ARTICLES

Diagonalization of multicomponent wave equations with a Born-Oppenheimer example

Stefan Weigert and Robert G. Littlejohn
Department of Physics and Lawrence Berkeley Laboratory, University of California, Berkeley, California 94720
(Received 28 September 1992)

A general method to decouple multicomponent linear wave equations is presented. First, the Weyl calculus is used to transform operator relations into relations between c-number valued matrices. Then it is shown that the symbol representing the wave operator can be diagonalized systematically up to arbitrary order in an appropriate expansion parameter. After transforming the symbols back to operators, the original problem is reduced to solving a set of linear uncoupled *scalar* wave equations. The procedure is exemplified for a particle with a Born-Oppenheimer-type Hamiltonian valid through second order in $\hbar$. The resulting effective scalar Hamiltonians are seen to contain an additional velocity-dependent potential. This contribution has not been reported in recent studies investigating the adiabatic motion of a neutral particle moving in an inhomogeneous magnetic field. Finally, the relation of the general method to standard quantum-mechanical perturbation theory is discussed.

PACS number(s): 03.65.Sq, 03.40.Kf

I. INTRODUCTION

The spatiotemporal evolution of many physical systems is governed by linear multicomponent wave equations, the electromagnetic radiation field, and quantum-mechanical spinor wave functions being familiar examples. It is common to study the approximate behavior of solutions with short wavelengths since under this assumption, typically, the problems simplify considerably without losing their essential features. Possibly a close relationship to an underlying, more familiar theory may emerge; such a situation can yield valuable insight into the original theory, as is the case for wave optics versus geometrical optics and for quantum mechanics versus classical mechanics. Going beyond the lowest order in some appropriate expansion parameter is more or less straightforward in problems involving scalar waves whereas in the case of multicomponent fields already the first nontrivial order tends to become laborious. In studies of the adiabatic motion of a neutral particle subjected to an external magnetic field the question of higher-order terms arises naturally.

It is the purpose of this paper to present a systematic method to diagonalize Hermitian multicomponent wave operators up to *arbitrary* order in an appropriate ordering parameter. These results extend work done by Littlejohn and Flynn [1,2], who showed by making use of the Weyl calculus how to achieve the diagonalization up to the first nontrivial order; an investigation of electromagnetic wave propagation along similar lines has been performed up to first order by Brent and Fishman [3]. This approach turns out to be so powerful that at least formally (in other words, putting aside questions of convergence) an exact diagonalization of arbitrary Hermitian wave operators can be achieved.

II. THEORY

To start, an outline of the approach to diagonalize multicomponent wave equations is given—for a more detailed account of the theory the reader is referred to Littlejohn's and Flynn's work, to which terminology and notation of the present work are adapted.

Let the wave equation at stake be given by

$$D(\hat{\mathbf{q}},\hat{\mathbf{k}})\Psi=0\ , \tag{1}$$

where the ψ_α are the components of the A-dimensional wave field or "spinor" Ψ, and the $(A\times A)$ matrix $\hat{D}$ has elements $D_{\alpha\beta}(\hat{\mathbf{q}},\hat{\mathbf{k}})$. The operators $\hat{\mathbf{q}}$ and $\hat{\mathbf{k}}$ correspond to position and momentum, and they are assumed to fulfill the commutation relations $[\hat{q}_n,\hat{k}_m]=i\epsilon\delta_{nm}$ $(m,n=1,2,\ldots)$. The quantity ϵ denotes the ordering parameter and coincides in quantum-mechanical problems with $\hbar$. Often the position representation is particularly convenient; with $\hat{\mathbf{q}}\rightarrow\mathbf{q}$ and $\hat{\mathbf{k}}\rightarrow-i\epsilon\partial/\partial\mathbf{q}$ the linear operator $\hat{D}$ turns into a matrix of (pseudo) differential operators [4] coupling the components of $\Psi(\mathbf{q})$.

By means of the Weyl calculus [5] a one-to-one correspondence between operators and symbols can be set up which in the position representation reads

$$A(\mathbf{q},\mathbf{p})=\int d\mathbf{s}\langle\mathbf{q}-s/2|\,\hat{A}\,|\mathbf{q}+\mathbf{s}/2\rangle\exp[i\mathbf{s}\cdot\mathbf{k}/\epsilon]\ , \tag{2}$$

where $|\mathbf{q}\rangle$ are the elements of the position basis and $\hat{A}$ is

any operator. The operators $\hat{\mathbf{q}}$, $\hat{\mathbf{k}}$, and $D(\hat{\mathbf{q}},\hat{\mathbf{k}})$, in particular, are mapped into uniquely defined phase-space functions $\mathbf{q}$, $\mathbf{k}$, and $D(\mathbf{q},\mathbf{k})$. Effectively, symbols are c-number representations of operators with modified rules for forming products and, *a fortiori,* commutators. These rules keep track of the noncommutativity of the basic operators $\hat{\mathbf{q}}$ and $\hat{\mathbf{k}}$.

The Weyl correspondence has a number of appealing properties. For example, the Hermiticity of the wave operator, $\hat{D}^{\dagger}_{\alpha\beta}=\hat{D}_{\beta\alpha}$, is reflected in the Hermiticity of its symbol matrix, also called "dispersion tensor": $D^{\dagger}_{\alpha\beta}=D_{\beta\alpha}$. In addition (quantum-mechanical) expectation values of operators $\hat{A}$ turn into averages of phase-space functions $A(\mathbf{q},\mathbf{k})$ with respect to appropriate densities in phase space.

Basic to the following development is the Moyal formula [6] which relates the symbol of a product of operators $\hat{A}\hat{B}$ to the symbols of the operators $\hat{A}$ and $\hat{B}$. Explicitly, one has

$$A(\hat{\mathbf{q}},\hat{\mathbf{k}})B(\hat{\mathbf{q}},\hat{\mathbf{k}})\rightarrow A(\mathbf{q},\mathbf{k})e^{\epsilon\overleftrightarrow{\mathcal{L}}}B(\mathbf{q},\mathbf{k}) , \tag{3}$$

where the operator $\exp[\epsilon\overleftrightarrow{\mathcal{L}}]$ is defined via its Taylor expansion

$$e^{\epsilon\overleftrightarrow{\mathcal{L}}}=\sum_{l=0}^{\infty}\frac{\epsilon^l}{l!}\overleftrightarrow{\mathcal{L}}^{\,l}=\sum_{l=0}^{\infty}\frac{\epsilon^l}{l!}\left[\frac{i}{2}\right]^l(\overleftarrow{\partial}_{\mathbf{q}}\cdot\overrightarrow{\partial}_{\mathbf{k}}-\overleftarrow{\partial}_{\mathbf{k}}\cdot\overrightarrow{\partial}_{\mathbf{q}})^l , \tag{4}$$

and the arrows indicate that the partial derivatives $\partial_{\mathbf{q}}\equiv\partial/\partial\mathbf{q}$, . . . act on the factors on the left or on the right of it, respectively; the operator $\overleftrightarrow{\mathcal{L}}$ is a useful shorthand for the ordinary Poisson bracket $\{\ ,\ \}$, and the factor $i/2$ has been included for convenience.

The strategy used to diagonalize the operator-valued matrix $\hat{D}$ consists in introducing an operator-valued unitary matrix $\hat{U}$ such that

$$\hat{U}^{\dagger}\hat{D}\,\hat{U}=\hat{\Lambda} \tag{5}$$

holds, where $\hat{\Lambda}$ is required to have nonzero elements on the diagonal only. Having decoupled the individual "polarizations" one can proceed to determine the solutions $\Phi=\hat{U}^{\dagger}\Psi$ of the transformed wave operator $\hat{\Lambda}$ by standard methods. Finally, one obtains the sought-after wave fields Ψ of the operator $\hat{D}$ by writing

$$\Psi=\hat{U}\Phi . \tag{6}$$

The hard part of this program, clearly, consists of finding the operator $\hat{U}$ which diagonalizes the wave operator $\hat{D}$.

By making use of the Weyl correspondence rules the problem becomes tractable: having transcribed the operator relation Eq. (5) into a relation between fields of matrices defined all over phase space, one can resort to methods of linear algebra. This approach is based on the assumption that the symbols of the operators $\hat{U}$ and $\hat{\Lambda}$ can be expanded into power series of the ordering parameter ϵ, that is,

$$\hat{U}\rightarrow U(\mathbf{q},\mathbf{k})=\sum_{u=0}^{\infty}\epsilon^u U_u , \tag{7}$$

and, correspondingly,

$$\hat{\Lambda}\rightarrow\Lambda(\mathbf{q},\mathbf{k})=\sum_{\lambda=0}^{\infty}\epsilon^{\lambda}\Lambda_{\lambda} . \tag{8}$$

In general, such expansions will hold in large parts of phase space; the ordering in ϵ, however, breaks down in regions of "mode conversion" or "Landau-Zener coupling" [7] (cf. below). From now on it is assumed that either there are no such regions in phase space or that one stays away from them.

The requirement that $\hat{U}$ be a unitary operator imposes the following condition on the symbol matrices U_u:

$$I=U^{\dagger}e^{\epsilon\overleftrightarrow{\mathcal{L}}}U\equiv\sum_{M=0}^{\infty}\epsilon^M\sum_{\substack{l,u,v\\(u+l+v=M)}}\frac{1}{l!}U_u^{\dagger}\overleftrightarrow{\mathcal{L}}^{\,l}U_v , \tag{9}$$

with $u,l,v\geq 0$, using Eqs. (4) and (7); the Mth power of ϵ is multiplied by a sum of $(M+1)(M+2)/2$ terms. Comparing the various powers of ϵ on the left and right of Eq. (9) one finds

$$\begin{aligned} I&=U_0^{\dagger}U_0 ,\\ 0&=\sum_{\substack{l,u,v\\(u+l+v=M)}}\frac{1}{l!}U_u^{\dagger}\overleftrightarrow{\mathcal{L}}^{\,l}U_v,\quad M=1,2,\ldots . \end{aligned} \tag{10}$$

As a result, the lowest-order symbol U_0 has to be a *unitary* matrix. Decompose the symbol matrices U_u into two parts (A_u and B_u are Hermitian matrices)

$$U_u=U_0(A_u-iB_u),\quad u=1,2,\ldots , \tag{11}$$

where for later convenience the matrix U_0 has been factored out. In other words, the matrices $-iU_0^{\dagger}U_u$ have been decomposed into their Hermitian and anti-Hermitian parts. For a given $u\equiv M$ one finds that the matrix A_M is determined by all matrices $U_0,\ldots,U_{M-1}$, since each of Eqs. (10) can be written as

$$A_M=\tfrac{1}{2}(U_M^{\dagger}U_0+U_0^{\dagger}U_M)=\sum_{\substack{l,u,v\\(u+l+v=M)}}{}'\frac{1}{l!}U_u^{\dagger}\overleftrightarrow{\mathcal{L}}^{\,l}U_v ,\qquad M=1,2,\ldots , \tag{12}$$

where the prime on the sum indicates that the terms with $u=M$ or $v=M$ have to be left out. For $M=1$ this result reproduces correctly the first-order calculation [2]

$$A_1=-\tfrac{1}{2}U_0^{\dagger}\overleftrightarrow{\mathcal{L}}U_0=\frac{1}{4i}\{U_0^{\dagger},U_0\} . \tag{13}$$

No condition whatsoever on the matrices B_u follows from Eqs. (9); it will be shown momentarily that this freedom is (more than) sufficient in order to diagonalize the symbol D.

Instead of transcribing Eq. (5) directly into symbols (as Littlejohn and Flynn did) here the equivalent equation

$$\hat{D}\,\hat{U}=\hat{U}\,\hat{\Lambda} \tag{14}$$

is chosen as a more convenient starting point: no triple products of operators occur which would require the twofold application of the Moyal formula (3). The symbol equivalent of Eq. (14) is given by

$$D(\mathbf{q},\mathbf{k})e^{\overleftarrow{\epsilon\mathcal{L}}}U(\mathbf{q},\mathbf{k})=U(\mathbf{q},\mathbf{k})e^{\overleftarrow{\epsilon\mathcal{L}}}\Lambda(\mathbf{q},\mathbf{k}) , \tag{15}$$

or, using Eqs. (4), (7), and (8),

$$\sum_{M=0}^{\infty}\epsilon^M\left\{\sum_{\substack{l,u\\(u+l=M)}}\frac{1}{l!}D\overleftarrow{\mathcal{L}}^{\,l}U_u-\sum_{\substack{l,u,\lambda\\(u+l+\lambda=M)}}\frac{1}{l!}U_u\overleftarrow{\mathcal{L}}^{\,l}\Lambda_\lambda\right\}=0 , \tag{16}$$

with in total $(M+1)(M+4)/2$ terms in the sum multiplying the Mth power of ϵ. The lowest-order term implies

$$U_0^\dagger D\, U_0=\Lambda_0 , \tag{17}$$

thus, the matrix U_0 is fixed by the condition that it diagonalize D. Consequently, the columns of the matrix U_0, that is, $\tau_\alpha^{(\mu)}\equiv U_{0,\alpha\mu}$, are eigenvectors of the dispersion tensor D,

$$D\tau^{(\mu)}=\Lambda_0^{(\mu)}\tau^{(\mu)} , \tag{18}$$

with eigenvalues $\Lambda_0^{(\mu)}\equiv\Lambda_{0,\mu\mu}(\mathbf{q},\mathbf{k})$ which are functions on phase space. There are A eigenstates of the matrix D each of which is labeled by an index (μ) in parentheses in order to distinguish it from other labels which refer to the *components* of the vectors $\tau^{(\mu)}$. The unitarity of the matrix U_0 guarantees the normalization of the states $\tau^{(\mu)}$. In these terms, regions of mode conversion are characterized by eigenvalues $\Lambda_0^{(\mu)}$ and $\Lambda_0^{(\mu')}$ separated by a "distance" ϵ or less, or by rapidly (over a length scale ϵ) varying eigenvectors $\tau^{(\mu)}$ and $\tau^{(\mu')}$.

Solving the Mth of Eqs. (16) for Λ_M (which occurs exactly once) one finds

$$\Lambda_M=U_0^\dagger\left\{\sum_{\substack{u,l\\(u+l=M)\\(u\neq M)}}\frac{1}{l!}D\overleftarrow{\mathcal{L}}^{\,l}U_u-\sum_{\substack{l,u,\lambda\\(u+l+\lambda=M)\\(\lambda,u\neq M)}}\frac{1}{l!}U_u\overleftarrow{\mathcal{L}}^{\,l}\Lambda_\lambda\right\}+[\Lambda_0,A_M-iB_M] , \tag{19}$$

where Eq. (17) has been used, and the bracket [,] stands for the matrix commutator. Since the commutator of any matrix with a diagonal matrix is equal to a matrix with zero diagonal elements one realizes that the nonvanishing diagonal terms of Λ_M are completely determined by lower-order quantities $U_0,\ldots,U_{M-1}$, and $\Lambda_0,\ldots,\Lambda_{M-1}$, and by $D=U_0\Lambda_0U_0^\dagger$. The requirement that Λ_M be diagonal, in fact, is a condition on the as yet undetermined off-diagonal elements of the matrix B_M: one can solve for these matrix elements by setting the off-diagonal elements of the left-hand side of Eq. (19) equal to zero,

$$B_{M,\alpha\beta}=\frac{i}{\Lambda_M^{(\alpha)}-\Lambda_M^{(\beta)}}\{[A_M,\Lambda_0]-U_0^\dagger P\}_{\alpha\beta},\quad \alpha\neq\beta , \tag{20}$$

where the matrix in curly brackets of Eq. (19) is denoted by P, and the matrix Λ_0 is assumed to have no degenerate eigenvalues. As a result, all the elements of the matrices A_M and B_M are determined by the condition that U be a unitary matrix which diagonalize the dispersion tensor D up to the Mth order—all elements except those on the diagonal of B_M. It has been remarked by Littlejohn and Flynn [2] that the diagonal elements of the matrix [8] B_1 effect a phase transformation of the states $\tau^{(\mu)}$ only and, therefore, are physically not relevant. Their argument, however, is correct only through terms of first order in ϵ and, thus, cannot be applied here. For the time being, the ambiguity of the diagonal elements of B_M, thus, has no physical explanation; to achieve the diagonalization of D, it is sufficient and convenient [9] to choose $B_{M,\alpha\alpha}=0$ for all values of M.

Having diagonalized the symbol D of the wave operator $\hat{D}$ up to the desired order one can determine, at least in principle, the *operators* associated with the symbols on the diagonal of Λ by inverting the Weyl correspondence, Eq. (2). Then, one has to find an exact or approximate solution of the resulting scalar wave equations, and from the knowledge of the symbol U—which fixes the diagonalizing operator $\hat{U}$—one eventually will find the solutions of the original set of Eq. (1), using Eq. (6).

III. A BORN-OPPENHEIMER-TYPE EXAMPLE

A quantum system with a wave operator $\hat{D}=\hat{H}-E\hat{I}$ is considered as an example. The Hamiltonian operator $\hat{H}$ reads

$$H(\hat{\mathbf{q}},\hat{\mathbf{p}})=\frac{\hat{\mathbf{p}}^2}{2m}I+V(\hat{\mathbf{q}}) , \tag{21}$$

and the $A\times A$ unit matrix is denoted by I. The components of Ψ are coupled by the matrix $V(\hat{\mathbf{q}})$ which depends on the operators $\hat{\mathbf{q}}$ only. Such Hamiltonians arise, for example, in the Born-Oppenheimer treatment of molecules [10] (with A corresponding to the number of effectively coupled electronic states), or in the study of neutral particles with nonzero magnetic moment in external magnetic fields [11,12] (with $A=2s+1$ being the number of spin states). In the following, the Hamiltonian $H(\hat{\mathbf{q}},\hat{\mathbf{p}})$ will be diagonalized up to second order in ϵ which is to be identified with Planck's constant $\hbar$. It is assumed that either there are no mode conversion regions in the problem at hand, or that one stays away from such regions in phase space.

The nontrivial part of the Weyl transform of the wave operator $\hat{D}$ is given by that of the Hamiltonian

$$H(\mathbf{q},\mathbf{p})=\frac{\mathbf{p}^2}{2m}I+V(\mathbf{q}) . \tag{22}$$

To lowest order in ϵ, one has to determine the matrix U_0 which diagonalizes D, Eq. (17). Since the kinetic energy term is already diagonal, the eigenvectors of D are identical to those of the potential matrix $V(\mathbf{q})$,

$$V\tau^{(\mu)}=V_D^{(\mu)}\tau^{(\mu)},\quad \mu=1,2,\ldots,A , \tag{23}$$

which will be assumed to be nondegenerate: $V_D^{(\mu)}\neq V_D^{(\mu')}$ for $\mu\neq\mu'$. Note that the eigenvectors $\tau^{(\mu)}=\tau^{(\mu)}(\mathbf{q})$ depend on $\mathbf{q}$ only [13]. Consequently, the terms on the diagonal of Λ_0 read

$$\Lambda_0^{(\mu)} \equiv \Lambda_{0,\mu\mu} = \frac{\mathbf{p}^2}{2m} + V_D^{(\mu)}(\mathbf{q}) - E \ , \tag{24}$$

and the dispersion surfaces (representing the physically relevant parts of phase space) are defined by the condition $\Lambda_0^{(\mu)}=0$.

In the following the second and third terms in the expansion of the symbol

$$\Lambda = \Lambda_0 + \epsilon\Lambda_1 + \epsilon^2\Lambda_2 + O(\epsilon^3) \tag{25}$$

are calculated explicitly by evaluating Eq. (19) for $M=1$ and 2. The result is displayed in Eq. (44).

It is straightforward to calculate the first-order corrections to the diagonal terms of Λ_0 by writing down Eq. (19) for $M=1$,

$$\Lambda_1 = U_0^\dagger (D\overleftrightarrow{\mathcal{L}} U_0 - U_0 \overleftrightarrow{\mathcal{L}} \Lambda_0) + [\Lambda_0, A_1 - iB_1] \ . \tag{26}$$

Since U_0 is composed of the eigenvectors $\tau^{(\mu)}$ it depends on $\mathbf{q}$ only, and one finds in agreement with previous results [7]

$$\Lambda_1^{(\mu)} \equiv \Lambda_{1,\mu\mu} = \frac{i}{2m}(U_0^\dagger \{\mathbf{p}^2 I, U_0\})_{\mu\mu} + [\Lambda_0, A_1 - iB_1]_{\mu\mu} = \frac{-1}{m}\mathbf{p}\cdot\mathbf{A}^{(\mu)} \ , \tag{27}$$

where $\mathbf{A}^{(\mu)} \equiv \mathbf{A}_{\mu\mu}$ denotes the elements on the diagonal of

$$\mathbf{A}_{\alpha\beta} = i\tau^{(\alpha)*}\cdot\nabla\tau^{(\beta)} \equiv i\langle\alpha|\nabla|\beta\rangle \ ; \tag{28}$$

here the Dirac notation $|\alpha\rangle$ for the states $\tau^{(\alpha)}$ has been introduced.

In order to proceed to second order one has to determine the matrix $U_1 = U_0(A_1 - iB_1)$. Since $U_0 = U_0(\mathbf{q})$ it follows from Eq. (13) that $A_1 = 0$. The off-diagonal elements of the matrix B_1 are given by

$$B_{1,\alpha\beta} = \frac{i}{m}\frac{\mathbf{p}\cdot\mathbf{A}_{\alpha\beta}}{\Lambda_0^{(\alpha)} - \Lambda_0^{(\beta)}}, \quad \alpha \neq \beta \ , \tag{29}$$

and the diagonal elements of this matrix had been chosen to be zero: $B_{1,\mu\mu}=0$.

The second-order correction Λ_2 is obtained from Eq. (19) for $M=2$,

$$\Lambda_2 = U_0^\dagger \{\tfrac{1}{2}(D\overleftrightarrow{\mathcal{L}}^2 U_0 - U_0\overleftrightarrow{\mathcal{L}}^2\Lambda_0) + (D\overleftrightarrow{\mathcal{L}} U_1 - U_1\overleftrightarrow{\mathcal{L}}\Lambda_0 - U_0\overleftrightarrow{\mathcal{L}}\Lambda_1) - U_1\Lambda_1\} + [\Lambda_0, A_2 - iB_2] \ . \tag{30}$$

Using the explicit form of the Hamiltonian, Eq. (22), and Eq. (24) one obtains

$$\Lambda_2 = U_0^\dagger \left\{ \frac{1}{4m}(\mathbf{p}^2 I\overleftrightarrow{\mathcal{L}}^2 U_0 - U_0\overleftrightarrow{\mathcal{L}}^2\mathbf{p}^2 I) + \frac{1}{2m}(\mathbf{p}^2 I\overleftrightarrow{\mathcal{L}} U_1 - U_1\overleftrightarrow{\mathcal{L}}\mathbf{p}^2 I) - U_0\overleftrightarrow{\mathcal{L}}\Lambda_1 + V\overleftrightarrow{\mathcal{L}} U_1 - U_1\overleftrightarrow{\mathcal{L}} V_D \right\} + iB_1\Lambda_1 + [\Lambda_0, A_2 - iB_2] \ , \tag{31}$$

where V_D is obtained from diagonalizing the potential matrix V, and $U_1 = -iU_0B_1$ has been used.

First of all, the terms containing the square of $\overleftrightarrow{\mathcal{L}}$, as a matter of fact, do not contribute since they exactly cancel each other,

$$\begin{aligned}(\mathbf{p}^2 I\overleftrightarrow{\mathcal{L}}^2 U_0)_{\alpha\beta} &= (\partial_{p_\rho}\partial_{p_\sigma}\mathbf{p}^2\delta_{\alpha\gamma})(\partial_{q_\rho}\partial_{q_\sigma}U_{0,\gamma\beta}) \\ &= (\partial_{q_\rho}\partial_{q_\sigma}U_{0,\alpha\gamma})(\partial_{p_\rho}\partial_{p_\sigma}\mathbf{p}^2\delta_{\gamma\beta}) \\ &= (U_0\overleftrightarrow{\mathcal{L}}^2\mathbf{p}^2 I)_{\alpha\beta} \ .\end{aligned} \tag{32}$$

Here and in the following summation over indices occurring twice is assumed implicitly; bracketed indices ( ), however, are excluded from the summation convention.

Next, a similar argument shows that the remaining terms with a factor $\mathbf{p}^2$ give *identical* contributions. Using $U_{1,\rho\sigma} = -iU_{0,\rho\tau}B_{1,\tau\sigma}$ one obtains

$$\begin{aligned}\frac{i}{2}\{\mathbf{p}^2 I, U_1\}_{\alpha\beta} &= \frac{1}{2i}\frac{\partial}{\partial p_\lambda}(\mathbf{p}^2\delta_{\alpha\gamma})\frac{\partial U_{1,\gamma\beta}}{\partial q_\lambda} \\ &= -B_{1,\sigma\beta}\mathbf{p}\cdot\nabla U_{0,\alpha\sigma} - U_{0,\alpha\sigma}\mathbf{p}\cdot\nabla B_{1,\sigma\beta} \ .\end{aligned} \tag{33}$$

Multiplication by $U_0^\dagger$ from the left yields

$$\frac{i}{2}(U_0^\dagger\{\mathbf{p}^2 I, U_1\})_{\alpha\beta} = \frac{-1}{m}\left[\sum_{\gamma\,(\neq\beta)}^{A}\frac{(\mathbf{p}\cdot\mathbf{A}_{\alpha\gamma})(\mathbf{p}\cdot\mathbf{A}_{\gamma\beta})}{\Lambda_0^{(\gamma)} - \Lambda_0^{(\beta)}} + m\mathbf{p}\cdot\nabla B_{1,\alpha\beta}\right] , \tag{34}$$

so that the contribution to the diagonal of Λ_2 is given by

$$\frac{i}{4m}[U_0^\dagger(\{\mathbf{p}^2 I, U_1\} - \{U_1, \mathbf{p}^2 I\})]_{\mu\mu} = \frac{-1}{m^2}\sum_{\gamma\,(\neq\mu)}^{A}\frac{|\mathbf{p}\cdot\mathbf{A}_{\gamma\mu}|^2}{\Lambda_0^{(\gamma)} - \Lambda_0^{(\mu)}} \ . \tag{35}$$

Then, the fifth term in Eq. (31) follows from multiplying

$$-\frac{i}{2}\{U_0, \Lambda_1\}_{\alpha\beta} = \frac{i}{2m}\mathbf{A}^{(\beta)}\cdot\nabla\tau_\alpha^{(\beta)} \tag{36}$$

from the left with $U_0^\dagger$ leading to

$$-\frac{i}{2}(U_0^\dagger\{U_0, \Lambda_1\})_{\alpha\beta} = \frac{1}{2m}\mathbf{A}^{(\beta)}\cdot\mathbf{A}_{\alpha\beta} \ , \tag{37}$$

with diagonal elements

$$-\frac{i}{2}(U_0^\dagger\{U_0, \Lambda_1\})_{\mu\mu} = \frac{1}{2m}\mathbf{A}^{(\mu)}\cdot\mathbf{A}^{(\mu)} \ . \tag{38}$$

Only one of the remaining two terms linear in $\mathbf{p}$ gives a nonzero contribution. The second one, stemming from

$$\frac{-i}{2}(U_0^\dagger\{U_1, V_D\})_{\alpha\beta} = \frac{-1}{2}\frac{\partial B_{1,\alpha\beta}}{\partial\mathbf{p}}\cdot\frac{\partial V_D^{(\beta)}}{\partial\mathbf{q}} \ , \tag{39}$$

is equal to zero since $B_{1,\mu\mu}=0$. The other term can be written as

$$\frac{i}{2}(U_0^\dagger\{V, U_1\})_{\alpha\beta} = \frac{1}{2}\tau_\gamma^{(\alpha)*}U_{0,\sigma\rho}\frac{\partial V_{\gamma\sigma}}{\partial\mathbf{q}}\frac{\partial B_{1,\rho\beta}}{\partial\mathbf{p}} \ . \tag{40}$$

Using Eq. (23) one finds

$$\tau_\gamma^{(\alpha)*}\tau_\sigma^{(\rho)}\nabla V_{\gamma\sigma}=\tau_\gamma^{(\alpha)*}[\nabla(V_{\gamma\sigma}\tau_\sigma^{(\rho)})-V_{\gamma\sigma}\nabla\tau_\sigma^{(\rho)}]$$
$$=\delta_{\alpha\rho}\nabla V_D^{(\rho)}+(V_D^{(\rho)}-V_D^{(\alpha)})\langle\alpha|\nabla|\rho\rangle\ , \quad (41)$$

so that

$$\frac{i}{2}(U_0^\dagger\{V,U_1\})_{\alpha\beta}=\frac{1}{2m}\frac{V_D^{(\gamma)}-V_D^{(\alpha)}}{V_D^{(\gamma)}-V_D^{(\beta)}}\mathbf{A}_{\alpha\gamma}\cdot\mathbf{A}_{\gamma\beta}$$
$$+\frac{1}{2}\nabla V_D^{(\alpha)}\frac{\partial B_{1,\alpha\beta}}{\partial\mathbf{p}}\ , \quad (42)$$

and thus for $\alpha=\beta=\mu$

$$\frac{i}{2}(U_0^\dagger\{V,U_1\})_{\mu\mu}=\frac{1}{2m}\sum_{\gamma\ (\neq\mu)}^{A}|\mathbf{A}_{\gamma\mu}|^2\ . \quad (43)$$

The eighth term of the right-hand side of Eq. (31) does not contribute since it is the product of the diagonal matrix Λ_1 with B_1 which was assumed to be zero on the diagonal, and the last term of Eq. (31) vanishes on the diagonal anyway.

The final result for the second-order dispersion tensor is obtained by collecting all the terms of Eqs. (24), (27), (35), (38), and (43),

$$\Lambda^{(\mu)}=\Lambda_0^{(\mu)}+\epsilon\Lambda_1^{(\mu)}+\epsilon^2\Lambda_2^{(\mu)}+O(\epsilon^3)$$
$$=\frac{1}{2m}(\mathbf{p}-\epsilon\mathbf{A}^{(\mu)})^2+V_D^{(\mu)}-E+\frac{\epsilon^2}{2m}\sum_{\gamma\ (\neq\mu)}^{A}|\mathbf{A}_{\gamma\mu}|^2$$
$$-\frac{\epsilon^2}{m^2}\sum_{\gamma\ (\neq\mu)}^{A}\frac{|\mathbf{p}\cdot\mathbf{A}_{\gamma\mu}|^2}{V_D^{(\gamma)}-V_D^{(\mu)}}+O(\epsilon^2)\ . \quad (44)$$

In this expression first- and second-order terms combine to formally reproduce the structure of a Hamiltonian of a particle in fictitious electromagnetic fields with vector and scalar potentials—apart from the last term which can be interpreted as an additional momentum-dependent potential. It originates from the first-order off-diagonal terms.

The transformation of the symbol $\Lambda^{(\mu)}$ back to an operator $\hat{\Lambda}^{(\mu)}$ is achieved by using the following correspondence rules [14], obtained from inverting Eq. (2):

$$\mathbf{q}\to\hat{\mathbf{q}},\quad \mathbf{p}\to\hat{\mathbf{p}}\ ,$$
$$\mathbf{p}f(\mathbf{q})\to\tfrac{1}{2}[\hat{\mathbf{p}}f(\hat{\mathbf{q}})+f(\hat{\mathbf{q}})\hat{\mathbf{p}}]\ , \quad (45)$$
$$\mathbf{p}^2f(\mathbf{q})\to\tfrac{1}{4}[\hat{\mathbf{p}}^2f(\hat{\mathbf{q}})+2\hat{\mathbf{p}}f(\hat{\mathbf{q}})\hat{\mathbf{p}}+f(\hat{\mathbf{q}})\hat{\mathbf{p}}^2]\ ,$$

which come down to "symmetrizing" the classical expressions in $\mathbf{q}$ and $\mathbf{p}$ and subsequently replacing the canonical variables by the operators $\hat{\mathbf{q}}$ and $\hat{\mathbf{p}}$. The operators on the diagonal of the transformed wave operator read

$$\hat{\Lambda}^{(\mu)}=\frac{1}{2m}[\hat{\mathbf{p}}-\epsilon\mathbf{A}^{(\mu)}(\hat{\mathbf{q}})]^2+V_D^{(\mu)}(\hat{\mathbf{q}})-E\hat{I}$$
$$+\frac{\epsilon^2}{2m}\sum_{\gamma\ (\neq\mu)}^{A}|\mathbf{A}_{\gamma\mu}(\hat{\mathbf{q}})|^2$$
$$-\frac{\epsilon^2}{m^2}\sum_{\gamma\ (\neq\mu)}^{A}G_{\gamma\mu}(\hat{\mathbf{q}},\hat{\mathbf{p}})+O(\epsilon^3)\ , \quad (46)$$

where

$$G_{\gamma\mu}(\hat{\mathbf{q}},\hat{\mathbf{p}})=\tfrac{1}{4}\sum_{j,k=1}^{3}[\hat{p}_j\hat{p}_kg_{jk}^{\gamma\mu}(\hat{\mathbf{q}})+\hat{p}_jg_{jk}^{\gamma\mu}(\hat{\mathbf{q}})\hat{p}_k$$
$$+\hat{p}_kg_{jk}^{\gamma\mu}(\hat{\mathbf{q}})\hat{p}_j+q_{jk}^{\gamma\mu}(\hat{\mathbf{q}})\hat{p}_j\hat{p}_k]\ , \quad (47)$$

with the abbreviation

$$g_{jk}^{\gamma\mu}(\hat{\mathbf{q}})=\frac{A_{\gamma\mu,j}(\hat{\mathbf{q}})A_{\gamma\mu,k}(\hat{\mathbf{q}})}{V_D^{(\gamma)}(\hat{\mathbf{q}})-V_D^{(\mu)}(\hat{\mathbf{q}})}\ . \quad (48)$$

It is remarkable that one obtains a formally identical second-order contribution from a perturbational analysis of a model where both spin and particle motion are treated completely classically [15]. Comparing this result with calculations done for a neutral particle with magnetic moment moving in an external magnetic field [11,12] shows that the last term of Eq. (46) in this context has not been reported earlier, although it follows necessarily in the present systematic procedure. The actual size of the various terms in Eq. (46) is discussed in detail for the motion of a neutral particle with magnetic moment in an inhomogeneous magnetic field [15].

IV. RELATION TO QUANTUM PERTURBATION THEORY

The structure of the first- and second-order terms in the final expression of the diagonalized symbol, Eq. (44), is similar to that of the familiar formula of quantum-mechanical perturbation theory according to Rayleigh and Schrödinger. In an M-dimensional Hilbert space the first corrections to the nth energy eigenvalue $E_n^{(0)}$ of the unperturbed Hamiltonian [16] $\hat{H}_0$ read

$$\tilde{E}_n(\epsilon)=\tilde{E}_n^{(0)}+\epsilon\tilde{E}_n^{(1)}+\epsilon^2\tilde{E}_n^{(2)}+O(\epsilon^3)$$
$$=\tilde{E}_n^{(0)}+\epsilon V_{nn}+\epsilon^2\sum_{m\ (\neq n)}^{M}\frac{|V_{mn}|^2}{\tilde{E}_m^{(0)}-\tilde{E}_n^{(0)}}+O(\epsilon^3)\ , \quad (49)$$

where V_{mn} are the matrix elements of a perturbing potential $\hat{V}$ in the unperturbed basis with states $|n\rangle$.

For a comparison of the expressions stemming from multicomponent wave equations and standard quantum-mechanical perturbation theory it is useful to present the latter one in terms of matrix notation [17]. In this variation of the common formulation all states are displayed at the same time [18], and one writes down Schrödinger's equation with respect to the eigenfunctions of the unperturbed Hamiltonian $\hat{H}_0$ in the compact form

$$H\,U=U\,E\ , \quad (50)$$

U being a unitary matrix as before and E being a diagonal matrix. Assuming that the Hamiltonian is given in the form

$$H=\sum_n\lambda^nH_n\ , \quad (51)$$

and that the expansions

$$E=\sum_n \lambda^n E_n \ , \qquad U=\sum_n \lambda^n U_n \tag{52}$$

hold, one finds the following expressions for the corrections of the eigenvalues:

$$E_0=U_0^\dagger H_0 U_0 \ ,$$
$$E_1=U_0^\dagger H_1 U_0+[E_0,U_0^\dagger U_1] \ , \tag{53}$$
$$E_2=U_0^\dagger H_2 U_0+U_0^\dagger H_1 U_1-U_0^\dagger U_1 E_1+[E_0,U_0^\dagger U_2] \ ,$$

etc., which should be compared to Eqs. (17), (26), and (30). One realizes that, actually, the same mechanism as in the multicomponent problem leads to the determination of the matrices U_n and E_n: writing $U_n=U_0(A_n-iB_n)$ one sees that (i) all matrices A_n are fixed by the unitarity of U, (ii) the off-diagonal elements of the matrix B_n will have factors $(\tilde{E}_m^{(0)}-\tilde{E}_n^{(0)})^{-1}$, and that (iii) the nth-order correction E_n is completely determined by matrices $U_{n'}$ and $E_{n'}$ with $n'<n$. Apart from this similarity, the remaining terms multiplying the various powers of ϵ in Eqs. (50) and (16), respectively, are quite different—for example, the number of terms is different, especially, if the nondiagonal part of the symbol equation (22) depends on $\mathbf{p}$, too. Furthermore, all terms of this equation typically depend on the phase-space variables throughout. Nevertheless, it is interesting to ask whether the established formal similarity is sufficient to transfer results concerning the convergence of Rayleigh-Schrödinger perturbation theory [19] to the present treatment of multicomponent wave equations. To do this, one has to write the relation (14) between symbols in the form of Eq. (50), that is,

$$D(\epsilon)U(\epsilon)=U(\epsilon)\Lambda(\epsilon) \ . \tag{54}$$

Unfortunately, due to the different origins of Eqs. (53) and their counterparts Eqs. (17), (26), and (30), this is possible only if in the expansion

$$D=\sum_n \lambda^n D_n \tag{55}$$

one would allow the terms D_n to depend on the matrices $U_{n'}$, $(n'<n)$, i.e.,

$$D_1=(D\overleftrightarrow{\mathcal{L}}U_0-U_0\overleftrightarrow{\mathcal{L}}\Lambda_0)U_0^\dagger \ ,$$
$$D_2=-D_1U_1U_0^\dagger+\tfrac{1}{2}(D_0\overleftrightarrow{\mathcal{L}}^2U_0-U_0\overleftrightarrow{\mathcal{L}}^2\Lambda_0)$$
$$+(D_0\overleftrightarrow{\mathcal{L}}U_1-U_1\overleftrightarrow{\mathcal{L}}\Lambda_0-U_0\overleftrightarrow{\mathcal{L}}\Lambda_1) \ . \tag{56}$$

This, however, is an artificial approach; and, as a most serious drawback, arguments pertaining to questions of convergence of the perturbation series *cannot* be settled simply by reference to Rayleigh-Schrödinger theory in this way, due to the explicit appearance of lower-order corrections in the higher-order perturbation terms.

Equations (45) are more general than Eqs. (53) in another, previously mentioned sense: since the matrices in Eq. (14) depend on the dynamical variables of phase space, $\mathbf{q}$ and $\mathbf{p}$, one actually deals with a family of similar equations parametrized by points of phase space, whereas usually the energy eigenvalues depend on a small number of parameters only (as, for example, on the components of the electric field $\mathbf{E}$ in the Stark effect). All quantities being now phase-space functions means that convergence properties will vary from point to point; in particular, the perturbational approach is seen once more to break down in regions of phase space where dispersion surfaces come close to each other.

V. SUMMARY

A general scheme is presented which allows one to diagonalize linear coupled wave equations up to arbitrary order in an appropriate expansion parameter. This procedure leads to an asymptotic series, as most perturbation expansions do.

The decoupling of multicomponent wave equations to first order also has been achieved by means of Feynman path integrals [22,23]. However, it seems to us that in this approach it is less straightforward to deal with higher-order corrections. Furthermore, an interesting problem which is specific to this method becomes ever more pronounced: is the erratic nature of typical paths (being continuous but nowhere differentiable) compatible with the assumption of adiabaticity? A brief discussion of this question can be found in [23].

The general results have been made explicit up to second order for the example of a particle described by a Hamiltonian of Born-Oppenheimer type. The systematic incorporation of first-order off-diagonal terms implies the existence of a velocity-dependent second-order correction of the effective Hamiltonians. This result is of interest for various physical systems all of which can be described by Hamiltonians of Born-Oppenheimer type. Recently, second-order calculations of the adiabatic motion of a neutral particle in an inhomogeneous magnetic field have been performed by Berry [11] and Aharonov and Stern [12]; it turns out that in these approaches the velocity-dependent second-order contribution has not been obtained. Another application where second-order effects turn out to be important is the calculation of hyperfine structure constants for the hydrogen ion H_2^+. In particular, Babb and Dalgarno [20] have determined the ground-state wave function for this system to first order by explicitly taking into account the off-diagonal elements of the matrix coupling the electron to the nuclei. The calculation of the corresponding second-order correction of the Hamiltonian, presumably, is equivalent to the result obtained here. Also, in the description of slow atomic collisions one has to go beyond the standard formulation of the Born-Oppenheimer approximation, as is discussed, e.g., in [21].

In conclusion, the problem of diagonalizing multicomponent wave operators is seen to be mapped by means of the Weyl calculus to the problem of diagonalizing finite-dimensional matrices similar to quantum-mechanical perturbation theory but in a somewhat more general setting. Eventually, after diagonalizing the symbol matrix up to the desired order, it is associated via the Weyl correspondence with a unique operator-valued matrix having zero off-diagonal elements. Consequently, the original problem has been reduced to solving a set of uncoupled scalar

wave equations; for this task, however, it is possible to rely on well-known procedures.

ACKNOWLEDGMENTS

The authors thank Greg Flynn for helpful discussions and Jim Morehead for carefully reading the manuscript. This work was supported by the Director, Office of Energy Research, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098. S.W. was supported by the Swiss National Science Foundation and the Freiwillige Akademische Gesellschaft (Basel).

[1] R. G. Littlejohn and W. G. Flynn, Phys. Rev. Lett. **66**, 2839 (1991).
[2] R. G. Littlejohn and W. G. Flynn, Phys. Rev. A **44**, 5239 (1991).
[3] R. I. Brent and L. Fishman, in *Directions in Electromagnetic Wave Modeling,* edited by H. L. Bertoni and L. B. Felsen (Plenum, New York, 1991), pp. 31–38.
[4] V. P. Maslov and M. V. Fedoriuk, *Semi-Classical Approximation in Quantum Mechanics* (Reidel, Boston, 1981).
[5] Detailed information about the Weyl calculus can be found in N. L. Balazs and B. K. Jennings, Phys. Rep. **104**, 347 (1984); or in S. W. McDonald, *ibid.* **158**, 337 (1988), for example.
[6] J. E. Moyal, Proc. Cambridge Philos. Soc. Math. Sci. **45**, 99 (1947).
[7] R. G. Littlejohn and W. G. Flynn, Chaos **2**, 149 (1992).
[8] In their notation this matrix is called *G*.
[9] Three different possibilities of choosing diagonal elements of matrices *U* which diagonalize a given Hamiltonian are investigated in the context of van Vleck perturbation theory in a paper by I. Shavitt and L. T. Redmon, J. Chem. Phys. **73**, 5711 (1980). In particular, the choice of vanishing diagonal elements is called "canonical." The terminology suggests that this might be a preferred choice; in view of an underlying gauge structure of the quantum-mechanical phases, however, no preferred choice seems to exist.
[10] M. Born and R. Oppenheimer, Ann. Phys. (Leipzig) **84**, 457 (1927).
[11] M. V. Berry, in *Geometric Phases in Physics,* edited by A. Shapere and F. Wilczek (World Scientific, Singapore, 1989), p. 7.
[12] Y. Aharonov and A. Stern, Phys. Rev. Lett. **69**, 3593 (1992).
[13] From now on, all greek indices will be understood to run from 1 to *A*.
[14] See, for example, the paper by N. L. Balazs and B. K. Jennings in Ref. [5].
[15] R. G. Littlejohn and St. Weigert (unpublished).
[16] E. Merzbacher, *Quantum Mechanics* (Wiley, New York, 1970).
[17] M. Born, W. Heisenberg, and P. Jordan, Z. Phys. **35**, 557 (1926).
[18] A. Dalgarno, in *Quantum Theory,* edited by D. R. Bates (Academic, New York, 1961), Vol. I, p. 199.
[19] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953), Vol. 2, Chap. IX.
[20] J. F. Babb and A. Dalgarno, Phys. Rev. Lett. **66**, 880 (1991).
[21] B. Zygelman, Phys. Lett. A **125**, 476 (1987).
[22] H. Kuratsuji and S. Iida, Prog. Theor. Phys. **74**, 439 (1985).
[23] D. Loss and P. M. Goldbart, Phys. Rev. A **45**, 13 544 (1992).