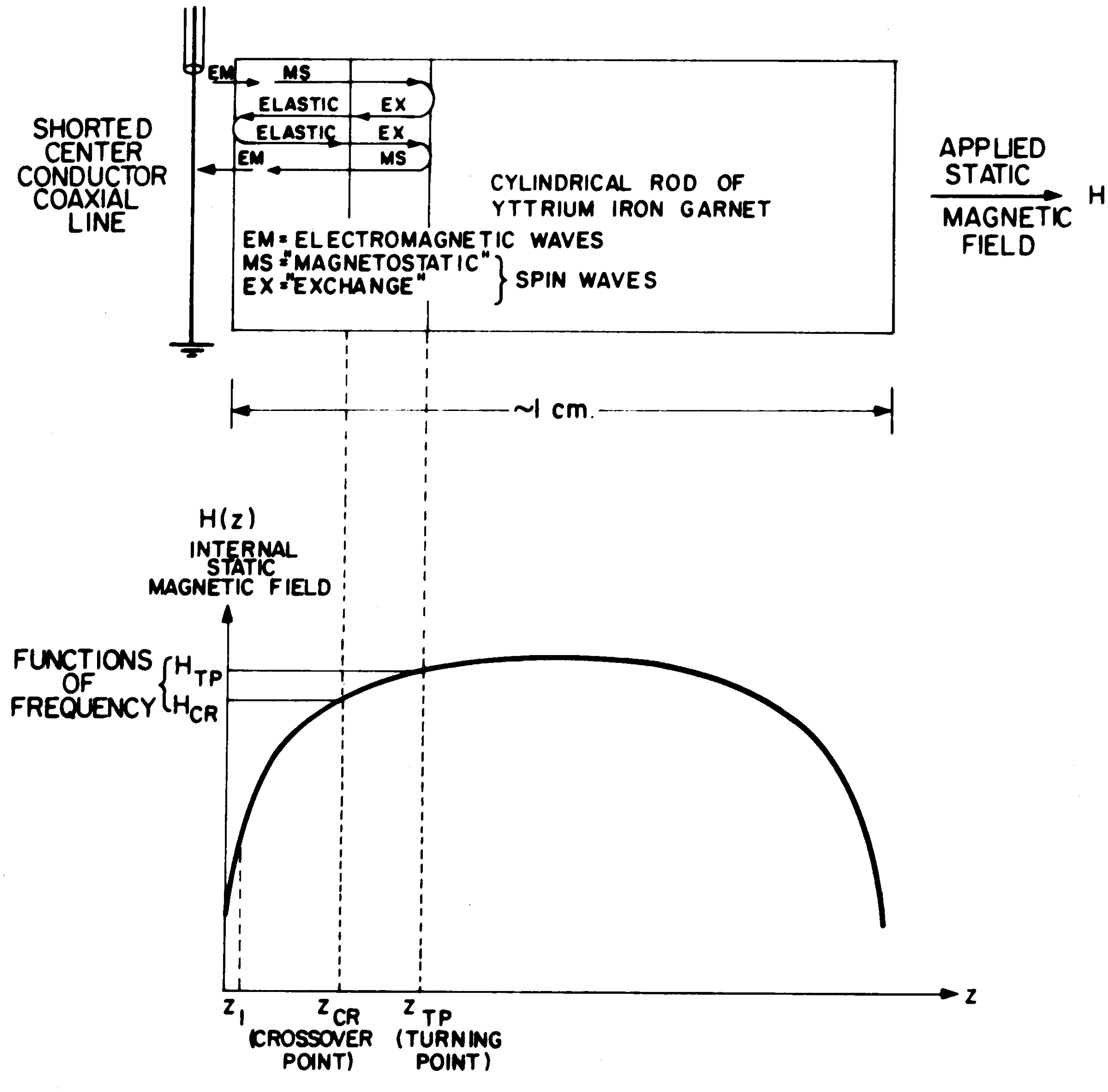


FIGURE 3

THE DISPERSION RELATION IN RODS OF YTTRIUM IRON GARNET

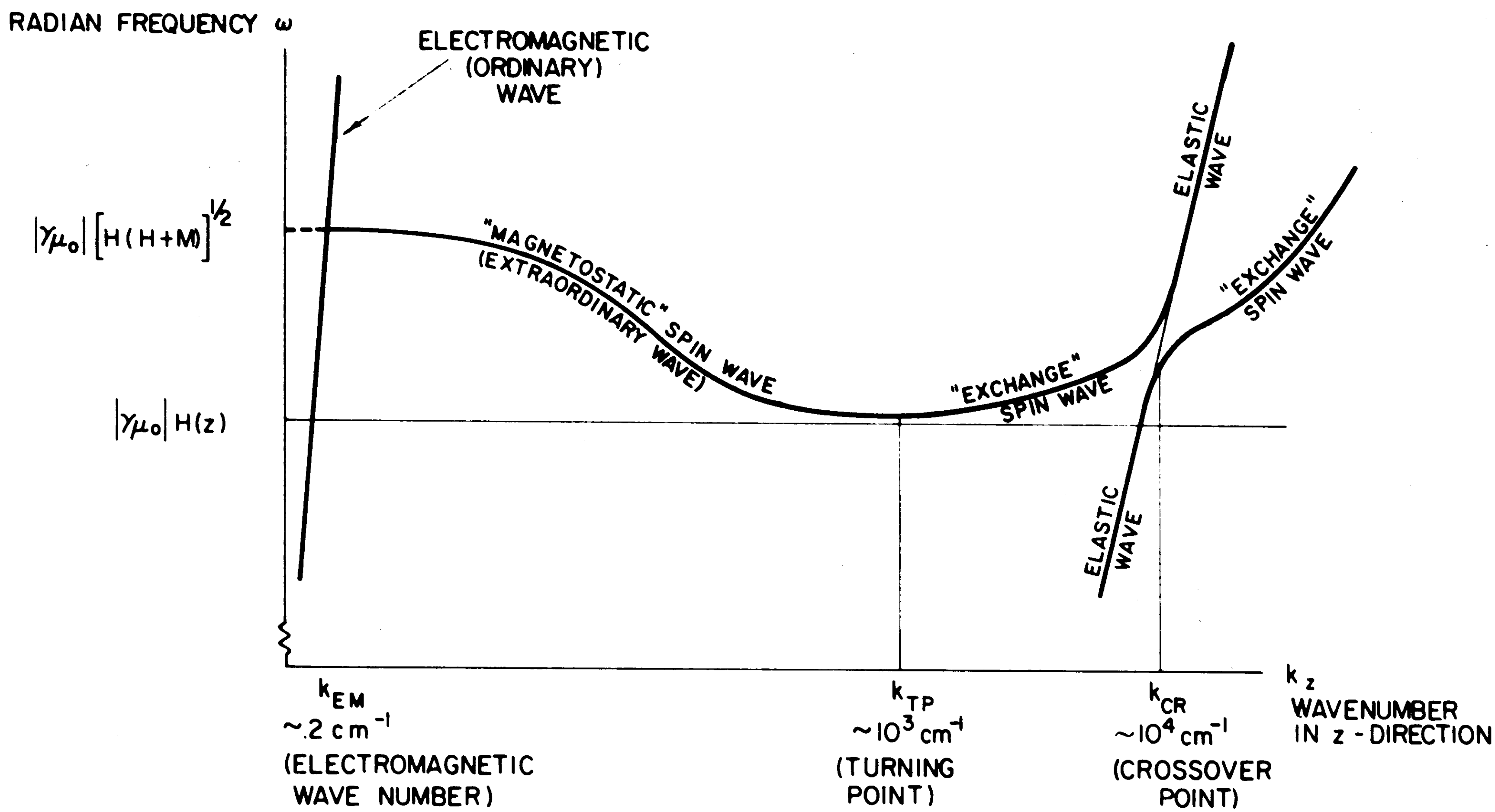
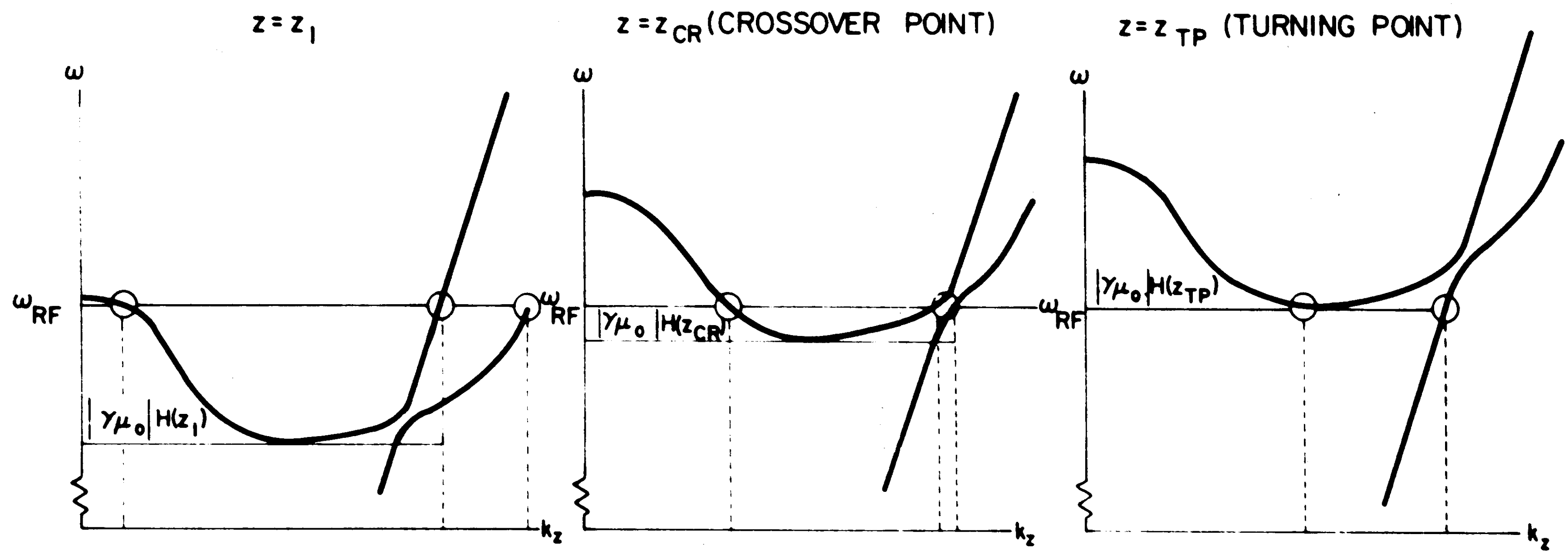


FIGURE 4

THE EFFECT OF A NON-UNIFORM FIELD H (z) ON THE DISPERSION RELATION

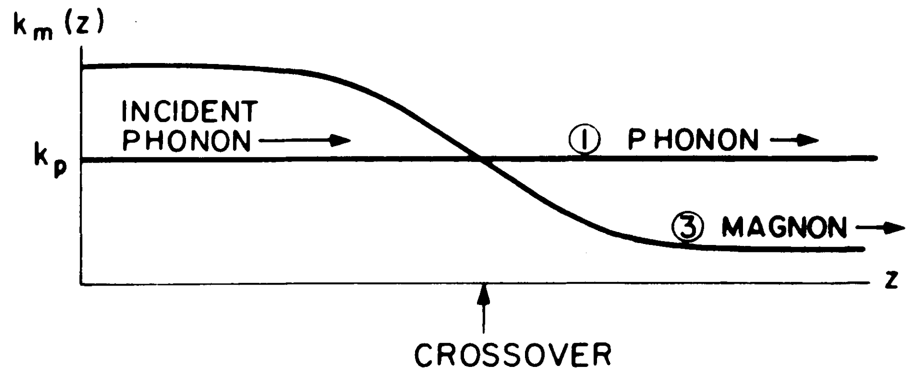


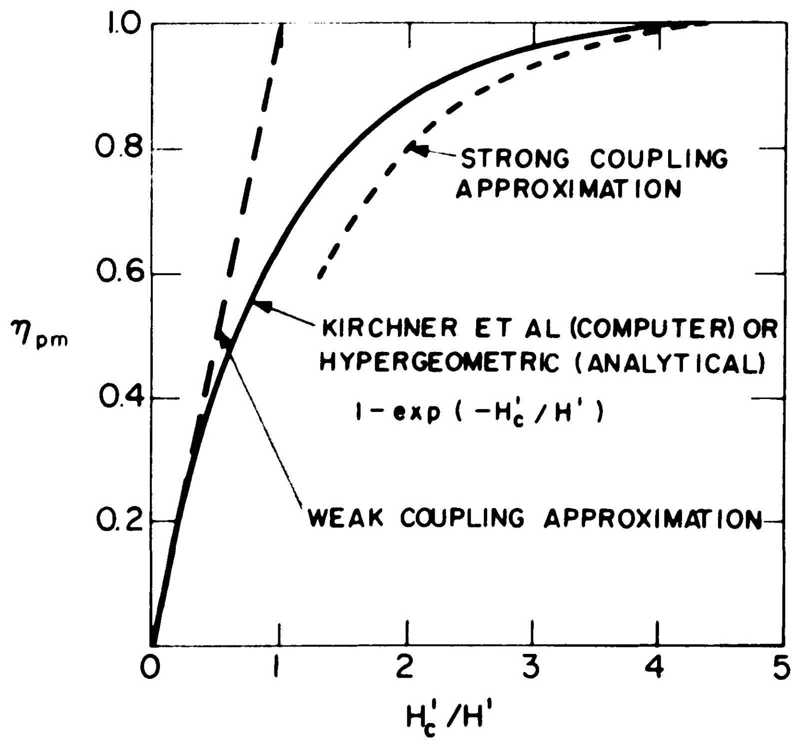
DOTTED VERTICAL LINES SHOW THE ALLOWED VALUES OF k_z AT EACH POINT
ω_{RF} IS THE FIXED FREQUENCY OF EXCITATION

FIGURE 5

SOLUTIONS FROM COUPLED MODE EQUATIONS
(NEGLECTING REFLECTIONS)

$$A_p'' + j(k_m(z) + k_p)A_p' + (\frac{a^2}{4} - k_m(z)k_p)A_p = 0$$

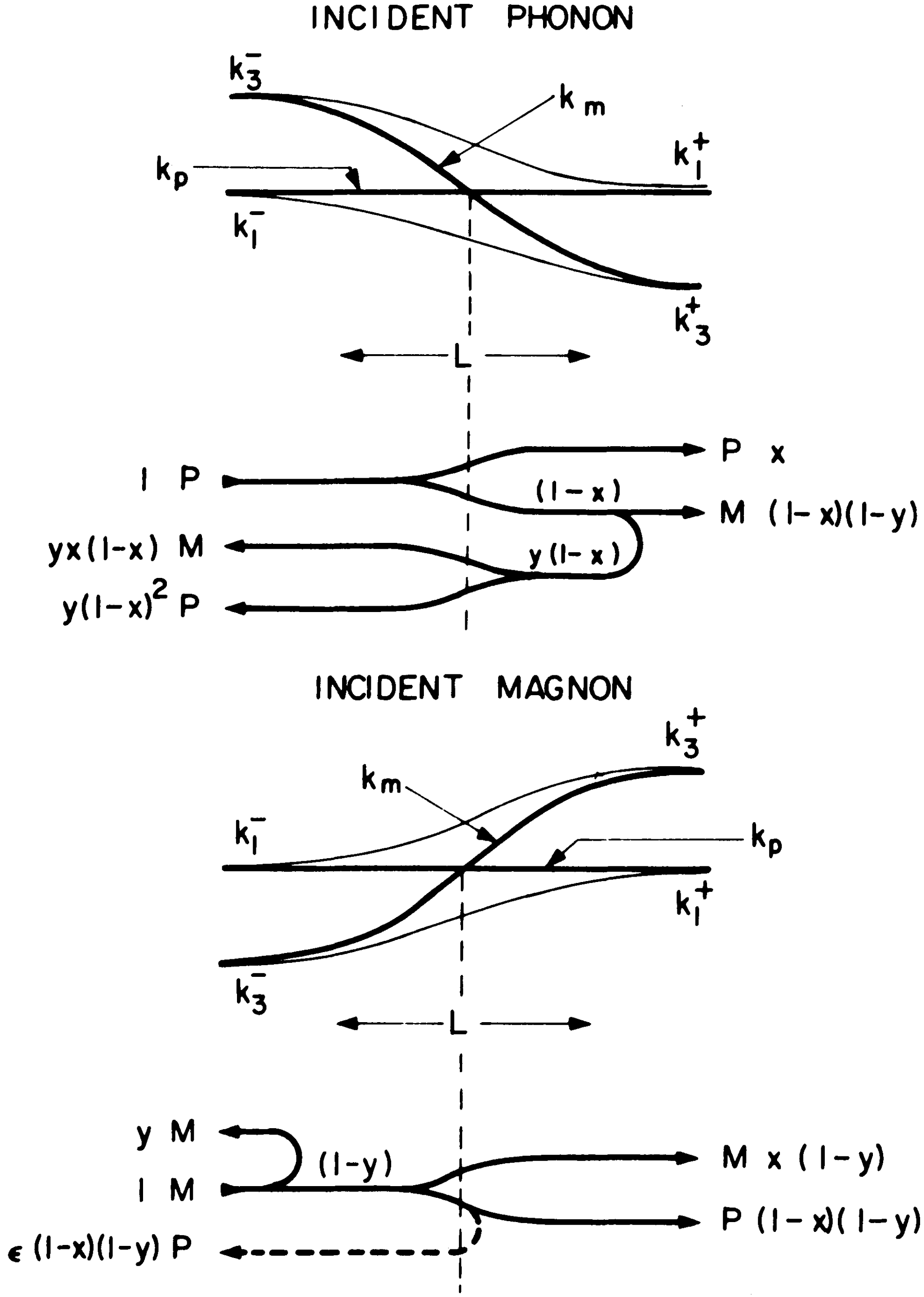




H_c' IS A CRITICAL FIELD GRADIENT; H' IS THE ACTUAL FIELD GRADIENT AT CROSSOVER

FIGURE 6

SCHEMATIC INTERPRETATION OF ANALYTICAL RESULTS FOR MAGNETOELASTIC REFLECTION AND TRANSMISSION POWER FRACTIONS



$x = \exp\left[-H'_{CRIT}/|H'|\right]$ $y = \exp\left[-\pi k_m^{MIN} L\right]$ $\epsilon = \exp\left[-\pi k_p L\right]$

FIGURE 7

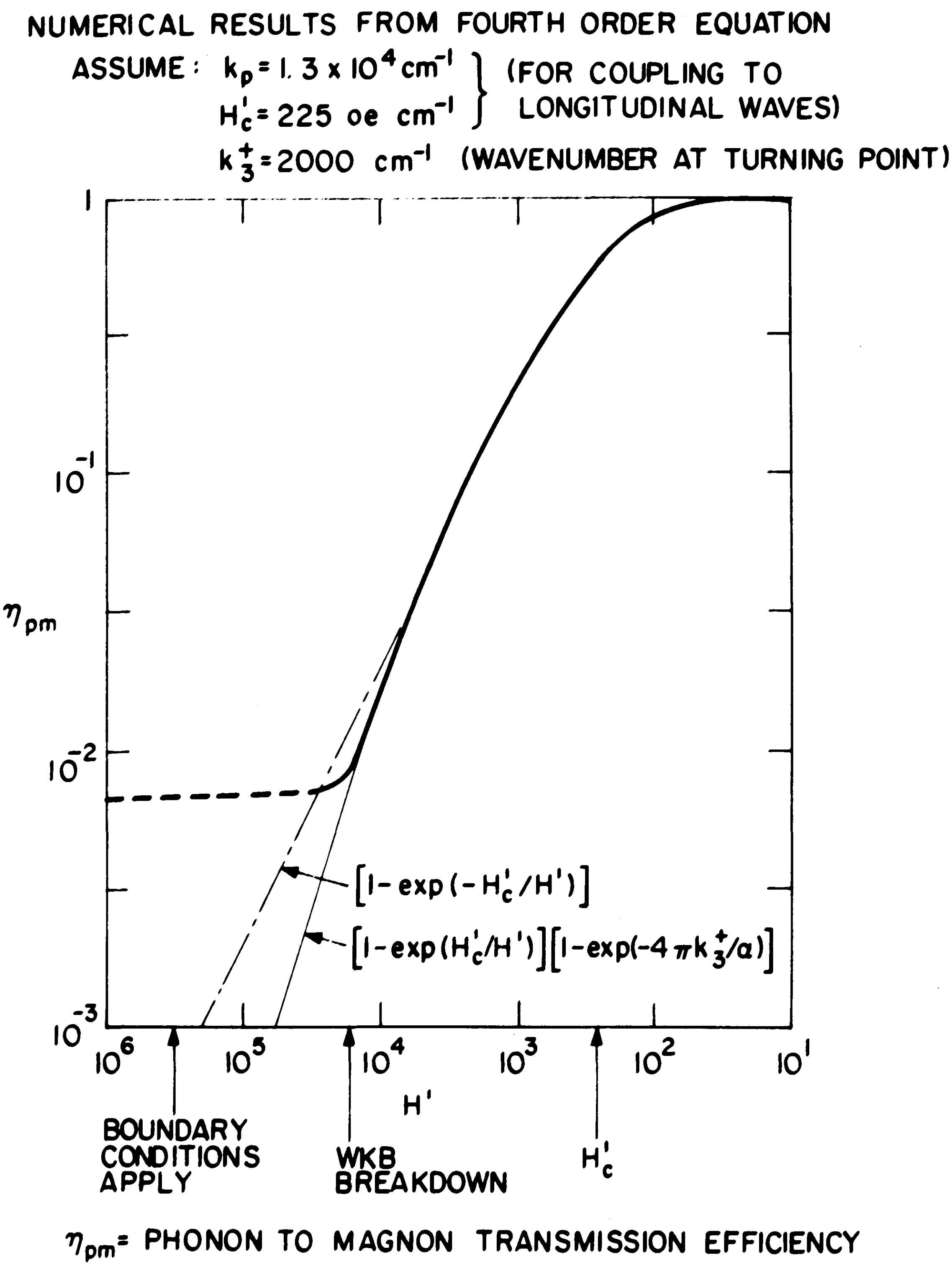
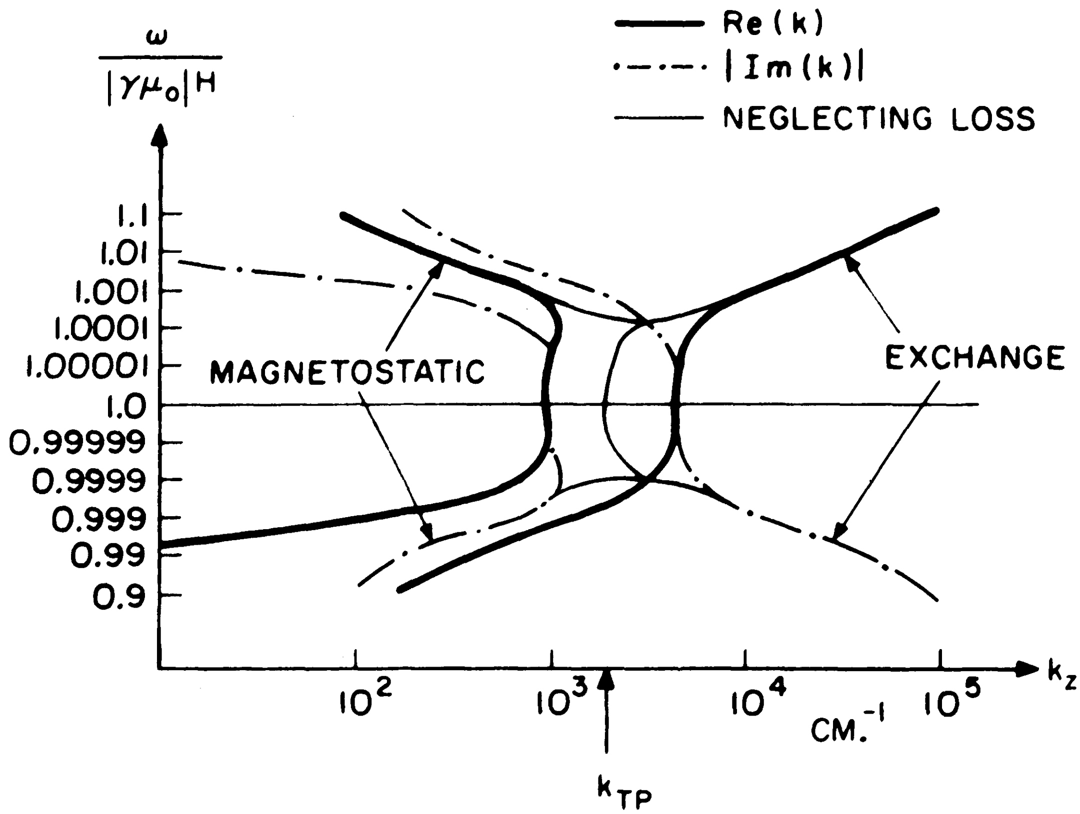

NUMERICAL RESULTS FROM FOURTH ORDER EQUATION
ASSUME: kp = 1.3 x 10^4 cm^-1
H'c = 225 oe cm^-1
(FOR COUPLING TO LONGITUDINAL WAVES)
k3+ = 2000 cm^-1 (WAVENUMBER AT TURNING POINT)
1
10^-1
10^-2
10^-3
ηpm
[1 - exp(-H'c/H')]
[1 - exp(H'c/H')][1 - exp(-4πk3+/a)]
10^6
10^5
10^4
10^3
10^2
10^1
H'
BOUNDARY CONDITIONS APPLY
WKB BREAKDOWN
H'c
ηpm = PHONON TO MAGNON TRANSMISSION EFFICIENCY

FIGURE 8

DISPERSION RELATION INCLUDING LOSS FOR SPIN WAVES NEAR THE TURNING POINT



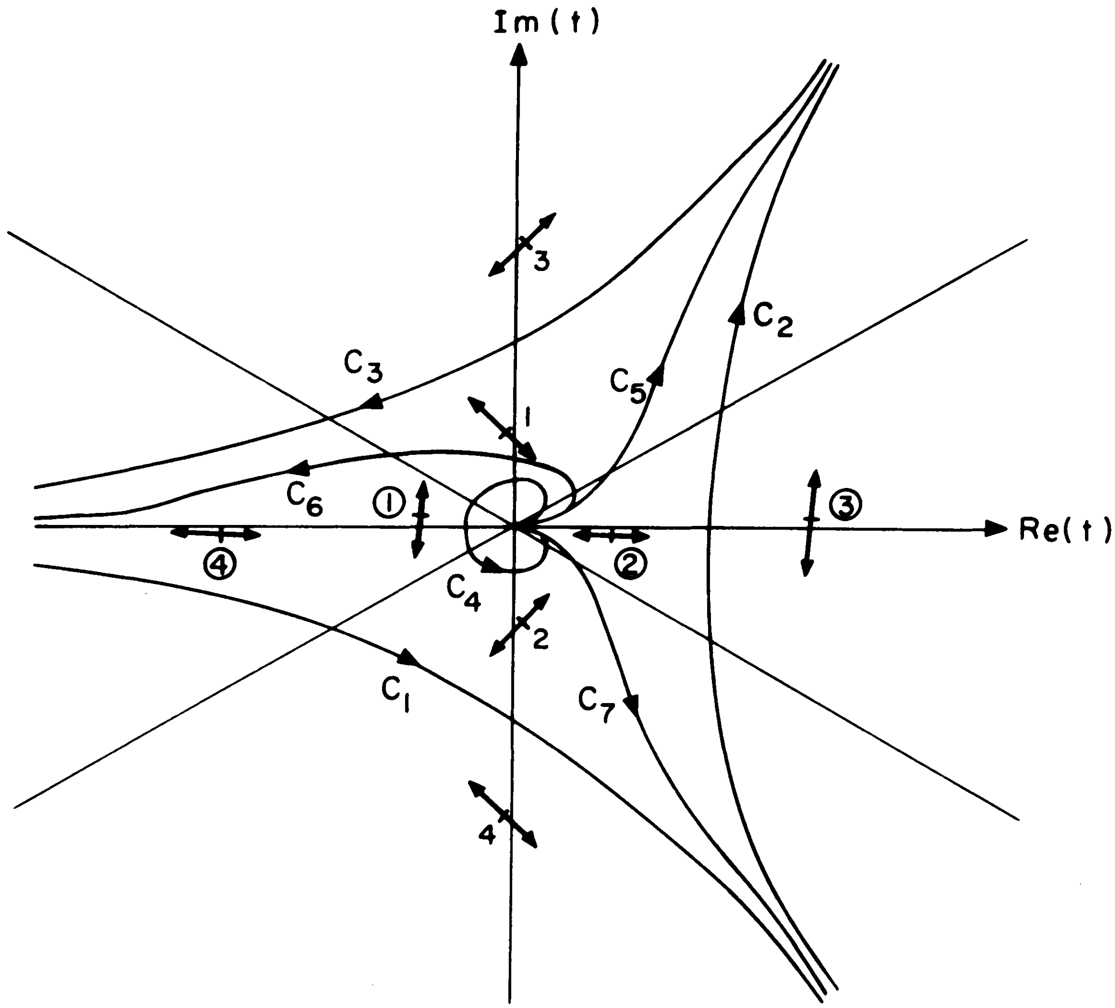
$k_t = 20\ \text{cm}^{-1}$; $M\lambda = 5 \times 10^{-9}\ \text{Oe cm}^2$

$\Delta H_k = 0.36$ Oe

$\Delta H_{crit} = 0.17$ Oe

FIGURE 9

UNDEFORMED CONTOURS USED IN INTEGRAL SOLUTIONS BY THE WKB METHOD AND THE DIRECTION OF STEEPEST DESCENT PATHS THROUGH THE SADDLE POINTS

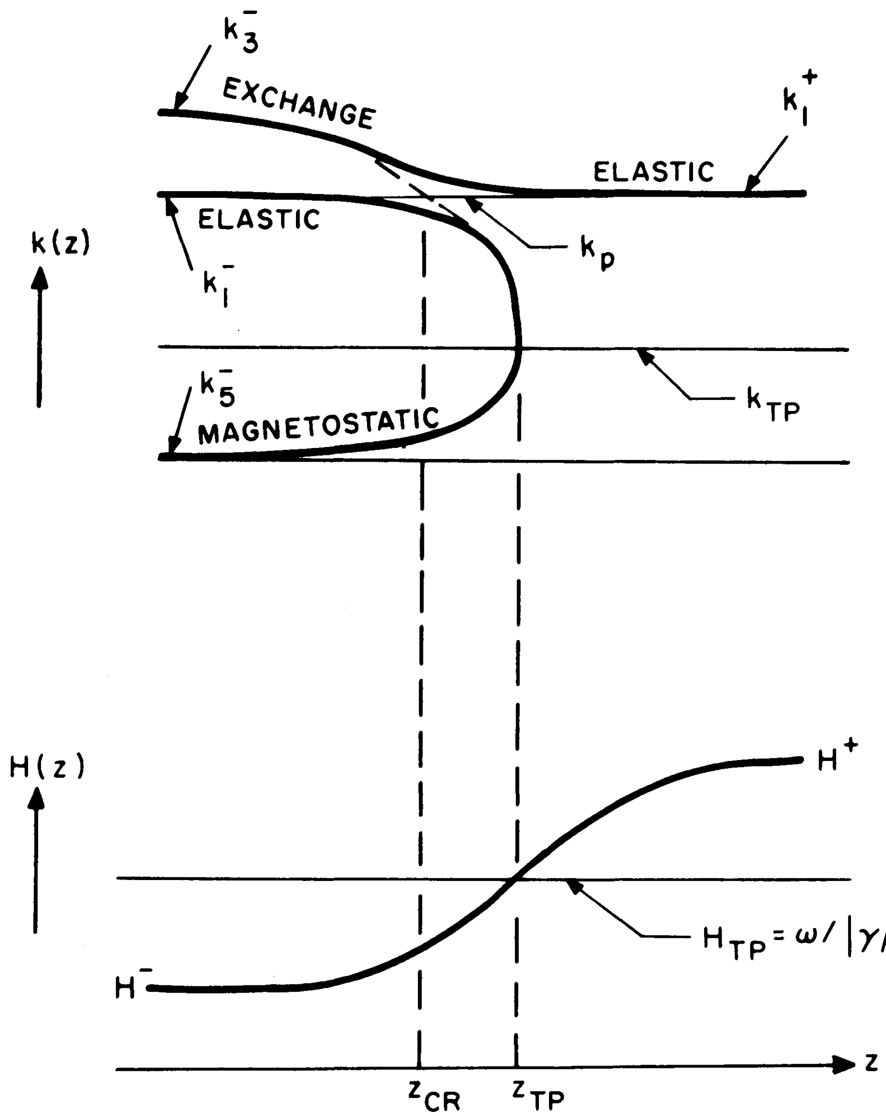


1, 2, 3, 4 : SADDLE POINTS AT $S \rightarrow +\infty$

①, ②, ③, ④ : SADDLE POINTS AT $S \rightarrow -\infty$

FIGURE 10

SCHEMATIC VARIATION OF WAVENUMBERS OF MAGNETOSTATIC, EXCHANGE, AND ELASTIC WAVES FOR THE MAGNETIC FIELD VARIATION SHOWN



z_{CR} : CROSSOVER POINT

z_{TP} : TURNING POINT

FIGURE 11

INDEX OF NOTATION

(Other ad hoc notation is explained in the text near where it occurs.)

a	=	characteristic wavenumber measuring the strength of the magnetoelastic interaction
$A_i(z)$	=	the amplitude of the i^{th} quasi-normal mode
b_2	=	magnetoelastic interaction constant
$B_{ii}(z)$, $b_i(z)$	=	WKB amplitudes for the i^{th} quasi-normal mode
c_{44}	=	elastic stiffness constant for transverse shear
$C_{i\ell}(z)$	=	matrix elements describing coupling from the i^{th} to the ℓ^{th} quasi-normal mode
$D(z)$	=	the matrix in the equation for coupled first-order differential equations
$\underline{E}_x(z)$, $\underline{E}_y(z)$	=	components of electric field for ordinary electromagnetic waves
$\underline{e}_x$, $\underline{e}_y$, $\underline{e}_z$	=	components of the electric field for spin waves
$E_{ii}(z)$	=	the exponential phase integral factors $\exp(-j\int^z k_i d\xi)$
E_{MS}, E_{EX}	=	the first error terms in the asymptotic expansions for "magnetostatic" and "exchange" waves, respectively, in the WKB method
$f(z)$	=	function transforming the dependent variable in differential equations
$F(z)$	=	function satisfying a differential equation which can be transformed to a generalized hypergeometric equation, or which can be solved by the WKB method
$G(\zeta)$	=	function satisfying a generalized hypergeometric equation in the transformed variable ζ
$G_g^-(\zeta)$, $G_h^+(\zeta)$	=	generalized hypergeometric functions identifiable as normal modes at $\zeta \to 0$ $(z \to -\infty)$ and $\zeta \to -\infty$ $(z \to +\infty)$, respectively
$H(z)$	=	the z-component of the internal static magnetic field

H_{TP}	=	the static magnetic field at the turning point
$\underline{H}_x(z), \underline{H}_y(z)$	=	components of magnetic field for ordinary electromagnetic waves
$\underline{h}_x, \underline{h}_y, \underline{h}_z, \underline{h}_r, \underline{h}_\varphi$	=	components of the dynamic magnetic field for spin waves
$H_c' = H_{crit}'$	=	critical gradient of the static magnetic field
j	=	the square root of (-1)
k_i	=	wavenumber for the i^{th} normal mode
$k_i(z)$	=	"local" wavenumber for the i^{th} quasi-normal mode
k_o	=	wavenumber of ordinary electromagnetic waves
$k_{MS}, k_m = k_{EX}, k_p$	=	z-components of the wavenumbers for the unperturbed "magnetostatic" spin wave (extraordinary electromagnetic wave), the exchange-dominated spin wave (magnon), and the elastic wave (phonon), respectively
k_t	=	the transverse wavenumber for bounded samples
k_{TP}	=	z-component of the common wavenumber for spin waves at the turning point
k_z	=	the z-component of a general wavenumber
L	=	effective transition width of a hyperbolic tangent variation
$L_{i\ell}$	=	elements of the linear combination matrix which giagonalizes D
M	=	the saturation magnetization
$\underline{m}_x, \underline{m}_y, \underline{m}_z$	=	components of the dynamic magnetization
M_{gh}	=	matrix element relating the g^{th} normal mode at $z = -\infty$ to the h^{th} normal mode at $z = +\infty$
N_i	=	the i^{th} normal mode
$p(z)$	=	the varying parameter in equations which can be transformed to a generalized hypergeometric differential equation

P_{hg}	=	matrix element arising from circuit relations for generalized hypergeometric functions, relating the h^{th} normal mode at $z = +\infty$ to the g^{th} normal mode at $z = -\infty$
Q_i	=	the i^{th} quasi-normal mode
r_m	=	symmetric sum of m-fold products of $(-\rho_i)$
$\underline{R}_x, \underline{R}_y, \underline{R}_z$	=	components of the dynamic lattice displacement for elastic waves
$R_{i\ell}$	=	reflection factor from wave i to wave ℓ
s	=	normalized distance z in the WKB method
s_m	=	symmetric sum of m-fold products of (σ_i)
S_{EM}, S_m, S_p	=	z-components of electromagnetic, exchange (magnon), and elastic (phonon) power flows, respectively
$T_{i\ell}$	=	transmission factor from wave i to wave ℓ
u(z)	=	normalized dynamic magnetization
U	=	column vector of fields related by coupled first-order differential equations
v(z)	=	normalized lattice displacement
w_m	=	coefficient of $(q - m)^{th}$ derivative term in a q^{th} order differential equation
x	=	as a subscript, a transverse direction
x	=	fraction of power unconverted at magnetoelastic crossover point (section 4.2)
y	=	as a subscript, a transverse direction
y	=	fraction of power reflected due to WKB breakdown for the "exchange" waves (section 4.2)
z	=	the direction of propagation, and direction of any applied static fields

α	=	the parameter in the transformation of the independent variable in transforming to a generalized hypergeometric differential equation
γ	=	the gyromagnetic ratio
δH	=	the difference in magnetic fields, $H_+ - H_-$, evaluated at $z = \pm\infty$
ΔH	=	the full spin-wave linewidth
ΔH_{crit}	=	the critical spin-wave linewidth
ϵ	=	the permittivity
ζ	=	the independent variable in generalized hypergeometric differential equations
η_{pm}	=	phonon/magnon conversion efficiency at the magnetoelastic crossover point
θ	=	$\zeta \frac{d}{d\zeta}$
λ	=	the exchange constant
μ	=	the permeability
ρ_i, σ_i	=	parameters in generalized hypergeometric equations related to the wavenumbers k_i at $z = -\infty$ and $z = +\infty$, respectively
ω	=	the radian frequency of the single frequency excitation
+, -	=	as superscripts and sometimes subscripts, these denote limiting forms at $z = +\infty$ and $z = -\infty$, respectively, as in k_i^+ and k_i^-, H_+ and H_-, μ_+ and μ_-, and z_+ and z_-.
+, -	=	as subscripts, these denote a circularly polarized component, as in $\underline{m}_{\pm} = \underline{m}_x \pm j\underline{m}_y$, or a characteristic of a circularly polarized wave, such as k_+ and k_- and ϵ_+ and ϵ_-
$\nabla_{\pm}$	=	$\frac{\partial}{\partial x} \pm j \frac{\partial}{\partial y}$
∇_t^2	=	transverse Laplacian $\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$
$\langle \ \rangle$	=	time average

BIBLIOGRAPHY: BY AUTHOR

ABRAMOWITZ, M. and STEGUN, I.A. (1966), eds., Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, National Bureau of Standards, Applied Mathematical Series 55, U.S. Government Printing Office, Washington, D.C., p. 9.

ADAIR, J.E. and HADDAD, G.I. (1969), "Coupled-mode Analysis of Nonuniform Coupled Transmission Lines", I.E.E.E. Trans. Microwave Theory and Techniques 17, 746.

ADAM, J.D., COLLINS, J.H., and RUBINO, N. (1968), "Electromagnetic Matching Considerations for YIG Delay Lines", Proc. I.E.E.E. 56, 1373.

ADDISON, R.C., AULD, B.A. and COLLINS, J.H. (1968), "Ray-Theory Analysis of Magnetoelastic Delay Lines", J. Appl. Phys. 39, 1828.

ALTMAN, C. and CORY, H. (1969a), "The Simple Thin-film Optical Method in Electromagnetic Wave Propagation", Radio Sci. 4, 449.

ALTMAN, C. and CORY, H. (1969b), "The Generalized Thin-film Optical Method in Electromagnetic Wave Propagation", Radio Sci. 4, 459.

ARUNASALAM, V., MESERVEY, E.B., GURNEF, M.N. and DAVIDSON, R.C. (1968), "Absorption in a Hot Plasma Near the Second Harmonic of the Electron Cyclotron Frequency", Phys. Fluids 11, 1076.

AULD, B.A. (1963), "Coupling of Electromagnetic and Magnetostatic Waves in Ferrite Waveguides", Symposium on Electromagnetic Theory and Antennas, E.C. Jordan, ed., Pergamon Press, Inc., London, p. 549.

AULD, B.A. (1965), "Geometrical Optics of Magnetoelastic Wave Propagation in a Nonuniform Magnetic Field", Bell Syst. Tech. J. 44, 495.

AULD, B.A., COLLINS, J.H. and WEBB, D.C. (1968), "Excitation of Magnetoelastic Waves in YIG Delay Lines", J. Appl. Phys. 39, 1598.

AZEVEDO, J.C. (1966), "The Oscillations of an Inhomogeneous Plasma", Ph.D. Thesis, Dept. of Physics, M.I.T.

BEKEFI, G. (1966), Radiation Processes in Plasmas, John Wiley & Sons, Inc., New York.

BELLMAN, R.E. (1964), Perturbation Techniques in Mathematics, Physics and Engineering, Holt, Rinehart and Winston, New York.

BERNSTEIN, I.B. (1958), "Waves in a Plasma in a Magnetic Field", Phys. Rev. 109, 10.

BLÖTEKJAER, K. and QUATE, C.F. (1964), "The Coupled Modes of Acoustic Waves and Drifting Carriers in Piezoelectric Crystals", Proc. I.E.E.E. 52, 360.

BREKHOVSKIKH, L.M. (1960), Waves in Layered Media, Academic Press, Inc., New York.

BREMMER, H. (1951), "The WKB Approximation as the First Term of a Geometric-optical Series" in Theory of Electromagnetic Waves, Interscience, New York, p. 125.

BREMMER, H. (1958), "Propagation of Electromagnetic Waves", Handbuch der Physik, vol. 16, Springer-Verlag, Berlin.

BUCHSBAUM, S.J. and HASEGAWA, A. (1966), "Longitudinal Plasma Oscillations near Electron Cyclotron Harmonics", Phys. Rev. 143, 303.

BUDDEN, K.G. (1961), Radio Waves in the Ionosphere, Cambridge Univ. Press, Cambridge.

BURKE, E.R. and BHAGAT, S.M. (1967), "The Propagation of Magnetostatic Spin Waves in Yttrium Iron Garnet", Ph.D. Thesis, Tech. Rep. No. 760, University of Maryland, Dept. of Phys.

BURMAN, R. and GOULD, R.N. (1965), "The Reflection of Waves in a Generalized Epstein Profile", Canad. J. Phys. 43, 921.

BURMAN, R. (1967), "Coupled Wave Equations for Propagation in Inhomogeneous Compressible Plasmas", Proc. Camb. Phil. Soc. 63, 1229.

CARRIER, G.F., KROOK, M. and PEARSON, C.E. (1966), Functions of a Complex Variable: Theory and Technique, McGraw-Hill Book Company, Inc., New York.

CLEMMOW, P.C. and HEADING, J. (1954), "Coupled Forms of the Differential Equations Governing Radio Propagation in the Ionosphere", Proc. Camb. Phil. Soc. 50, 319.

CODDINGTON, E.A. An Introduction to Differential Equations, Prentice-Hall, Inc., Englewood Cliffs, 1961.

CODDINGTON, E.A. and LEVINSON, N. (1955), Theory of Ordinary Differential Equations, McGraw-Hill Book Company, Inc., New York.

COLLINS, J.H. (1967), "Simultaneous Propagation of Magnetostatic and Spin-Elastic Waves in Nonellipsoidal Ferrimagnets", Electronics Lett. 3, 59.

COLLINS, J.H. and WILSON, D.A. (1968), "Optical Probing of Magnetostatic Modes in YIG Delay Lines", Appl. Phys. Lett. 12, 331.

COMSTOCK, R.L. (1965), "Wave Number Dependence of Spin Wave Relaxation in YIG", Appl. Phys. Lett. 6, 29.

COOK, J.S. (1955), "Tapered Velocity Couplers", Bell Syst. Tech. J. 34, 807.

COPSON, E.T. (1935), Theory of Functions of a Complex Variable, Oxford Univ. Press, Oxford.

DESORMIERE, B. and LEGALL, H. (1969), "Magnetostatic Mode Excitation in YIG Rods Containing a Turning Point", J. Appl. Phys. 40, 1191.

DILLON, J.F., JR. (1968), "Origin and Uses of the Faraday Rotation in Magnetic Crystals", J. Appl. Phys. 39, 922.

DILLON, J.F., JR., REMEIKA, J.R. and STATON, C.R. (1969), "Linear Magnetic Birefringence in Cubic Magnetic Crystals", J. Appl. Phys. 40, 1510.

ECKERSLEY, T.L. (1950), "Coupling of the Ordinary and Extraordinary Rays in the Ionosphere", Proc. Phys. Soc. B 63, 49.

EPSTEIN, P.S. (1930a), "Geometrical Optics in Absorbing Media", Proc. Nat. Acad. Sci. 16, 37.

EPSTEIN, P.S. (1930b), "Reflection of Waves in an Inhomogeneous Absorbing Medium", Proc. Nat. Acad. Sci. 16, 627.

ERDELYI, A., MAGNUS, W., OBERHETTINGER, F. and TRICOMI, F.G. (1953), Higher Transcendental Functions, Vols. I,II,III, McGraw-Hill Book Company, Inc., New York.

FELSEN, L.B. (1969), "Transients in Dispersive Media, Part I: Theory", I.E.E.E. Trans. on Antennas and Propagation 17, 191.

FISHER, R.K. and GOULD, R.W. (1969), "Resonance Cones in the Field Pattern of a Short Antenna in an Anisotropic Plasma", Phys. Rev. Lett. 22, 1093.

FLETCHER, P.C. and KITTEL, C. (1960), "Considerations on the Propagation and Generation of Magnetostatic Waves and Spin Waves", Phys. Rev. 120, 2004.

FOX, A.G. (1955), "Wave Coupling by Warped Normal Modes", Bell Syst. Tech. J. 34, 823.

FREDRICKS, R.W. and SCARF, F.L. (1969), "Electromagnetic and Electrostatic Bernstein Modes in a Warm Collisionless Maxwellian Plasma", Plasma Waves in Space and the Laboratory, Univ. of Edinburgh Press, Edinburgh.

FRIEDRICHS, K.O. (1965), Advanced Ordinary Differential Equations, Gordon & Breach Science Publishers, New York.

GANTMACHER, F.R. (1959), The Theory of Matrices, Vols. I,II, Chelsea Publishing Co., New York.

GINZBURG, V.L. (1961), Propagation of Electromagnetic Waves in Plasma, Gordon & Breach Science Publishers, New York.

GORMAN, D. (1966), "Plasma Waves at Singular Turning Points", Phys. Fluids 9, 1262.

GRUBER, S.and BEKEFI, G. (1968), "Excitation of Longitudinal Waves near Electron-Cyclotron Harmonics", Phys. Fluids 11, 122.

HADDAD, G.I. and ADAIR, J.E. (1965), "Coupled-Mode Description of Space-Charge Waves on Nonuniform Beams", I.E.E.E. Trans. Electron Devices 12, 536.

HADDAD, G.I. and BEVENSEE, R.M. (1963), "Start-Oscillation Conditions of Tapered Backward-Wave Oscillators", I.E.E.E. Trans. Electron Devices 10, 389.

HAUS, H.A. (April 8, 1958), "Electron Beam Waves in Microwave Tubes", M.I.T.-R.L.E. Tech. Rep. #316.

HEADING, J. (1955), "The Reflexion of Vertically-Incident Long Radio Waves from the Ionosphere when the Earth's Magnetic Field is Oblique", Proc. Roy. Soc.(London) A231, 414.

HEADING, J. (1961), "The Nonsingular Embedding of Transition Processes within a More General Framework of Coupled Variables", J. Res. N.B.S. 65D (Radio Propagation), 595.

HEADING, J. (1962), An Introduction to Phase Integral Methods, Methuen, London.

HEADING, J. (1963), "Composition of Reflection and Transmission Formulae", J. Res. N.B.S. 67D (Radio Propagation), 65.

HEADING, J. (1965), "Refractive Index Profiles Based on the Hypergeometric Equation and the Confluent Hypergeometric Equation", Proc. Camb. Phil. Soc. 61, 897.

HEADING, J. (1967a), "Investigations into a New Stratified Hyperbolic Profile", Proc. Camb. Phil. Soc. 63, 439.

HEADING, J. (1967b), "Coupled Electroacoustic Waves", Radio Sci. 2, 1401.

HEADING, J. (1968), "Differential and Integral Equations Arising at Coupling Points in Propagation Theory", J. Diff. Eq. 4, 519.

HEADING, J. (1969a), "Propagation in an Inhomogeneous Gyrational Medium", Quart. J. of Mech and Appl. Math. 22, 75.

HEADING, J. (1969b), "Polarization of Obliquely Reflected Waves from an Isotropic Plane-stratified Plasma", Radio Sci. 4, 441.

HEADING, J. and WHIPPLE, R.T.P. (1952), "The Oblique Reflexion of Long Wireless Waves from the Ionosphere at Places where the Earth's Magnetic Field is Regarded as Vertical", Phil. Trans. Roy. Soc.(London) A244, 469.

HILDEBRAND, F.B. (1965), Methods of Applied Mathematics, 2nd Ed., Prentice-Hall, Inc., Englewood Cliffs, N.J.

HOLT, A.G.J. and AHMED, K.U. (1968), "Generalized Distributions of Nonuniformity of Hypergeometric Transmission Line", Electronics Lett. 4, 167.

HORTON, C.W., JR. (1966), "Coupling of Transverse and Longitudinal Waves at the Upper Hybrid Frequency", Phys. Fluids 9, 815.

HOUGARDY, R.W. and SAXON, D.S. (1963), "Propagation of Electromagnetic Waves in an Anisotropic Stratified Medium", Symposium on Electromagnetic Theory and Antennas, E.C. Jordan, Ed., Pergamon Press, Inc., New York, p. 635.

HU, H.L., REZENDE, S.M. and MORGENTHALER, F.R. (1970), "Measurements of Longitudinal Elastic-Wave/Spin-Wave Conversion Efficiencies in an Axially Magnetized YIG Rod", J. Appl. Phys. 41, 1417.

JACKSON, J.D. (1963), Classical Electrodynamics, John Wiley & Sons, Inc., New York.

JOHNSON, C.C. (1965), Field and Wave Electrodynamics, McGraw-Hill Book Company, Inc., New York.

JOHNSON, C.C. (1968), "Coupled-mode Description of a Nonuniform Transmission Line", Proc. I.E.E.E. 56, 1209.

JOSEPH, R.I. and SCHLÖMANN, E. (1965), "Demagnetizing Field in Nonellipsoidal Bodies", J. Appl. Phys. 36, 1579.

KARBOWIAK, A.E. (1967), "Wave Propagation in Inhomogeneous Media", Electromagnetic Wave Theory, J. Brown, ed., Pergamon Press, Inc., New York, p. 1.

KEDZIE, R.W. (1966), "New Mode of Magnetostatic-Wave Propagation in Yttrium Iron Garnet Rods", Bull. Am. Phys. Soc. 11, 221.

KEDZIE, R.W. (1968), "Magnetostatic Mode Propagation in Axially Magnetized YIG Rods Containing a Turning Point", J. Appl. Phys. 39, 2731.

KELLER, H.G. and KELLER, J.B. (1962), "Exponential-like Solutions of Systems of Linear Ordinary Differential Equations", J. Soc. Indust. Appl. Math. 10, 246.

KIRCHNER, E.K., DONAGHY, L.F., OLSON, F.A. and AULD, B.A. (1966), "Analysis of Magnetoelastic Conversion Efficiency", J. Appl. Phys. 37, 988.

KITTEL, C. (1958a), "Interaction of Spin Waves and Ultrasonic Waves in Ferromagnetic Crystals", Phys. Rev. 110, 836.

KITTEL, C. (1958b), "Excitation of Spin Waves in a Ferromagnet by a Uniform RF Field", Phys. Rev. 110, 1295.

KOHANE, T., SCHLÖMANN, and JOSEPH, R.I. (1965), "Microwave Magnetoelastic Resonances in a Nonuniform Magnetic Field", J. Appl. Phys. 36, 1267.

KOPECKY, V. and PREINHAELTER, J. (1969), "Transformation and Absorption of Waves in an Inhomogeneous Beam-Plasma System", Plasma Phys. 11, 333.

KOPECKY, V., PREINHALTER, J. and VACLAVIK, J. (1969), "Transformation of Waves and Electron Heating in a Radially Inhomogeneous Plasma", J. Plasma Phys. 3, 179.

KORN, G.A. and KORN, T.M. (1961), Mathematical Handbook for Scientists and Engineers, McGraw-Hill Book Company, Inc., New York.

KUCKES, A.F. (1968), "Resonant Absorption of Electromagnetic Waves in a Non-uniformly Magnetized Plasma", Plasma Phys. 10, 367.

KUCKES, A.F. and DAWSON, J.M. (1965), "Electron Cyclotron Harmonic Radiation from a Plasma", Phys. Fluids 8, 1007.

KUEHL, H.H. (1967), "Coupling of Transverse and Longitudinal Waves Below the Second Electron Cyclotron Harmonic", Phys. Rev. 154, 124.

LANDAU, L.D. and LIFSHITZ, E.M. (1958), Quantum Mechanics - Non-relativistic Theory, Addison-Wesley, Reading, Mass.

LANGER, R.E. (1957), "On the Asymptotic Solutions of a Class of Ordinary Differential Equations of the Fourth Order with a Special Reference to an Equation of Hydrodynamics", Trans. Amer. Math. Soc. 84, 144.

LANGER, R.E. (1959a), "The Asymptotic Solutions of a Linear Differential Equation of the Second Order with Two Turning Points", Trans. Am. Math. Soc. 90, 113.

LANGER, R.E. (1959b), "Formal Solutions and a Related Equation for a Class of Fourth Order Differential Equations of a Hydrodynamic Type", Trans. Amer. Math. Soc. 92, 371.

LEWIS, M.F. and LACKLISON, D.E. (1966), "A New Mode in Axially Magnetized YIG Rods", Appl. Phys. Lett. 9, 414.

LEWIS, M.F. and SCOTTER, D.G. (1969), "Magnetostatic Waves in Short, Axially Magnetized YIG Rods", J. Appl. Phys. 40, 3618.

LIGHTHILL, M.J. (1960), "Studies on Magneto-hydrodynamic Waves and other Anisotropic Wave Motions", Phil. Trans. Roy. Soc.(London) ser. A 252, 397.

LIN, C.C. and RABENSTEIN, A.L. (1960), "On the Asymptotic Solutions of a Class of Ordinary Differential Equations of the Fourth Order. I. Existence of Regular Formal Solutions", Trans. Amer. Math. Soc. 94, 24.

LONNGREN, K.E., SJÖLUND, A. and WEISSGLAS, P. (1966), "Coupling of Waves in a Hot Inhomogeneous Anisotropic Electron Plasma", Plasma Phys. (J. Nucl. Energy Part C) 8, 657.

LOUISELL, W.H. (1955), "Analysis of the Single Tapered Mode Coupler", Bell Syst. Tech. J. 34, 853.

LOUISELL, W.H. (1960), Coupled Modes and Parametric Electronics, John Wiley & Sons, Inc., New York.

LUKE, Y.L. (1969), The Special Functions and Their Approximations, 2 vols., Academic Press, New York.

MASON, W.P. (1966), Crystal Physics of Interaction Processes, Academic Press, Inc., New York.

MATTHEWS, H. and LE CRAW, R.C. (1962), "Acoustic Wave Rotation by Magnon-Phonon Interaction", Phys. Rev. Lett. 8, 397.

MILLINGTON, G. (1969), "Stokes Phenomenon", Radio Sci. 4, 95.

MORGENTHALER, F.R. (1967), "Photon-Magnon Conversion Near a Material Interface", Electronics Lett. 3, 299.

MORGENTHALER, F.R., REZENDE, S.M., HU, H.L. and PLATZKER, A. (1970), "Spin-longitudinal and Spin-shear Elastic Wave Conversion in an Axially Magnetized YIG Rod", Appl. Phys. Lett. 16, 133.

NUTTALL, J. (1965), "Singularities of the Green's Function for a Collisionless Magnetoplasma", Phys. Fluids 8, 286.

PEARLSTEIN, L.D. (1965), "Remarks on an Exact Solution of a Universal Instability", Phys. Fluids 8, 1743.

PEARLSTEIN, L.D. and BHADRA, D. (1969), "Dispersion of the Upper Hybrid Mode in a Spatially Inhomogeneous Plasma", Phys. Fluids 12, 213.

PEARSON, G.A. (1966), "Wave Propagation near Electron Cyclotron Harmonics in Nonuniform Plasmas", Phys. Fluids 9, 2454.

PRICE, G.H. (1968), "The Characteristics of Impulses Reflected from a sech^2 Electron-density Profile and Bandpass Filtered", Proc. I.E.E.E. 56, 1341.

RABENSTEIN, A.L. (1958), "Asymptotic Solutions of $u^{iv} + \lambda^2(zu'' + \alpha u' + \beta u) = 0$ for Large $|\lambda|$ ", Arch. Rational Mech. and Anal. 1, 418.

RADIO SCIENCE (1966), September issue.

RADO, G.T. and WEERTMAN, J.F. (1959), "Spin Wave Resonance in a Ferromagnetic Metal", J. Phys. Chem. Solids 11, 315.

RAMO, S., WHINNERY, J.R. and VAN DUZER, T. (1965), Fields and Waves in Communication Electronics, John Wiley & Sons, Inc., New York.

RAWER, K. (1939), "Elektrische Wellen in einem Geschichteten Medium", Annalen der Physik, 5 Folge., 35, 385.

REM, J., SLUIJTER, F.W. and WEENINK, M.P.H. (1967), "Waves in an Axially Symmetric Inhomogeneous Warm Plasma with Frequencies near the Second Harmonic of the Electron Cyclotron Frequency", Physica 37, 393.

REZENDE, S.M., and MORGENTHALER, F.R. (1969a), "Magnetoelastic Waves in Time-Varying Magnetic Fields. I. Theory", J. Appl. Phys. 40, 524.

RYDBECK, O.E.H. (1960), "The Propagation and Coupling of Waves in Inhomogeneous Media with Applications to Electromagnetic and Space Charge Waves", Res. Rep. No. 7, Res. Lab. of Electronics, Chalmers Univ. of Tech., Sweden.

RYDBECK, O.E.H. (1967), "Wave Coupling in Multi-mode Inhomogeneous Ionized Media", Electromagnetic Wave Theory, J. Brown, ed., Pergamon Press, Inc., London, p. 45.

SCHELKUNOFF, S.A. (1955), "Conversion of Maxwell's Equations into Generalized Telegraphist's Equations", Bell Syst. Tech. J. 34, 995.

SCHLÖMANN, E. (1960), "Generation of Phonons in High-Power Ferromagnetic Resonance Experiments", J. Appl. Phys. 31, 1647.

SCHLÖMANN, E. (1961), "Spin-Wave Spectroscopy", Adv. in Quantum Electronics, Columbia Univ. Press, New York, p. 437.

SCHLÖMANN, E. (1964), "Generation of Spin-Waves in Nonuniform Magnetic Fields I. Conversion of Electromagnetic Power into Spin-Wave Power and Vice Versa", J. Appl. Phys. 35, 159.

SCHLÖMANN, E. and JOSEPH, R.I. (1964), "Generation of Spin Waves in Nonuniform Magnetic Fields. III. Magnetoelastic Interaction", J. Appl. Phys. 35, 2382; Erratum: J. Appl. Phys. 36, 875(1965).

SESHADRI, S.R. (1965), "Wave Propagation in a Two-component Warm Plasma", J. Res. N.B.S. 69D (Radio Science), 579.

SLUIJTER, F.W. (1967), "Wave Propagation through an Epstein Profile across a Static Magnetic Field", Electromagnetic Wave Theory, J. Brown, ed., Pergamon Press, Inc., London.

SNOW, C. (1952),"Hypergeometric and Legendre Functions with Applications to Integral Equations of Potential Theory", N.B.S. Applied Mathematics Series 19, chapter 7.

SOMMERFELD, A. (1964), Electrodynamics, Academic Press, Inc. New York.

SORBELLO, R.S. (1965), "Wave Propagation in Dispersive Anisotropic Media", S.M. Thesis, M.I.T., Dept. of Elec. Eng.

STIX, T.H. (1960), "Absorption of Plasma Waves", Phys. Fluids 3, 19.

STIX, T.H. (1962), Theory of Plasma Waves, McGraw-Hill Book Company, Inc., New York.

STIX, T.H. (1965a), "Coupling of Bernstein Modes to the Radiation Field in an Inhomogeneous Plasma", Bull. Am. Phys. Soc. 10, 230.

STIX, T.H. (1965b), "Radiation and Absorption via Mode Conversion in an Inhomogeneous Collision-Free Plasma", Phys. Rev. Lett. 15, 878.

THOMPSON, G.H.B. (1963), "Backward Waves in Longitudinally Magnetized Ferrite Filled Guides", Symposium on Electromagnetic Theory and Antennas, E.C. Jordan, ed., Pergamon Press, Inc., London, p. 603.

TIMOFEEV, A.V. (1968), "Resonance Interaction of Flute Oscillations with Ion Larmor Drift", Nucl. Fusion 8, 99.

VASILE, C.F. (1967), "Simplifying Maxwell's Equations in Gyrotropic Media", I.E.E.E. Trans. Microwave Theory and Techniques 15, 324.

VASILE, C.F. and LA ROSA, R. (1968a), "Guided Wave Propagation in Gyromagnetic Media as Applied to the Theory of Exchange Spin-Wave Excitation", J. Appl. Phys. 39, 1863.

VASILE, C.F. and LA ROSA, R. (1968b), "The Character of Modes in Small Axially Magnetized Ferrite-Filled Waveguides", J. Appl. Phys. 39, 2380.

VASILE, C.F. and LA ROSA, R. (1968c). "Method of Testing the Quality of Bars of Single Crystal Yttrium Iron Garnet", Electronics Lett. 4, 400.

VASILE, C.F. and LA ROSA, R. (1969), "Two-Port Magnetoelastic Delay Line Utilizing Off-Axis Propagation in Single Crystal Yttrium Iron Garnet", I.E.E.E. Trans. Sonics and Ultrasonics 16, 22.

VASILE, C.F. and FELSEN, L.B. (to be published), "Radiation from a Localized Source in a Gyromagnetic Medium".

WALDRON, R.A. (1965), "The Theory of Coupled Modes", Quart. J. Mech. and Appl. Math. 18, 385.

WALDRON, R.A. (1966), "The Theory of Reflections in a Tapered Waveguide", Radio and Electronic Engr. 32, 245.

WALDRON, R.A. (1967), "On the General Solution of the Coupled Mode Problem", Proc. I.E.E.E. 55, 2029.

WASOW, W. (1950), "A Study of the Solutions of the Differential Equation $y^{(4)} + \lambda^2(xy'' + y) = 0$ for Large Values of λ ", Annals of Mathematics 52, 350.

WASOW, W. (1965), Asymptotic Expansions for Ordinary Differential Equations, Interscience Publishers, New York.

WESTCOTT, B.S. (1968), "Exact Solutions for Electromagnetic Wave Propagation in Spherically Stratified Isotropic Media", Proc. Camb. Phil. Soc. 64, 227.

WESTCOTT, B.S. (1969a), "Exact Solutions for Vertically Polarized Electromagnetic Waves in Horizontally Stratified Isotropic Media", Proc. Camb. Phil. Soc. 66, 675.

WESTCOTT, B.S. (1969b), "Exact Solutions for Transverse Electromagnetic Wave Propagation in Cylindrically Stratified Axially Magnetized Plasma", Proc. Camb. Phil. Soc. 66, 129.

WESTCOTT, B.S. (1969c), "Generalized Confluent Hypergeometric and Hypergeometric Transmission Lines", I.E.E.E. Trans. Circuit Theory 16, 289.

WILCOX, C.H., ed., Asymptotic Solutions of Differential Equations and their Applications , John Wiley & Sons, Inc., New York, 1964.

BIBLIOGRAPHY: BY SUBJECT

This listing is not intended to be complete. The emphasis in the listing is on representative and/or recent references. Consult the listing by author for the complete information on each reference.

General Mathematical References:

ABRAMOWITZ, M. and STEGUN, I.A. (1966)
BELLMAN, R.E. (1964)
CARRIER, G.F., KROOK, M. and PEARSON, C.E. (1966)
CODDINGTON, E.A. (1961)
CODDINGTON, E.A. and LEVINSON, N. (1955)
COPSON, E.T. (1935)
ERDELYI, A., MAGNUS, W., OBERHETTINGER, F. and TRICOMI, F.G. (1953)
FRIEDRICHS, K.O. (1965)
GANTMACHER, F.R. (1959)
HILDEBRAND, F.B. (1965)
JOSEPH, R.I. and SCHLÖMANN, E. (1965)
KORN, G.A. and KORN, T.M. (1961).
WASOW, W. (1965)
WILCOX, C.H. (1964)

General Physics References:

BEKEFI, G. (1966)
BREKHOVSKIKH, L.M. (1960)
BREMMER, H. (1958)
BUDDEN, K.G. (1961)
GINZBURG, V.L. (1961)
JACKSON, J.D. (1963)
JOHNSON, C.C. (1965)
KARBOWIAK, A.E. (1967)
LANDAU, L.D. and LIFSHITZ, E.M. (1958)
MASON, W.P. (1966)
RAMO, S., WHINNERY, J.R. and VAN DUZER, T. (1965)
SOMMERFELD, A. (1964)

Coupled Waves: General and Miscellaneous:

ADAIR, J.E. and HADDAD, G.I. (1969)
BLOTEKJAER, K. and QUATE, C.F. (1964)
COOK, J.S. (1955)
DILLON, J.F., JR. (1968)
DILLON, J.F., JR., REMEIKA, J.R. and STATON, C.R. (1969)
FOX, A.G. (1955)
HADDAD, G.I. and ADAIR, J.E. (1965)
HAUS, H.A. (1958)
HEADING, J. (1961)
HEADING, J. (1969a)
HOUGARDY, R.W. and SAXON, D.S. (1963)
JOHNSON, C.C. (1968)
LOUISELL, W.H. (1960)
MATTHEWS, H. and LECRAW, R.C. (1962)

RYDBECK, O.E.H. (1960)
SCHELKUNOFF, S.A. (1955)
SLUIJTER, F.W. (1967)
WALDRON, R.A. (1965)
WALDRON, R.A. (1966)
WALDRON, R.A. (1967)

WKB Quasi-normal Modes:
BREMMER, H. (1951)
CLEMMOW, P.C. and HEADING, J. (1954)
HEADING, J. (1967b)
HEADING, J. (1968)
KELLER, H.G. and KELLER, J.B. (1962)
LOUISELL, W.H. (1955)
PEARLSTEIN, L.D. and BHADRA, D. (1969)
RYDBECK, O.E.H. (1960)
RYDBECK, O.E.H. (1967)
VASILE, C.F. and LA ROSA, R. (1968a)

Successive Approximations and Phase Integral Methods, and the General Extraction of Approximating Equations:
BURMAN, R. (1967)
ECKERSLEY, T.L. (1950)
HADDAD, G.I. and BEVENSEE, R.M. (1963)
HEADING, J. (1961)
HEADING, J. (1962)
HEADING, J. (1963)
HEADING, J. (1968)
HORTON, C.W., JR. (1966)
HOUGARDY, R.W. and SAXON, D.S. (1963)
LOUISELL, W.H. (1955)
RYDBÉCK, O.E.H. (1967)
SCHLÖMANN, E. and JOSEPH, R.I. (1964)

WKB Method:
AZEVEDO, J.C. (1966)
GORMAN, D. (1966)
HEADING, J. (1962)
HEADING, J. (1967)
HEADING, J. (1968)
KOPECKY, V. and PREINHAELTER, J. (1969)
KOPECKY, V., PREINHALTER, J. and VACLAVIK, J. (1969)
KUEHL, H.H. (1967a)
LANDAU, L.D. and LIFSHITZ, E.M. (1958)
LANGER, R.E. (1957)
LANGER, R.E. (1959a)
LANGER, R.E. (1959b)
LIN, C.C. and RABENSTEIN, A.L. (1960)
MILLINGTON, G. (1969)
PEARLSTEIN, L.D. and BHADRA, D. (1969)
RABENSTEIN, A.L. (1958)
STIX, T.H. (1962)
STIX, T.H. (1965a)
STIX, T.H. (1965b)

TIMOFEEV, A.V. (1968)
WASOW, W. (1950)
WASOW, W. (1965)
WILCOX, C.H. (1964)

Hypergeometric Functions, Their Generalizations and Applications:

BURMAN, R. and GOULD, R.N. (1965)
EPSTEIN, P.S. (1930b)
ERDELYI, A., MAGNUS, W., OBERHETTINGER, F. and TRICOMI, F.G. (1953)
HEADING, J. (1955)
HEADING, J. (1967a)
HEADING, J. (1965)
HEADING, J. (1969b)
HEADING, J. and WHIPPLE, R.T.P. (1952)
HOLT, A.G.J. and AHMED, K.U. (1968)
LUKE, Y.L. (1969)
PEARLSTEIN, L.D. (1965)
PRICE, G.H. (1968)
RAWER, K. (1939)
SLUIJTER, F.W. (1967)
SNOW, C. (1952)
WESTCOTT, B.S. (1968)
WESTCOTT, B.S. (1969a)
WESTCOTT, B.S. (1969b)
WESTCOTT, B.S. (1969c)

"Magnetostatic" Spin Waves (Extraordinary Electromagnetic Waves): General:

AULD, B.A. (1963)
BURKE, E.R. and BHAGAT, S.M. (1967)
COLLINS, J.H. and WILSON, D.A. (1968)
FLETCHER, P.C. and KITTEL, C. (1960)
LEWIS, M.F. and SCOTTER, D.G. (1969)
THOMPSON, G.H.B. (1963)
VASILE, C.F. (1967)
VASILE, C.F. and LA ROSA, R. (1968b)
VASILE, C.F. and LA ROSA, R. (1968c)

"Exchange" Spin Waves: General:

COMSTOCK, R.L. (1965)
KITTEL, C. (1958b)
RADO, G.T. and WEERTMAN, J.F. (1959)
VASILE, C.F. and LA ROSA, R. (1968c)

Spin-elastic Wave Conversion and Magnetoelastic Coupling:

ADDISON, R.C., AULD, B.A. and COLLINS, J.H. (1968)
AULD, B.A., COLLINS, J.H. and WEBB, D.C. (1968)
HU, H.L., REZENDE, S.M. and MORGENTHALER, F.R. (1970)
KIRCHNER, E.K., DONAGHY, L.F., OLSON, F.A. and AULD, B.A. (1966)
KITTEL, C. (1958a)
KOHANE, T., SCHLOMANN, E. and JOSEPH, R.I. (1965)

MORGENTHALER, F.R., REZENDE, S.M., HU, H.L. and PLATZKER, A. (1970)
REZENDE, S.M. and MORGENTHALER, F.R. (1969a)
SCHLÖMANN, E. (1960)
SCHLOMANN, E. and JOSEPH, R.I. (1964)

Conversions and Reflections of "Magnetostatic" (medium-k) and "Exchange" (high-k) Spin Waves at a Turning Point:
AULD, B.A., COLLINS, J.H. and WEBB, D.C. (1968)
COLLINS, J.H. (1967)
COLLINS, J.H. and WILSON, D.A. (1968)
DESORMIERE, B. and LEGALL, H. (1969)
KEDZIE, R.W. (1966)
KEDZIE, R.W. (1968)
KOHANE, T., SCHLÖMANN, E. and JOSEPH, R.I. (1965)
LEWIS, M.F. and LACKLISON, D.E. (1966)
LEWIS, M.F. and SCOTTER, D.G. (1969)
VASILE, C.F. and LA ROSA, R. (1968a)
VASILE, C.F. and LA ROSA, R. (1969)

Kinetic or Longitudinal (high-k) Waves in Ionized Media and Conversion from Extraordinary Electromagnetic or other Medium-k Waves:
ARUNASALAM, V., MERSERVEY, E.B., GURNEF, M.N. and DAVIDSON, R.C. (1968)
AZEVEDO, J.C. (1966)
BERNSTEIN, I.B. (1958)
BUCHSBAUM, S.J. and HASEGAWA, A. (1966)
BURMAN, R. (1967)
FREDRICKS, R.W. and SCARF, F.L. (1969)
GINZBURG, V.L. (1961)
GORMAN, D. (1966)
GRUBER, S. and BEKEFI, G. (1968)
HEADING, J. (1967b)
HORTON, C.W., JR. (1966)
KOPECKY, V. and PREINHAELTER, J. (1969)
KOPECKY, V., PREINHALTER, J. and VACLAVIK, J. (1969)
KUCKES, A.F. (1968)
KUCKES, A.F. and DAWSON, J.M. (1965)
KUEHL, H.H. (1967a)
LONNGREN, K. E., SJÖLUND, A., and WEISSGLAS, P. (1966)
PEARLSTEIN, L.D. and BHADRA, D. (1969)
PEARSON, G.A. (1966)
REM, J., SLUIJTER, F.W. and WEENINK, M.P.H. (1967)
RYDBECK, O.E.H. (1967)
SESHADRI, S.R. (1965)
STIX, T.H. (1960)
STIX, T.H. (1962)
STIX, T.H. (1965a)
STIX, T.H. (1965b)

Coupling to Waves within Materials or Ionized Media from External Electromagnetic Fields:
AULD, B.A. (1963)
ADAM, J.D., COLLINS, J.H. and RUBINO, N. (1968)

AULD, B.A., COLLINS, J.H. and WEBB, D.C. (1968)
BURKE, E.R. and BHAGAT, S.M. (1967)
DESORMIERE, B. and LEGALL, H. (1969)
FISHER, R.K. and GOULD, R.W. (1969)
FLETCHER, P.C. and KITTEL, C. (1960)
KITTEL, C. (1958b)
MORGENTHALER, F.R. (1967)
RADIO SCIENCE (1966)
SCHLÖMANN, E. (1961)
SCHLÖMANN, E. (1964)
VASILE, C.F. and FELSEN, L.B. (to be published)

Treatments of Transverse Variations by Ray Theory:
ADDISON, R.C., AULD, B.A. and COLLINS, J.H. (1968)
AULD, B.A. (1965)
EPSTEIN, P.S. (1930a)

Propagation of Pulses and Transients: General (not Including Calculations of Delay Time by Group Velocity Integration):
FELSEN, L.B. (1969)
LIGHTHILL, M.J. (1960)
NUTTALL, J. (1965)
PRICE, G.H. (1968)
SORBELLO, R.S. (1965)

Numerical Methods:
ALTMAN, C. and CORY, H. (1969a)
ALTMAN, C. and CORY, H. (1969b)
KIRCHNER, E.K., DONAGHY, L.F., OLSON, F.A. and AULD, B.A. (1966)

BIOGRAPHICAL NOTE

The author was born in New York City on June 20, 1942. After attending public schools in Stamford, Connecticut, he graduated from Stamford High School in June 1960. While attending Yale University in New Haven, Connecticut he resided at Jonathan Edwards College and entered the Department of Electrical Engineering of the School of Engineering. In June 1964, he was awarded the degree of Bachelor of Engineering with Highest Honors.

In September 1964 he entered the Department of Electrical Engineering at M. I. T. as a Hertz Fellow. He joined the Noise in Electron Devices group of the Research Laboratory of Electronics and received the degrees of Master of Science and Electrical Engineer in September 1965 and June 1966, respectively. After employment by the Raytheon Research Division in Waltham, Massachusetts, in the summer of 1966, he joined the Microwave and Quantum Magnetics Group of the Center for Materials Science and Engineering at M. I. T.

The author is a member of Tau Beta Pi and the Institute of Electrical and Electronics Engineers.