Congruent reduction in geometric optics and mode conversion

Lazar Friedland[a] and Allan N. Kaufman
Lawrence Berkeley Laboratory and Physics Department, University of California, Berkeley, California 94720

(Received 7 April 1987; accepted 14 July 1987)

Standard eikonal theory reduces, to $N = 1$, the order of the system of equations underlying wave propagation in inhomogeneous plasmas. The condition for this remarkable reducibility is that only one eigenvalue of the *unreduced* $N \times N$ dispersion matrix $\mathbf{D}(k,x)$ vanishes at a time. If, in contrast, two or more eigenvalues of $\mathbf{D}$ become simultaneously small, the geometric optics reduction scheme becomes singular. These regions are associated with linear mode conversion and are described by higher-order systems. A new reduction scheme is developed based on congruent transformations of $\mathbf{D}$, and it is shown that, in *degenerate* regions, a partial reduction of order is still possible. The method comprises a constructive step-by-step procedure, which, in the most frequent (doubly degenerate) case, yields a second-order system, describing the pairwise mode conversion problem in four-dimensional plasmas. This $N = 2$ case is considered in detail, and dimensionality arguments are used in studying the characteristic ordering of the elements of the reduced dispersion tensor in mode conversion regions. The congruent reduction procedure is illustrated by classifying pairwise degeneracies in cold multispecies magnetized plasmas.

I. INTRODUCTION

Small amplitude waves in inhomogeneous plasmas are usually described by systems of linear equations for such quantities as electromagnetic fields $\mathbf{E}$, $\mathbf{B}$, perturbed average velocities $\mathbf{v}_\alpha$ of various species in fluid models, etc. The mathematical complexity of the problem caused by the multicomponent structure of the waves is typically resolved by using some sort of reduction (elimination) of several of the wave components from the problem. The conventional geometric optics theory for weakly varying plasmas,[1] for example, makes use of the reduced dielectric tensor, a 3×3 matrix describing the components of the electric field $\mathbf{E}$ alone. Such issues as the properties of the reduction scheme itself, its validity and uniqueness, are not addressed in most of the studies, the assumption being that no effects are introduced or omitted in the reduction process. This common opinion, however, is unjustified in some applications. Indeed, a simple observation shows that partial information is, in fact, lost during the reduction, since the eliminated wave components can be found only by reversing the reduction procedure, i.e., by using information clearly missing from the reduced equations. In addition, a more subtle fact is that the reduction of some of the components of the unreduced problem may result in singularities in reduced systems[2]; therefore, in avoiding the singularities, the final reduced wave components may not necessarily be those of the electric field. Such singular situations are characteristic of near degeneracies of the unreduced problem and are associated with such phenomena as resonances and mode conversion. Thus the basic goal of any general reduction theory must be the development of an algorithm for finding the optimally reduced system, describing the smallest number of the "irreducible" wave components. The ultimate system must avoid singular coefficients, have the lowest possible order, and at the same time preserve the important properties of the unreduced wave, such as, for example, the conservation of wave action.

The first general analysis of order reduction in the geometric optics of plasmas has been reported recently.[3] A constructive reduction scheme was suggested and applied to streaming magnetized plasmas. Although the proposed algorithm in many cases identifies the form and order of the reduced system with the smallest possible number of components, certain limitations are still present in the method. For example, that theory is one dimensional and is inapplicable to plasma regions characterized by small diagonal elements of the unreduced (or partially reduced) dispersion matrix. In the present work, we further develop the reduction theory, generalize it to arbitrary space and time varying plasmas and, by using a variational principle and congruent transformations, put it on a more solid mathematical basis, removing the above-mentioned limitation on the form of the dispersion tensor.

The scope of the paper is as follows. In Sec. II we shall consider a linear, homogeneous, integral *evolution equation* [see Eq. (4)] describing an N-component wave field $\mathbf{Z}(x)$ on space-time in a weakly varying plasma. On using a variational formulation of this equation and an eikonal representation for $\mathbf{Z}$, we shall obtain an Nth-order system of first-order PDE's (the *transport equation*) for the slow amplitude $\mathbf{A}(x)$ of the wave [Eq. (28)]. The remainder of Sec. II studies how a linear integral transformation $\mathbf{Q}(x,x')$ on $\mathbf{Z}$ affects the form and order of the transport equation. It will be shown that, to lowest order in the geometric optics expansion, this transformation leads to a transport equation for the transformed amplitude similar in form to the original trans-

[a] Permanent address: Center for Plasma Physics, Racah Institute of Physics, Hebrew University, Jerusalem, Israel.

port equation, but with the original dispersion tensor $\mathbf{D}(x,x')$ replaced by its *congruence transform*.

In Sec. III we shall prove the *reduction theorem*, i.e., we shall find the transformation that will accomplish the desired reduction of the order of the system. The basic idea is to transform $\mathbf{Z}$, systematically, in such a way that at each step one of the components of the transformed amplitude becomes small and is, in effect, *eliminated* from the problem. We shall find the desired *nonsingular* transformation, which, when applied, systematically annihilates the reducible wave components, thus reducing the order of the transport equation governing wave propagation in weakly varying plasmas. It will also be shown that this reduction procedure preserves the important physical information on the processes involved in the wave propagation and, in particular, both conserves and preserves wave-action flux, usually associated with energy conservation. Thus, because of its generality, this theory eliminates the need for making *ad hoc* assumptions (of questionable validity in certain parameter regions) when dealing with a specific problem.

In Sec. IV we shall discuss an important example of pairwise mode conversion, which occurs when two eigenvalues of the unreduced system are *near degenerate*. We shall demonstrate that the reduction method in this case (here denoted *normal degeneracy*) automatically describes the modes participating in the conversion process. The successive reduction steps yield in this case a characteristic, irreducible, second-order system of *coupled mode equations*. The reduction theory is thus of use in many plasma physics applications. Section V presents an example of such an application. The reduction algorithm is applied to classifying the possible normal degeneracies (pairwise mode conversion situations) in a cold multispecies magnetized plasma.

II. GENERAL CONGRUENCE TRANSFORMATIONS

In this section, we shall consider *general* linear transformations of the (complex valued) N-vector field $\mathbf{Z}(x)$ defined on $x=(\mathbf{r},t)$ and the resulting transformations of the evolution equation. In Sec. III, we shall then select that transformation which accomplishes the desired reduction.

Whatever physics is contained in the set of N components $\{Z_i(x)\}$ is equally contained in a different representation $\{\overline{Z}_i(x)\}$ if the two sets are related by an invertible integral transformation:

$$Z_i(x)=\int d^4x'\, Q_{ij}(x,x')\overline{Z}_j(x'), \qquad (1)$$

$$\overline{Z}_i(x)=\int d^4x'\, \widetilde{Q}_{ij}(x,x')Z_j(x'), \qquad (2)$$

with matrices $\mathbf{Q}$ and $\widetilde{\mathbf{Q}}$ satisfying

$$\int d^4x'\, Q_{ij}(x,x')\widetilde{Q}_{jm}(x',x'')=\delta_{im}\delta^4(x-x''). \qquad (3)$$

We now suppose that the original field $\mathbf{Z}(x)$ satisfies the homogeneous linear "evolution" equation

$$\int d^4x'\, D_{ij}(x,x')Z_j(x')=0, \qquad (4)$$

with the given Hermitian "dispersion" matrix $\mathbf{D}$:

$$D_{ij}(x,x')=D^*_{ji}(x',x). \qquad (5)$$

The evolution equation is equivalent to the variational principle $\delta S=0$ for the action functional

$$S(\mathbf{Z})\equiv\int d^4x\int d^4x'\, Z^*_i(x)D_{ij}(x,x')Z_j(x'). \qquad (6)$$

Let us now substitute the transformation (1) into the functional S, obtaining

$$S=\int d^4x\int d^4x'\, \overline{Z}^*_i(x)\overline{D}_{ij}(x,x')\overline{Z}_j(x'), \qquad (7)$$

where

$$\overline{D}_{ij}(x,x')\equiv\int d^4x''\int d^4x'''\, Q^*_{im}(x,x'')\times D_{mn}(x'',x''')Q_{nj}(x''',x'). \qquad (8)$$

This is a generalized congruence transformation. Since stationarity of S with respect to $\{Z_i(x)\}$ is equivalent to stationarity with respect to $\{\overline{Z}_i(x)\}$, we see that the form (7) yields the evolution equation

$$\int d^4x'\, \overline{D}_{ij}(x,x')\overline{Z}_j(x')=0, \qquad (9)$$

in place of (4). The purpose of the transformation $\mathbf{Q}$ is to obtain a transformed dispersion matrix $\overline{\mathbf{D}}$ such that (9) is easier to deal with than (4).

We now specialize to the case of a weakly nonuniform medium, such that the Weyl transform of the dispersion matrix,

$$D_{ij}(k,x)\equiv\int d^4s\, D_{ij}(x_1=x+\tfrac{1}{2}s,\, x_2=x-\tfrac{1}{2}s)\exp(-ik\cdot s), \qquad (10)$$

is a slowly varying function of x. [Note that we use the same notation for the two-point kernel $D_{ij}(x_1,x_2)$ and its transform $D_{ij}(k,x)$, distinguishing them by their different arguments.] By the rules of the Weyl calculus,[4] the transform of the congruence transformation (8) is

$$\overline{D}_{ij}(k,x)=Q^*_{im}(k,x)\exp[(i/2)\overleftrightarrow{L}]\times D_{mn}(k,x)\exp[(i/2)\overleftrightarrow{L}]Q_{nj}(k,x), \qquad (11)$$

in terms of the Weyl transform of $\mathbf{Q}$ and the Janus operator

$$\overleftrightarrow{L}\equiv\frac{\overleftarrow{\partial}}{\partial x}\cdot\frac{\overrightarrow{\partial}}{\partial k}-\frac{\overleftarrow{\partial}}{\partial k}\cdot\frac{\overrightarrow{\partial}}{\partial x}. \qquad (12)$$

Since we wish the transformed dispersion matrix $\overline{\mathbf{D}}(k,x)$ to inherit the desirable property of being slowly varying in x, we must require that the transformation matrix $\mathbf{Q}(k,x)$ itself be slowly varying. Let δ denote the small parameter associated with $\partial/\partial x$. Then to zeroth order in δ, (11) yields the *algebraic* transformation

$$\overline{D}_{ij}(k,x)=Q^*_{im}(k,x)D_{mn}(k,x)Q_{nj}(k,x) \qquad (13)$$

as the equivalent of the integral transformation (8). Note that (13) is *local* to the general phase-space point $z=(k,x)$.

If we should want the relation (11) to *first* order in δ, we find

$$\overline{D}_{ij}(z)=\overline{D}^{(0)}_{ij}(z)+(i/2)(\{Q^*_{im},D_{mn}\}Q_{nj}+Q^*_{im}\{D_{mn},Q_{nj}\}+\{Q^*_{im},Q_{nj}\}D_{mn}), \quad (14)$$

where $\overline{\mathbf{D}}^{(0)}$ is the right-hand side (rhs) of (13) and the Poisson bracket on phase-space functions $g(z)$ is

$$\{g_1,g_2\}\equiv\frac{\partial g_1}{\partial x}\cdot\frac{\partial g_2}{\partial k}-\frac{\partial g_1}{\partial k}\cdot\frac{\partial g_2}{\partial x}. \quad (15)$$

From the theory of Hermitian forms, we know that a Hermitian matrix $\mathbf{D}$ can be diagonalized by a unitary transformation, which is a special case of a congruence transformation (13). The diagonalizing matrix in this case is constructed from the eigenvectors of $\mathbf{D}$. The eigenvectors, however, may vary *rapidly* in cases when the eigenvalues of $\mathbf{D}$ are near degenerate. Then (13) is an invalid approximation to (11), even when the elements of $\mathbf{D}$ are slowly varying in x. Thus, if the ordering (the slow variation) is to be preserved, it is no longer possible to fully diagonalize $\mathbf{D}$ in general.

We now proceed to introduce the eikonal assumption, as a restriction on the allowed form of the field. In the original representation, we set

$$Z_i(x)=A_i(x)\exp[i\psi(x)], \quad (16)$$

where the amplitude $\mathbf{A}(x)$ and the wave vector $\kappa_\mu(x)=\partial\psi/\partial x^\mu$ are both slowly varying in x. Before substituting (16) into the action functional (6), it is convenient to express the latter in terms of the Wigner tensor

$$W_{ji}(k,x)\equiv\int d^4s\,Z_j(x_1=x+\tfrac{1}{2}s)\times Z^*_i(x_2=x-\tfrac{1}{2}s)\exp(-ik\cdot s); \quad (17)$$

then (6) reads as

$$S=\int d^4x\int\frac{d^4k}{(2\pi)^4}D_{ij}(k,x)W_{ji}(k,x). \quad (18)$$

We now express the Wigner tensor (17) in terms of the eikonal phase and amplitude, obtaining[5]

$$W_{ij}(k,x)=(2\pi)^4\delta^4[k-\kappa(x)]A_j(x)A^*_i(x)+\frac{i}{2}(2\pi)^4\frac{\partial\delta^4[k-\kappa(x)]}{\partial k_\mu}\times[(\partial_\mu A_j)A^*_i-A_j(\partial_\mu A^*_i)]+O(\delta^2) \quad (19)$$

as an asymptotic series in δ (we use $\partial_\mu\equiv\partial/\partial x^\mu$).

Using (19) in (18), we obtain the action functional S as an integral

$$S=\int d^4x\,\mathscr{L}(x) \quad (20)$$

of the Lagrangian density

$$\mathscr{L}(x)=\mathscr{L}^{(0)}(x)+\mathscr{L}^{(1)}(x)+O(\delta^2), \quad (21)$$

with

$$\mathscr{L}^{(0)}(x)=A^*_i(x)D_{ij}[\kappa(x),x]A_j(x), \quad (22)$$

$$\mathscr{L}^{(1)}(x)=\frac{i}{2}\frac{\partial D_{ij}}{\partial\kappa_\mu}[\kappa(x),x]\times[(\partial_\mu A^*_i)A_j-A^*_i(\partial_\mu A_j)]. \quad (23)$$

Stationarity of S with respect to $\{Z_i(x)\}$ implies stationarity with respect to $\{A_i(x)\}$ and $\psi(x)$, independently. Because the phase enters $\mathscr{L}$ only through its gradient $\kappa(x)$, the variation $\delta\psi(x)$ yields

$$\delta S=\int d^4x[\delta\kappa_\mu(x)]\frac{\partial\mathscr{L}}{\partial\kappa_\mu}=\int d^4x[\delta\psi(x)]d_\mu J^\mu(x), \quad (24)$$

where we have defined the wave-action four-flux

$$J^\mu(x)\equiv-\frac{\partial\mathscr{L}(x)}{\partial\kappa_\mu} \quad (25)$$

and use the notation

$$d_\mu=\frac{\partial}{\partial x^\mu}+\frac{\partial\kappa_\nu}{\partial x^\mu}\frac{\partial}{\partial\kappa_\nu}. \quad (26)$$

Stationarity then yields the wave-action conservation law

$$d_\mu J^\mu(x)=0 \quad (27)$$

associated, in time-independent cases, with energy conservation.

Variation of S with respect to the amplitude yields the *transport equation*

$$D_{ij}[\kappa(x),x]A_j(x)=\frac{i}{2}\left[\frac{\partial D_{ij}}{\partial\kappa_\mu}(\partial_\mu A_j)+d_\mu\left(\frac{\partial D_{ij}}{\partial\kappa_\mu}A_j\right)\right]+O(\delta^2). \quad (28)$$

It can be verified that the law (27), with the definition (25), is a consequence of (28); thus (28) can be considered as evolving both the amplitude and the phase.

We now consider the consequences of the transformation $\mathbf{Q}$. Because (7) has the same form as (6), *all* our results (16)–(28) are valid for the new barred fields. It remains to relate A_j, ψ, and J^μ to their barred counterparts.

We write $\overline{\mathbf{Z}}$ in the form

$$\overline{Z}_i(x)=\overline{A}_i(x)\exp[i\overline{\psi}(x)] \quad (29)$$

and substitute into (1):

$$A_i(x)\exp[i\psi(x)]=\int d^4x'\,Q_{ij}(x,x')\times\overline{A}_j(x')\exp[i\overline{\psi}(x')]. \quad (30)$$

Because, by assumption, the rapid variation is only in the phase factor, we may set

$$\overline{\psi}(x)=\psi(x), \quad (31)$$

i.e., to impose the invariance of the phase function [*and* its gradient $\kappa_\mu(x)$] under the transformation and use (31) to all orders, absorbing higher-order corrections in the transformation of the amplitude.

For the amplitude transformation, we use the inverse Weyl transform

$$Q_{ij}(x,x')=\int\frac{d^4\kappa}{(2\pi)^4}Q_{ij}\left[\kappa,\frac{1}{2}(x+x')\right]\times\exp[i\kappa\cdot(x-x')] \quad (32)$$

in (30). After some algebra, we obtain

$$A_i(x) = Q_{ij}[\kappa(x),x]\bar{A}_j(x) - \left(\frac{i}{2}\right)\left[d_\mu\left(\frac{\partial Q_{ij}}{\partial \kappa_\mu}\bar{A}_j\right) + \frac{\partial Q_{ij}}{\partial \kappa_\mu}\partial_\mu \bar{A}_j\right] + O(\delta^2). \quad (33)$$

[The resemblance of the rhs of (33) to (28) is not accidental; this formula is a property of the use of pseudodifferential operators.]

Next we examine the transformation of the wave-action four-flux ($\overline{25}$) (a barred equation number means that equation with all fields barred). The transformed zeroth-order wave-action is

$$\bar{J}^{(0)\mu}(x) \equiv -\frac{\partial \bar{\mathscr{L}}^{(0)}(x)}{\partial \kappa_\mu} = -\bar{A}_i^*\bar{A}_j \frac{\partial \bar{D}_{ij}^{(0)}}{\partial \kappa_\mu} \quad [\text{by } (\overline{28})].$$

Then, by (13) and (33), we have

$$\bar{J}^{(0)\mu}(x) = J^{(0)\mu}(x) - \frac{\partial Q_{im}^*}{\partial \kappa_\mu}\bar{A}_i^* D_{mn} A_n - \frac{\partial Q_{nj}}{\partial \kappa_\mu}\bar{A}_j A_m^* D_{mn}. \quad (34)$$

The last two terms in (34) are of $O(\delta)$ [see Eq. (28)] and thus, to zeroth order, the wave-action flux is invariant under the transformation. Therefore, the transformed wave-action flux not only satisfies the conservation law (27), but also preserves, to lowest order, its value. This latter feature is especially important when the transformation yields a reduced system (see Sec. III) since *fewer* wave components can then be used in evaluating the flux.

III. REDUCTION THEOREM

In this section we shall show how to select a transformation $\mathbf{Q}$ such that the dimensionality N of the wave field $\mathbf{A}$ and the transport equation (28) are effectively reduced to $N-1$. This process can then be repeated until the system is irreducible. The condition for reducibility is that at least one element of D_{ij} is of $O(1)$.

To illustrate the reduction idea, suppose $N=3$ (later, in Sec. V, we will use this example in an actual application) and let $D_{33} = O(1)$, while all other elements may be $O(1)$ or $O(\delta)$. Choose

$$\mathbf{Q} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -D_{31}/D_{33} & -D_{32}/D_{33} & 1 \end{pmatrix}. \quad (35)$$

Then, to zero order, by (13), we obtain

$$\bar{\mathbf{D}}^{(0)} = \begin{pmatrix} \bar{D}_{11}^{(0)} & \bar{D}_{12}^{(0)} & 0 \\ \bar{D}_{12}^{(0)*} & \bar{D}_{22} & 0 \\ 0 & 0 & D_{33} \end{pmatrix}, \quad (36)$$

which is block-diagonal, with

$$\bar{D}_{11}^{(0)} = D_{11} - |D_{13}|^2/D_{33},$$
$$\bar{D}_{22}^{(0)} = D_{22} - |D_{23}|^2/D_{33}, \quad (37)$$
$$\bar{D}_{12}^{(0)} = D_{12} - D_{13}D_{32}/D_{33}.$$

The transport equation ($\overline{28}$) then yields, for $i=3$,

$$D_{33}\bar{A}_3 = O(\delta), \quad (38)$$

from which we conclude that $A_3 = O(\delta)$. On the other hand, Eq. (33) (for $i=1,2$) yields $A_1 = \bar{A}_1 + O(\delta^2)$ and $A_2 = \bar{A}_2 + O(\delta^2)$. Thus the transformation (35) annihilates A_3 without affecting (to first order) the remaining two components of $\mathbf{A}$. Furthermore, since the transformed matrix (36), to zero order, is block diagonal, the transformed transport equation for components A_1, A_2 *decouples* from A_3. Indeed, if one defines a *reduced* vector

$$\mathbf{A}^r = \begin{pmatrix} \bar{A}_1 \\ \bar{A}_2 \end{pmatrix} = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} + O(\delta^2), \quad (39)$$

it is described by a transport equation similar in form to (28), with $\mathbf{D}$ replaced by the (2×2) *reduced* dispersion matrix

$$\mathbf{D}^r = \begin{pmatrix} \bar{D}_{11} & \bar{D}_{12} \\ \bar{D}_{12}^* & \bar{D}_{22} \end{pmatrix}. \quad (40)$$

Thus the problem is effectively reduced from $N=3$ to $N=2$.

Next, we suppose that all the diagonal elements D_{11}, D_{22}, D_{33} are $O(\delta)$, so that the preceding method cannot be used, but that at least one pair of off-diagonal elements, say D_{32}, D_{23}, is $O(1)$. Then we choose the constant transformation

$$\mathbf{Q} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \alpha \\ 0 & 0 & 1 \end{pmatrix}, \quad (41)$$

with the constant α chosen below. The transformed dispersion matrix has

$$\bar{D}_{33} = 2\,\mathrm{Re}(\alpha D_{32}) + D_{22} + |\alpha|^2 D_{33}. \quad (42)$$

If $\mathrm{Re}\,D_{32} = O(1)$, we choose $\alpha = 1$; if $\mathrm{Re}\,D_{32} = O(\delta)$, but $\mathrm{Im}\,D_{32} = O(1)$, we choose $\alpha = i$. Then $\bar{D}_{33} = O(1)$ and the prior procedure can be used, with $\mathbf{Q}$ given by Eq. (35) in terms of $\bar{\mathbf{D}}$, to reduce from $N=3$ to $N=2$. After this *second* transformation, the reduced dispersion matrix $\mathbf{D}^r$ has, to lowest order, $D_{22}^r = -|D_{23}|^2/\bar{D}_{33}$. Therefore, a *third* transformation, based on D_{22}^r being of $O(1)$, reduces the system from $N=2$ to $N=1$. Thus an off-diagonal element pair of $O(1)$ [with all diagonal elements of $O(\delta)$] allows reduction in N by 2.

Now, from case $N=3$, we proceed to general N and formulate the reduction theorem.

Reduction theorem: If there exists at least one large [of $O(1)$] element in the matrix $\mathbf{D}(k,x)$, characterizing the unreduced N-component (Nth-order) system, then one of the components of $\mathbf{A}$ can be eliminated from the problem in such a way that the remaining $N-1$ components of the wave are fully described by a reduced, Hermitian, $(N-1)\times(N-1)$ dispersion matrix $\mathbf{D}^r$, with the reduced transport equation of the form (28) and $\mathbf{D}$ replaced by $\mathbf{D}^r$.

Proof: (1) For general N suppose that a *diagonal* element D_{kk} of $\mathbf{D}$ is of $O(1)$. Then choose $\mathbf{Q}$ to be[6]

$$Q_{ij} = \delta_{ij} - \delta_{ik}(D_{kj}/D_{kk}) + \delta_{ik}\delta_{jk} \quad (43)$$

[see (35) for $N=3$]. From (33) we have

$$A_i(x) = \bar{A}_i(x) \quad (i \neq k) \quad (44)$$

and

$$A_k(x) = -\sum_{i\neq k} \frac{D_{ki}}{D_{kk}} A_i + \overline{A}_k + O(\delta), \tag{45}$$

from which we see that

$$\overline{A}_k = \sum_i \frac{D_{ki}}{D_{kk}} A_i + O(\delta). \tag{46}$$

However, from Eq. (28), $\Sigma_i D_{ki} A_i = O(\delta)$, so we see that the k th component $\overline{A}_k$ is of $O(\delta)$ in the new representation, while all the other components (44) are invariant under the transformation.

The formula for $\overline{\mathbf{D}}$ is, by (14),

$$\overline{D}_{ij} = D_{ij} - D_{ik}D_{kj}/D_{kk} + \delta_{ik}\delta_{jk}D_{kk} + O(\delta) \tag{47}$$

to zeroth order; thus $\overline{D}_{ik} = O(\delta)$ $(i\neq k)$ and the sum over i, j in the transformed Lagrangian density [Eqs. (21)–(23)], to first order, can be restricted to $i\neq k, j\neq k$. Therefore, the component $\overline{A}_k$ of the wave is effectively eliminated. The variation of the transformed Lagrangian then again yields the transport Eq. (28), describing the *reduced* $(N-1)$-component wave amplitude $\mathbf{A}^r = \{A_i,\ i\neq k\}$, characterized by the reduced $(N-1)\times(N-1)$ dispersion matrix

$$\mathbf{D}^r = \{\overline{D}_{ij};\ i,j\neq k\}$$
$$= D_{ij} - D_{ik}D_{kj}/D_{kk} - iD_{kk}P_{ij},\ i,j\neq k\ \text{[by (14)]}, \tag{48}$$

where P_{ij} is the Poisson bracket

$$P_{ij} = \{D_{ik}/D_{kk},\ D_{kj}/D_{kk}\}. \tag{49}$$

This result, for the 1-D case, was derived in Ref. 3. Since the wave-action flux is invariant to zeroth order, under the congruence transformation [see (34)] we have

$$J^{(0)\mu} = \overline{J}^{(0)\mu} = -\overline{A}_i^*\overline{A}_j \frac{\partial \overline{D}_{ij}}{\partial \kappa_\mu}$$
$$= -\sum_{i,j\neq k} A_i^* A_j \frac{\partial D^r_{ij}}{\partial \kappa_\mu} + O(\delta); \tag{50}$$

therefore, to lowest order, the *reduced* system preserves the information on the unreduced wave-action flux. The flux can be found from the reduced amplitude $\mathbf{A}^r$, while the effect of the eliminated component A_k on the flux is now carried by the reduced dispersion matrix $\mathbf{D}^r$.

(2) When *all* diagonal elements D_{ii} are of $O(\delta)$, but an off-diagonal pair, say D_{qr} and D_{rq} $(r\neq q)$, is of $O(1)$, the generalization of (41) is

$$Q_{ij} = \delta_{ij} + \alpha\delta_{ir}\delta_{jq}, \tag{51}$$

while $\overline{\mathbf{D}}$ is found to be

$$\overline{D}_{ij} = D_{ij} + \alpha D_{ir}\delta_{jq} + \alpha^* D_{rj}\delta_{iq} + |\alpha|^2\delta_{jq}\delta_{iq}D_{rr}. \tag{52}$$

Thus

$$\overline{D}_{qq} = D_{qq} + 2\,\mathrm{Re}(\alpha D_{qr}) + |\alpha|^2 D_{rr}, \tag{53}$$

which is $O(1)$, with α chosen as previously in case $N=3$. We then proceed as above to reduce first to $N-1$ (by eliminating A_q) and then further (by eliminating A_r) to $N-2$. These successive transformations annihilate the qth and rth components of $\mathbf{A}$, while the remaining $N-2$ components are invariant under the transformations. This completes the proof of the theorem.

The usefulness of the result just derived is that it provides a constructive step-by-step reduction scheme in weakly varying plasmas of arbitrary geometry. The method is consistent with the eikonal approximation in that it avoids singularly varying coefficients in eliminating "reducible" wave components, while at the same time keeping the basic first-order differential structure of the system. The successive application of the algorithm yields the final reduced Hermitian matrix $\mathbf{D}^f$ of rank $M\leqslant N$ such that *all* of its elements are of $O(\delta)$. In this case all M eigenvalues of the reduced tensor are small; any further attempt to reduce the system would yield singular coefficients, so that the system is *irreducible* within the geometric optics approximation. In the simplest and most frequent case, $M=1$ and thus all but one of the components of $\mathbf{A}$ are eliminated from the problem. This case describes the nondegenerate plasma, where only one of the eigenvalues of $\mathbf{D}$ is small. The final transport equation (28) in this situation is a *single*, first-order PDE for the remaining wave component A^f:

$$D^f A^f = \frac{i}{2}\left[\frac{\partial D^f}{\partial \kappa_\mu}(\partial_\mu A^f) + d_\mu\left(\frac{\partial D^f}{\partial \kappa_\mu}A^f\right)\right]. \tag{54}$$

This *scalar* equation can be solved perturbatively by the usual methods,[1] i.e., by integrating along the rays of geometric optics: $dx^\mu/d\sigma = -\partial D^f/\partial\kappa_\mu$, $d\kappa_\mu/d\sigma = \partial D^f/\partial x^\mu$.[1] Less frequent, but nevertheless important in applications, is the situation when $M=2$, in which case the final system comprises a set of *two coupled PDE's*. This corresponds to the pairwise linear mode conversion problem, the solution of which has been found recently[7] for a general geometry. Section IV describes this problem in more detail in the context of the reduction procedure just developed. The case $M>2$ seems to be less realistic for systems of finite degree of freedom. Nevertheless, multiple linear mode interaction may be important in kinetic problems, since they have infinite degrees of freedom. We shall consider that problem in future studies.

IV. NORMAL DEGENERACY AND PAIRWISE MODE CONVERSION

Because of both its importance and its complexity, the case when the described reduction procedure yields *two* coupled "irreducible" PDE's ($M=2$, see the end of Sec. III) requires further discussion. Thus we suppose that, after a number of reduction steps, the matrix $\mathbf{D}^f$ can be written as

$$\mathbf{D}^f = \begin{pmatrix} D_a & \eta \\ \eta^* & D_b \end{pmatrix}, \tag{55}$$

where all the elements are of $O(\delta)$, while we denote the reduced amplitude by $\mathbf{A}^f = (A_a, A_b)$. We shall now argue that the degeneracy of (55) is a rare situation, typically taking place in *small* plasma regions, where the elements D_a, D_b, and η of $\mathbf{D}^f$ usually have certain characteristic properties. Indeed, the degeneracy implies a simultaneous satisfaction of *three* conditions, i.e., D_a, D_b, and η to be of $O(\delta)$. Since the three elements are, typically, *independent* functions on 8-D phase space (k,x), satisfaction of all three conditions takes place on a 4-D subspace (η, in general, is complex) and therefore is a rare event. In order to illustrate the argument, a typical picture of the regions of smallness [of $O(\delta)$] of the matrix elements D_a, D_b, and η in the phase space is shown schematically in Fig. 1(a). The width of the regions in the x space is shown large compared to that in the k space in order to emphasize the weak *spatial* variation of the plasma parameters. It can be seen in Fig. 1 that, generally, the three smallness regions intersect in different locations in the phase space, basically because of their relative narrowness in the k space. This illustrates the improbability of a full degeneracy of the dispersion matrix in general. Similarly, a complete degeneracy of a 3×3 Hermitian dispersion matrix (case $M=3$), requiring a simultaneous satisfaction of *nine* smallness conditions on the 8-D phase space, is practically impossible.

Returning to the more realistic case of $M=2$, we now argue that the most probable scenario of such a degeneracy is that because of some special physical conditions as, for example, the existence of a global small parameter in the problem, one of the elements of $\mathbf{D}^f$ is of $O(\delta)$ in the *extended* region of the phase space. The other two become degenerate as before, essentially at a point away from which, being locally linear functions of k and x, they rapidly become of $O(1)$. In other words, the full degeneracy of the matrix may take place in regions, where one of the elements of $\mathbf{D}^f$ is *small* in its magnitude and a *weak* function of both k and x. Such a situation is illustrated in Fig. 1(b), where we have plotted, as in Fig. 1(a), the narrow (in k space) regions of smallness of D_a and D_b, and the wide (in k space) region of smallness of η (a more precise ordering of η and its derivatives, in this case, will be given later). The complete degeneracy of the dispersion matrix is more probable in this case [like the intersection of the three regions in Fig. 1(b)].

At this stage we shall assume that indeed one of the elements D_a, D_b, or η is small over an extended region of the phase space. Let us show that then there exists a remarkable difference between the cases when such a small element of $\mathbf{D}^f$ is off diagonal [element η (case A)] or diagonal [element D_a or D_b (case B)]. To illustrate the argument we refer to Figs. 2(a) and 2(b). The regions of smallness of the elements of $\mathbf{D}^f$ are shown in these illustrations for cases A and B. The shaded areas in Fig. 2 represent the phase-space regions, where Det $(\mathbf{D}^f) = D_aD_b - \eta^2$ is of $O(\delta)$. We can see that in case A, one finds *two* possible channels for satisfying the dispersion relation Det $(\mathbf{D}^f)=0$ in the nondegenerate regions, i.e., the regions where D_a and D_b are of $O(\delta)$. Each of the channels represents the possibility of propagation of *a distinct mode.* Indeed, since η is small over an extended region of the phase space, the dispersion relation implies that away from the degenerate point, one should have either

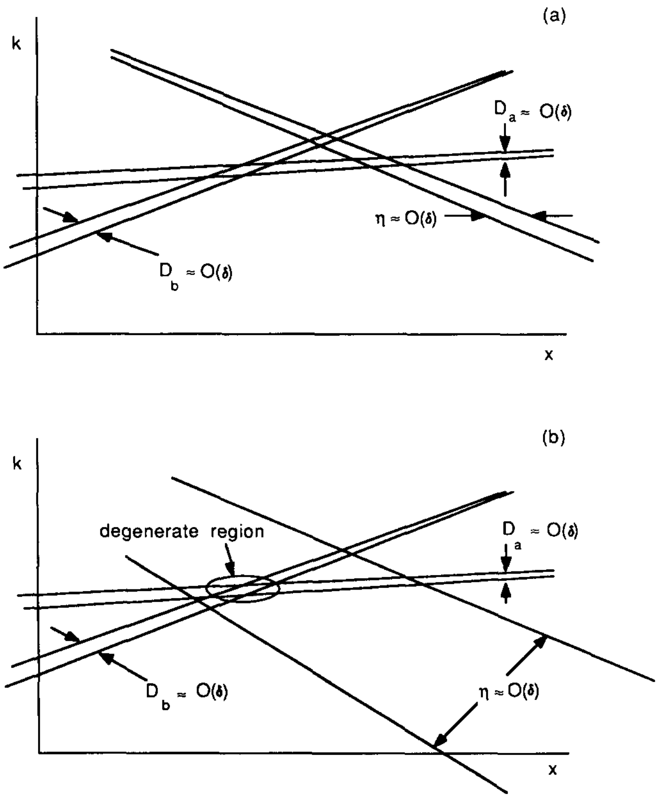

FIG. 1. Regions of smallness of the elements of the reduced dispersion matrix for case $M=2$. (a) All three elements of the matrix vary rapidly with **k**; complete degeneracy is not a characteristic of this case. (b) One of the elements of the dispersion matrix (η) is small in an extended region of the phase space. The possibility of a complete degeneracy is greatly enhanced.

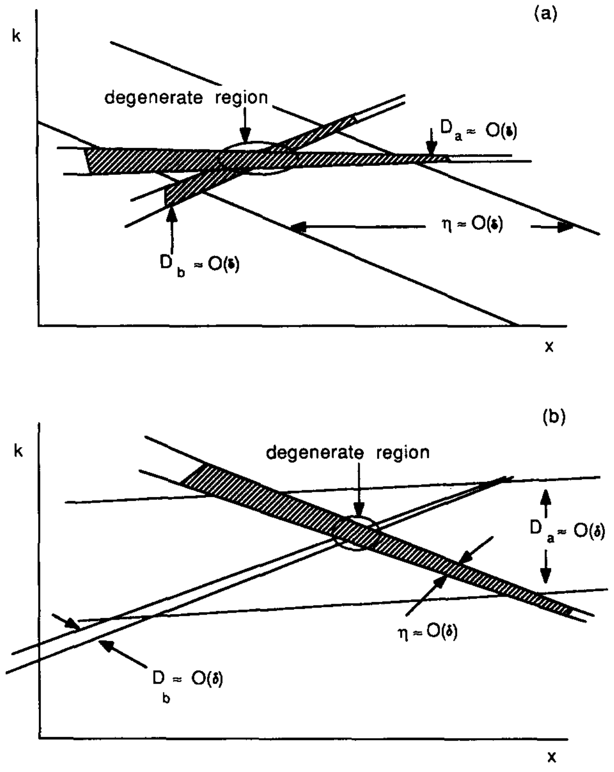

FIG. 2. Regions of smallness of the elements of the reduced dispersion matrix. The dashed areas correspond to the regions in phase space where the determinant of the dispersion matrix is $O(\delta)$. (a) Case A: η is small over an extended region of phase space. Two possible channels exist for satisfying the dispersion relation outside the degenerate region (the dashed areas). This is the normal degeneracy situation. (b) Case B: D_a is $O(\delta)$ over an extended region. Only one propagation channel in the nondegenerate region is available in this case.

$D_a = O(\delta)$, $D_b = O(1)$ or $D_a = O(1)$, $D_b = O(\delta)$. Thus one has a nondegenerate situation with two possible modes described essentially by the zeros of either D_a or D_b, with η serving as a small, nearly constant, mode coupling coefficient. This coupling is important only in the vicinity of the "crossing point" k_0, x_0 defined by $D_a(k_0,x_0) = D_b(k_0,x_0) = 0$. In case B [Fig. 2(b)], in contrast, only *one* channel for satisfying the dispersion relation $\mathrm{Det}(\mathbf{D}^f) = 0$ in the nondegenerate region exists, i.e., the region η of $O(\delta)$, allowing for propagation of only a *single mode*.

Thus we conclude that the most probable scenario for double degeneracy, accompanied by a possible local action flux redistribution between two modes (mode conversion), corresponds to case A. We shall refer to this scenario in the following as the "*normal degeneracy*."

The solution of the system of the two transport equations for the wave amplitude, in the case of the normal scenario of the pairwise degeneracy just described, has been presented in Ref. 7. It was shown that the action flux J_a associated with the dispersion relation $D_a = 0$ in the nondegenerate region is only partially transmitted through the degenerate region (the neighborhood of the crossing point); the transmission coefficient is

$$T = \exp(-2\pi|\eta|^2/|B|), \tag{56}$$

where B is the Poisson bracket

$$B = \frac{\partial D_a}{\partial k_\mu}\frac{\partial D_b}{\partial x^\mu} - \frac{\partial D_a}{\partial k_\mu}\frac{\partial D_b}{\partial x^\mu}. \tag{57}$$

At this point, we can refine the smallness conditions on η corresponding to the normal degeneracy. In such cases one typically follows a nondegenerate mode along a ray in the phase space, generated by the dispersion function $D_a = 0$. The normal degeneracy occurs in cases when η remains small along this nondegenerate ray during a time sufficiently long for crossing with the second mode (given by $D_b = 0$). Thus we require η to remain small along the rays, and so

$$\frac{d\eta}{d\sigma_{a,b}} = \{\eta, D_{a,b}\} = O(\delta^2), \tag{58}$$

where $\sigma_{a,b}$ are scalars parametrizing the rays and $\{\cdot,\cdot\}$ is the usual Poisson bracket. Equation (58) is the *normal degeneracy condition*.

Summarizing, we have shown that pairwise mode conversion events are typically associated with the *normal degeneracy* of the final 2×2 dispersion matrix ($M = 2$) and that the reduction algorithm, described in Sec. III, automatically provides the characteristic form of the dispersion matrix, describing two easily identifiable coupled modes associated with the diagonal elements of the matrix, while its nondiagonal element serves as the small mode coupling coefficient. These objects can then be used directly in the mode conversion theory[7] for calculating the transmission and mode conversion coefficients in cases of interest. Section V presents an example of the application of the reduction algorithm in classifying possible pairwise mode conversion situations in a cold plasma model.

V. CLASSIFICATION OF PAIRWISE MODE CONVERSION EVENTS IN COLD MULTISPECIES MAGNETIZED PLASMAS

We proceed from the conventional local plasma 3×3 dielectric tensor,[8] describing the three components of the electric field $\mathbf{E} = (E_x, E_y, E_z)$ in the wave:

$$\mathbf{D} = \begin{pmatrix} A & -iD & \gamma \\ iD & B & 0 \\ \gamma & 0 & C \end{pmatrix}, \tag{59}$$

where

$$\gamma = n^2 \cos\theta \sin\theta, \quad A = S - n^2\cos^2\theta, \\ B = S - n^2, \quad C = P - n^2 \sin^2\theta. \tag{60}$$

Here θ is the angle between the background magnetic field $\mathbf{B}_0 = B_0\mathbf{e}_z$ and $\mathbf{n} = c\mathbf{k}/\omega$; the axes are locally oriented so that $\mathbf{n}$ is in the xz plane. The coefficients in these equations are

$$S = (R+L)/2, \quad D = (R-L)/2,$$

$$R = 1 - \sum_k \frac{\omega_k^2}{\omega(\omega + \epsilon_k\Omega_k)}, \quad L = 1 - \sum_k \frac{\omega_k^2}{\omega(\omega - \epsilon_k\Omega_k)}, \tag{61}$$

$$P = 1 - \sum_k \frac{\omega_k^2}{\omega^2}, \quad \omega_k^2 = \frac{4\pi n_k Z_k^2 e^2}{m_k},$$

where Z_k, ϵ_k, m_k, and Ω_k are the charge number, its sign, the mass, and the absolute value of the local gyrofrequency for the several plasma species.

It should be emphasized at this point that the dispersion tensor (59) already describes a partially reduced problem. Indeed, the wave magnetic component was reduced by using Faraday's law $\mathbf{B} = c(\mathbf{k}\times\mathbf{E})/\omega$, which is a nonsingular step provided that $\mathbf{n}$ can be viewed as an object of $O(1)$, which we shall assume in the following. Furthermore, the denominators $\omega + \epsilon_k\Omega_k$ in Eq. (61) are the result of the elimination of the perturbed fluid velocities $\mathbf{v}_k$ of the various species. When some of these denominators are small [of $O(\delta)$] this reduction is unjustified within the eikonal approximation. The corresponding velocity components of the wave are irreducible and the elimination procedure should be applied to other wave components, such as the components of the electric field. Examples of the reduction in such cyclotron resonance situations, by proceeding from the unreduced dispersion matrix (a necessity in this case), can be found in Ref. 3. Furthermore, kinetic effects may be important at resonances; the study of these effects, however, is outside the scope of the present work. Thus we shall assume here that $\mathbf{n}$, R, L, and P are of $O(1)$, so that further reduction can indeed proceed from the partially reduced tensor (59).

We are interested in reducing the problem to a (2×2) case and thus make only one reduction step by using the algorithm of the reduction theorem. There are, in general, the following *four* possibilities of the reduction in the case of interest.

Case 1: A is of $O(1)$, so that the component E_x can be eliminated.

Case 2: B is of $O(1)$, so that one can eliminate E_y.

Case 3: C is of $O(1)$, so one can eliminate E_z.

Case 4: A, B, and C are all of $O(\delta)$, but either D or γ, or both, are of $O(1)$.

Let us proceed to cases 2–4 first.

Case 2: In this case the reduced dispersion matrix [see Eq. (48)] is

$$\mathbf{D}_2^r = \begin{pmatrix} A - D^2/B & \gamma \\ \gamma & C \end{pmatrix}. \tag{62}$$

Since we are interested in pairwise mode coupling situations, we shall now assume that matrix (62) is irreducible and therefore that all its elements are $O(\delta)$. Thus, following the discussion of Sec. IV, we interpret the situation as a coupling between the two modes

$$\begin{aligned} D_a &= A - D^2/B = 0 \\ &\Rightarrow (S - n^2)(S - n^2\cos^2\theta) - D^2 = 0, \\ D_b &= C = 0 \Rightarrow P - n^2\sin^2\theta = 0, \end{aligned} \tag{63}$$

while $\gamma = n^2\cos\theta\sin\theta$ is viewed as a small coupling coefficient. The normal degeneracy scenario implies then that γ is small in an extended region of the phase space, which corresponds to the following two situations. For case 2a, $\theta \approx O(\delta)$, with the plasma parameters varying mainly in the direction of the magnetic field $\mathbf{B}_0$ [then θ remains small in an extended plasma region and the normal degeneracy condition (58) is satisfied]. For case 2b, $\theta \approx \pi/2 + O(\delta)$ in plasmas varying primarily perpendicularly to $\mathbf{B}_0$. Thus a 1-D model can be used in describing case 2a, while case 2b generally requires a two-dimensional treatment. In case 2a, the local dispersion relations of the coupled modes are

$$n^2 = S \pm D = R,L, \quad P = 0, \tag{64}$$

describing parallel-propagating whistler modes of opposite circular polarizations and the electrostatic plasma mode. The coupling between the modes is caused by a small deviation from parallel propagation. This linear mode conversion phenomenon had been studied extensively in connection with the tripling effect in the ionosphere.[9] Case 2b corresponds to almost perpendicular propagation and the coupled modes are given by

$$n^2 = RL/S, \quad n^2 = P. \tag{65}$$

These are the ordinary and extraordinary modes; their coupling is caused by a small, but finite, departure from the perpendicular propagation. According to (65), the coupling takes place in plasma regions where $RL \approx PS$ and $P > 0$. These conditions are satisfied at frequencies just above Ω_i and at low plasma densities ($\omega_e < \omega$), corresponding to the boundary between regions 6a and 6b on the CMA diagram in Ref. 10.

Case 3: Here the reduced dispersion matrix [Eq. (48)] is

$$\mathbf{D}_3^r = \begin{pmatrix} A - \gamma^2/C & -iD \\ iD & B \end{pmatrix}. \tag{66}$$

Again, assuming the irreducibility of the matrix and the normal degeneracy scenario, we have two coupled modes described by the dispersion functions

$$\begin{aligned} D_a &= A - \gamma^2/C = 0 \Rightarrow n^2 = PS/(S\sin^2\theta + P\cos^2\theta), \\ D_b &= B = 0 \Rightarrow n^2 = S, \end{aligned} \tag{67}$$

while the coupling is caused by the small parameter D, i.e.,

$$\frac{R - L}{2} = \sum_k \frac{\epsilon_k \Omega_k \omega_k^2}{\omega(\omega^2 - \Omega_k^2)} \approx O(\delta). \tag{68}$$

Since this condition must be satisfied in an extended plasma region, we conclude that the three following possibilities exist.

Case 3a: $\omega/\min(\Omega_k) \approx O(\delta)$ and $\min(\omega_k)/\omega \approx O(1)$. For a two-component plasma, this situation corresponds to hydromagnetic (Alfvén) waves.

Case 3b: $\max(\Omega_k)/\omega = \Omega_e/\omega \approx O(\delta)$ and $\omega_e/\omega \approx O(1)$. This is the weak field case, describing almost isotropic and thus doubly degenerate plasma.

Case 3c: $\omega_e/\omega \approx O(\delta)$ and $\Omega_e/\omega \approx O(1)$. This is the low plasma density case characteristic of the edges of magnetized plasmas.

Note that, generally in case 3, we assume that in the first equation in (67), $S\sin^2\theta + P\cos^2\theta \neq 0$, which otherwise becomes the cold plasma resonance condition. Note, also, that at the crossing point ($D_a = D_b = 0$), the first equation in (67) becomes $n^2 = P$. Thus we conclude that case 3 is characteristic of plasma regions where two conditions are satisfied:

$$R = L + O(\delta), \quad P = S > 0. \tag{69}$$

We see that in this case the coupling coefficient does not depend on $\mathbf{k}$, so that in contrast to the basically 1-D or 2-D case 2, case 3 may describe a fully 3-D mode conversion situation.

Case 4: This is the simplest case, since the system can be reduced twice (see the end of the proof of the reduction theorem in Sec. III). For our 3×3 unreduced matrix, we thus arrive at a scalar, i.e., case 4 corresponds to a nondegenerate situation. Mode conversion is impossible in this case.

Finally, we return to case 1.

Case 1: Here the reduced dispersion matrix is

$$\mathbf{D}_1^r = \begin{pmatrix} B - D^2/A & -i\gamma D/A \\ i\gamma D/A & C - \gamma^2/A \end{pmatrix}. \tag{70}$$

The two coupled modes in this case are

$$D_a = BA - D^2 = 0, \quad D_b = CA - \gamma^2 = 0, \tag{71}$$

and the coupling is caused by the small coupling coefficient $\gamma D/A \approx O(\delta)$. Then two possibilities exist.

Case 1a: $\gamma \approx O(\delta)$ and $D \approx O(1)$. The second equation in (71) then yields $C = 0$ and therefore case 1a is identical to case 2 considered previously.

Case 1b: $D \approx O(\delta)$ and $\gamma \approx O(1)$. The first equation in (71) then yields $B = 0$, so that this case coincides with case 3.

Thus, case 1 does not introduce new mode conversion situations and we conclude that generally, for $\mathbf{n}$, R, L, P of $O(1)$ and away from the cold plasma resonances (the assumptions used in our analysis), there exist only three distinct normal pairwise degeneracies, each corresponding to a mode conversion situation, i.e., cases 2a, 2b, and 3. This number of degenerate situations is expected, of course, since there are only three pairwise degeneracies for the three eigenvalues of $\mathbf{D}^r$.

ACKNOWLEDGMENT

This research was supported by the U.S. Department of Energy, under Contract No. DE-AC03-76SF00098.

[1]I. B. Bernstein and L. Friedland, in *Handbook of Plasma Physics*, edited by M. Rosenbluth and R. Sagdeev (North-Holland, Amsterdam, 1983), Vol. 1, Chap. 2.5.

[2]L. Friedland, Phys. Fluids **28**, 3260 (1985).

[3]L. Friedland and G. Goldner, Phys. Fluids **29**, 4073 (1986).

[4]S. W. McDonald, Ph.D. thesis, University of California, Berkeley, 1983.

[5]A. N. Kaufman, H. Ye, and Y. Hui, Phys. Lett. A **120**, 327 (1987).

[6]Similar transformations are used in diagonalizing quadratic forms; see H. Margenau and G. M. Murphy, *The Mathematics of Physics and Chemistry* (Van Nostrand, New York, 1947), p. 307.

[7]L. Friedland, G. Goldner, and A. N. Kaufman, Phys. Rev. Lett. **58**, 1392 (1987).

[8]T. H. Stix, *The Theory of Plasma Waves* (McGraw-Hill, New York, 1962), p. 11.

[9]V. L. Ginzburg, *The Propagation of Electromagnetic Waves in Plasmas* (Pergamon, Oxford, 1964), p. 295.

[10]Reference 8, p. 30.