Block diagonalisation of Hermitian matrices

1989 J. Phys. A: Math. Gen. 22 2427

(http://iopscience.iop.org/0305-4470/22/13/035)

J. Phys. A: Math. Gen. **22** (1989) 2427-2439. Printed in the UK

Block diagonalisation of Hermitian matrices

L S Cederbaum, J Schirmer and H-D Meyer

Theoretische Chemie, Institut für Physikalische Chemie, Universität Heidelberg, D-6900 Heidelberg, Federal Republic of Germany

Received 3 June 1988, in final form 10 March 1989

Abstract. Block diagonalisation of the Hamiltonian by an unitary transformation is an important theoretical tool, e.g., for deriving the effective Hamiltonian of the quasidegenerate perturbation theory or for determining diabatic molecular electronic states. There are infinitely many different unitary transformations which bring a given Hermitian matrix into block diagonal form. It is, therefore, important to investigate under which conditions the transformation becomes unique. The explicit construction of such a transformation and its properties is discussed in detail. An illustrative example is presented. The non-Hermitian case is briefly discussed as well.

1. Introduction

Much attention has been paid to the diagonalisation of matrices. Most of the work has been devoted to applications in numerous fields and to the development of efficient numerical procedures [1, 2]. Recently interest has arisen in the transformation which block diagonalises a matrix. By a block diagonal matrix we mean a matrix which consists of square matrices (blocks) along its diagonal and is zero elsewhere. In the following we briefly discuss a few examples where block diagonalisation of matrices is of interest. The Born-Oppenheimer approximation leads to a basic concept for molecules, liquids and solids. It allows for the introduction of (adiabatic) electronic states and of nuclear vibrations in these states. Although valid in many cases, this approximation may fail in particular if the energies of two or more electronic states are close to each other. The Jahn-Teller [3] and Renner-Teller [4] effects are well known examples for such a failure. In these and other situations it is convenient and extremely useful to introduce so-called diabatic [5] electronic states which simplify the treatment and more naturally reflect the physics of the problem. Diabatic states can be obtained from the adiabatic ones by an orthogonal transformation which block diagonalises the Hamiltonian of the system (see [6, 7] and references therein).

Another interesting example is the construction of the effective interaction between 'particles'. Using a block diagonalisation procedure the interaction of the one-particle states can be replaced by the effective interaction of only those one-particle states which are relevant for the problem under consideration. Such an effective Hamiltonian approach may simplify the problem, give additional insight and, most importantly, lead to effective interaction elements which can be transferable to other related systems [8]. Furthermore, the method may be used to justify and possibly extend semiempirical approaches [9] which have been successfully applied to molecules and solids. Closely related to the effective Hamiltonian approach is a branch of perturbation theory called quasidegenerate perturbation theory [10-13]. Here, a matrix representation of the

0305-4470/89/132427+13$02.50

Hamiltonian is transformed into two blocks, one of small dimension containing the relevant information about the states of interest and the other referring to all other states. The blocks are computed via perturbation theory. The quasidegenerate perturbation theory has enjoyed applications in nuclear, atomic and molecular physics. Very recently the method of block diagonalisation has been used to derive an approximation scheme to evaluate the particle-particle propagator [14]. The resulting scheme offers new insights and is considerably simpler than those determined from the diagramatic expansion (Feynman diagrams) of the propagator.

The last example we would like to mention is related to the subject of random matrix ensembles. In physics, random matrix ensembles have been introduced to study spectral properties of complicated systems in particular nuclei [15-17]. The Gaussian orthogonal ensemble (GOE) is the most important one. Its members are symmetric matrices, the elements of which are Gaussian random variables. The ensemble is invariant under orthogonal transformations. Many spectral properties of this ensemble can be calculated analytically. In the meantime it has become clear that complex spectra of nuclei [17, 18], atoms [19] and molecules [20] follow the predictions of the GOE. It can be shown that the matrices (blocks) determined by block diagonalising GOE matrices again constitute a GOE [21]. This enables us to establish an interrelation between matrix Hamiltonians of realistic systems and the GOE.

In contrast to the diagonalisation of a Hermitian matrix, there are infinitely many transformations which bring such a matrix into block diagonal form (even if the dimensions of the blocks and the assignment of the eigenvalues to each block are specified). The relevant question immediately arises whether there exist simple physically appealing conditions under which the transformation to block diagonal form is uniquely defined. A major part of this work is devoted to the detailed answer of this question. Particular attention is paid to the explicit construction of the transformation (using these conditions only) and to the discussion of its properties. An illustrative example is also presented. The main emphasis is put on unitary transformations, but non-unitary ones are discussed as well.

2. Theory

2.1. General

We consider a Hermitian matrix $\mathbf{H}$ which is block diagonalised by a unitary matrix $\mathbf{T}$

$$\mathscr{H} = \mathbf{T}^{\dagger}\mathbf{H}\mathbf{T}. \tag{1}$$

The Hermitian block diagonal matrix $\mathscr{H}$ has square $d_n \times d_n$ matrices $\mathscr{H}_{nn}$, $n = 1, 2, \ldots, N$, along its diagonal and is zero elsewhere. The main aim of this work is to determine conditions for selecting $\mathbf{T}$ for a given, but arbitrary, choice of the d_n. To be precise we put, without loss of generality, the eigenvalues λ_i, $i = 1, 2, \ldots, \dim(H)$, of $\mathbf{H}$ in a desired specific order, for instance in the order of increasing size: $\lambda_{i+1} \geqslant \lambda_i$. By choosing the same ordering of eigenvalues also for the block diagonal matrix $\mathscr{H}$, it is then uniquely defined which eigenvalues of $\mathbf{H}$ are contained in each of the blocks $\mathscr{H}_{nn}$.

The secular equation of $\mathbf{H}$ in matrix notation

$$\mathbf{H}\mathbf{S} = \mathbf{S}\boldsymbol{\Lambda} \tag{2}$$

defines the unitary eigenvector matrix $\mathbf{S}$ and the diagonal matrix Λ of eigenvalues λ_i. The unitarity conditions $\mathbf{S}^\dagger\mathbf{S} = \mathbf{S}\mathbf{S}^\dagger = \mathbf{1}$, where $\mathbf{1}$ is the unit matrix, together with (2), do not allow the unique definition of $\mathbf{S}$. Each eigenvector (column of $\mathbf{S}$) can still be multiplied by an arbitrary phase factor. In particular, if $\mathbf{H}$ or a part of it is diagonal from the outset, we may, without loss of generality, choose $\mathbf{S}$ or the corresponding part of it to be a unit matrix. Clearly, the same phase convention should also apply to the eigenvector matrix $\mathbf{F}^\dagger$ of the transformed matrix $\mathscr{H}$. By writing

$$\mathscr{H}\mathbf{F}^\dagger = \mathbf{F}^\dagger\Lambda \tag{3}$$

one easily finds, using (1) and (2), that

$$\mathbf{T} = \mathbf{S}\mathbf{F} \tag{4}$$

where $\mathbf{F}$ is a unitary block diagonal matrix. Similarly to $\mathscr{H}$ we view $\mathbf{F}$ and generally $\mathbf{A}$ ($\mathbf{A} = \mathbf{S}, \mathbf{H}, \Lambda, \mathbf{T}, \ldots$) as a supermatrix, the elements of which are $d_n \times d_m$ matrices F_{nm} and A_{nm}, respectively. Since $\mathbf{F}$ is unitary and block diagonal, we have $F_{nn}F_{nn}^\dagger = F_{nn}^\dagger F_{nn} = \mathbf{1}_{nn}$ and $F_{nm} = 0$ for $n \neq m$. Whenever it is unambiguous we shall drop in the following the index n from the $d_n \times d_n$ unit matrix $\mathbf{1}_{nn}$. The secular equation (3) actually gives rise to a set of secular equations $\mathscr{H}_{nn}F_{nn}^\dagger = F_{nn}^\dagger\Lambda_{nn}$ for $n = 1, 2, \ldots, N$. For later use we define here the matrix $\mathbf{S}_{\mathrm{BD}}$ which is the block diagonal part of $\mathbf{S}$, i.e.,

$$\mathbf{S}_{\mathrm{BD}} = \begin{bmatrix} S_{11} & & & \\ & S_{22} & & 0 \\ & & \ddots & \\ 0 & & & S_{NN} \end{bmatrix} \tag{5}$$

Equation (4) describes the most general transformation which block diagonalises $\mathbf{H}$. The term $\mathbf{S}$ first brings $\mathbf{H}$ into the diagonal form and the subsequent application of $\mathbf{F}$ yields any of the possible block diagonal forms. Apart from being unitary, the matrices F_{nn}, $n = 1, 2, \ldots, N$, are arbitrary. Here, we are led to the question whether there exist *elementary conditions* which *uniquely* determine the F_{nn} and thus the transformation $\mathbf{T}$. By an elementary condition we mean a condition put forward by a very simple and transparent requirement which, in meaningful applications, will usually be considered a 'must'. In this work we shall consider two very different types of elementary conditions. The first one naturally arises from the field of quasidegenerate perturbation theory and effective Hamiltonians and is discussed in § 2.2. The second one, discussed in § 2.3, is of more general relevance, being related to common transformation properties of operators.

2.2. Least action of the unitary transformation

The unitary matrix $\mathbf{T}$ brings the Hermitian matrix $\mathbf{H}$ into block diagonal form. Once the required *block structure* is defined, i.e. for each block the dimension d_n, a space of d_n eigenvectors of $\mathbf{H}$ and a space of d_n basis set vectors, (4) gives the general form of the transformation matrix $\mathbf{T}$. Beyond this form, we want $\mathbf{T}$ to change the original matrix $\mathbf{H}$ as little as possible. In other words, the resulting block diagonal matrix $\mathbf{H}$ should be, except for being block diagonal, similar to the original matrix $\mathbf{H}$ in the sense that $\mathbf{T}$ is as close as possible to the unit matrix $\mathbf{1}$. Mathematically we may require

$$\|\mathbf{T} - \mathbf{1}\| = \text{minimum} \tag{6}$$

where $\|\mathbf{A}\|$ denotes the Euclidean norm of $\mathbf{A}$. (The use of norms other than the Euclidean one is discussed at the end of § 2.3.) The condition (6) is a very weak condition motivated by the idea that the *only* action that $\mathbf{T}$ should perform is to bring $\mathbf{H}$ into block diagonal form. Equation (4) ensures the block diagonal form and the condition (6) guarantees that $\mathbf{T}$ 'does nothing' otherwise.

There are infinitely many transformations $\mathbf{T}$ which block diagonalise $\mathbf{H}$. We shall see below that the condition (6) suffices to uniquely determine $\mathbf{T}$. Before proving this statement we briefly discuss the consequences of (6) from the point of view of quasidegenerate perturbation theory. In perturbation theory the matrix $\mathbf{H}$ is decomposed into the sum of a simple, usually diagonal, matrix and an interaction matrix. A specific eigenvalue and the corresponding eigenvector are expanded in power series in the matrix elements of the interaction matrix. To first order in the interaction this eigenvalue is identical to the corresponding diagonal element of the matrix $\mathbf{H}$. This common textbook approach may be termed single-reference perturbation theory. In many situations one prefers to use the less common but highly interesting multireference perturbation theory, which is usually called quasidegenerate perturbation theory. Instead of considering the expansion of a single eigenvalue, one may expand simultaneously a set of eigenvalues. This is done by formally first block diagonalising $\mathbf{H}$ and subsequently expanding $\mathscr{H}_{nn}$, or more precisely, its elements, in a power series. Diagonalisation of $\mathscr{H}_{nn}$ then provides the eigenvalues of interest. The convergence of the perturbation expansion is the main concern of quasidegenerate perturbation theory. To first order in the interaction the matrix $\mathscr{H}_{nn}$ is identical to the corresponding block of the original matrix $\mathbf{H}$. Second- and higher-order terms provide corrections to this block of $\mathbf{H}$ which eventually, in the case of convergence, will add up to give $\mathscr{H}_{nn}$. Clearly, the best convergence is obtained if these correction terms are as small as possible. Equation (6) presents the best condition on the transformation $\mathbf{H} \rightarrow \mathscr{H}$ which achieves this goal since $\mathbf{T}$ is as close as possible to unity.

We may now present the following theorem.

Theorem 1. Let $\mathbf{S}$ denote the eigenvector matrix of the Hermitian matrix $\mathbf{H}$ and assume that the block diagonal part of $\mathbf{S}$, $\mathbf{S}_{\mathrm{BD}}$, is non-singular. The unitary matrix $\mathbf{T}$ which block diagonalises $\mathbf{H}$ is then uniquely defined by the condition that it is as close as possible to a unit matrix, i.e.

$$\|\mathbf{T}-\mathbf{1}\| = \text{minimum}.$$

The result is

$$\mathbf{T} = \mathbf{S}\mathbf{S}_{\mathrm{BD}}^{\dagger}(\mathbf{S}_{\mathrm{BD}}\mathbf{S}_{\mathrm{BD}}^{\dagger})^{-1/2}. \tag{7}$$

Proof. Using (4) one finds

$$\|\mathbf{T}-\mathbf{1}\|^2 = \mathrm{Tr}[(\mathbf{T}^{\dagger}-\mathbf{1})(\mathbf{T}-\mathbf{1})] = 2\,\mathrm{Tr}\,\mathbf{1} - \mathrm{Tr}(\mathbf{F}\mathbf{S}) - \mathrm{Tr}(\mathbf{F}^{\dagger}\mathbf{S}^{\dagger}) \tag{8}$$

where Tr $\mathbf{A}$ is the trace of $\mathbf{A}$. Since $\mathbf{F}$ is a block diagonal matrix, we have

$$\mathrm{Tr}(\mathbf{F}\mathbf{S}) = \mathrm{Tr}(\mathbf{F}\mathbf{S}_{\mathrm{BD}}) = \sum_{n=1}^{N} \mathrm{Tr}(\boldsymbol{F}_{nn}\boldsymbol{S}_{nn}) \tag{9}$$

and the condition (6) splits into a set of analogous conditions on each block of $\mathbf{F}$:

$$\|\boldsymbol{F}_{nn}\boldsymbol{S}_{nn} - \mathbf{1}\| = \text{minimum}. \tag{10a}$$

It is also clear that the minimum of the RHS of (8) is equivalent to the maximum of $\mathrm{Tr}(\mathbf{FS})+\mathrm{Tr}(\mathbf{F}^{\dagger}\mathbf{S}^{\dagger})$ and analogously for each block of $\mathbf{F}$:

$$\mathrm{Tr}(\boldsymbol{F}_{nn}\boldsymbol{S}_{nn})+\mathrm{Tr}(\boldsymbol{F}_{nn}^{\dagger}\boldsymbol{S}_{nn}^{\dagger})=\text{maximum}. \tag{10b}$$

We decompose $\boldsymbol{S}_{nn}$ according to the polar decomposition [1]

$$\boldsymbol{S}_{nn}=\boldsymbol{U}\boldsymbol{D}\boldsymbol{V}^{\dagger} \tag{11a}$$

where $\boldsymbol{U}$ and $\boldsymbol{V}$ are the unitary eigenvector matrices of $\boldsymbol{S}_{nn}\boldsymbol{S}_{nn}^{\dagger}$ and $\boldsymbol{S}_{nn}^{\dagger}\boldsymbol{S}_{nn}$, respectively,

$$\boldsymbol{S}_{nn}\boldsymbol{S}_{nn}^{\dagger}\boldsymbol{U}=\boldsymbol{U}\boldsymbol{D}^{2} \tag{11b}$$

$$\boldsymbol{S}_{nn}^{\dagger}\boldsymbol{S}_{nn}\boldsymbol{V}=\boldsymbol{V}\boldsymbol{D}^{2} \tag{11c}$$

and $\boldsymbol{D}$ is a diagonal matrix with elements D_i, $i=1,2,\ldots,d_n$, where the D_i^2 are the common eigenvalues of the above matrices. Since $\boldsymbol{S}_{nn}$ is assumed to be a non-singular matrix, we may choose the phases of the column vectors of $\boldsymbol{U}$ and $\boldsymbol{V}$ such that all the D_i are positive real numbers. In this way $\boldsymbol{U}\boldsymbol{V}^{\dagger}$ is uniquely defined.

We now have

$$\mathrm{Tr}(\boldsymbol{F}_{nn}\boldsymbol{S}_{nn})=\mathrm{Tr}(\boldsymbol{V}^{\dagger}\boldsymbol{F}_{nn}\boldsymbol{U}\boldsymbol{D})=\sum_{i=1}^{d_n}A_{ii}D_i \tag{12a}$$

$$\boldsymbol{A}=\boldsymbol{V}^{\dagger}\boldsymbol{F}_{nn}\boldsymbol{U} \tag{12b}$$

where the A_{ii} are the diagonal elements of the matrix $\boldsymbol{A}$. Since $\boldsymbol{A}$ is unitary, i.e. $|A_{ii}|\leqslant 1$ and $D_i>0$, it follows from the condition (10) that $\boldsymbol{A}=\mathbf{1}$. Hence, $\boldsymbol{F}_{nn}$ takes on the appearance

$$\boldsymbol{F}_{nn}=\boldsymbol{V}\boldsymbol{U}^{\dagger}=\boldsymbol{S}_{nn}^{\dagger}(\boldsymbol{S}_{nn}\boldsymbol{S}_{nn}^{\dagger})^{-1/2}. \tag{13}$$

This relation, together with (4), concludes the proof of theorem 1.

The above theorem, and in particular the result (7), deserves some comments. If $\mathbf{S}_{\mathrm{BD}}$ and thus at least one of the $\boldsymbol{S}_{nn}$ is a singular matrix, i.e. one or several eigenvalues $D_i=0$, then the result $\boldsymbol{F}_{nn}=\boldsymbol{V}\boldsymbol{U}^{\dagger}$ shown on the LHS of (13) still fulfils the condition (6) and is, in fact, the general solution. However, the product $\boldsymbol{V}\boldsymbol{U}^{\dagger}$ is now no longer uniquely defined, since the choice of the phases of the eigenvectors corresponding to the vanishing eigenvalues is arbitrary. In practical situations one should try to avoid singular $\boldsymbol{S}_{nn}$, since such matrices do not contain sufficient information on the eigenvectors of $\mathbf{H}$ (see discussion below). A perturbative expansion of $\mathbf{T}$ certainly diverges if $\mathbf{S}_{\mathrm{BD}}$ is singular.

Several different unitary transformations $\mathbf{T}$ and correspondingly block diagonal matrices $\mathscr{H}$ have been suggested in the literature, all in the context of quasidegenerate perturbation theory. Some straightforward algebra shows that our result (7) coincides with that of des Cloizeaux [11] and with that used by Brandow [22]. A series of different results have been discussed which can be obtained via generalised Van Vleck transformation (see a recent paper by Svercek and Hubac [23] and references therein). Clearly, these results do not fulfil the elementary condition that $\mathbf{T}$ is as close as possible to unity (see also the next section).

The eigenvector matrices of $\mathbf{H}$ and $\mathscr{H}$ are $\mathbf{S}$ and $\mathbf{F}^{\dagger}$, respectively. Since the transformation $\mathbf{T}$ fulfils the condition (6) which implies that it brings $\mathbf{H}$ into block diagonal form but has as little as possible effect otherwise, it can be expected that $\mathbf{S}$ and $\mathbf{F}^{\dagger}$ are as similar as possible. Indeed, Klein [24] has shown that for des Cloizeaux's

result the eigenvectors of $\mathbf{H}$ and $\mathscr{H}$ are 'as similar as possible'. In the following we briefly discuss this relevant finding in the context of the present work. That an eigenvector of $\mathbf{H}$ and the corresponding one of $\mathscr{H}$ are as similar as possible means that the scalar product of eigenvectors of both matrices to the same eigenvalue takes on its maximum. We refer to this scalar product as the overlap of the vectors in question. All eigenvalues should be treated on the same footing and, therefore, we ask for the *mean* overlap of the eigenvectors of $\mathscr{H}$ and the corresponding ones of $\mathbf{H}$ to take on its maximum. The mean overlap, MO, of the eigenvectors of $\mathscr{H}$, which are columns of $\mathbf{F}^+$, and those of $\mathbf{H}$, which are columns of $\mathbf{S}$, is given by

$$\mathrm{MO}(\mathbf{F}^+, \mathbf{S}) = \frac{1}{\dim(H)} \mathrm{Tr}(\mathbf{FS}). \tag{14}$$

From the proof of theorem 1 one readily finds that the maximum of the mean overlap of $\mathbf{F}^+$ and $\mathbf{S}$ (if assumed to be real) follows from the condition (6). Moreover, as $\|\mathbf{F}^+ - \mathbf{S}\| = \|\mathbf{T} - \mathbf{1}\|$ holds, the condition (6) of least action of the unitary block diagonalising transformation is equivalent to the requirement that the eigenvectors of $\mathbf{H}$ and $\mathscr{H}$ are as similar as possible. Again, since only the mean value and not individual overlaps are involved, the latter requirement is a very weak one. Hence, it is satisfactory to see that this requirement suffices for uniquely determining the transformation $\mathbf{T}$ in a *constructive* way as shown in the proof of theorem 1.

We have introduced the condition (6) of least action of the unitary transformation $\mathbf{T}$. One might assume that then $\mathbf{H}$ and $\mathscr{H}$ are as similar as possible according to

$$\|\mathbf{H} - \mathscr{H}\| = \text{minimum} \tag{15}$$

or, instead, that (15) stated as a condition is superior to condition (6). Without going into details we note that both assumptions are wrong and present a simple example illustrating this point. In our example the blocks $\boldsymbol{H}_{nn}$ of $\mathbf{H}$ are diagonal matrices ($\boldsymbol{H}_{nn} = c_n \mathbf{1}$, where the c_n are real constants) and the $\boldsymbol{H}_{nm}$, $n \neq m$, are non-vanishing arbitrary matrices. This leads to an expression for $\|\boldsymbol{H} - \mathscr{H}\|$ which is totally independent of $\mathbf{T}$ or $\mathscr{H}$ and hence gives no hint how to construct $\mathbf{T}$.

2.3. Least information for the unitary transformation

The elementary condition discussed in the preceding subsection has been motivated by the idea of least action of the unitary transformation $\mathbf{T}$ and quasidegenerate perturbation theory. On the other hand, one may ask whether a transformation to block diagonal form can be uniquely derived on the ground of general transformation properties of $\mathbf{T}$ alone. We first write the matrix equation (1) in operator notation:

$$\hat{\mathscr{H}} = \hat{T}^+ \hat{H} \hat{T} \tag{16}$$

which is always possible, since the matrices $\mathbf{T}$, $\mathbf{H}$ and $\mathscr{H}$ uniquely define the operators $\hat{T}$, $\hat{H}$ and $\hat{\mathscr{H}}$ once a basis $\phi^0 = \{\phi_i^0\}$ is specified. The matrix $\mathbf{H}$ can now be written as $\mathbf{H} = \{H_{ij} = \langle \phi_i^0 | \hat{H} | \phi_j^0 \rangle\}$, and similar expressions hold for $\mathbf{T}$ and $\mathscr{H}$. Let $\psi^0 = \{\psi_k^0\}$ be a basis of eigenstates of $\hat{H}$. The eigenvector matrix $\mathbf{S}$ of $\mathbf{H}$ is the representation of ψ^0 with respect to ϕ^0. Since $\mathbf{S} = \{S_{ik} = \langle \phi_i^0 | \psi_k^0 \rangle = \langle \phi_i^0 | \hat{1} | \psi_k^0 \rangle\}$, we may equivalently consider $\mathbf{S}$ as the representation of the unit operator $\hat{1}$ with respect to the bases ϕ^0 and ψ^0 and indicate this by writing $\mathbf{S}(\phi^0, \psi^0)$. From $\mathbf{T} = \mathbf{SF}$ it immediately follows that $\mathbf{F} = \{F_{ki} = \langle \psi_k^0 | \hat{T} | \phi_i^0 \rangle\}$ which implies that $\mathbf{F}$ is the representation of $\hat{T}$ with respect to the bases ψ^0 and ϕ^0. We indicate this by writing $\mathbf{F}(\psi^0, \phi^0)$ and note that $\mathbf{T} = \mathbf{F}(\phi^0, \phi^0)$.

Definition. Let $\mathbf{F}(\psi, \phi)$ and $\mathbf{S}(\phi, \psi)$ be the matrix representations of the operators $\hat{T}$ and $\hat{1}$, respectively, with respect to bases ϕ and ψ (see above). $\mathbf{T}$ is *fully determined* by $\mathbf{S}$ if a prescription g exists such that

$$\mathbf{F}(\psi, \phi) = g(\mathbf{S}(\phi, \psi)) \tag{17}$$

is valid for all bases ϕ and ψ for which $\mathbf{F}(\psi, \phi)$ is block diagonal.

Let us now consider a basis change $\phi \to \tilde{\phi}$ and $\psi \to \tilde{\psi}$ defined as usual by unitary transformations $\mathbf{X}$ and $\mathbf{Y}$, respectively†. From the definitions of $\mathbf{F}$ and $\mathbf{S}$ we immediately find

$$\mathbf{F}(\tilde{\psi}, \tilde{\phi}) = \mathbf{Y}^{\dagger}\mathbf{F}(\psi, \phi)\mathbf{X} \tag{18a}$$

$$\mathbf{S}(\tilde{\phi}, \tilde{\psi}) = \mathbf{X}^{\dagger}\mathbf{S}(\phi, \psi)\mathbf{Y}. \tag{18b}$$

If $\mathbf{T}$ is fully determined by $\mathbf{S}$, we may insert (18) into (17) and obtain

$$Y^{\dagger} g(\mathbf{S}(\phi, \psi))\mathbf{X} = g(\mathbf{X}^{\dagger}\mathbf{S}(\phi, \psi)\mathbf{Y}). \tag{19}$$

Note that this relation is valid for all block diagonal unitary matrices $\mathbf{X}$ and $\mathbf{Y}$ and, as we shall prove below, suffices to uniquely determine the prescription g and hence the block diagonalising transformation matrix $\mathbf{T}$.

Before stating our second theorem, we would like to briefly explain why we consider the condition that $\mathbf{T}$ is fully determined by $\mathbf{S}$ to be an elementary condition. $\mathbf{T}$ should block diagonalise $\mathbf{H}$. The only information given to us is the knowledge of $\mathbf{H}$ and our choice of the block structure. Hence, we would like to construct $\mathbf{T}$ using *only* this information. All information concerning $\mathbf{H}$ is contained in its eigenvalues and eigenvectors. The eigenvector matrix $\mathbf{S}$ contains the information on the eigenvectors as well as the information on the block structure (ordering and grouping of the eigenstates $\{\psi_k^0\}$ and basis states $\{\phi_k^0\}$ into blocks is described by the columns and rows of $\mathbf{S}$, respectively). While the eigenvalues of $\mathbf{H}$ are invariant quantities and do not depend on the specification of an underlying basis ϕ^0 to define $\hat{H}$, the eigenstates $\{\psi_k^0\}$ do depend on this specification. We may thus consider the eigenvalues as fixed and $\mathbf{S}$ as the only remaining quantity which determines $\mathbf{T}$.

Theorem 2. Let $\mathbf{S}$ be the eigenvector matrix of a Hermitian matrix $\mathbf{H}$. For a given block structure let $\mathbf{S}_{\mathrm{BD}}$, which is the block diagonal part of $\mathbf{S}$, be non-singular. Assume the phase convention that the eigenvector matrix of a diagonal matrix is a unit matrix.

For each $\mathbf{H}$ there exists exactly one matrix $\mathbf{T}$ which transforms it into the block diagonal matrix $\mathscr{H} = \mathbf{T}^{\dagger}\mathbf{H}\mathbf{T}$ with the prescribed block structure such that $\mathbf{T}$ is fully determined by $\mathbf{S}$.

Proof. The transformation property (19) is the keystone in this proof. Since $\mathbf{F}$, $\mathbf{X}$ and $\mathbf{Y}$ have identical block diagonal form, we may consider each block n separately: $\boldsymbol{F}_{nn}$, $\boldsymbol{X}_{nn}$ and $\boldsymbol{Y}_{nn}$. In particular it follows that $\boldsymbol{F}_{nn}$ depends on $\boldsymbol{S}_{nn}$ and not on the other blocks of $\mathbf{S}$. Introducing for convenience the shorthand notation $\mathbf{F} = \mathbf{F}(\mathbf{S})$ for (17), we may restate (19) in the shorthand notation

$$\boldsymbol{Y}_{nn}^{\dagger}\boldsymbol{F}_{nn}(\boldsymbol{S}_{nn})\boldsymbol{X}_{nn} = \boldsymbol{F}_{nn}(\boldsymbol{X}_{nn}^{\dagger}\boldsymbol{S}_{nn}\boldsymbol{Y}_{nn}). \tag{20}$$

† $\mathbf{F}$ is a unitary block diagonal matrix. Only those bases ϕ and ψ are allowed which preserve this block diagonal structure of $\mathbf{F}$. It follows that the transformation matrices $\mathbf{X}$ and $\mathbf{Y}$ in (18) must be block diagonal as well.

Choosing $\boldsymbol{X}_{nn} = \boldsymbol{U}$ and $\boldsymbol{Y}_{nn} = \boldsymbol{V}$ taken from the polar decomposition of $\boldsymbol{S}_{nn} = \boldsymbol{UDV}^{\dagger}$, see (11), we obtain

$$\boldsymbol{F}_{nn}(\boldsymbol{S}_{nn}) = \boldsymbol{V}\boldsymbol{F}_{nn}(\boldsymbol{D})\boldsymbol{U}^{\dagger}. \tag{21}$$

The matrix $\boldsymbol{F}^{\dagger}_{nn}(\boldsymbol{D})$ is to be considered as the eigenvector matrix of a matrix $\mathscr{H}_{nn}$ originating from a matrix $\mathbf{H}$ which has an eigenvector matrix with a diagonal block $\boldsymbol{S}_{nn}$. Inserting a diagonal unitary matrix $\boldsymbol{\Delta}$ with diagonal elements $\exp(\mathrm{i}\delta_k)$ for $\boldsymbol{X}^{\dagger}_{nn}$ or $\boldsymbol{Y}_{nn}$ and setting the other matrix equal to $\mathbf{1}$ leads, together with $\boldsymbol{S}_{nn} = \boldsymbol{D}$, to

$$\boldsymbol{F}_{nn}(\boldsymbol{D\Delta}) = \boldsymbol{\Delta}^{\dagger}\boldsymbol{F}_{nn}(\boldsymbol{D}) \tag{22}$$

$$\boldsymbol{F}_{nn}(\boldsymbol{\Delta D}) = F_{nn}(\boldsymbol{D})\boldsymbol{\Delta}^{\dagger}. \tag{23}$$

Since $\boldsymbol{\Delta D} = \boldsymbol{D\Delta}$, it is clear that $\boldsymbol{F}_{nn}(\boldsymbol{D})$ is also a diagonal matrix, and being unitary its diagonal elements are phase factors. As an eigenvector matrix of some $\mathscr{H}_{nn}$, we may set $\boldsymbol{F}_{nn}(\boldsymbol{D}) = \mathbf{1}_{nn}$ by our phase convention. It follows from (21) that $\boldsymbol{F}_{nn}(\boldsymbol{S}_{nn}) = \boldsymbol{VU}^{\dagger}$ which completes our proof.

Remark. Although theorems 1 and 2 start from very different points of view, both lead to the same unique block diagonal matrix and

$$\mathbf{T} = \mathbf{S}\mathbf{S}^{\dagger}_{\mathrm{BD}}(\mathbf{S}_{\mathrm{BD}}\mathbf{S}^{\dagger}_{\mathrm{BD}})^{-1/2}.$$

Theorem 1 can now easily be understood in the context of the more general theorem 2. It was shown in the proof of theorem 1 that the minimum principle (6) leads to a *unique* $\mathbf{F}$ minimising $\|\mathbf{SF}-\mathbf{1}\|$ for the Euclidean norm. Since $\mathbf{S}$ is fixed in this process, the resulting matrix $\mathbf{T} = \mathbf{SF}$ is fully determined by $\mathbf{S}$. In combining both theorems we can now show that the minimum principle (6) is *not restricted to the use of the Euclidean norm.* This principle yields the result (7) for any norm, provided minimising $\|\mathbf{T}-\mathbf{1}\|$ uniquely defines $\mathbf{T}$ and the norm $\|A\|$ chosen is invariant under unitary transformations of the matrix $\mathbf{A}$. If the norm chosen is such that there is more than one $\mathbf{T}$ minimising $\|\mathbf{T}-\mathbf{1}\|$, then the result (7) is one of the solutions of (6).

Withouth resorting to the phase convention, the transformation property (19) leads to a set of possible transformation matrices which can compactly be classified by (21) with a diagonal and unitary matrix $\boldsymbol{F}_{nn}(\boldsymbol{D})$. Using (19) we may rewrite (21) to take on the appearance

$$\mathbf{F}(\mathbf{S}) = \mathbf{S}^{\dagger}_{\mathrm{BD}}(\mathbf{S}_{\mathrm{BD}}\mathbf{S}^{\dagger}_{\mathrm{BD}})^{-1/2}\mathbf{F}((\mathbf{S}_{\mathrm{BD}}\mathbf{S}^{\dagger}_{\mathrm{BD}})^{1/2}). \tag{24}$$

The set of matrices $\mathbf{F}(\mathbf{S})$ which fulfil the transformation property (19) is obtained by multiplying the unique result of theorem 2 from the right by the same matrices, but now as a function of $(\mathbf{S}_{\mathrm{BD}}\mathbf{S}^{\dagger}_{\mathrm{BD}})^{1/2}$. Introducing the phase convention that the eigenvector matrix of a diagonal matrix is unity, one arrives at $\mathbf{F}((\mathbf{S}_{\mathrm{BD}}\mathbf{S}^{\dagger}_{\mathrm{BD}})^{1/2}) = \mathbf{1}$.

It is somewhat surprising that a trivial and non-restricting phase convention suffices to make $\mathbf{F}$ unique. To have a better understanding of this point we show that, if only real matrices are involved, the phase convention can be dropped. For this purpose we assume that, if $\mathbf{H}$ is already in a block diagonal form as required, the transformation leaves this matrix as it is, i.e. $\mathbf{T} = \mathbf{1}$ and $\mathbf{H} = \mathscr{H}$. Consequently, $\mathbf{F}(\mathbf{1}) = \mathbf{1}$. Since $\boldsymbol{F}_{nn}(\boldsymbol{D})$ is an orthogonal and diagonal matrix, its elements are 1 or -1, i.e $(\boldsymbol{F}_{nn}(\boldsymbol{D}))_{ij} = \delta_{ij}\,\mathrm{sgn}[f(D_1, D_2, \ldots, D_{d_n})]$, where f is some unknown function. Assuming $\boldsymbol{F}_{nn}(\boldsymbol{D})$ or, more precisely, its elements to be continuous functions of the D_i, it immediately follows that $\boldsymbol{F}_{nn}(\boldsymbol{D}) = \mathbf{1}$ and thus, from (24), $\mathbf{F}((\mathbf{S}_{\mathrm{BD}}\mathbf{S}^{\dagger}_{\mathrm{BD}})^{1/2}) = \mathbf{1}$. If the functions are non-continuous, $\mathbf{F}$ is non-continuous in $\mathbf{S}$ and, in particular, quasidegenerate perturbation

theory is inapplicable. For complex matrices the situation is more subtle and $\mathbf{F}(\mathbf{S})$ is not uniquely determined when the phase convention is dropped. Then (21) or, equivalently, (24) is the general result. A simple example for the additional matrix is $\mathbf{F}((\mathbf{S}_{BD}\mathbf{S}_{BD}^{\dagger})^{1/2}) = \exp[i\alpha(\mathbf{S}_{BD}\mathbf{S}_{BD}^{\dagger} - \mathbf{1})]$, where α is an arbitary real number.

2.4. A few useful properties

The elements of the block diagonal matrix calculated via theorems 1 and 2 are given by

$$\mathscr{H}_{nn} = \boldsymbol{F}_{nn}^{\dagger}(\boldsymbol{S}_{nn})\Lambda_{nn}\boldsymbol{F}_{nn}(\boldsymbol{S}_{nn}) \tag{25a}$$

$$\boldsymbol{F}_{nn}(\boldsymbol{S}_{nn}) = \boldsymbol{S}_{nn}^{\dagger}(\boldsymbol{S}_{nn}\boldsymbol{S}_{nn}^{\dagger})^{-1/2}. \tag{25b}$$

Every non-singular matrix can be uniquely decomposed [1, 2] into a product $\boldsymbol{A}\,\boldsymbol{B}$ of a unitary matrix $\boldsymbol{A}$ and a Hermitian matrix $\boldsymbol{B}$. Interestingly, the matrix $\boldsymbol{F}_{nn}^{\dagger}$ is the unitary matrix resulting from the decomposition of the block $\boldsymbol{S}_{nn}^{\dagger}$: $\boldsymbol{S}_{nn}^{\dagger} = \boldsymbol{F}_{nn}(\boldsymbol{S}_{nn}\boldsymbol{S}_{nn}^{\dagger})^{1/2}$.

The matrix $\boldsymbol{F}_{nn}$ can be rewritten to have many different appearances. One form of particular usefulness is obtained from (25*b*) by multiplying it from the right with $\mathbf{1} = (\boldsymbol{S}_{nn}\boldsymbol{S}_{nn}^{\dagger})^{-1/2}(\boldsymbol{S}_{nn}\boldsymbol{S}_{nn}^{\dagger})^{1/2}$:

$$\boldsymbol{F}_{nn}(\boldsymbol{S}_{nn}) = \boldsymbol{S}_{nn}^{-1}(\boldsymbol{S}_{nn}\boldsymbol{S}_{nn}^{\dagger})^{1/2}. \tag{25c}$$

With this form the transformation matrix $\mathbf{T}$ takes on a particularly simple appearance in the important case of $N = 2$. Writing $\mathbf{S}\mathbf{S}^{\dagger} = \mathbf{1}$ in block form we find that

$$\mathbf{T} = \mathbf{U}(\mathbf{U}^{\dagger}\mathbf{U})^{-1/2} \tag{26a}$$

where

$$\mathbf{U} = \begin{pmatrix} \mathbf{1} & \boldsymbol{X} \\ -\boldsymbol{X}^{\dagger} & \mathbf{1} \end{pmatrix} \qquad \mathbf{U}^{\dagger}\mathbf{U} = \begin{pmatrix} \mathbf{1} + \boldsymbol{X}\boldsymbol{X}^{\dagger} & 0 \\ 0 & \mathbf{1} + \boldsymbol{X}^{\dagger}\boldsymbol{X} \end{pmatrix} \tag{26b}$$

and $\boldsymbol{X} = -(\boldsymbol{S}_{21}\boldsymbol{S}_{11}^{-1})^{\dagger} = \boldsymbol{S}_{12}\boldsymbol{S}_{22}^{-1}$. It should be noted that in order to compute $\mathbf{T}$ and thus $\mathscr{H}$ it is not necessary to know all the eigenvectors of $\mathbf{H}$ as may be concluded by expression (7). According to (26) $\mathbf{T}$ can be obtained from the matrix $\boldsymbol{X}$ which is expressible by the eigenvectors corresponding to the eigenvalues of only one of the blocks of $\mathscr{H}$. This is of practical relevance in particular if the dimension of one of $\mathscr{H}_{11}$ and $\mathscr{H}_{22}$ is small. We refer to § 3 for an illustrative example.

2.5. The non-Hermitian case

Until now we have discussed the situation of Hermitian matrix $\mathbf{H}$ to be block diagonalised by a unitary transformation $\mathbf{T}$. If $\mathscr{H}$ should not be necessarily Hermitian, $\mathbf{T}$ need not be unitary and we may define $\mathscr{H}$ according to the following transformation:

$$\mathbf{T}^{-1}\mathbf{H}\mathbf{T} = \mathscr{H}. \tag{27a}$$

The eigenvector matrix $\mathbf{F}^{\dagger}$ of $\mathscr{H}$ in the Hermitian situation cannot become identical to $\mathbf{S}_{BD}$ which is, in general, a non-unitary matrix. The best possible result was produced by theorem 1 and given in (7). In the non-Hermitian case the eigenvector matrix $\mathbf{F}^{-1}$ of $\mathscr{H}$ can become identical to $\mathbf{S}_{BD}$. The best result along the ideas behind theorem 1 is to put the eigenvector matrix of $\mathscr{H}$ equal to $\mathbf{S}_{BD}$. This immediately leads via (4) to

$$\mathbf{T} = \mathbf{S}\mathbf{S}_{BD}^{-1}. \tag{27b}$$

More precisely: because of (4) we have $\|\mathbf{T}-\mathbf{1}\| = \|\mathbf{F}-\mathbf{S}^{\dagger}\|$ which obviously takes on its minimum at $\mathbf{F}=\mathbf{S}^{\dagger}_{\mathrm{BD}}$. Inspection of (27*a*) shows that $\mathbf{F}^{-1}$ is the eigenvector matrix of $\mathscr{H}$ and hence the result (27*b*) strictly follows from theorem 1 if $\mathscr{H}$ need not be Hermitian. Although of different appearance, the result (27*b*) coincides with that first introduced by des Cloiseaux [11]. Here, this result is easily derived from the elementary condition $\|\mathbf{T}-\mathbf{1}\| =$ minimum. In the case of $N=2$ the explicit form of $\mathbf{T}$ in (27*b*) is helpful in straightforwardly obtaining $\mathbf{T}=\mathbf{U}$, where $\mathbf{U}$ has been given in (26*b*) (see also § 3).

Comparing the result (27*b*) with that in (25*c*) for a unitary $\mathbf{T}$, we may write

$$\mathbf{T}_u = \mathbf{T}_{nu}(\mathbf{T}^{\dagger}_{nu}\mathbf{T}_{nu})^{-1/2}. \tag{28}$$

The unitary transformation matrix $\mathbf{T}_u$ is uniquely obtained from the $A \cdot B$ decomposition (see text above (25*c*)) of the non-unitary one $\mathbf{T}_{nu}$. This finding can, of course, be stated only after $\mathbf{T}_u$ has been uniquely derived via theorem 1 or/and 2.

It is worth noting that $\mathbf{T}$ in (27*b*) and thus also $\mathbf{F}=\mathbf{S}^{-1}_{\mathrm{BD}}$ fulfil the transformation property (19) as is the situation in the Hermitian case. Contrary to the Hermitian case, where (19) uniquely defines the block diagonalising transformation, this transformation is not uniquely obtained by requiring (19) to be valid if $\mathbf{F}$ is not unitary. Equation (21) is still valid, but since the diagonal matrix $\boldsymbol{F}_{nn}(\boldsymbol{D})$ appearing there is not unitary, it can have infinitely many appearances. Choosing $\boldsymbol{F}_{nn}(\boldsymbol{D}) = \boldsymbol{D}^{-1}$ readily leads to $\mathbf{F}=\mathbf{S}^{-1}_{\mathrm{BD}}$. For $\boldsymbol{F}_{nn}(\boldsymbol{D}) = \boldsymbol{D}^{-2}$, for instance, we find $\mathbf{F}=\mathbf{S}^{-1}_{\mathrm{BD}}(\mathbf{S}_{\mathrm{BD}}\mathbf{S}^{\dagger}_{\mathrm{BD}})^{-1/2}$ which also fulfils the invariance conditions (19). Another interesting non-unitary example is $\mathbf{F}=\mathbf{S}^{\dagger}_{\mathrm{BD}}$, which is obtained for the choice $\boldsymbol{F}_{nn}(\boldsymbol{D}) = \boldsymbol{D}$. Any choice of $\boldsymbol{F}_{nn}(\boldsymbol{D})$ leads to a transformation $\mathbf{T}$ which fulfils (19). In view of this finding for the non-Hermitian case it is interesting to see that the phase convention (see § 2.1 or theorem 2) suffices to uniquely define the transformation in the Hermitian case (once the transformation property (19) is used, of course).

3. An illustrative example

The investigation of a simple non-trivial example facilitates some of the findings of the preceding section. Schucan and Weidenmüller [25] studied the explicit construction of an energy-independent effective interaction and paid special attention to the three-dimensional case. The effective interaction obtained is non-Hermitian. We shall study the same example in the framework of the present theory, determining thereby the non-Hermitian as well as the Hermitian results.

The matrix

$$\mathbf{H} = \begin{bmatrix} E_1 & 0 & V_1 \\ 0 & E_2 & V_2 \\ V_1^* & V_2^* & E_3 \end{bmatrix} \tag{29}$$

is to be transformed into a block diagonal matrix $\mathscr{H}$, where $\mathscr{H}_{11}$ is a two-dimensional matrix and $\mathscr{H}_{22}$ is a one-dimensional matrix. Of course, $\lambda = \mathscr{H}_{22}$ is an exact eigenvalue of $\mathbf{H}$. We choose λ to be the eigenvalue originating from E_3, i.e. $\lambda \to E_3$ for $V_1, V_2 \to 0$. The matrix $\mathscr{H}_{11}$ will be expressed in terms of the elements of the original matrix $\mathbf{H}$ and λ. The unitary matrix $\mathbf{T}=\mathbf{U}(\mathbf{U}^{\dagger}\mathbf{U})^{-1/2}$ which transforms $\mathbf{H}$ into $\mathscr{H}$ is given in (26). In particular we notice that the matrix $\boldsymbol{X}$, which gives $\mathbf{U}$ and thus $\mathbf{T}$, is fully determined

by the eigenvector of $\mathbf{H}$ corresponding to the eigenvalue λ. We define

$$\boldsymbol{X} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \tag{30a}$$

and notice that the third column of $\mathbf{U}$ is proportional to the above-mentioned eigenvector of $\mathbf{H}$:

$$\begin{bmatrix} E_1 & 0 & V_1 \\ 0 & E_2 & V_2 \\ V_1^* & V_2^* & E_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix} \tag{30b}$$

One readily finds

$$x_1 = \frac{V_1}{\lambda - E_1} \qquad x_2 = \frac{V_2}{\lambda - E_2} \tag{30c}$$

which completes the determination of $\mathbf{U}$, see (26*b*).

The non-Hermitian block diagonal matrix may be obtained via (27*a*), which in this case is $\mathscr{H} = \mathbf{U}^{-1}\mathbf{H}\mathbf{U}$. Instead of evaluating $\mathbf{U}^{-1}$, we make use of the interesting fact that the blocks along the diagonal of $\mathbf{U}^{-1}\mathbf{H}\mathbf{U}$ and of $\mathbf{H}\mathbf{U}$ are identical. This immediately gives

$$\mathscr{H}_{\mathrm{NH}} = \begin{bmatrix} E_1 + \dfrac{|V_1|^2}{E_1 - \lambda} & \dfrac{V_1 V_2^*}{E_2 - \lambda} & 0 \\ \dfrac{V_1^* V_2}{E_1 - \lambda} & E_2 + \dfrac{|V_2|^2}{E_2 - \lambda} & 0 \\ 0 & 0 & \lambda \end{bmatrix} \tag{31}$$

where the subscript 'NH' indicates the non-Hermitian case. The result (31) is equivalent to the one found by Schucan and Wiedenmüller [25] who considered real matrix elements V_1 and V_2.

The Hermitian block diagonal matrix $\mathscr{H} = \mathbf{T}^{\dagger}\mathbf{H}\mathbf{T}$ can, of course, be found using the results (30) and (31) and the general relation

$$\mathscr{H} = (\mathbf{U}^{\dagger}\mathbf{U})^{1/2} \mathscr{H}_{\mathrm{NH}} (\mathbf{U}^{\dagger}\mathbf{U})^{-1/2}. \tag{32}$$

The calculation, although straightforward, is quite lengthy in spite of the simple structure of $(\mathbf{U}^{\dagger}\mathbf{U})^{\pm 1/2}$:

$$(\mathbf{U}^{\dagger}\mathbf{U})^{\pm 1/2} = \mathbf{1} + \frac{q^{\pm 1} - 1}{|x_1|^2 + |x_2|^2} \begin{bmatrix} |x_1|^2 & x_1 x_2^* & 0 \\ x_1^* x_2 & |x_2|^2 & 0 \\ 0 & 0 & |x_1|^2 + |x_2|^2 \end{bmatrix} \tag{33a}$$

where

$$q = (1 + |x_1|^2 + |x_2|^2)^{1/2}. \tag{33b}$$

The calculation becomes very simple once it is noticed that the non-Hermitian matrix (31) can be written as

$$\mathscr{H}_{\mathrm{NH}} = \lambda(\mathbf{1} - \mathbf{U}^{\dagger}\mathbf{U}) + \begin{bmatrix} E_1 & 0 & 0 \\ 0 & E_2 & 0 \\ 0 & 0 & \lambda \end{bmatrix} \mathbf{U}^{\dagger}\mathbf{U}. \tag{33c}$$

This leads to the final result which takes on the following appearance:

$$\mathscr{H} = \tfrac{1}{2}(\mathscr{H}_{\mathrm{NH}} + \mathscr{H}_{\mathrm{NH}}^{\dagger}) + \frac{E_2 - E_1}{(1+q)^2}\mathbf{C} \tag{34a}$$

where the 'correction' matrix $\mathbf{C}$ is

$$\mathbf{C} = \begin{bmatrix} |x_1 x_2|^2 & \tfrac{1}{2}x_1 x_2^*(|x_2|^2 - |x_1|^2) & 0 \\ \tfrac{1}{2}x_1^* x_2(|x_2|^2 - |x_1|^2) & -|x_1 x_2|^2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{34b}$$

The Hermitian block diagonal matrix $\mathscr{H}$ is obtained as half the sum of the non-Hermitian version $\mathscr{H}_{\mathrm{NH}}$ and its HC plus a correction matrix. While $\mathscr{H}_{\mathrm{NH}}$ is, for a fixed value of λ, of quadratic order in V, $\mathscr{H}$ contains terms up to infinite order in V. It should be remembered, however, that λ itself actually is a complicated function of the V_1 and V_2. Remark: since the matrices $\mathbf{U}^{\dagger}\mathbf{U}$, $\mathscr{H}_{\mathrm{NH}}$ and $\mathscr{H}$ are block diagonal matrices, one can perform all calculations for each block separately. For the present example we have perferred to perform the computation with the full matrices in order to show explicitly their block diagonal form.

4. Summary

The unitary matrix $\mathbf{T}$ which transforms the matrix $\mathbf{H}$ into block diagonal matrix $\mathscr{H}$ can be uniquely determined by imposing elementary conditions. If we have no special requirements concerning the transformation, it is natural to demand that the only action that $\mathbf{T}$ should perform is to bring $\mathbf{H}$ into block diagonal form, i.e. $\mathbf{T}$ should be as close to unity as possible. Interestingly, this very weak condition already defines uniquely the matrix $\mathbf{T}$ and thus also $\mathscr{H}$. To fulfil this condition is also of relevance for perturbation theory. Since $\mathbf{T}$ is as close as possible to unity, the correction terms which are to be calculated by a perturbation expansion of $\mathbf{T}$ are as small as possible.

In a second more general approach the consequences of another elementary condition are studied. The condition is that only the information contained in $\mathbf{H}$ and, of course, the given block form should be used to construct $\mathbf{T}$ and thus the block diagonal matrix $\mathscr{H}$. This very weak condition alone does not uniquely determine $\mathbf{T}$. We rather find an interesting limited set of possible transformations which fulfil this condition. The trivial phase convention, that the eigenvector matrix of a diagonal matrix is put equal to $\mathbf{1}$, then uniquely determines $\mathbf{T}$ out of this set of possible transformations. The resulting $\mathbf{T}$ coincides with that of the first approach. It is remarkable that two very different elementary conditions give rise to the same result for the case of unitary transformations. This also underlines the significance of the transformation derived.

If $\mathbf{T}$ is not unitary, the first condition again leads to a unique result while the second condition is fulfilled for a wider class of matrices. Several interesting properties of $\mathbf{T}$ are investigated. In particular a useful representation of this matrix is presented for the case where $\mathscr{H}$ consists of two blocks only. This representation is explicitly used to compute $\mathscr{H}$ in closed form for a simple but non-trivial example.

Acknowledgment

The authors thank Professor Dr E M Majster for very useful discussions.

References

[1] Parlett B N 1980 *The Symmetric Eigenvalue Problem* (Englewood Cliffs, NJ: Prentice-Hall)
[2] Wilkinson J H 1965 *the Algebraic Eigenvalue Problem* (Oxford: Clarendon)
[3] Jahn H A and Teller E 1938 *Proc. R. Soc.* A **161** 220
Englman R 1972 *The Jahn-Teller Effect* (New York: Wiley)
[4] Renner R 1934 *Z. Phys.* **92** 172
Jungen Ch and Merer A J 1976 *Molecular Spectroscopy: Modern Research* vol 2, ed K N Rao (New York: Academic)
[5] Landau L D 1932 *Phys. Z. Sowjetunion* **2** 46
Lichten W 1967 *Phys. Rev.* **164** 131
Smith F T 1969 *Phys. Rev.* **179** 111
[6] Özkan I and Goodman L 1979 *Chem. Rev.* **79** 275
[7] Cederbaum L S, Köppel H and Domcke W 1981 *Int. J. Quantum Chem.* S **15** 251
Pacher T, Cederbaum L S and Köppel H 1988 *J. Chem. Phys.* **89** 7367
[8] Oleksik J J, Takada T and Freed K F 1985 *Chem. Phys. Lett.* **113** 249
[9] Segal G A (ed) 1977 *Semiempirical Methods of Electronic Structure Calculation* (New York: Plenum)
[10] Bloch C 1958 *Nucl. Phys.* **6** 329
[11] des Cloizeaux J 1960 *Nucl. Phys.* **20** 321
[12] Brandow B H 1967 *Rev. Mod. Phys.* **39** 771
[13] Kvasnicka V 1977 *Adv. Chem. Phys.* **36** 345
[14] Tarantelli A and Cederbaum L S 1989 *Phys. Rev.* A **39** 1639
[15] Porter C E (ed) 1965 *Statistical Theories of Spectra: Fluctuations* (New York: Academic)
[16] Mehta M 1967 *Random Matrices and the Statistical Theory of Energy Levels* (New York: Academic)
[17] Brody T A, Flores J, French J B, Mello P A, Pandey A and Wong S S M 1981 *Rev. Mod. Phys.* **53** 385
[18] Bohigas O, Hag R U and Pandey A 1985 *Phys. Rev. Lett.* **54** 1645
[19] Rosenzweig N and Porter C F 1960 *Phys. Rev.* **120** 1689
Camarda H S and Georgopulos P D 1983 *Phys. Rev. Lett.* **50** 492
[20] Haller E, Köppel H and Cederbaum L S 1983 *Chem. Phys. Lett.* **101** 215
Abramson E, Field R W, Imre D, Innes K K and Kinsey J L 1984 *J. Chem. Phys.* **80** 2298
[21] Zimmermann Th and Cederbaum L S 1987 *Phys. Rev. Lett.* **59** 1496
[22] Brandow B H 1970 *Ann. Phys., NY* **57** 214
[23] Svercek M and Hubac I 1987 *Int. J. Quantum Chem.* **31** 625
[24] Klein D J 1974 *J. Chem. Phys.* **61** 786
[25] Schucan T H and Weidenmüller H A 1972 *Ann. Phys., NY* **73** 108