J. Plasma Physics (1997), *vol.* 58, *part* 2, *pp.* 259–275. Printed in the United Kingdom
© 1997 Cambridge University Press

Mode coupling in space- and time-varying anisotropic absorbing plasmas

B. Sh. SABZEVARI†
Institute for Studies in Theoretical Physics and Mathematics PO Box 19395-5531, Tehran, Iran

(Received 29 June 1996)

A four-dimensional systematic mathematical approach for investigating propagation and coupling of wave modes in a slowly varying (in all space directions and time) anisotropic, absorbing plasma is represented. The formalism is especially useful for energy considerations of the waves. It is applicable to general cases of mode couplings in plasmas with general geometries of space–time and magnetic field configurations. A simple example of how this formalism can be applied to practical cases is given.

1. Introduction

Almost all investigations of wave coupling in plasmas have been for cases where the inhomogeneity of the medium is only in one direction, i.e. plane-stratified plasmas (Stix and Swanson 1983; Budden 1985; Swanson 1989; Stix 1992). The propagation of electromagnetic waves in four-dimensional space- and time-varying plasmas was studied by Bernstein and Friedland (1983) and Kravtsov (1969) in the context of geometrical optics. However, in these studies the coupling of modes was not considered, and attention was concentrated on other themes, such as energy relations and the absorption of independent waves. There exist papers on four-dimensional mode conversion (Friedland *et al.* 1987; Kaufman and Friedland 1987), generalizing previous work by Cairns and Lashmore-Davies (1983) on mode conversion in one-dimensionally varying plasmas. However, mode conversion is only a special case of mode coupling where two isolated coupling points begin to coalesce (Budden and Smith 1974; Budden 1985, Chap. 17), i.e. cases where the coupling point is not a branch point of the eigen values of the dispersion matrix. Three-dimensional transmission of the fast wave around the ion cyclotron resonance was considered by Friedland (1990) using reduction techniques.

The propagation and coupling of electromagnetic waves in media that vary in more than one dimension have many applications. The most important of these is perhaps in radiofrequency heating and investigations of plasmas in tokamaks (Cairns 1991) and in the physics of the ionosphere (Budden 1985). The plasmas there are considered as plane-stratified. However, the variation of the curvature of the plasma should be taken into account. In almost all cases when wave coupling has been under consideration, this curvature has been assumed to be negligible, or at

† Present address: Academy Adelbert von Chamisso, PO Box 10 44 55, D-40035 Düsseldorf, Germany.

https://doi.org/10.1017/S0022377897005709 Published online by Cambridge University Press

most to give only small and unimportant effects. But experience in plasma physics has always shown that small variations can give huge effects, because generally plasmas are unstable media. In the case of stability, it has been shown that small variations can lead to essentially new results, for example relativistic effects in wave absorption (McDonald *et al.* 1994) or the presence of a non-negligible Larmor radius (Cairns *et al.* 1995). Another point is that in papers on mode coupling the magnetic field in the plasma has been taken to vary in only one direction. But we know that in the ionosphere the magnetic field is curved and also that in a tokamak the magnetic field is a superposition of toroidal and poloidal configurations. Mode-coupling calculations taking into account the curvature of the magnetic field could give new insight in the physics of such of plasmas.

It is the aim of this paper to give a systematic derivation of coupled wave equations in slowly varying plasmas that are three-dimensionally inhomogeneous, non-stationary, anisotropic and absorbing. In order to be self-contained, we also include a systematic representation of some results previously presented at conferences (Suchy and Sabzevari 1992*a*,*b*) and refer to these papers for the explicit calculations. Some coupled wave equations were also represented there, but they must satisfy special conditions and are not applicable in general, unlike the equations derived in the present paper.

2. The eikonal–Maxwell system of differential equations

In a plasma with time and space dispersion the electric current density, dielectric displacement and magnetic induction must be written as integral equations

$$\left.\begin{aligned} \mathbf{J}(\mathbf{X}) &= \int d^4\mathbf{X}'\, \boldsymbol{\sigma}^{\mathrm{ker}}(\mathbf{X}-\mathbf{X}';\mathbf{X})\cdot\mathbf{E}(\mathbf{X}'),\\ \mathbf{D}(\mathbf{X}) &= \int d^4\mathbf{X}'\, \boldsymbol{\varepsilon}^{\mathrm{ker}}(\mathbf{X}-\mathbf{X}';\mathbf{X})\cdot\mathbf{E}(\mathbf{X}'),\\ \mathbf{B}(\mathbf{X}) &= \int d^4\mathbf{X}'\, \boldsymbol{\mu}^{\mathrm{ker}}(\mathbf{X}-\mathbf{X}';\mathbf{X})\cdot\mathbf{H}(\mathbf{X}'), \end{aligned}\right\} \tag{1}$$

where kernels are the electric conductivity, dielectric permittivity and magnetic permeability tensors. They have two kinds of dependences on the space–time four-vectors $\mathbf{X}$ and $\mathbf{X}'$. Those before the semicolon are rapidly varying and due to the collective effects of the medium, and those after the semicolon are slowly varying and due to the features of the medium itself, such as density. In the integration over time, causality must of course be taken into account, but we redefine it in the kernels so that the equations can be written in the above compact form, where all integrations are from $-\infty$ to ∞. These equations must be solved together with the Maxwell equations

$$\left.\begin{aligned} \boldsymbol{\nabla}\times\mathbf{H}(\mathbf{X}) &= \frac{\partial}{\partial t}\mathbf{D}(\mathbf{X})+\mathbf{j}(\mathbf{X}),\\ \boldsymbol{\nabla}\times\mathbf{E}(\mathbf{X}) &= -\frac{\partial}{\partial t}\mathbf{B}(\mathbf{X}). \end{aligned}\right\} \tag{2}$$

To do this, we first transform the integral equations (1) into differential equations with the help of geometrical optics (Kravstov 1969; Suchy and Sabzevari

https://doi.org/10.1017/S0022377897005709 Published online by Cambridge University Press

1992*a*), by writing $\mathbf{E}, \mathbf{H}$ and $\mathbf{j}$ in eikonal form

$$\left.\begin{aligned}\mathbf{E}(\mathbf{X}) &= \mathbf{E}(;\mathbf{X})\exp[i\phi(\mathbf{X})],\\ \mathbf{H}(\mathbf{X}) &= \mathbf{H}(;\mathbf{X})\exp[i\phi(\mathbf{X})],\\ \mathbf{j}(\mathbf{X}) &= \mathbf{j}(;\mathbf{X})\exp[i\phi(\mathbf{X})],\end{aligned}\right\} \tag{3}$$

where $\phi(\mathbf{X})$ is the eikonal, which in geometrical optics is a rapidly varying function of space and time, and $\mathbf{E}(;\mathbf{X}), \mathbf{H}(;\mathbf{X})$ and $\mathbf{j}(;\mathbf{X})$ are the amplitudes, which are slowly varying, because of the slowness of variation of the medium. The first two equations of (3) can also be considered as the wave entering the plasma, and the last as the movement of the charged particles in response to the waves. In principle, constant external fields could be added to (3), but it turns out that they do not give rise to any new physical results.

Three different scales are considered here. First, we suppose that on each particle only the particles in its vicinity in space and time have an essential influence. We denote the relevant scale by Δ_1, which is small compared with the size of the plasma in space and time. Hence in (1)

$$\mathbf{X} - \mathbf{X}' = O(\Delta_1) \ll 1. \tag{4}$$

Then it is sufficient to integrate only over the vicinity of the point $\mathbf{X}$. The second scale arises from the slowness of the variation of the medium, we denote this (small) scale by Δ_2. Then we can write

$$\frac{\partial \mathbf{E}(;\mathbf{X})}{\partial \mathbf{X}}, \frac{\partial \mathbf{H}(;\mathbf{X})}{\partial \mathbf{X}}, \frac{\partial \mathbf{G}(;\mathbf{X})}{\partial \mathbf{X}} = O(\Delta_2) \ll 1, \tag{5}$$

where $\mathbf{G}(;\mathbf{X})$ is an arbitrary function of the medium. The third scale, which we denote by Λ, is due to the largeness of the frequency and wavenumber, which are defined by the four-dimensional wave vector

$$\mathbf{K}(\mathbf{X}) = \frac{\partial \phi(\mathbf{X})}{\partial \mathbf{X}}. \tag{6}$$

In geometrical optics, this can be written as

$$\mathbf{K}(\mathbf{X}) = \Lambda\, \mathbf{G}(;\mathbf{X}), \quad \Lambda \gg 1, \tag{7}$$

where $\mathbf{G}(;\mathbf{X})$ is a slowly varying function because of the slowness of variation of the medium, and is $O(1)$. From (6) and (7),

$$\phi(\mathbf{X}) = \int^{\mathbf{X}} \mathbf{K}(\mathbf{X}') \cdot d\mathbf{X}' = \Lambda \int^{\mathbf{X}} \mathbf{G}(;\mathbf{X}') \cdot d\mathbf{X}' = O(\Lambda). \tag{8}$$

Hence, because of (5),

$$\frac{\partial}{\partial \mathbf{X}}\mathbf{K}(\mathbf{X}) = \Lambda \frac{\partial}{\partial \mathbf{X}}\mathbf{G}(;\mathbf{X}) = O(\Lambda\Delta_2). \tag{9}$$

In general, these three scales do not have to be of the same order:

$$O(\Delta_1) \neq O(\Delta_2) \neq O\left(\frac{1}{\Lambda}\right). \tag{10}$$

In particular, the variation of the wave vector does not have to be as small along the ray path as is often believed and assumed, i.e. strong absorption can be assumed.

https://doi.org/10.1017/S0022377897005709 Published online by Cambridge University Press

Now expanding $\phi(\mathbf{X}'), \mathbf{E}(\mathbf{X}')$ and $\mathbf{H}(\mathbf{X}')$ around the point $\mathbf{X}$, we obtain

$$\mathbf{E}(\mathbf{X}') = e^{i\phi(\mathbf{X})}e^{i(\mathbf{X}'-\mathbf{X})\cdot\mathbf{K}(\mathbf{X})}\left\{\mathbf{E}(;\mathbf{X}) + (\mathbf{X}'-\mathbf{X})\cdot\frac{\partial}{\partial\mathbf{X}}\mathbf{E}(;\mathbf{X}) + \frac{i}{2}\left[(\mathbf{X}'-\mathbf{X})(\mathbf{X}'-\mathbf{X}):\frac{\partial}{\partial\mathbf{X}}\mathbf{K}\right]\mathbf{E}(;\mathbf{X}) + O(\Delta_1^2\Delta_2) + O(\Lambda\Delta^3\Delta_2^2)\right\}, \tag{11}$$

where we have used dyadic and double scalar products. A similar expression can be obtained for $\mathbf{H}(\mathbf{X}')$. Fourier-transforming the rapidly varying part of $\boldsymbol{\sigma}^{\ker}(\mathbf{X}-\mathbf{X}';\mathbf{X}), \boldsymbol{\varepsilon}^{\ker}(\mathbf{X}-\mathbf{X}';\mathbf{X})$ and $\boldsymbol{\mu}^{\ker}(\mathbf{X}-\mathbf{X}';\mathbf{X})$ in (1) and using the properties of the delta function $\delta(\mathbf{K}'-\mathbf{K})$ and its derivatives, we can transform (1) into (Kravtsov 1969; Suchy and Sabzevari 1992*a*)

$$\left.\begin{aligned}\mathbf{j}(\mathbf{X}) &= e^{i\phi(\mathbf{X})}[\boldsymbol{\sigma}(\mathbf{K};\mathbf{X})(1 - i\overleftrightarrow{\partial} - \tfrac{1}{2}i\overleftarrow{\partial})]\cdot\mathbf{E}(;\mathbf{X}),\\ \mathbf{D}(\mathbf{X}) &= e^{i\phi(\mathbf{X})}[\boldsymbol{\varepsilon}(\mathbf{K};\mathbf{X})(1 - i\overleftrightarrow{\partial} - \tfrac{1}{2}i\overleftarrow{\partial})]\cdot\mathbf{E}(;\mathbf{X}),\\ \mathbf{B}(\mathbf{X}) &= e^{i\phi(\mathbf{X})}[\boldsymbol{\mu}(\mathbf{K};\mathbf{X})(1 - i\overleftrightarrow{\partial} - \tfrac{1}{2}i\overleftarrow{\partial})]\cdot\mathbf{H}(;\mathbf{X}),\end{aligned}\right\} \tag{12}$$

where terms $O(\Delta_1^2\Delta_2) + O(\Lambda\Delta^3\Delta_2^2)$ terms have been neglected and the Janus and Spreading operators are defined as

$$\overleftrightarrow{\partial} = \frac{\overleftarrow{\partial}}{\partial\mathbf{K}}\cdot\frac{\overrightarrow{\partial}}{\partial\mathbf{X}}, \qquad \overleftarrow{\partial} = \frac{\overleftarrow{\partial^2}}{\partial\mathbf{K}\,\partial\mathbf{K}}:\frac{\partial}{\partial\mathbf{X}}\mathbf{K}. \tag{13}$$

To combine (12) with the Maxwell equations, we first have to calculate the time derivatives of $\mathbf{D}(\mathbf{X})$ and $\mathbf{B}(\mathbf{X})$. To do this, we first write $\mathbf{D}(\mathbf{X})$ (and correspondingly $\mathbf{B}(\mathbf{X})$) in the form (Kravtsov 1969; Suchy and Sabzevari 1992*a*)

$$\mathbf{D} = \mathbf{D}(\mathbf{K};\mathbf{X})\exp[i\phi(\mathbf{X})] \tag{14}$$

with

$$\mathbf{D}(\mathbf{K};\mathbf{X}) = [\boldsymbol{\varepsilon}(\mathbf{K};\mathbf{X})(1 - i\overleftrightarrow{\partial} - \tfrac{1}{2}i\overleftarrow{\partial})]\cdot\mathbf{E}(;\mathbf{X}) + O(\Delta_1^2\Delta_2) + O(\Lambda\Delta_1^3\Delta_2^2) \tag{15}$$

From (7), we can write

$$\frac{\partial}{\partial\mathbf{K}} = \frac{1}{\Lambda}\frac{\partial}{\partial\mathbf{G}} = O\left(\frac{1}{\Lambda}\right), \tag{16}$$

$$\frac{\partial^2}{\partial\mathbf{K}\,\partial\mathbf{K}} = \frac{1}{\Lambda^2}\frac{\partial^2}{\partial\mathbf{G}\,\partial\mathbf{G}} = O\left(\frac{1}{\Lambda^2}\right), \tag{17}$$

and it follows that from these equations

$$\frac{\partial}{\partial t}\mathbf{D}(\mathbf{K};\mathbf{X}) = \frac{\partial}{\partial t}[\boldsymbol{\varepsilon}(\mathbf{K};\mathbf{X})\cdot\mathbf{E}(;\mathbf{X})] + O\left(\frac{\Delta_2}{\Lambda}\right). \tag{18}$$

In all equations until now, we have kept terms $O(\Delta_1\Delta_2)$, i.e. terms that are products of two small numbers. The last term in (18) is also the product of two small numbers. But we can neglect it compared with the previous term, since $O(1/\Lambda) \ll O(\Delta_1)$. This follows from the fact that in geometrical optics the scale of the wavelength and period can be considered as much smaller than the scale of the distance and time, at which different parts of the medium have an essential influence on each other. Using the result (18) and the corresponding expressions for the magnetic

https://doi.org/10.1017/S0022377897005709 Published online by Cambridge University Press

induction, we finally obtain

$$\frac{\partial}{\partial t}\mathbf{D}(\mathbf{X}) = e^{i\phi(\mathbf{X})}\Big\{ -i\omega(\mathbf{X})\boldsymbol{\varepsilon}(\mathbf{K};\mathbf{X}) - \omega(\mathbf{X})\boldsymbol{\varepsilon}(\mathbf{K};\mathbf{X})(\overleftrightarrow{\partial} + \tfrac{1}{2}\overleftarrow{\partial})X + \left[\frac{\partial}{\partial t}\boldsymbol{\varepsilon}(\mathbf{K};\mathbf{X})\right]_{\mathbf{K}}\Big\}\cdot\mathbf{E}(;\mathbf{X}), \tag{19}$$

and a corresponding expression for $\partial\mathbf{B}(\mathbf{X})/\partial t$, where $\boldsymbol{\varepsilon}$ is interchanged with $\boldsymbol{\mu}$ and $\mathbf{E}(;\mathbf{X})$ with $\mathbf{H}(;\mathbf{X})$. The time derivatives of $\mathbf{D}(\mathbf{X})$ and $\mathbf{B}(\mathbf{X})$ together with $\mathbf{j}(\mathbf{X})$ from (12) can now be combined with the Maxwell equations (2) to give a system of partial differential equations for the components of the six-dimensional electromagnetic field amplitude

$$\mathbf{f}(;\mathbf{X}) = \begin{bmatrix}\mathbf{E}(;\mathbf{X})\\ \mathbf{H}(;\mathbf{X})\end{bmatrix}. \tag{20}$$

Since we have different modes propagating in the medium, we first expand the electromagnetic field in terms of these modes:

$$\mathbf{f}(\mathbf{X}) = \sum_{\alpha} e^{\phi_\alpha(\mathbf{X})}\,\mathbf{f}_\alpha(;\mathbf{X}), \tag{21}$$

where $\phi_\alpha(\mathbf{X})$ is the eikonal and $\mathbf{f}_\alpha(;\mathbf{X})$ is the electromagnetic amplitude for each mode and each mode has a separate wave vector

$$\mathbf{K}_\alpha(\mathbf{X}) = \frac{\partial\phi_\alpha(\mathbf{X})}{\partial\mathbf{X}}. \tag{22}$$

Then the system of partial differential equations for the electromagnetic amplitudes of the modes can be represented as

$$\sum_{\alpha}\exp[i\phi_\alpha(\mathbf{X})]\Big\{\mathbf{M}(\mathbf{K}_\alpha;\mathbf{X})(i + \overleftrightarrow{\partial}_\alpha + \tfrac{1}{2}\overleftarrow{\partial}_\alpha) - \left[\frac{\partial}{\partial t}\mathbf{C}(\mathbf{K}_\alpha;\mathbf{X})\right]_{\mathbf{K}_\alpha}\Big\}\cdot\mathbf{f}_\alpha(;\mathbf{X}) = 0, \tag{23}$$

which we call the eikonal–Maxwell system (Suchy and Sabzevari 1992*a*), with the Maxwell tensor

$$\mathbf{M}(\mathbf{K}_\alpha;\mathbf{X}) = \begin{bmatrix}\omega_\alpha(\mathbf{k};\mathbf{X})\,\boldsymbol{\varepsilon}(\mathbf{K}_\alpha;\mathbf{X}) + i\,\boldsymbol{\sigma}(\mathbf{K}_\alpha;\mathbf{X}) & \mathbf{k}\times\mathbf{I}\\ -\mathbf{k}\times\mathbf{I} & \omega_\alpha(\mathbf{k};\mathbf{X})\,\boldsymbol{\mu}(\mathbf{K}_\alpha;\mathbf{X})\end{bmatrix}, \tag{24}$$

where $\mathbf{I}$ is the unit tensor, $\mathbf{k}$ is the three-dimensional spatial part of the wave vector (22) and ω_α is the frequency†. $\mathbf{C}(\mathbf{K}_\alpha;\mathbf{X})$ is the material tensor

$$\mathbf{C}(\mathbf{K}_\alpha;\mathbf{X}) = \begin{bmatrix}\boldsymbol{\varepsilon}(\mathbf{K}_\alpha;\mathbf{X}) & 0\\ 0 & \boldsymbol{\mu}(\mathbf{K}_\alpha;\mathbf{X})\end{bmatrix}. \tag{25}$$

$\overleftrightarrow{\partial}_\alpha$ and $\overleftarrow{\partial}_\alpha$ are the operators (13) corresponding to each mode. The derivatives with respect to each mode are defined as

$$\frac{\partial}{\partial\mathbf{K}_\alpha}\mathbf{M}(\mathbf{K}_\alpha;\mathbf{X}) = \frac{\partial}{\partial\mathbf{K}}\mathbf{M}(\mathbf{K};\mathbf{X})\Big|_{\mathbf{K}=\mathbf{K}_\alpha}. \tag{26}$$

† In $\mathbf{K}_\alpha$, it is sufficent to attach the index α to one component only, because of the dispersion relation (30).

https://doi.org/10.1017/S0022377897005709 Published online by Cambridge University Press

3. The eigenvalue spectrum of the modes and polarization vectors

It can be shown (Suchy and Sabzevari 1992*a*) that the Maxwell tensor (24) can be written as the sum of two parts:

$$\mathbf{M}(\mathbf{K};\mathbf{X}) = \mathbf{M}_T(\mathbf{K}_T;\mathbf{X}) - \lambda\,\mathbf{M}_L(\mathbf{K}_T;\mathbf{X}), \tag{27}$$

where we have divided the wave vector into longitudinal and transverse parts,

$$\mathbf{K} = \mathbf{K}_T + \lambda\,\mathbf{g}^{\lambda}, \tag{28}$$

with $\mathbf{g}^{\lambda}$ the axis in the direction of the longitudinal part. For example, λ can be chosen as minus the frequency ω. Then $\mathbf{K}_T$ is the spatial wave vector $\mathbf{k}$, and we call these the frequency modes. For stratified (not necessarily plane) media, λ can be chosen as the component of $\mathbf{k}$ perpendicular to the stratification surfaces. For a thorough discussion of various possibilities, see Suchy and Sabzevari (1992*a*). It should be noted that using the method represented here restricts the dispersion of $\boldsymbol{\varepsilon}$, $\boldsymbol{\sigma}$ and $\boldsymbol{\mu}$ to the three components of $\mathbf{K}_T$.

The form (27) of the Maxwell tensor is called a generalized characteristic matrix. We can define right and left eigenvectors:

$$\left.\begin{aligned}\mathbf{M}(\mathbf{K}_\alpha;\mathbf{X})\cdot\mathbf{g}_\alpha &= [\mathbf{M}_T(\mathbf{K}_T;\mathbf{X}) - \lambda_\alpha(\mathbf{K}_T;\mathbf{X})\,\mathbf{M}_L(\mathbf{K}_T;\mathbf{X})]\cdot\mathbf{g}_\alpha = 0,\\ \bar{\mathbf{g}}_\alpha\cdot\mathbf{M}(\mathbf{K}_\alpha;\mathbf{X}) &= \bar{\mathbf{g}}_\alpha\cdot[\mathbf{M}_T(\mathbf{K}_T;\mathbf{X}) - \lambda_\alpha(\mathbf{K}_T;\mathbf{X})\,\mathbf{M}_L(\mathbf{K}_T;\mathbf{X})] = 0.\end{aligned}\right\} \tag{29}$$

Equations (29) have non-trivial solutions only when

$$\operatorname{Det}\mathbf{M}(\mathbf{K}_\alpha;\mathbf{X}) = 0. \tag{30}$$

This gives us the dispersion relation. We choose the $\mathbf{g}_\alpha$ as the polarization vectors. The advantage of this choice is that the right eigenvectors are parallel to the electromagnetic field in a homogeneous medium, as can be seen immediately from (23). We can then divide the space–time into small cells, and in each cell the medium can be considered as homogeneous. When we go from one cell to the other, the $\mathbf{g}_\alpha$ will change direction and size. So the first equation of (29) gives us the direction (but not the magnitude) of the wave field at each space–time point. The left eigenvectors in general have no real physical meaning. They are only useful for further calculations. But they can be considered as waves in a concomitant space (Suchy and Altman 1975).

It is easy to prove the following bi-orthogonality relation:

$$\bar{\mathbf{g}}_\alpha\cdot\mathbf{M}_L\cdot\mathbf{g}_\beta = \delta_{\alpha\beta}\,\bar{\mathbf{g}}_\alpha\cdot\mathbf{M}_L\cdot\mathbf{g}_\alpha. \tag{31}$$

The right and left eigenvectors are identical when the Maxwell tensor is equal to its own transpose. In the case of a Hermitian Maxwell tensor, the left eigenvectors become the complex conjugate of the right eigenvectors.

Since the right (left) eigenvectors can be defined as an arbitrary linear combination of the columns (rows) of the adjoint of $\mathbf{M}(\mathbf{K}_\alpha;\mathbf{X})$, we can define

$$\left.\begin{aligned}\mathbf{g}_\alpha &= \frac{\operatorname{Adj}\mathbf{M}(\mathbf{K}_\alpha;\mathbf{X})\cdot\mathbf{c}}{[\bar{\mathbf{c}}\cdot\operatorname{Adj}\mathbf{M}(\mathbf{K}_\alpha;\mathbf{X})\cdot\mathbf{c}]^{1/2}},\\ \bar{\mathbf{g}}_\alpha &= \frac{\bar{\mathbf{c}}\cdot\operatorname{Adj}\mathbf{M}(\mathbf{K}_\alpha;\mathbf{X})}{[\bar{\mathbf{c}}\cdot\operatorname{Adj}\mathbf{M}(\mathbf{K}_\alpha;\mathbf{X})\cdot\mathbf{c}]^{1/2}},\end{aligned}\right\} \tag{32}$$

https://doi.org/10.1017/S0022377897005709 Published online by Cambridge University Press

where $\bar{\mathbf{c}}$ and $\mathbf{c}$ are arbitrary constant column and row vectors. An expression for the biorthogonality relation (31) can be obtained (Suchy and Sabzevari 1992*b*):

$$\bar{\mathbf{g}}_\alpha \cdot \mathbf{M}_L \cdot \mathbf{g}_\alpha = \Theta \prod_{\gamma \neq \alpha} (\lambda_\alpha - \lambda_\gamma). \tag{33}$$

The factor $\Theta = \Theta(\mathbf{K}_T; \mathbf{X})$ occurs because we have a generalized characteristic matrix, namely the matrix (27). With this expression, it is possible to normalize the polarization vectors in a suitable form:

$$\left.\begin{aligned} \widehat{\mathbf{f}}_\alpha &= \frac{\mathbf{g}_\alpha}{(\bar{\mathbf{g}}_\alpha \cdot \mathbf{M}_L \cdot \mathbf{g}_\alpha)^{1/2}} = \frac{\mathbf{g}_\alpha}{\left[\Theta \prod_{\gamma \neq \alpha} (\lambda_\alpha - \lambda_\gamma)\right]^{1/2}}, \\ \widehat{\bar{\mathbf{f}}}_\alpha &= \frac{\bar{\mathbf{g}}_\alpha}{(\bar{\mathbf{g}}_\alpha \cdot \mathbf{M}_L \cdot \mathbf{g}_\alpha)^{1/2}} = \frac{\bar{\mathbf{g}}_\alpha}{\left[\Theta \prod_{\gamma \neq \alpha} (\lambda_\alpha - \lambda_\gamma)\right]^{1/2}}, \end{aligned}\right\} \tag{34}$$

and from the bi-orthogonality relations (31) there follow the bi-orthonormality relations

$$\left.\begin{aligned} \widehat{\bar{\mathbf{f}}}_\alpha \cdot \mathbf{M}_L \cdot \widehat{\mathbf{f}}_\alpha &= 1 \qquad \text{always}, \\ \widehat{\bar{\mathbf{f}}}_\alpha \cdot \mathbf{M}_L \cdot \widehat{\mathbf{f}}_\beta &= 0 \qquad (\lambda_\alpha \neq \lambda_\beta). \end{aligned}\right\} \tag{35}$$

We now define the complete electromagnetic wave amplitude for each mode by

$$\mathbf{f}_\alpha = a_\alpha(;\mathbf{X})\, \widehat{\mathbf{f}}_\alpha(\mathbf{K}_\alpha; \mathbf{X}), \tag{36}$$

where a_α is the scalar amplitude of the wave. To write the modes in this form is of great advantage. The reason is that $\bar{\mathbf{f}}_\alpha \cdot \mathbf{M}_L \cdot \mathbf{f}_\alpha$ is equal to the energy of the wave propagating through the medium (Suchy and Sabzevari 1992*a*) when collisions are neglected, i.e. the Maxwell tensor is Hermitian. Hence, because of the first of the bi-orthonormality rations (35), $|a_\alpha|^2$ is identical to the wave energy of the mode α. In the case of a non-Hermitian Maxwell tensor, this last quantity is proportional to the wave energy at each space–time point; hence $|a_\alpha|^2/|a_\beta|^2$ gives us directly the transmission, reflection or absortption coefficients of different modes.

4. Mode-coupling point

Since the λs are functions of space and time, there are some points $\mathbf{X}_0$, where $\lambda_\alpha(\mathbf{K}_T(\mathbf{X}_0); \mathbf{X}_0) = \lambda_\beta(\mathbf{K}_T(\mathbf{X}_0); \mathbf{X}_0)$. Then, from (29) and (34) the polarization vectors $\widehat{\mathbf{f}}_\alpha$ and $\widehat{\mathbf{f}}_\beta$ become parallel when the rank of the matrix $\mathbf{M}(\mathbf{K}_\alpha; \mathbf{X})$ $(= \mathbf{M}(\mathbf{K}_\beta; \mathbf{X}))$ is equal to 5. This is because if the rank of a matrix is equal to the dimension minus the number of linearly independent eigencolumns of the matrix then we have only one linearly independent $\widehat{\mathbf{f}}_\alpha$ $(= \widehat{\mathbf{f}}_\beta)$. Otherwise, there would be more than one linearly independent eigencolumn for each mode. For example, if the rank were 4 then we would have two linearly independent eigencolumns for each α and β, i.e. $\mathbf{f}_{\alpha 1}$ $(= \mathbf{f}_{\beta 1})$ and $\mathbf{f}_{\alpha 2}$ $(= \mathbf{f}_{\beta 2})$. Hence $\mathbf{f}_{\alpha 1}$ would be linearly independent of $\mathbf{f}_{\beta 2}$. In this case, at the point $\mathbf{X}_0$, we could find two vectors $\widehat{\mathbf{f}}_\alpha$ and $\widehat{\mathbf{f}}_\beta$, corresponding to different modes, which were linearly independent, i.e. not parallel. No mode coupling would occur in this case. The two modes would pass by without any influence on each other (Budden & Smith 1974; Budden 1985, Secs 17.6 and 17.7). Hence the

https://doi.org/10.1017/S0022377897005709 Published online by Cambridge University Press

exact definition of a coupling point is

coupling point $\Leftrightarrow$

$$\{\lambda_\alpha = \lambda_\beta \ \text{ for } \ \mathbf{X} = \mathbf{X}_0 \ \text{ and } \ \operatorname{Rank}\mathbf{M}(\mathbf{K}_\alpha;\mathbf{X}) = \operatorname{Rank}\mathbf{M}(\mathbf{K}_\beta;\mathbf{X}) = 5\}\,. \quad (37)$$

It should be noted that at the coupling points the scalar amplitudes a_α and a_β are generally not equal. Physically, coupling does not occur at one mathematical point. Actually there is a small area around the coupling point where coupling takes place. This coupling area can be defined as

coupling area $\Leftrightarrow$

$$\{|\mathbf{X}-\mathbf{X}_0| < \epsilon \iff |\lambda_\alpha - \lambda_\beta| < \delta \ \mid \ \epsilon,\delta \ll 1, \text{ and } \lambda_\alpha(\mathbf{X}_0) = \lambda_\beta(\mathbf{X}_0)\}\,. \quad (38)$$

It is possible to estimate the size of ϵ and δ (see Sabzevari 1996). When two waves reach the border of this area, the coupling is weak. The more they penetrate the area, the stronger the coupling will be. At the coupling point, the coupling is maximum. Usually, a singularity is to be expected there (for the example of a stratified medium, Budden 1972; Budden 1985, Sec. 16.3; see Sabzevari 1992, 1993).

From (34) and (35), there follow the bi-orthonormality relations

$$\left.\begin{aligned} \widehat{\overline{\mathbf{f}}}_\alpha \cdot \mathbf{M}_L \cdot \widehat{\mathbf{f}}_\beta = 1 \quad &\text{for } \ \mathbf{X} = \mathbf{X}_0, \\ \widehat{\overline{\mathbf{f}}}_\alpha \cdot \mathbf{M}_L \cdot \widehat{\mathbf{f}}_\beta = 0 \quad &\text{for } \ \mathbf{X} \neq \mathbf{X}_0. \end{aligned}\right\} \quad (39)$$

As we can see from (39), the point $\mathbf{X}_0$ is a 'jump singularity', because the scalar product is discontinuous there.

5. Transport equations for the amplitude

Since the polarization vectors $\widehat{\mathbf{f}}_\alpha$ can be calculated from the local eigenvector equation (29) at each space–time point $\mathbf{X}$, we need another equation to calculate the scalar amplitudes a_α along the rays to obtain the complete wave field (36) at each space–time point. Putting (36) into the system (23), a system of partial differential equations for the amplitudes can be obtained (for the technical details of the calculations see Suchy and Sabzevari 1992*a*):

$$\sum_\gamma e^{\phi_\gamma(\mathbf{X})}\left(\mathbf{W}_{\rho\gamma}\cdot\frac{\partial a_\gamma}{\partial \mathbf{X}} + \Gamma_{\rho\gamma}a_\gamma\right) = 0. \quad (40)$$

The transport vectors $\mathbf{W}_{\rho\gamma}$ give the change in the amplitude in the direction of this vector, as can be seen from (40). They have the form

$$\mathbf{W}_{\rho\gamma} = \delta_{\rho\gamma}\frac{d\mathbf{X}}{d\tau_\gamma} + (\lambda_\rho - \lambda_\gamma)\,\widehat{\overline{\mathbf{f}}}_\rho \cdot \mathbf{M}_L \cdot \left(\widehat{\mathbf{f}}_\gamma \frac{\overleftarrow{d}}{d\mathbf{K}_T}\right), \quad (41)$$

and the coefficients of the system are

$$\Gamma_{\rho\gamma} = \delta_{\rho\gamma}\frac{\overrightarrow{\partial}_\gamma \mathbf{D}_\gamma}{2} + (\lambda_\rho - \lambda_\gamma)\,\widehat{\overline{\mathbf{f}}}_\rho \cdot \mathbf{M}_L \cdot \left(\widehat{\mathbf{f}}_\gamma \frac{\overleftarrow{d^2}}{d\mathbf{K}_T d\mathbf{K}_T} \ : \ \frac{\partial}{\partial \mathbf{X}_T}\mathbf{K}_T\right) + \hat{\Gamma}_{\rho\gamma}, \quad (42)$$

https://doi.org/10.1017/S0022377897005709 Published online by Cambridge University Press

with

$$\hat{\Gamma}_{\rho\gamma} = \hat{\bar{\mathbf{f}}}_{\rho} \cdot \left[\mathbf{M}_L \cdot \frac{d\hat{\mathbf{f}}_{\gamma}}{d\tau_{\gamma}} + \left(\frac{d\mathbf{M}_L}{d\tau_{\gamma}} \right)_{\mathbf{X}} \cdot \hat{\mathbf{f}}_{\gamma} - \mathbf{M}_{\gamma} \frac{\overleftarrow{d}}{d\mathbf{K}_T} : \left(\frac{\partial}{\partial \mathbf{X}_T} \hat{\mathbf{f}}_{\gamma} \right)_{\mathbf{K}_T} + \left(\frac{\partial \mathbf{C}}{\partial t} \right)_{\mathbf{K}_{\gamma}} \cdot \hat{\mathbf{f}}_{\gamma} \right]. \tag{43}$$

In these relations $\mathbf{X} = \mathbf{X}_T + \mathbf{g}_{\lambda}\mathbf{X}_L$, with $\mathbf{g}_{\lambda} \cdot \mathbf{g}^{\lambda} = 1$ (cf. (28)). τ_{γ} is the parametrization of the ray γ, and is defined via Hamilton's equations (Bernstein and Friedland 1983; Suchy and Sabzevari 1992*a*)

$$\left.\begin{aligned} \frac{d\mathbf{X}}{d\tau_{\gamma}} &= \frac{\partial D_{\gamma}(\mathbf{K};\mathbf{X})}{\partial \mathbf{K}}, \\ \frac{d\mathbf{K}}{d\tau_{\gamma}} &= -\left[\frac{\partial D_{\gamma}(\mathbf{K};\mathbf{X})}{\partial \mathbf{X}} \right]_{\mathbf{K}}, \end{aligned}\right\} \tag{44}$$

where D_{γ} is defined as

$$\operatorname{Det} \mathbf{M}(\mathbf{K};\mathbf{X}) = -\Theta(\mathbf{K}_T;\mathbf{X}) \prod_{\rho} D_{\gamma}(\mathbf{K};\mathbf{X}), \tag{45}$$

with

$$D_{\gamma}(\mathbf{K};\mathbf{X}) = \lambda - \lambda_{\gamma}(\mathbf{K}_T;\mathbf{X}). \tag{46}$$

In the case where all the modes propagate independently, i.e. when there is no mode coupling, there are no points where the ray paths coincide and the paths are independent of each other. In this case the functions

$$\phi_{\gamma}(\mathbf{X}) = \int_{\text{ray path}}^{\mathbf{X}} d\mathbf{X}' \cdot \mathbf{K}_{\gamma}(\mathbf{X}') \tag{47}$$

are independent of each other. The second terms in (41) and (42) become identical to zero and the Kronecker symbol becomes unity. From (40), there then follows an ordinary differential equation for the scalar amplitude of each mode α along its ray path:

$$\frac{da_{\alpha}}{d\tau_{\alpha}} + \Gamma_{\alpha\alpha}\, a_{\alpha} = 0 \tag{48}$$

with

$$\Gamma_{\alpha\alpha} = \tfrac{1}{2}\, \vec{\partial}_{\alpha}\, D_{\alpha} + \hat{\Gamma}_{\alpha\alpha}. \tag{49}$$

The solution is

$$a_{\alpha}(\tau_{\alpha}) = \text{const} \times \exp\left(-\int^{\tau_{\alpha}} d\tau'_{\alpha}\, \Gamma_{\alpha\alpha} \right), \tag{50}$$

which gives us the amplitude along its path. For more discussions about the solutions of independent modes see Suchy and Sabzevari (1992*a*, Sec. 7) and Sabzevari (1992, 1993).

6. Coupled wave equations

In this section we explicitly derive coupled ordinary differential equations for the amplitudes of coupled waves near their coupling point (except for the coupling point

https://doi.org/10.1017/S0022377897005709 Published online by Cambridge University Press

itself, because there is a jump singularity there – but we can approach the coupling point arbitrarily closely), i. e. where $\lambda_\alpha \approx \lambda_\beta$.

In the case of the coupling of the two modes α and β, we can see from (47) that ϕ_α and ϕ_β can no longer be considered as independent, because their ray paths coincide in $\mathbf{X}$ and $\mathbf{K}$ space at the coupling point $\mathbf{X}_0$. Since an arbitrary constant phase can be added to the eikonals, we can write, near the coupling point in $\mathbf{K}$ space,

$$\phi_\alpha(\mathbf{X}) = \int_{\mathbf{X}_0}^{\mathbf{X}} d\mathbf{X}' \cdot \mathbf{K}_\alpha(\mathbf{X}') \approx \int_{\mathbf{X}_0}^{\mathbf{X}} d\mathbf{X}' \cdot \mathbf{K}_\beta(\mathbf{X}') = \phi_\beta(\mathbf{X}). \tag{51}$$

When only the two modes couple and are decoupled from the other modes, ϕ_α and ϕ_β are independent of the eikonals corresponding to the other modes. From (40), it then follows that

$$\left.\begin{aligned} \mathbf{W}_{\alpha\alpha} \cdot \frac{\partial a_\alpha}{\partial \mathbf{X}} + \mathbf{W}_{\alpha\beta} \cdot \frac{\partial a_\beta}{\partial \mathbf{X}} + \Gamma_{\alpha\alpha} a_\alpha + \Gamma_{\alpha\beta} a_\beta = 0, \\ \mathbf{W}_{\beta\alpha} \cdot \frac{\partial a_\alpha}{\partial \mathbf{X}} + \mathbf{W}_{\beta\beta} \cdot \frac{\partial a_\beta}{\partial \mathbf{X}} + \Gamma_{\beta\alpha} a_\alpha + \Gamma_{\beta\beta} a_\beta = 0. \end{aligned}\right\} \tag{52}$$

For the transport vectors, from (41),

$$\mathbf{W}_{\alpha\beta} = (\lambda_\alpha - \lambda_\beta)\, \hat{\bar{\mathbf{f}}}_\alpha \cdot \mathbf{M}_L \cdot \left(\hat{\mathbf{f}}_\beta \frac{\overleftarrow{d}}{d\mathbf{K}_T} \right) \quad \text{for } \alpha \neq \beta. \tag{53}$$

It may seem at first, that this term is small, because $\lambda_\alpha \approx \lambda_\beta$. But this assumption is invalid, as we shall soon see. To calculate the transport vector, we use (33) to write (34) as

$$\left.\begin{aligned} \hat{\mathbf{f}}_\beta &= (\lambda_\beta - \lambda_\alpha)^{-1/2} \frac{\mathbf{g}_\beta}{\left[\Theta \prod_{\gamma \neq \alpha,\beta} (\lambda_\beta - \lambda_\gamma)\right]^{1/2}}, \\ \hat{\bar{\mathbf{f}}}_\alpha &= (\lambda_\alpha - \lambda_\beta)^{-1/2} \frac{\bar{\mathbf{g}}_\alpha}{\left[\Theta \prod_{\gamma \neq \alpha,\beta} (\lambda_\alpha - \lambda_\gamma)\right]^{1/2}}. \end{aligned}\right\} \tag{54}$$

It is clear from (31) and (33) that

$$\bar{\mathbf{g}}_\alpha \cdot \mathbf{M}_L \cdot \mathbf{g}_\beta = 0 \quad \text{for all } \lambda_\alpha \text{ and } \lambda_\beta \neq \lambda_\alpha. \tag{55}$$

We differentiate $\hat{\mathbf{f}}_\beta$ by $\mathbf{K}_T$ and multiply the derivative from the right by $\hat{\bar{\mathbf{f}}}_\alpha \cdot \mathbf{M}_L(\mathbf{K}_T; \mathbf{X})$ to obtain, with (55), for (53)

$$\mathbf{W}_{\alpha\beta} = i \frac{\bar{\mathbf{g}}_\alpha \cdot \mathbf{M}_L \cdot (\mathbf{g}_\beta \overleftarrow{d} / d\mathbf{K}_T)}{\Theta \left[\prod_{\gamma \neq \alpha,\beta} (\lambda_\alpha - \lambda_\gamma)(\lambda_\beta - \lambda_\gamma) \right]^{1/2}}. \tag{56}$$

Since $\mathbf{g}_\beta$ is an arbitrary linear combination of the columns of $\operatorname{Adj} \mathbf{M}(\mathbf{K}_\beta; \mathbf{X})$, the elements of this column vector are polynomials in λ_β. Hence

$$\mathbf{g}_\beta \frac{\overleftarrow{d}}{d\mathbf{K}_T} = \left(\frac{\partial}{\partial \lambda_\beta} \mathbf{g}_\beta \right) \frac{\partial \lambda_\beta}{\partial \mathbf{K}_T} + \left(\mathbf{g}_\beta \frac{\overleftarrow{\partial}}{\partial \mathbf{K}_T} \right)_{\lambda_\beta}. \tag{57}$$

https://doi.org/10.1017/S0022377897005709 Published online by Cambridge University Press

It can be shown (see the appendix) that near the coupling point, in general,

$$\frac{\partial \lambda_\beta}{\partial \mathbf{K}_T} \propto \frac{1}{\lambda_\alpha - \lambda_\beta}. \tag{58}$$

This relation is valid not only for the derivatives to $\mathbf{K}_T$ but also for the derivatives of all kinds of variables. The coefficients of the polynomials in $\mathbf{g}_\beta$ can in general be assumed as analytic, since they depend only on the elements of the matrices $\mathbf{M}_T(\mathbf{K}_T;\mathbf{X})$ and $\mathbf{M}_L(\mathbf{K}_T;\mathbf{X})$, which depend only on the $\boldsymbol{\varepsilon}, \boldsymbol{\mu}, \boldsymbol{\sigma}, \mathbf{K}_T$ and $\mathbf{X}$. Therefore, in general, the first term in (57) is growing relative to the second as we approach the coupling point. Hence we can write

$$\mathbf{g}_\beta \frac{\overleftarrow{d}}{d\mathbf{K}_T} \approx \left(\frac{\partial}{\partial \lambda_\beta}\mathbf{g}_\beta\right)\frac{\partial \lambda_\beta}{\partial \mathbf{K}_T}. \tag{59}$$

Now from the first of Hamilton's equations (44) and the definition of D_β from (46) it follows that

$$\frac{d\mathbf{X}}{d\tau_\beta} = -\frac{\partial \lambda_\beta}{\partial \mathbf{K}_T} + \mathbf{g}_\lambda. \tag{60}$$

Near the coupling point, (58) can be used to write (60) as

$$\frac{d\mathbf{X}}{d\tau_\beta} \approx -\frac{\partial \lambda_\beta}{\partial \mathbf{K}_T}. \tag{61}$$

Putting this into (59), it follows from (56) that

$$\left.\begin{aligned} \mathbf{W}_{\alpha\beta}\cdot\frac{\partial a_\beta}{\partial \mathbf{X}} &= \Upsilon_1 \frac{da_\beta}{d\tau_\beta}, & \mathbf{W}_{\alpha\alpha}\cdot\frac{\partial}{\partial \mathbf{X}} &= \frac{d}{d\tau_\alpha}, \\ \mathbf{W}_{\beta\alpha}\cdot\frac{\partial a_\alpha}{\partial \mathbf{X}} &= \Upsilon_2 \frac{da_\alpha}{d\tau_\alpha}, & \mathbf{W}_{\beta\beta}\cdot\frac{\partial}{\partial \mathbf{X}} &= \frac{d}{d\tau_\beta}, \end{aligned}\right\} \tag{62}$$

where

$$\left.\begin{aligned} \Upsilon_1 &= \frac{(\lambda_\alpha - \lambda_\beta)\bar{\mathbf{g}}_\alpha \cdot \mathbf{M}_L \cdot (\partial \mathbf{g}_\beta/\partial\lambda_\beta)}{(\bar{\mathbf{g}}_\alpha \cdot \mathbf{M}_L \cdot \mathbf{g}_\alpha)^{1/2}(\bar{\mathbf{g}}_\beta \cdot \mathbf{M}_L \cdot \mathbf{g}_\beta)^{1/2}} = -i\frac{\bar{\mathbf{g}}_\alpha \cdot \mathbf{M}_L \cdot (\partial \mathbf{g}_\beta/\partial\lambda_\beta)}{\Theta\left[\prod_{\gamma\neq\alpha,\beta}(\lambda_\alpha - \lambda_\gamma)(\lambda_\beta - \lambda_\gamma)\right]}, \\ \Upsilon_2 &= \frac{-(\lambda_\alpha - \lambda_\beta)\bar{\mathbf{g}}_\beta \cdot \mathbf{M}_L \cdot (\partial \mathbf{g}_\alpha/\partial\lambda_\alpha)}{(\bar{\mathbf{g}}_\alpha \cdot \mathbf{M}_L \cdot \mathbf{g}_\alpha)^{1/2}(\bar{\mathbf{g}}_\beta \cdot \mathbf{M}_L \cdot \mathbf{g}_\beta)^{1/2}} = +i\frac{\bar{\mathbf{g}}_\beta \cdot \mathbf{M}_L \cdot (\partial \mathbf{g}_\alpha/\partial\lambda_\alpha)}{\Theta\left[\prod_{\gamma\neq\alpha,\beta}(\lambda_\alpha - \lambda_\gamma)(\lambda_\beta - \lambda_\gamma)\right]^{1/2}} \end{aligned}\right\} \tag{63}$$

for $\lambda_\alpha \approx \lambda_\beta$ and $\lambda_\alpha > \lambda_\beta$ (when $\lambda_\alpha < \lambda_\beta$, the signs of Υ_1 and Υ_2 should be interchanged).

We are now ready to construct the coupled equations. Near the coupling points, the paths almost coincide, and we can write $d\tau_\alpha \approx d\tau_\beta = d\tau$. Putting (62) into (52), we obtain, after some subtractions and additions, the system

$$\frac{d}{d\tau}\mathbf{a} - \boldsymbol{\Gamma}\cdot\mathbf{a} = 0, \tag{64}$$

where

$$\mathbf{a} = \begin{bmatrix} a_\alpha \\ a_\beta \end{bmatrix}, \qquad \boldsymbol{\Gamma} = \begin{bmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{bmatrix}, \tag{65}$$

https://doi.org/10.1017/S0022377897005709 Published online by Cambridge University Press

with

$$\left.\begin{aligned} \Gamma_{11} &= \frac{\Upsilon_1\Gamma_{\beta\alpha} - \Gamma_{\alpha\alpha}}{1 - \Upsilon_1\Upsilon_2}, \qquad \Gamma_{12} = \frac{\Upsilon_1\Gamma_{\beta\beta} - \Gamma_{\alpha\beta}}{1 - \Upsilon_1\Upsilon_2}, \\ \Gamma_{21} &= \frac{\Upsilon_2\Gamma_{\alpha\alpha} - \Gamma_{\beta\alpha}}{1 - \Upsilon_1\Upsilon_2}, \qquad \Gamma_{22} = \frac{\Upsilon_2\Gamma_{\alpha\beta} - \Gamma_{\beta\beta}}{1 - \Upsilon_1\Upsilon_2}. \end{aligned}\right\} \tag{66}$$

The solution of (64) is

$$\mathbf{a} = \exp\left(\int^{\tau} \mathbf{\Gamma}\, d\tau'\right) \cdot \mathbf{a}_0, \tag{67}$$

where $\mathbf{a}_0$ is an arbitrary constant column vector.

In general, Υ_1 and Υ_2 are expected to be finite near the coupling point, since $\mathbf{g}_\alpha$ and $\bar{\mathbf{g}}_\alpha$ are polynomials in λ. By the same methods presented above, it can be shown that, in general, $\Gamma_{\alpha\beta}$,$\Gamma_{\beta\alpha}$, $\Gamma_{\alpha\alpha}$ and $\Gamma_{\beta\beta}$ are growing to infinity as we approach the coupling point. This is due to the second derivative of the second term on the right-hand side of (42). From (58), this second derivative is $O(\lambda_\alpha - \lambda_\beta)^{-2}$, and the singularity cannot be removed when multiplied by $\lambda_\alpha - \lambda_\beta$, so that a term $O(\lambda_\alpha - \lambda_\beta)^{-1}$ remains. Also, $\hat{\Gamma}_{\alpha\beta}$ is singular in general, because of the derivatives of the polarization vector in the first term and the two derivatives of the second term on the right-hand side, since the polarization vector changes very rapidly around the coupling points. For a stratified medium, a different formalism than that presented here, constructed for the coupling of an arbitrary number of coupled waves, results in coupled wave equations with singular coefficients (Sabzevari 1992, 1993). The form of the equations in such that the singularities first have to be removed in order to solve the problem (see also Budden 1985, Chaps 16 and 17; Friedland 1985), because singular coefficients can result in a large variation of the amplitude near the coupling point, and hence in the breakdown of the geometrical optics conditions. In those papers the remedy takes the form of introducing a substitute basis for the polarization vectors. Perhaps these singularities would not appear if the problem were reconsidered in the context of nonlinearity (D. D. Tskhakaya 1996, personal communication). It is not necessary to concern ourselves with this point when we are not interested in the solutions around the coupling points, but only in reflection, transmission, absorption and conversion coefficients (i.e. asymptotic solutions). The best way to investigate these coefficients using the system of equations (64) is to employ S-matrix methods (Volland 1962*a*,*b*; Budden 1985, Secs. 18.5). With these methods, there is no need to bother about the singularities at the coupling point. For an example, see below. S-matrix methods have also been used for the case of mode-conversion problems in one dimension by Littlejohn and Flynn (1993).

7. The example of a plane-stratified plasma

To give a simple example of how to handle the above formulae in practical cases and also to have a check of the correctness of the formalism, we consider the well-known case of a plane-stratified isotropic collisionless cold plasma with an arbitrary number of ion components. This is only to show how this formalism can be applied. Calculations for more complicated plasmas are then straight-

https://doi.org/10.1017/S0022377897005709 Published online by Cambridge University Press

forward. By plane-stratified, we mean that the inhomogenity is in one direction, so that each infinitesimally thin stratification layer can be considered as homogeneous, and the densities in adjacent layers differ only infinitesimally. The wave is in the (x, z) plane, and θ is the angle between the $\mathbf{k}$ vector and the z direction. In this case the Maxwell tensor has the simple form (Swanson 1989, Sec. 2.1.2)

$$\mathbf{M} = \begin{bmatrix} \omega\epsilon_0 P & 0 & 0 & 0 & -k\cos\theta & 0 \\ 0 & \omega\epsilon_0 P & 0 & k\cos\theta & 0 & -k\sin\theta \\ 0 & 0 & \omega\epsilon_0 P & 0 & k\sin\theta & 0 \\ 0 & k\cos\theta & 0 & \omega\mu_0 & 0 & 0 \\ -k\cos\theta & 0 & k\sin\theta & 0 & \omega\mu_0 & 0 \\ 0 & -k\sin\theta & 0 & 0 & 0 & \omega\mu_0 \end{bmatrix}, \tag{68}$$

where

$$P = 1 - \sum_j \frac{\omega_{pj}^2}{\omega^2}. \tag{69}$$

We consider monochromatic waveguide modes, i.e. with the component of the wave vector in the z direction. Hence $\lambda = k_z = k\cos\theta$, and, from Snell's law, $\boldsymbol{K}_T = k_x = k\sin\theta = (\omega/c_0)\sin\theta_0 =$const, where c_0 is the speed of light in vacuum and θ_0 is the angle that the wave makes with the z direction before entering the plasma, i.e. in the vacuum. Then we have

$$\mathbf{M}_L = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{70}$$

The coefficients (42) reduce to

$$\Gamma_{\rho\gamma} = \widehat{\bar{\mathbf{f}}}_\rho \cdot \mathbf{M}_L \cdot \left(\frac{d}{dz}\widehat{\mathbf{f}}_\gamma\right). \tag{71}$$

To derive the right and left eigenvectors $\mathbf{g}$ and $\bar{\mathbf{g}}$, it suffices first to calculate the second row and column of Adj $\mathbf{M}$ respectively. We then define

$$\bar{\mathbf{c}} = [0 \quad 1 \quad 0 \quad 0 \quad 0 \quad 0], \qquad \mathbf{c} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \tag{72}$$

https://doi.org/10.1017/S0022377897005709 Published online by Cambridge University Press

and obtain from (32)

$$\mathbf{g}_\alpha = \bar{\mathbf{g}}_\alpha = \frac{\omega^3}{c_0^2}\left(\frac{\mu_0}{\omega}\right)^{1/2}\begin{bmatrix} 0 \\ [P(P-n_\alpha^2)]^{1/2} \\ 0 \\ \left(\frac{\epsilon_0}{\mu_0}\right)^{1/2}[P(P-n_\alpha^2)]^{1/2} n_\alpha \cos\theta \\ 0 \\ \left(\frac{\epsilon_0}{\mu_0}\right)^{1/2}[P(P-n_\alpha^2)]^{1/2} n_\alpha \sin\theta \end{bmatrix}, \tag{73}$$

where $n_\alpha = (c_0/\omega)k_\alpha$ is the refractive index. From (34), we obtain

$$\hat{\mathbf{f}}_\alpha = \hat{\bar{\mathbf{f}}}_\alpha = \left(\frac{c_0\mu_0}{2}\right)^{1/2}\begin{bmatrix} 0 \\ (n_\alpha\cos\theta)^{-1/2} \\ 0 \\ \left(\frac{\epsilon_0}{\mu_0}\right)^{1/2}(n_\alpha\cos\theta)^{1/2} \\ 0 \\ \left(\frac{\epsilon_0}{\mu_0}\right)^{1/2} n_\alpha^{1/2}\sin\theta(\cos\theta)^{-1/2} \end{bmatrix}. \tag{74}$$

Using (73), (63) can be calculated as

$$\Upsilon_{\substack{1\\2}} = \pm\frac{k_{z\alpha}-k_{z\beta}}{2(k_{z\alpha}k_{z\beta})^{1/2}}\left\{1-\frac{(c_0^2/\omega^2)\left(k_{z\beta,\,\alpha}^2+k_{z\alpha}k_{z\beta}\right)}{P\left[P-(c_0^2/\omega^2)k_x^2-(c_0^2/\omega^2)k_{z\beta,\,\alpha}^2\right]}\right\}, \tag{75}$$

and, from (74) and (71)

$$\Gamma_{\alpha,\beta} = \frac{1}{4(k_{z\alpha}k_{z\beta})^{1/2}}\left(1-\frac{k_{z\alpha}}{k_{z\beta}}\right)\frac{d}{dz}k_{z\beta}. \tag{76}$$

When we consider decoupled modes, i.e. cases where mode coupling occurs nowhere, it follows from (76) that $\Gamma_{\alpha\alpha} = \Gamma_{\beta\beta} = 0$, and then it follows from (48) that the scalar part of the amplitude (but not the polarization vector) does not change when the wave propagates through the medium. This is to be expected, since wave absorption is a kinetic effect and does not occur in a cold collisionless plasma, which can also be concluded immediately from the fact that the entropy cannot change when temperature is absent. In a cold collisionless plasma mode coupling can occur only in the form of reflection and transmission, without absorption. Hence

$$\frac{d}{dz}|a_\alpha|^2 = 0 \tag{77}$$

is locally fulfilled for decoupled modes. We now investigate pure reflection. There,

https://doi.org/10.1017/S0022377897005709 Published online by Cambridge University Press

we have (Swanson 1989, Sec. 2.1.2)

$$k_{z\alpha} = -k_{z\beta} = \left(\frac{\omega^2}{c_0^2}P - k_x^2\right)^{1/2} \to 0 \tag{78}$$

as we approach the reflection point. Hence (64) can be used, since (58) is valid. For (75) and (76), it follows from (78) that

$$\Upsilon_{\substack{1,\\2}} = \pm i\left(\frac{\omega^2}{c_0^2 k_x^2} - 1\right) \tag{79}$$

$$\Gamma_{\alpha\beta} = -\Gamma_{\beta\alpha} = \frac{i}{2k_{z\alpha}}\frac{d}{dz}k_{z\alpha}, \tag{80}$$

and for (64) it follows that from (76), (71) and (66)

$$\frac{d}{dz}\begin{bmatrix} a_\alpha \\ a_\beta \end{bmatrix} + \frac{\Gamma_{\alpha\beta}}{\Upsilon_1^2 + 1}\begin{bmatrix} \Upsilon_1 & 1 \\ -1 & \Upsilon_1 \end{bmatrix}\begin{bmatrix} a_\alpha \\ a_\beta \end{bmatrix} = 0. \tag{81}$$

Solutions for the amplitudes a_α and a_β can be obtained by S-matrix methods. We define the reflection point by z_0, i.e. $P(z_0) = 0$. z_2 is a point above, and z_1 a point below z_0 (imagine for example the ionosphere for a cold stratified plasma with z as the hight). The solutions connecting the waves at z_1 through the reflection point with the waves at z_2 can then be written as

$$\mathbf{a}(z_2) = \mathbf{S}(z_2, z_1) \cdot \mathbf{a}(z_1), \tag{82}$$

where $\mathbf{S}$ is the S-matrix. The diagonal elements are the reflection coefficients and the off-diagonal elements the transmission coefficients (Volland 1962*a*, (21); Budden 1985, (18.21)). In matrix theory the S-matrix is the matrizant, which is defined as (Gantmacher 1960)

$$\mathbf{S}(z_2, z_1) = \mathbf{I} + \int_{z_1}^{z_2} d\tau_1\, \mathbf{\Gamma}(\tau_1) + \int_{z_1}^{z_2} d\tau_1\, \mathbf{\Gamma}(\tau_1) \cdot \int_{z_1}^{\tau_1} d\tau_2\, \mathbf{\Gamma}(\tau_2) + \dots . \tag{83}$$

Since z_0 is a reflection point, $a_\alpha(z_2)$ and $a_\beta(z_2)$ are respectively the downgoing and upgoing modes above z_0, and $a_\alpha(z_1)$ and $a_\beta(z_1)$ are respectevely the upgoing and downgoing modes below z_0 (cf. Budden 1985, (18.20)).

For our purposes, it is sufficient to calculate the matrizant up to first order. We obtain

$$\mathbf{S}(z_2, z_1) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \frac{iA(z_2, z_1)}{4(1 + \Upsilon_1^2)}\begin{bmatrix} \Upsilon_1 & 1 \\ -1 & \Upsilon_1 \end{bmatrix}, \tag{84}$$

with

$$A(z_2, z_1) = \ln \frac{P(z_2) - (c_0^2/\omega^2)k_x^2}{P(z_1) - (c_0^2/\omega^2)k_x^2}. \tag{85}$$

If z_1 and z_2 are two points on the border of the coupling region, and the coupling region is assumed to be small, then $k_{z\alpha}^2(z_1) = k_{z\beta}^2(z_1) \approx k_{z\alpha}^2(z_2) = k_{z\beta}^2(z_2)$, and (85) can be written as

$$A(z_2, z_1) = \ln \frac{k_{z\alpha}^2(z_2)}{k_{z\alpha}^2(z_1)} = \ln \frac{k_{z\beta}^2(z_2)}{k_{z\beta}^2(z_1)} \approx \ln 1 = 0. \tag{86}$$

https://doi.org/10.1017/S0022377897005709 Published online by Cambridge University Press

We see from (84) that the reflection coefficients become 1 and the transmission coefficients 0, as is to be expected for the case of pure reflection.

To have an additional check on the general formalism represented in this paper, we want to investigate another effect, although we are in general not interested in what happens to the wave fields near the coupling points. The results obtained here are only qualitatively correct, because the geometrical optic conditions break down when we are too close to the coupling point. From the theory of propagation of waves in inhomogeneous media, the following curious behaviour is well known (Landau *et al.* 1984, Sec. 88). Assume that the reflection point is close to a point where $P = 0$. If the latter is defined at $z = 0$ then the wave penetrates beyond the reflection point z_0 (i.e. the point where (78) is satisfied), such that the wave field becomes singular as $z \to 0$. Then, if dissipation of energy is completely absent, the energy of the waves is infinite at $z = 0$. Hence, physically, an infinite amount of energy is transmitted through the point $z = z_0$ towards the point $z = 0$. We now want to investigate this effect. Multiplying (81) from the left by $(a_\alpha^*, -b_\beta^*)$, we obtain

$$\frac{d}{dz}\ln(|a_\alpha|^2 - |a_\beta|^2) = \frac{-2\Upsilon_1}{1+\Upsilon^2}\Gamma_{\alpha\beta}. \tag{87}$$

Since $P \geqslant 0$ in a plasma without absorption, for small P we obtain from (87)

$$(|a_\alpha|^2 - |a_\beta|^2) = \frac{\text{const}}{(k_{z\alpha}^2)^P}, \tag{88}$$

which is the difference between the incoming and reflected energies, i.e. the transmitted energy near the reflection point. Hence, at the reflection point, where $k_{z\alpha}^2 \to 0$, the transmitted energy becomes infinite, as is to be expected when there is no absorption (Landau *et al.* 1984, end of Sec. 88).

Appendix A

Equation (45) can be written with the help of (46) as

$$\mathrm{Det}\mathbf{M}(\mathbf{K};\mathbf{X}) = -\Theta(\lambda-\lambda_\alpha)(\lambda-\lambda_\beta)\prod_{\gamma\neq\alpha,\beta}(\lambda-\lambda_\gamma) = -\Theta(\lambda^2+B\lambda+C)\prod_{\gamma\neq\alpha,\beta}(\lambda-\lambda_\gamma), \tag{A 1}$$

where $B = B(\mathbf{K}_T;\mathbf{X})$ and $C = C(\mathbf{K}_T;\mathbf{X})$ are analytic. The roots λ_α and λ_β are

$$\lambda_{\substack{\alpha\\\beta}} = \tfrac{1}{2}[-B \pm (B^2-4C)^{1/2}]. \tag{A 2}$$

At the coupling point, we have

$$B^2 - 4C = 0. \tag{A 3}$$

Hence the coupling point is a branch point (for the example of a plane-stratified medium, see Budden 1972; Budden and Smith 1974; Budden 1985, Sec. 16.3). We have

$$\frac{\partial}{\partial\mathbf{K}_T}\lambda_{\substack{\alpha\\\beta}} = -\frac{1}{2}\frac{\partial B}{\partial\mathbf{K}_T} \pm \frac{1}{4}\left[\frac{\partial}{\partial\mathbf{K}_T}(B^2-4C)\right]\frac{1}{\lambda_\alpha-\lambda_\beta}, \tag{A 4}$$

from which (58) follows near the coupling point.

https://doi.org/10.1017/S0022377897005709 Published online by Cambridge University Press

References

Bernstein, I. B. and Friedland, L. 1983 *Basic Plasma Physics* I (ed. M. N. Rosenbluth and R. Z. Sagdeev), p. 367. North-Holland, Amsterdam.

Budden, K. G. 1972 *J. Atmos. Terr. Phys.* **34**, 1909–1921.

Budden, K. G. 1985 *The Propagation of Radio Waves*. Cambridge University Press.

Budden, K. G. and Smith, M. S. 1974 *Proc. R. Soc. Lond.* **A341**, 1–30.

Cairns, R. A. 1991 *Radiofrequency Heating of Plasma*. Adam Hilger, Bristol.

Cairns, R. A., Holt, H., McDonald, D. C., Taylor, M. and Lashmore-Davies, C. N. 1995 *Phys. Plasmas* **2**, 3702–3710.

Cairns, R. A. and Lashmore-Davies, C. N. 1983 *Phys. Fluids* **26**, 1268–1274.

Friedland, L. 1985 *Phys. Fluids* **28**, 3260–3268.

Friedland, L. 1990 *Phys. Fluids* **B2**, 1204–1209.

Friedland, L., Goldner, G. and Kaufman, A. N. 1987 *Phys. Rev. Lett.* **58**, 1392-1394.

Gantmacher, F. R. 1960 *Matrix Theory*, Vol. 2. Chelsea, New York.

Kaufman, A. N. and Friedland, L. 1987 *Phys. Lett.* **A123**, 387–389.

Kravtsov, Yu. A. 1969 *Soviet Phys. JETP* **28**, 769-772.

Landau, L. D., Lifschitz, E. M. and Pitaevski, L. P. 1984 *Electrodynamics of Continuous Media*. Pergamon, Oxford.

Littlejohn, R. G. and Flynn, W. G. 1993 *Phys. Rev. Lett.* **70**, 1799–1802.

McDonald, D. C., Cairns, R. A. and Lashmore-Davies, C. N. 1994 *Phys. Plasmas* **1**, 842–849.

Sabzevari, B. 1992 *J. Plasma Phys.* **47**, 49–60.

Sabzevari, B. 1993 *J. Plasma Phys.* **49**, 413–427.

Sabzevari, B. 1996 IPM-Report 96–175.

Stix, T. H. 1992 *Waves in Plasmas*. American Institute of Physics, New York.

Stix, T. H. & Swanson, D. G. 1983 *Basic Plasma Physics* I (ed. M. N. Rosenbluth and R. Z. Sagdeev), p. 335–366. North-Holland, Amsterdam.

Suchy, K. and Altman, C. 1975 *J. Plasma Phys.* **13**, 299–316.

Suchy, K. and Sabzevari, B. 1992*a* *Proceedings of URSI National Committee Germany. Kleinheubacher Berichte 1991*, Vol. 35, pp. 191-206. Telekom.

Suchy, K. and Sabzevari, B. 1992*b* *Proceedings of 2nd Symposium on Plasma Dynamics: Theory and Applications, Trieste* (ed. M. Tessarotto), pp. 147–156. University of Trieste.

Swanson, D. G. 1989 *Plasma Waves*. Academic Press, New York.

Volland, H. 1962*a* *Arch. Elektr. Ubertragung* **16**, 328–334.

Volland, H. 1962*b* *J. Atmos. Terr. Phys.* **24**, 853–856.

https://doi.org/10.1017/S0022377897005709 Published online by Cambridge University Press