Physica **93A** (1978) 327–353 © *North-Holland Publishing Co.*

DIAGONALIZATION OF THE QUADRATIC BOSON HAMILTONIAN

J.H.P. COLPA

Natuurkundig Laboratorium, Universiteit van Amsterdam, Nederland

Received 9 August 1977
Revised 30 January 1978

A general treatment has been given of the problem of diagonalizing a hamiltonian which is a homogeneous quadratic expression in a finite number of boson construction operators. The treatment provides us with several systematic procedures to diagonalize such hamiltonians in practice; each algorithm in principle requires only a single unitary diagonalization of a hermitian matrix.

1. Introduction

We consider the diagonalization of a hamiltonian of the general form (m arbitrary positive integer)

$$\mathbf{H} \equiv \sum_{r',r=1}^{m} \{\boldsymbol{\alpha}_{r'}^{\dagger} D_{1r'r} \boldsymbol{\alpha}_r + \boldsymbol{\alpha}_{r'}^{\dagger} D_{2r'r} \boldsymbol{\alpha}_r^{\dagger} + \boldsymbol{\alpha}_{r'} D_{3r'r} \boldsymbol{\alpha}_r + \boldsymbol{\alpha}_{r'} D_{4r'r} \boldsymbol{\alpha}_r^{\dagger}\}, \tag{1.1}$$

which is a homogeneous quadratic expression in a finite number of construction operators $\boldsymbol{\alpha}_r, \boldsymbol{\alpha}_r^{\dagger}$. These operators satisfy the well-known commutation relations

$$[\boldsymbol{\alpha}_{r'}, \boldsymbol{\alpha}_r^{\dagger}]_{-\eta} \equiv \boldsymbol{\alpha}_{r'} \boldsymbol{\alpha}_r^{\dagger} - \eta \boldsymbol{\alpha}_r^{\dagger} \boldsymbol{\alpha}_{r'} = \delta_{r'r}, \quad [\boldsymbol{\alpha}_{r'}, \boldsymbol{\alpha}_r]_{-\eta} = [\boldsymbol{\alpha}_{r'}^{\dagger}, \boldsymbol{\alpha}_r^{\dagger}]_{-\eta} = 0, \tag{1.2}$$

where $\eta = +1$ or $= -1$ depending on whether the m operator pairs are of the boson or fermion type, respectively. The coefficients in eq. (1.1) have been written as elements of the m-square matrices D_i ($i = 1, 2, 3, 4$) and can be complex numbers. The diagonalization of the hamiltonian (1.1) amounts to the application of a homogeneous linear transformation of the set of m pairs $\boldsymbol{\alpha}_r, \boldsymbol{\alpha}_r^{\dagger}$ into another set of m construction-operator pairs $\boldsymbol{\gamma}_r, \boldsymbol{\gamma}_r^{\dagger}$, satisfying commutation relations (1.2) with $\boldsymbol{\alpha}$ replaced by $\boldsymbol{\gamma}$, which rewrites eq. (1.1) into the form

$$\mathbf{H} = \sum_{r=1}^{m} \hbar\omega_r \boldsymbol{\gamma}_r^{\dagger} \boldsymbol{\gamma}_r + \text{constant}. \tag{1.3}$$

As is well known the spectrum of $\mathbf{H}$ can be characterized just by the values of the m mode energies $\hbar\omega_r$ which occur in the right-hand side of eq. (1.3).

A linear transformation of a set of construction operators into another such set is called a Bogoliubov transformation after Bogoliubov[1]) who carried out a diagonalization by means of such a linear transformation. The same diagonalization procedure had been used earlier by Holstein and Primakoff[2]). Sometimes, but not in this paper, the name Bogoliubov transformation is restricted to a *diagonalizing* linear transformation.

In the fermion case the diagonalization of the hamiltonian is rather easy: only unitary diagonalizations of hermitian matrices are involved. In this connection we refer to the very neat and elegant method presented by Lieb et al.[3]) which is applicable in the case of *real* coefficients $D_{ir'r}$. See also subsection 6.4 for the more general case of complex coefficients; usually the order of the matrices to be diagonalized is then two times larger.

In the boson case the situation is entirely different. Like in the fermion case one can write down the "equations of motion" for the $\boldsymbol{\gamma}_r$: $[\boldsymbol{\gamma}_r, \mathbf{H}] = \hbar\omega_r\boldsymbol{\gamma}_r$ $(r = 1, 2, \ldots, m)$. After substituting for $\boldsymbol{\gamma}_r$ a linear expression in the $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}^\dagger$ with unknown coefficients and for $\mathbf{H}$ eq. (1.1), we easily derive using eq. (1.2) an eigenvalue problem in $\hbar\omega_r$ ‡ in which the eigenvectors give the coefficients in the diagonalizing Bogoliubov transformation. At this point difficulties come in because in general the problem does not appear as a unitary diagonalization of a hermitian matrix so that one cannot use the theorems which exist for this well-known problem. Many examples of diagonalizations have been worked out in the literature but hardly any general approach can be found: no systematic procedure is followed and usually the diagonalization takes place in several steps, the whole procedure being very specific for the problem under consideration§. Almost always the result is a *hermitian*, so-called dynamical, matrix, the eigenvalues of which are the *squares* of the mode energies. In accordance with this many treatments seem to suggest[9]) that the dynamical matrix is the natural entity to look for when solving a boson diagonalization problem. Although the knowledge of the dynamical matrix may have great advantages for calculations so that it is worth looking for it, it will be made plausible in this paper that its existence is not guaranteed. In appendix A a theorem is proved which gives an impression of the strong conditions on the coefficients in the hamiltonian (1.1) for the existence of the dynamical matrix and which often can be applied for its construction.

Obviously a general approach to the boson diagonalization problem would

‡ Essentially the same discussion can be given in the language of the Green-function formalism leading to the same eigenvalue problem[4]). For the relatively simple problem under consideration in this paper, use of Green functions does not seem to have any advantage. There exist other ways than from the equations of motion to obtain the same eigenvalue problem: the lines of argument which we shall follow in this paper to arrive at the eigenvalue problem is also very familiar[5]).

§ Compare e.g. the diagonalizations in the theory of phonons[6,7]), in refs. 4 and 8 and in ref. 5.

be of great advantage. However, the most general approaches we could find are contained in ref. 5 and in the book of Berezin[10]). Lindgård et al.[5]) solve the problem only for rather special types of hamiltonians, whereas Berezin does not completely diagonalize the hamiltonian but indeed gives a method for solving the problem‡. This method, however, does not seem to be very useful in practice (see subsection 6.6 of this paper) and can only be applied to hamiltonians with all coefficients *real*.

In the present paper we give an elementary but detailed matrix-algebraic analysis of the whole problem of the diagonalization of a hamiltonian of the form (1.1) in the boson case, i.e. under which conditions the diagonalization is actually possible and how one can carry it out in practice. The algorithms presented all involve only a single unitary diagonalization of a hermitian matrix, for the computer solution of which problem techniques are available which are much more advanced than those for the solution of a general eigenvalue problem. In subsection 6.2 it will be indicated which of the algorithms in an actual case is most efficient.

In sections 2–3 we treat the particular kind of quadratic hamiltonian which occurs in solid-state physics; after that the most general hamiltonian (1.1) is easily handled in section 4. A simpler diagonalization procedure, which can be applied in many cases, is presented in section 5. We conclude this paper with a discussion (section 6), including also the remark (subsection 6.4) that all considerations of this paper are easily translated into the fermion case so that for example also general *fermion* hamiltonians with *complex* coefficients can be treated (cf. ref. 3). In appendix A a theorem concerning dynamical matrices is proved and in appendix B we give some additional comments on the matrix-algebraic terminology of this paper.

2. Diagonalization of the so-called Bogoliubov boson hamiltonian. Introduction

In this and the subsequent section we investigate the special case in which the hamiltonian $\mathbf{H}$ has the form:

$$\mathbf{H} = \sum_{r',r=1}^{m} \{\boldsymbol{\alpha}_{r'}^{\dagger} \Delta_{1r'r} \boldsymbol{\alpha}_r + \boldsymbol{\alpha}_{r'}^{\dagger} \Delta_{2r'r} \boldsymbol{\alpha}_{m+r}^{\dagger} + \boldsymbol{\alpha}_{m+r'} \Delta_{3r'r} \boldsymbol{\alpha}_r + \boldsymbol{\alpha}_{m+r'} \Delta_{4r'r} \boldsymbol{\alpha}_{m+r}^{\dagger}\}, \tag{2.1}$$

where Δ_i $(i = 1, 2, 3, 4)$ are m-square matrices with complex numbers as elements. $\mathbf{H}$ contains m boson-operator pairs $\boldsymbol{\alpha}_r$, $\boldsymbol{\alpha}_r^{\dagger}$ and m pairs $\boldsymbol{\alpha}_{m+r}$, $\boldsymbol{\alpha}_{m+r}^{\dagger}$,

‡ Berezin was only interested in the proof of the *existence* of a diagonalizing transformation for the much harder case of an *infinite* number of construction operators.

i.e. $2m$ pairs in total, which satisfy commutation relations like eq. (1.2) ($\eta = 1$)‡:

$$[\boldsymbol{\alpha}_{\rho'}, \boldsymbol{\alpha}^{\dagger}_{\rho}] = \delta_{\rho'\rho}, \qquad [\boldsymbol{\alpha}_{\rho'}, \boldsymbol{\alpha}_{\rho}] = [\boldsymbol{\alpha}^{\dagger}_{\rho'}, \boldsymbol{\alpha}^{\dagger}_{\rho}] = 0. \tag{2.2}$$

Comparison of eq. (2.1) – involving a set of $2m$ instead of m boson-operator pairs – with the more general class of hamiltonians (1.1), shows that the hamiltonian (2.1) can be considered as a bilinear expression [terms like $(\boldsymbol{\alpha}_\rho)^2$ where one operator is squared, do not occur] of a special type. A hamiltonian of the form (2.1) is called a Bogoliubov hamiltonian. The reason why we consider first eq. (2.1) is that it often occurs in practice [in solid-state physics e.g. all operators with index r correspond to the same wave vector $\boldsymbol{k}$, those with $m + r$ to $-\boldsymbol{k}$; m denotes the number of degrees of freedom in the unit cell (or less)]. Furthermore, having solved the diagonalization problem for eq. (2.1), the diagonalization of the general form (1.1), though following a somewhat different scheme, is treated very easily (section 4).

The hermiticity of the operator (2.1) means that the $2m$-square matrix§

$$\mathscr{D} \equiv \begin{pmatrix} \Delta_1 & \Delta_2 \\ \Delta_3 & \Delta_4 \end{pmatrix} \tag{2.3}$$

is hermitian. This matrix $\mathscr{D}$ plays a dominant role in the rest of the paper and will be called the grand-dynamical matrix of the Bogoliubov hamiltonian (2.1). We rewrite eq. (2.1) in matrix notation:

$$\mathbf{H} \equiv \boldsymbol{a}^{\dagger}\mathscr{D}\boldsymbol{a}, \qquad \boldsymbol{a}^{\dagger} \equiv (\boldsymbol{\alpha}^{\dagger}\ \boldsymbol{\alpha}_{m+}), \quad \boldsymbol{a} \equiv \begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\alpha}^{\dagger}_{m+} \end{pmatrix}, \tag{2.4}$$

where $\boldsymbol{a}^{\dagger}$ and $\boldsymbol{a}$ are a row and column vector, respectively, each with $2m$ construction operators as elements, and where $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}^{\dagger}_{m+}$ are column m-vectors.

In the introduction we indicated that the diagonalization of $\mathbf{H}$ amounts to a homogeneous linear transformation of the set of construction operators $\boldsymbol{\alpha}^{\dagger}_1, \ldots, \boldsymbol{\alpha}^{\dagger}_m, \boldsymbol{\alpha}^{\dagger}_{m+1}, \ldots, \boldsymbol{\alpha}^{\dagger}_{2m}, \boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_m, \boldsymbol{\alpha}_{m+1}, \ldots, \boldsymbol{\alpha}_{2m}$ into a suitable analogous set of construction operators $\boldsymbol{\gamma}$. In section 3 we will prove that in the case of eq. (2.1) we can choose a transformation

$$(\boldsymbol{\gamma}^{\dagger}\ \boldsymbol{\gamma}_{m+}) \equiv \boldsymbol{c}^{\dagger} = \boldsymbol{a}^{\dagger}\mathscr{T}^{\dagger} \equiv (\boldsymbol{\alpha}^{\dagger}\ \boldsymbol{\alpha}_{m+})\mathscr{T}^{\dagger} \text{ and, equivalently,}$$
$$\begin{pmatrix} \boldsymbol{\gamma} \\ \boldsymbol{\gamma}^{\dagger}_{m+} \end{pmatrix} \equiv \boldsymbol{c} = \mathscr{T}\boldsymbol{a} \equiv \mathscr{T}\begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\alpha}^{\dagger}_{m+} \end{pmatrix}, \tag{2.5}$$

‡ In this paper indices ρ', ρ can assume the values $1, 2, \ldots, 2m$ and indices r', r the values $1, 2, \ldots, m$.

§ In this paper matrices often are represented in partitioned form: a $2m$-square matrix is represented by four m-square matrices and accordingly $2m$-vectors by two m-vectors. Script letters are used to indicate $2m$-square matrices and $2m$-vectors, Latin or Greek letters are used for m-square matrices and m-vectors.

so that the $2m$-square transformation matrix $\mathcal{T}$ transforms $\mathbf{H}$ into a diagonal form according to the scheme

$$\mathbf{H} \equiv \boldsymbol{\alpha}^{\dagger}\mathcal{D}\boldsymbol{\alpha} = \boldsymbol{\alpha}^{\dagger}\mathcal{T}^{\dagger}(\mathcal{T}^{\dagger})^{-1}\mathcal{D}\mathcal{T}^{-1}\mathcal{T}\boldsymbol{\alpha} = \boldsymbol{c}^{\dagger}\mathscr{E}\boldsymbol{c}, \tag{2.6}$$

where $\mathscr{E}$ is a diagonal matrix

$$(\mathcal{T}^{\dagger})^{-1}\mathcal{D}\mathcal{T}^{-1} = \mathscr{E} \equiv \hbar\,\mathrm{diag}(\omega_1, \omega_2, \ldots, \omega_m, \omega_{m+1}, \omega_{m+2}, \ldots, \omega_{2m}). \tag{2.7}$$

The diagonalized hamiltonian is

$$\mathbf{H} = \sum_{r=1}^{m} (\hbar\omega_r \boldsymbol{\gamma}_r^{\dagger}\boldsymbol{\gamma}_r + \hbar\omega_{m+r}\boldsymbol{\gamma}_{m+r}\boldsymbol{\gamma}_{m+r}^{\dagger}) = \sum_{\rho=1}^{2m} \hbar\omega_\rho(\boldsymbol{\gamma}_\rho^{\dagger}\boldsymbol{\gamma}_\rho + \tfrac{1}{2}) - \tfrac{1}{2}\left[\sum_{\rho=1}^{m}\hbar\omega_\rho - \sum_{\rho=m+1}^{2m}\hbar\omega_\rho\right]. \tag{2.8}$$

We conclude this section by some introductory remarks for the treatment in section 3, where the conditions will be given under which the diagonalization of $\mathbf{H}$ is possible and where an algorithm will be presented to perform the diagonalization in practice‡. Defining the "para Kronecker symbol"* $\hat{\delta}_{\rho'\rho}$ by

$$\begin{aligned} &\hat{\delta}_{\rho\rho} = 1 && \text{if } 1 \leqq \rho \leqq m, \\ &\hat{\delta}_{\rho\rho} = -1 && \text{if } m+1 \leqq \rho \leqq 2m, \\ &\hat{\delta}_{\rho'\rho} = 0 && \text{if } \rho' \neq \rho, \end{aligned} \tag{2.9}$$

we derive the conditions which should hold for the elements of $\mathcal{T}$ in order that the $\boldsymbol{\gamma}$ constitute a set of boson operators provided the $\boldsymbol{\alpha}$ do too:

$$\hat{\delta}_{\rho'\rho} = [\boldsymbol{c}_{\rho'}, \boldsymbol{c}_\rho^{\dagger}] = [(\mathcal{T}\boldsymbol{\alpha})_{\rho'}, (\boldsymbol{\alpha}^{\dagger}\mathcal{T}^{\dagger})_\rho] = \sum_{\rho''=1}^{m} \mathcal{T}_{\rho'\rho''}\mathcal{T}^*_{\rho\rho''} - \sum_{\rho''=m+1}^{2m} \mathcal{T}_{\rho'\rho''}\mathcal{T}^*_{\rho\rho''}, \tag{2.10a}$$

or in matrix notation:

$$\mathcal{T}\hat{\mathcal{I}}\mathcal{T}^{\dagger} = \hat{\mathcal{I}} \quad \text{or} \quad \mathcal{T}^{\dagger}\hat{\mathcal{I}}\mathcal{T} = \hat{\mathcal{I}} \quad \text{or} \quad \mathcal{T}^{\dagger}\hat{\mathcal{I}} = \hat{\mathcal{I}}\mathcal{T}^{-1}, \tag{2.10b}$$

where the "para unit matrix" $\hat{\mathcal{I}} \equiv \mathrm{diag}(1, 1, \ldots, 1, -1, -1, \ldots, -1)$ (note: $\hat{\mathcal{I}}_{\rho'\rho} = \hat{\delta}_{\rho'\rho}$). Here the third and the second equation are most easily derived from the first and third equation, respectively. In view of the close resemblance of eqs. (2.10a) and (2.10b) to the definition of unitarity we shall call a

‡ In case one is only interested in the prescriptions of the algorithm (not in the argumentation), see section 3, remark 1, which can be read at this moment if one does not bother about the meaning of the concept "para-unitary" and if one defines the $2m$-square matrix $\hat{\mathcal{I}}$ according to: $\hat{\mathcal{I}} \equiv \mathrm{diag}(1, 1, \ldots, 1, -1, -1, \ldots, -1)$.

* In the treatment of the fermion case on this place the Kronecker symbol $\delta_{\rho'\rho}$ is encountered. In situations like this we use the word or prefix "para" to provide a comparison with the fermion treatment (see also subsection 6.4).

matrix para-unitary‡ if its elements satisfy one of the four equivalent equations (2.10a) and (2.10b). Finally, by taking the inverse of the first eq. (2.10b), we see from the second definition eq. (2.10b) that the inverse $\mathcal{T}^{-1}$ of a para-unitary matrix $\mathcal{T}$ is also para-unitary.

Note that for the fermion case we would have obtained instead of eqs. (2.10a) and (2.10b) the unitarity condition for the matrix $\mathcal{T}$. In view of this the problem of diagonalizing a fermion Bogoliubov hamiltonian can be considered as solved completely (see subsection 6.4). To see more clearly the specific difficulties related to the diagonalization problem (2.7) in the boson case, we rewrite eq. (2.7) as:

$$(\mathcal{T}^{\dagger})^{-1}\mathcal{D}\mathcal{T}^{-1} = \mathcal{E} \equiv \hbar\ \mathrm{diag}(\omega_1, \omega_2, \dots, \omega_m, \omega_{m+1}, \omega_{m+2}, \dots, \omega_{2m})$$
$$\equiv \mathrm{diag}(\lambda_1, \lambda_2, \dots, \lambda_m, -\lambda_{m+1}, -\lambda_{m+2}, \dots, -\lambda_{2m}) \equiv \hat{\mathcal{I}}\mathcal{L}, \tag{2.11}$$

where the diagonal matrix $\mathcal{L}$ has been introduced to simplify the notation. Multiplying this equation on the left by $\mathcal{T}^{\dagger}$ we obtain according to the third eq. (2.10b):

$$\mathcal{D}\mathcal{T}^{-1} = \mathcal{T}^{\dagger}\hat{\mathcal{I}}\mathcal{L} = \hat{\mathcal{I}}\mathcal{T}^{-1}\mathcal{L}. \tag{2.12}$$

From eq. (2.12) we see that the ρth column $w_\rho \equiv (u_\rho\, v_\rho)'$ ($\rho = 1, 2, \dots, 2m$; u_ρ and v_ρ are column m-vectors and the prime denotes a column vector) of $\mathcal{T}^{-1}$ satisfies

$$\mathcal{D}w_\rho = \lambda_\rho\hat{\mathcal{I}}w_\rho \quad \text{or} \quad (\mathcal{D} - \lambda_\rho\hat{\mathcal{I}})w_\rho = 0 \quad \text{or} \quad \begin{pmatrix} \Delta_1 & \Delta_2 \\ \Delta_3 & \Delta_4 \end{pmatrix}\begin{pmatrix} u_\rho \\ v_\rho \end{pmatrix} = \lambda_\rho\begin{pmatrix} u_\rho \\ -v_\rho \end{pmatrix}, \tag{2.13}$$

and for the diagonal elements λ_ρ of $\mathcal{L}$ we obtain the equation§:

$$\det(\mathcal{D} - \lambda_\rho\hat{\mathcal{I}}) = 0 \quad \text{or} \quad \begin{vmatrix} \Delta_1 - \lambda_\rho I & \Delta_2 \\ \Delta_3 & \Delta_4 + \lambda_\rho I \end{vmatrix} = 0, \tag{2.14}$$

in which I is the m-square unit matrix. The difficulties connected with the diagonalization problem (2.7) are now easily recognized by multiplying eq. (2.13) on the left by $\hat{\mathcal{I}}$:

$$(\hat{\mathcal{D}} - \lambda_\rho\mathcal{I})w_\rho \equiv (\hat{\mathcal{I}}\mathcal{D} - \lambda_\rho\mathcal{I})w_\rho = 0 \quad \text{or} \quad \begin{pmatrix} \Delta_1 - \lambda_\rho I & \Delta_2 \\ -\Delta_3 & -\Delta_4 - \lambda_\rho I \end{pmatrix} w_\rho = 0, \tag{2.15}$$

‡ In the existing literature other prefixes are used for this concept, e.g. pseudo-unitary. In the present treatment we use the prefix to indicate counterpart concepts so that it appears in many words: para-eigenvector, paranorm, etc. but not in others: the counterpart (matrix) of a hermitian matrix is also hermitian, *not* para-hermitian. To avoid unwanted interference with other concepts we use the in matrix algebra unusual prefix "para-".

§ In this or almost this form the boson diagonalization problem is often formulated in the literature[4,5]).

where $\mathcal{I}$ is the $2m$-square unit matrix. The main questions to be solved can now be formulated. (i) Are the eigenvalues of the (in general non-hermitian) matrix $\hat{\mathcal{D}} \equiv \hat{\mathcal{I}}\mathcal{D}$ real? (ii) Does a para-unitary matrix $\mathcal{T}^{-1}$ exist with the eigenvectors of $\hat{\mathcal{D}}$ as columns? (iii) How can a para-unitary matrix $\mathcal{T}^{-1}$ be calculated in practice which diagonalizes the matrix $\mathcal{D}$ according to eq. (2.7)? In section 3 answers to these questions are given and a practical diagonalization algorithm is presented which involves only a single unitary diagonalization of a hermitian matrix.

3. Diagonalization of the Bogoliubov boson hamiltonian. Conditions and algorithm

In this section we investigate the conditions under which the diagonalization of the boson Bogoliubov hamiltonian (2.1) is possible according to the scheme (2.6). Since a hamiltonian has a lowest energy level the mode energies $\hbar\omega_\rho$ in eq. (2.7) should be non-negative [cf. eq. (2.8)]. Now it will turn out (see the last two examples in remark 4 at the end of this section) that mathematical difficulties occur in case we allow zero to be a mode energy (soft mode). Therefore we exclude this case from our treatment and give only a brief discussion on this restriction in subsection 6.5. The remaining problem is to find the conditions under which the para-unitary diagonalization of $\mathcal{D}$ according to eq. (2.7) is possible with all diagonal elements in $\mathcal{E}$ positive. The answer is given by the following theorem.

Theorem. A $2m$-square hermitian matrix $\mathcal{D}$ can be para-unitarily diagonalized into a matrix with all diagonal elements positive if and only if $\mathcal{D}$ is positive definite.

Proof of the necessity of the positive definiteness. Since $(\mathcal{T}^\dagger)^{-1} = (\mathcal{T}^{-1})^\dagger$, eq. (2.7) has the form of a conjunctivity transformation [i.e. the form $S^\dagger PS = Q$, P (and Q) being hermitian matrices and S a non-singular matrix]. For such transformations we have Sylvester's law of inertia which states that all conjunctivity diagonalizations of a certain (hermitian) matrix P lead to (diagonal) matrices for which the numbers p, n and z of positive, negative and zero diagonal elements, respectively, depend only on the matrix P and not on the particular choice of the diagonalizing matrix S (the diagonal elements are real because the matrix Q is hermitian). Since according to our assumptions eq. (2.7) requires that the matrix $\mathcal{D}$ can be diagonalized by a conjunctivity transformation into a matrix with all diagonal elements positive, we conclude that the familiar unitary diagonalization of $\mathcal{D}$ too should lead to a matrix with all diagonal elements positive. This means that the matrix $\mathcal{D}$ should be positive definite.

Proof of the sufficiency of the positive definiteness. We prove this by presenting a construction for the diagonalizing para-unitary matrix $\mathcal{T}^{-1}$ in eq. (2.7).

Since $\mathcal{D}$ is positive definite it is possible (see for example remark 1 after the proof) to find a (non-singular) $2m$-square matrix $\mathcal{K}$ which satisfies $\mathcal{D} = \mathcal{K}^\dagger\mathcal{K}$. Let $\mathcal{T}^{-1}$ be a para-unitary matrix which diagonalizes $\mathcal{D}$ according to eq. (2.7); we don't know yet whether such a matrix exists. If we denote in what follows by $\mathcal{E}^{\pm 1/2}$ the matrix with diagonal elements $\epsilon_\rho^{\pm 1/2}$ in case $\mathcal{E}$ is a diagonal matrix with all diagonal elements $\epsilon_\rho > 0$, then the matrix $\mathcal{U}$ defined by

$$\mathcal{U} \equiv \mathcal{K}\mathcal{T}^{-1}\mathcal{E}^{-1/2} \tag{3.1}$$

is unitary since ($\mathcal{I}$ is the $2m$-square unit matrix)

$$\mathcal{U}^\dagger\mathcal{U} = \mathcal{E}^{-1/2}(\mathcal{T}^{-1})^\dagger\mathcal{K}^\dagger\mathcal{K}\mathcal{T}^{-1}\mathcal{E}^{-1/2} = \mathcal{E}^{-1/2}(\mathcal{T}^{-1})^\dagger\mathcal{D}\mathcal{T}^{-1}\mathcal{E}^{-1/2} = \mathcal{E}^{-1/2}\mathcal{E}\mathcal{E}^{-1/2} = \mathcal{I}. \tag{3.2}$$

Moreover $\mathcal{U}$ diagonalizes the (hermitian) matrix $\mathcal{K}\hat{\mathcal{I}}\mathcal{K}^\dagger$, i.e.

$$\begin{aligned}\mathcal{U}^\dagger[\mathcal{K}\hat{\mathcal{I}}\mathcal{K}^\dagger]\mathcal{U} = \mathcal{U}^{-1}[\mathcal{K}\hat{\mathcal{I}}\mathcal{K}^\dagger](\mathcal{U}^\dagger)^{-1} &= \mathcal{E}^{1/2}\mathcal{T}\mathcal{K}^{-1}[\mathcal{K}\hat{\mathcal{I}}\mathcal{K}^\dagger](\mathcal{K}^\dagger)^{-1}\mathcal{T}^\dagger\mathcal{E}^{1/2} \\ &= \mathcal{E}^{1/2}\mathcal{T}\hat{\mathcal{I}}\mathcal{T}^\dagger\mathcal{E}^{1/2} = \mathcal{E}^{1/2}\hat{\mathcal{I}}\mathcal{E}^{1/2} = \mathcal{L},\end{aligned} \tag{3.3}$$

where the first eq. (2.10b) has been used and where the diagonal matrix $\mathcal{L}$ has been defined by eq. (2.11).

Let us conversely choose a unitary matrix $\mathcal{U}$ which diagonalizes the (hermitian) matrix $\mathcal{K}\hat{\mathcal{I}}\mathcal{K}^\dagger$ according to eq. (3.3) in such a way that the first half of the diagonal of $\mathcal{L}$ contains only positive elements and the second half only negative elements [Sylvester's law of inertia, applied to the first and last member of eq. (3.3), tells us that the number of positive elements and also the number of negative elements in $\mathcal{L}$ and $\hat{\mathcal{I}}$ is the same]. Such a matrix $\mathcal{U}$ provides us by eq. (3.1) with a para-unitary matrix $\mathcal{T}^{-1}$ which diagonalizes $\mathcal{D}$ according to eq. (2.7). Indeed, since by eqs. (3.1) and (3.3):

$$(\mathcal{T}^{-1})^\dagger\hat{\mathcal{I}}\mathcal{T}^{-1} = \mathcal{E}^{1/2}\mathcal{U}^\dagger(\mathcal{K}^\dagger)^{-1}\hat{\mathcal{I}}\mathcal{K}^{-1}\mathcal{U}\mathcal{E}^{1/2} = \mathcal{E}^{1/2}\mathcal{L}^{-1}\mathcal{E}^{1/2} = \hat{\mathcal{I}}, \tag{3.4}$$

the matrix $\mathcal{T}^{-1}$ is para-unitary according to the second eq. (2.10b). Furthermore, $\mathcal{T}^{-1}$ diagonalizes $\mathcal{D}$:

$$(\mathcal{T}^{-1})^\dagger\mathcal{D}\mathcal{T}^{-1} = \mathcal{E}^{1/2}\mathcal{U}^\dagger(\mathcal{K}^\dagger)^{-1}(\mathcal{K}^\dagger\mathcal{K})\mathcal{K}^{-1}\mathcal{U}\mathcal{E}^{1/2} = \mathcal{E}. \tag{3.5}$$

With the existence of a matrix $\mathcal{T}^{-1}$ with these properties we have shown the possibility of the para-unitary diagonalization of $\mathcal{D}$ and the theorem has been proved completely.

Remark 1. *Algorithm.* The proof given above suggests an algorithm – essentially amounting to a single unitary diagonalization of a hermitian matrix – for finding a para-unitary matrix $\mathcal{T}^{-1}$, which diagonalizes a given positive-definite hermitian matrix $\mathcal{D}$ according to eq. (2.7). Firstly we have to find a decomposition of the $2m$-square positive-definite hermitian matrix $\mathcal{D}$ in the form $\mathcal{D} = \mathcal{K}^{\dagger}\mathcal{K}$, where $\mathcal{K}$ is a $2m$-square matrix. In computer programs it is convenient to apply the so-called Cholesky decomposition (see for example ref. 11, where the considerations for real matrices are easily generalized to complex matrices) and find a (non-singular) complex matrix $\mathcal{K}$ which is upper-triangular (i.e. an element vanishes if the column index is smaller than the row index). The failure of the algorithm (by which we also understand the case that only a *singular* matrix $\mathcal{K}$ can be found, see also subsection 6.5) in case $\mathcal{D}$ is not positive definite can be used as a check on this property of $\mathcal{D}$. The computer time needed for a Cholesky decomposition is entirely negligible in comparison with the time entailed in a unitary diagonalization of a hermitian matrix, which also can be used to find a decomposition (see the end of subsection 6.5). Secondly we diagonalize unitarily the hermitian matrix $\mathcal{K}\hat{\mathcal{J}}\mathcal{K}^{\dagger}$ and find a unitary matrix $\mathcal{U}$, in which the columns are arranged in such a way that [see eq. (3.3)] the first m diagonal elements of the diagonalized matrix $\mathcal{U}^{\dagger}(\mathcal{K}\hat{\mathcal{J}}\mathcal{K}^{\dagger})\mathcal{U} = \mathcal{L}$ are positive, the last m negative; the diagonal matrix $\mathcal{E}$ in eq. (2.7) is then given by $\mathcal{E} = \hat{\mathcal{J}}\mathcal{L}$. Finally we calculate the matrix $\mathcal{T}^{-1}$ row by row, beginning with the last row, from eq. (3.1), written in the form $\mathcal{U}\mathcal{E}^{1/2} = \mathcal{K}\mathcal{T}^{-1}$. For testing computer programs it may be useful to realize that $2m$-square para-unitary matrices $\mathcal{T}$ are easily constructed from arbitrary m-square unitary matrices T. For example, if T is such a matrix and if c and s are any complex numbers with $|c|^2 - |s|^2 = 1$, then the matrix

$$\mathcal{T} \equiv \begin{pmatrix} cT & s^*T^* \\ sT & c^*T^* \end{pmatrix} \tag{3.6}$$

is para-unitary (see the first footnote in section 4 for the meaning of T^*). We further remark that no additional calculation of $\mathcal{T}$ is needed once $\mathcal{T}^{-1}$ is known. For with the help of the last eq. (2.10b) we see that for a para-unitary matrix $\mathcal{T}$ holds:

$$\mathcal{T}^{-1} \equiv \begin{pmatrix} U & W \\ V & X \end{pmatrix}, \quad \mathcal{T} \equiv \begin{pmatrix} U^{\dagger} & -V^{\dagger} \\ -W^{\dagger} & X^{\dagger} \end{pmatrix}. \tag{3.7}$$

Remark 2. *The generalized eigenvalue problem.* Eq. (2.13) has the form of what is known in matrix algebra as the "generalized eigenvalue problem":

$$(P - \lambda Q)w = 0, \tag{3.8}$$

where P and Q are n-square matrices; the complex constant λ and the n-vector w are called a generalized eigenvalue and a corresponding generalized eigenvector, respectively. In general the way of solving this problem (i.e.

determining values for λ and w) depends on the properties of P and Q. Our proof of the sufficiency of the positive definiteness above is based on a well-known solution of the problem (3.8) in the case that P and Q are hermitian matrices and moreover P is positive definite. One just performs the decomposition $P = K^\dagger K$ and finds the matrix R with $Q = K^\dagger R K$. Eq. (3.8) can then be written

$$K^\dagger(\lambda^{-1} I - R) K w = 0 \quad \text{or} \quad (R - \lambda^{-1} I) K w = 0, \tag{3.9}$$

which obviously is the simple eigenvalue problem connected with the hermitian matrix R (eigenvalues λ^{-1}, eigenvectors Kw).

Remark 3. *Another point of view.* In the proof above it was not necessary for the construction of the matrix $\mathcal{T}^{-1}$ to consider its individual columns [cf. question (ii) at the end of section 2]. In many cases [for example in case $m = 1$ so that the problem (2.13) can be solved by hand (see examples in remark 4 below), in section 4 and in appendix A] another point of view is convenient. For ρ' and ρ fixed eq. (2.10a) expresses a relation which holds between the elements of the ρ'th and ρth row in a para-unitary matrix. Expressing the second eq. (2.10b) in terms of the matrix elements we see that similar relations hold between the columns. Now we call two $2m$-vectors ${}_1 \equiv (u_1\, v_1)'$ and $w_2 \equiv (u_2\, v_2)'$ (u and v are column m-vectors and the prime denotes a column vector) paraperpendicular (to each other) if $w_1^\dagger \hat{\mathcal{J}} w_2 \equiv u_1^\dagger u_2 - v_1^\dagger v_2 = 0$ and denote the real value ${}^\dagger \hat{\mathcal{J}} w \equiv u^\dagger u - v^\dagger v$ as the paranorm of a vector $w \equiv (u\, v)'$ (note that one cannot change the sign of a paranorm by multiplication of the vector by a scalar). Then we can characterize a $2m$-square para-unitary matrix as a matrix in which the columns (or rows) are paraperpendicular and paranormalized, i.e. the paranorm of each of the first m columns equals 1, that of each of the last m columns -1. Calling furthermore the vectors w_ρ in eq. (2.13) para-eigenvectors of $\mathcal{D}$ at the para-eigenvalue λ_ρ (in what follows short: paravectors and paravalue), a comparison with eq. (2.11) shows that the problem in section 2 was to find $2m$ paravalues of $\mathcal{D}$, the first m of which should be positive, the last m negative. In order to find the para-unitary matrix $\mathcal{T}^{-1}$ we need $2m$ paravectors of $\mathcal{D}$ which correspond to the paravalues and which are paraperpendicular. The positive definiteness of $\mathcal{D}$ ensures that the first (last) m of them have a positive (negative) paranorm and after (para)normalization the paravectors can serve as the first (last) m columns of $\mathcal{T}^{-1}$. For we have for a paravector $w \equiv (u\, v)'$ of $\mathcal{D}$ at paravalue λ:

$$0 < (u^\dagger\ v^\dagger) \begin{pmatrix} \Delta_1 & \Delta_2 \\ \Delta_3 & \Delta_4 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = (u^\dagger\ v^\dagger) \lambda \begin{pmatrix} u \\ -v \end{pmatrix} = \lambda (u^\dagger u - v^\dagger v), \tag{3.10}$$

so that the paravalues of such $\mathcal{D}$ are real (this was seen already in the proof of the theorem above) and positive and negative paravalues correspond to paravectors with positive and negative paranorms, respectively. Note finally that paravectors w_1 and w_2 of a positive-definite hermitian matrix $\mathcal{D}$ at different paravalues λ_1 and λ_2 are paraperpendicular, which is easily seen by subtracting from

$$(u_2^\dagger\, v_2^\dagger)\mathcal{D}\binom{u_1}{v_1} = (u_2^\dagger\, v_2^\dagger)\lambda_1\binom{u_1}{-v_1} = \lambda_1(u_2^\dagger u_1 - v_2^\dagger v_1) \tag{3.11}$$

the complex conjugate with all subscripts 1 and 2 interchanged.

Remark 4. *Counterexamples.* We have seen that a large class of hermitian bilinear expressions of the form (2.1) can be diagonalized, viz. all expressions with a positive- (or of course negative-)definite hermitian grand-dynamical matrix (2.3). It is evident that this class does not exhaust all boson Bogoliubov operators which can be diagonalized. A trivial example is the expression $\mathbf{H} \equiv \boldsymbol{\alpha}_1^\dagger\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_2\boldsymbol{\alpha}_2^\dagger$ which has the form (2.1) (with $m = 1$) and obviously can be diagonalized [i.e. written into the form (1.3)], whereas the grand-dynamical matrix $\mathcal{D} \equiv \mathrm{diag}(1, -1)$ is not positive (or negative) definite. However, not all hermitian operators of the form (2.1) can be diagonalized (note that all $\hbar\omega_p$ should be real because in eq. (2.7) together with $\mathcal{D}$ also $\mathcal{C}$ is hermitian). For example, it is not possible to diagonalize the expression (again $m = 1$) $\mathbf{H} \equiv \boldsymbol{\alpha}_1^\dagger\boldsymbol{\alpha}_1 + \sqrt{2}(\boldsymbol{\alpha}_1^\dagger\boldsymbol{\alpha}_2^\dagger + \boldsymbol{\alpha}_2\boldsymbol{\alpha}_1) + \boldsymbol{\alpha}_2\boldsymbol{\alpha}_2^\dagger$, since the grand-dynamical matrix [$\Delta_1 = \Delta_4 = (1)$, $\Delta_2 = \Delta_3 = (\sqrt{2})$ has the non-real paravalues i and $-$i (see for "para" concepts used here and in what follows remark 3 above). Note that in this example the paranorms $|1+\mathrm{i}|^2 - |-\sqrt{2}|^2$ and $|1-\mathrm{i}|^2 - |-\sqrt{2}|^2$ of both paravectors $(1+\mathrm{i}\ -\sqrt{2})'$ and $(1-\mathrm{i}\ -\sqrt{2})'$ vanish whereas for diagonalizability of the expression (2.1) paranorms of $+1$ and -1 are required (see remark 3). Even in cases where the grand-dynamical matrix is positive *semi*definite, the diagonalization is not always possible. The expressions $\mathbf{H} \equiv c^2\boldsymbol{\alpha}_1^\dagger\boldsymbol{\alpha}_1 + cs(\boldsymbol{\alpha}_1^\dagger\boldsymbol{\alpha}_2^\dagger + \boldsymbol{\alpha}_2\boldsymbol{\alpha}_1) + s^2\boldsymbol{\alpha}_2\boldsymbol{\alpha}_2^\dagger$ (c and s any real numbers with $c^2 - s^2 = 1$) and $\mathbf{H} \equiv \boldsymbol{\alpha}_1^\dagger\boldsymbol{\alpha}_1 + (\boldsymbol{\alpha}_1^\dagger\boldsymbol{\alpha}_2^\dagger + \boldsymbol{\alpha}_2\boldsymbol{\alpha}_1) + \boldsymbol{\alpha}_2\boldsymbol{\alpha}_2^\dagger$ are examples of Bogoliubov operators with positive-semidefinite grand-dynamical matrices, the first of which can be diagonalized [into $\boldsymbol{\gamma}_1^\dagger\boldsymbol{\gamma}_1(+0\boldsymbol{\gamma}_2^\dagger\boldsymbol{\gamma}_2)$] whereas the second cannot since the corresponding grand-dynamical matrix has only one paravector $(1\ -1)'$. From the examples given above we consider the restriction to positive-definite grand-dynamical matrices quite natural from a mathematical point of view and satisfactory from a physical point of view (see subsection 6.5 for a calculation in connection with positive-*semi*definite grand-dynamical matrices).

4. Diagonalization of the general quadratic boson hamiltonian

We now turn to the diagonalization of the general quadratic hamiltonian (1.1):

$$\mathbf{H} \equiv \sum_{r',r=1}^{m} \{\boldsymbol{\alpha}_{r'}^{\dagger} D_{1r'r} \boldsymbol{\alpha}_r + \boldsymbol{\alpha}_{r'}^{\dagger} D_{2r'r} \boldsymbol{\alpha}_r^{\dagger} + \boldsymbol{\alpha}_{r'} D_{3r'r} \boldsymbol{\alpha}_r + \boldsymbol{\alpha}_{r'} D_{4r'r} \boldsymbol{\alpha}_r^{\dagger}\}, \tag{4.1}$$

where the m boson-operator pairs $\boldsymbol{\alpha}_r, \boldsymbol{\alpha}_r^{\dagger}$ satisfy the commutation relations (1.2) (with $\eta = 1$). An operator of the form (4.1) is hermitian if and only if the following relations hold:

$$\begin{aligned} &D_{1r'r} + D_{4rr'} = D_{1rr'}^{*} + D_{4r'r}^{*}, \qquad D_{1rr} = D_{1rr}^{*},\\ &D_{2r'r} + D_{2rr'} = D_{3rr'}^{*} + D_{3r'r}^{*}, \qquad 1 \leqq r', r \leqslant m \end{aligned} \tag{4.2}$$

(note: the first line of this equation implies that the diagonal elements of D_1 and D_4 are real). For the subsequent analysis it is essential that

$$\begin{aligned} &D_{1r'r} = D_{1rr'}^{*} = D_{4rr'}, \qquad D_{2r'r} = D_{2rr'} = D_{3rr'}^{*},\\ &D_{3r'r} = D_{3rr'} = D_{2rr'}^{*}, \qquad D_{4r'r} = D_{4rr'}^{*} = D_{1rr'}, \end{aligned} \qquad 1 \leqslant r', r \leqq m, \tag{4.3}$$

which obviously can be assumed without restricting the generality: a possibly necessary rewriting in this sense modifies the expression (4.1) – which is supposed to represent a hamiltonian and thus a hermitian operator – at most by an additive (real) constant. Then the $2m$-square matrix‡

$$\mathfrak{D} \equiv \begin{pmatrix} D_1 & D_2 \\ D_3 & D_4 \end{pmatrix} \equiv \begin{pmatrix} A & B \\ B^{*} & A^{*} \end{pmatrix} \tag{4.4}$$

has the form shown in eq. (4.4) and is hermitian (which implies that A is hermitian and B symmetric). We call $\mathfrak{D}$ the grand-dynamical matrix of the hamiltonian (4.1) and use it to rewrite eq. (4.1) in matrix notation:

$$\mathbf{H} \equiv \boldsymbol{a}^{\dagger} \mathfrak{D} \boldsymbol{a}, \qquad \boldsymbol{a}^{\dagger} \equiv (\boldsymbol{\alpha}^{\dagger}\ \boldsymbol{\alpha}), \quad \boldsymbol{a} \equiv \begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\alpha}^{\dagger} \end{pmatrix} \tag{4.5}$$

[cf. eq. (2.4); no confusion is to be expected that no extra symbols indicate that $\boldsymbol{\alpha}^{\dagger}$ and $\boldsymbol{\alpha}$ in $\boldsymbol{a}^{\dagger}$ are row m-vectors and in $\boldsymbol{a}$ column m-vectors]. To emphasize that the meaning of the concept grand-dynamical matrix depends on whether it belongs to a Bogoliubov hamiltonian of the form (2.1) or to a general hamiltonian of the form (4.1), we rewrite the Bogoliubov hamiltonian

‡ By A^{*} we denote the complex conjugate of the matrix A [elements: $(A^{*})_{r'r} = (A_{r'r})^{*}$].

(2.1) in the form (4.1) where in eq. (4.1) m is replaced by $2m$. Then the hamiltonian (2.1), which we denote by $\mathbf{H}_{\mathrm{Bog}}$ for the moment, becomes in matrix notation:

$$\mathbf{H}_{\mathrm{Bog}} \equiv (\boldsymbol{\alpha}^\dagger\ \boldsymbol{\alpha}^\dagger_{m+}\ \boldsymbol{\alpha}\ \boldsymbol{\alpha}_{m+})\tfrac{1}{2}\begin{pmatrix} \Delta_1 & 0 & 0 & \Delta_2 \\ 0 & \Delta_4^* & \Delta_3^* & 0 \\ 0 & \Delta_2^* & \Delta_1^* & 0 \\ \Delta_3 & 0 & 0 & \Delta_4 \end{pmatrix}\begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\alpha}_{m+} \\ \boldsymbol{\alpha}^\dagger \\ \boldsymbol{\alpha}^\dagger_{m+} \end{pmatrix} - \tfrac{1}{2}(\mathrm{Tr}\,\Delta_1 - \mathrm{Tr}\,\Delta_4), \tag{4.6}$$

in which the square matrix, including the factor $\frac{1}{2}$, is the grand-dynamical matrix of the general quadratic hamiltonian which matrix by definition should have the form of the right-hand side of eq. (4.4). The additive constant in eq. (4.6) is easily expressed in terms of the mode energies $\hbar\omega_\rho$ [cf. eq. (2.8)]:

$$\mathrm{Tr}\,\Delta_1 - \mathrm{Tr}\,\Delta_4 = \sum_{\rho=1}^{m} \hbar\omega_\rho - \sum_{\rho=m+1}^{2m} \hbar\omega_\rho. \tag{4.7}$$

One readily verifies, indeed, that the left-hand side of eq. (4.7) is an invariant of $\mathcal{D}_{\mathrm{Bog}}$, by which we denote the matrix in eq. (2.3) for the moment, under para-unitary transformations; the right-hand side of eq. (4.7) is the value of $\mathrm{Tr}\,\Delta_1 - \mathrm{Tr}\,\Delta_4$ for the diagonal matrix in the right-hand side of eq. (2.7).

We now proceed with the diagonalization of the general quadratic hamiltonian (4.5). Taking into account that the matrix $\mathcal{T}$ in a transformation to new construction operators $\boldsymbol{\gamma}$:

$$\boldsymbol{c} = \mathcal{T}\boldsymbol{a} \quad \text{or} \quad \begin{pmatrix} \boldsymbol{\gamma} \\ \boldsymbol{\gamma}^\dagger \end{pmatrix} = \begin{pmatrix} P & Q \\ Q^* & P^* \end{pmatrix}\begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\alpha}^\dagger \end{pmatrix} \tag{4.8}$$

has the partitioned square form shown in eq. (4.8) and that just like in eq. (2.10a) we can prove the para-unitarity of $\mathcal{T}$, we have according to eq. (3.7):

$$\mathcal{T} \equiv \begin{pmatrix} P & Q \\ Q^* & P^* \end{pmatrix}, \qquad \mathcal{T}^{-1} = \begin{pmatrix} P^\dagger & -(Q^*)^\dagger \\ -Q^\dagger & (P^*)^\dagger \end{pmatrix} \equiv \begin{pmatrix} U & V^* \\ V & U^* \end{pmatrix}. \tag{4.9}$$

Now in a diagonalization according to the scheme‡

$$\mathbf{H} \equiv \boldsymbol{a}^\dagger\mathcal{D}\boldsymbol{a} = \boldsymbol{a}^\dagger\mathcal{T}^\dagger(\mathcal{T}^\dagger)^{-1}\mathcal{D}\mathcal{T}^{-1}\mathcal{T}\boldsymbol{a} = \boldsymbol{c}^\dagger\mathcal{E}\boldsymbol{c} \tag{4.10}$$

with a matrix $\mathcal{T}$ of the form (4.9), the diagonal matrix $\mathcal{E}$ necessarily has the form:

$$(\mathcal{T}^\dagger)^{-1}\mathcal{D}\mathcal{T}^{-1} = \mathcal{E} \equiv \tfrac{1}{2}\hbar\,\mathrm{diag}(\omega_1, \omega_2, \ldots, \omega_m, \omega_1, \omega_2, \ldots, \omega_m). \tag{4.11}$$

‡ Note that, although eq. (4.10) has completely the same form as eq. (2.6), the two schemes are different: unlike the $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}^\dagger_{m+}$ in eq. (2.5), the two m-vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}^\dagger$ in $\boldsymbol{a}$ [eq. (4.5)] essentially contain the same construction operators.

The special form in the right-hand side, because the matrix $(\mathcal{T}^\dagger)^{-1}\mathcal{D}\mathcal{T}^{-1}$ is hermitian (so the diagonal elements are real), while its left-upper and right-lower m-square matrices are each other's complex conjugate if $\mathcal{D}$ and $\mathcal{T}^{-1}$ have the partitioned forms (4.4) and (4.9), respectively.

After these introductory remarks it is very easy to show that again the necessary and sufficient condition that a diagonalization according to the scheme (4.10) *with all ω positive* is possible, is that the grand-dynamical matrix $\mathcal{D}$ is positive definite‡. The necessity follows from eq. (4.11), in a similar way as in section 3. The sufficiency is seen from the following construction, also suitable for actual calculations, of a $\mathcal{T}^{-1}$ with the required properties. Let us choose a matrix $\mathcal{T}^{-1}$ which diagonalizes $\mathcal{D}$ para-unitarily. According to the theorem in section 3 such a matrix exists but the one chosen possibly has not the form (4.9). By comparing eqs. (4.11) and (2.11), we see from eq. (2.13) that the first m columns $(u_r\, v_r)'$ of $\mathcal{T}^{-1}$ satisfy:

$$\begin{aligned} Au_r + Bv_r &= \tfrac{1}{2}\hbar\omega_r u_r \\ B^*u_r + A^*v_r &= -\tfrac{1}{2}\hbar\omega_r v_r \end{aligned} \quad \text{or} \quad \begin{aligned} Av_r^* + Bu_r^* &= -\tfrac{1}{2}\hbar\omega_r v_r^* \\ B^*v_r^* + A^*u_r^* &= \tfrac{1}{2}\hbar\omega_r u_r^*, \quad r = 1, 2, \ldots, m, \end{aligned} \tag{4.12}$$

so that $(v_r^*\, u_r^*)'$ is a paravector at the paravalue $-\frac{1}{2}\hbar\omega_r$ if $(u_r\, v_r)'$ is a paravector at the paravalue $\frac{1}{2}\hbar\omega_r$ (see for "para" concepts section 3, remark 3). Since obviously together with the $(u_r\, v_r)'$ also the $(v_r^*\, u_r^*)'$ $(r = 1, 2, \ldots, m)$ are paraperpendicular to each other and moreover – according to the discussion around eq. (3.11) (different paravalues – the $(v_r^*\, u_r^*)'$ are paraperpendicular to the m vectors $(u_r\, v_r)'$, the matrix arising from the replacement of the last m columns of the chosen $\mathcal{T}^{-1}$ by $(v_1^*\, u_1^*)', (v_2^*\, u_2^*)', \ldots, (v_m^*\, u_m^*)'$, is a para-unitary matrix with the required properties: it has the form (4.9) and the diagonalization property (4.11).

To summarize, we have shown that the general quadratic boson hamiltonian (4.1) can be diagonalized with all mode energies positive if and only if the grand-dynamical matrix (4.4) is positive definite. The diagonalized hamiltonian is obtained by substitution of eqs. (4.11) and (4.8) into eq. (4.10):

$$\mathbf{H} = \sum_{r=1}^{m} \hbar\omega_r(\boldsymbol{\gamma}_r^\dagger\boldsymbol{\gamma}_r + \tfrac{1}{2}). \tag{4.13}$$

5. An often applicable simpler procedure for the diagonalization of quadratic boson hamiltonians

In sections 3 and 4 the possibility has been proved of the diagonalization of hamiltonians of the form (2.1) and (4.1), respectively. The proofs immediately

‡ The arrangement of the elements in the matrix $\mathcal{D}$ [eq. (4.4)] is essential. For example, Berezin[10]) makes a different choice and consequently has to require the positive definiteness of two matrices instead of one (see also subsection 6.6).

indicated the methods to perform the diagonalizations in practice. Although these methods were completely general, in special cases simpler procedures may be preferable. Lieb et al.[3]) described a procedure for diagonalizing a general quadratic hamiltonian of the form (4.1) in the case that the $\boldsymbol{\alpha}$ are fermion operators and the coefficients $D_{ir'r}$ are *real*. Like most of the considerations in the fermion case (see subsection 6.4) also this procedure can be translated into the boson case. This will be shown in this section because the case of real coefficients often occurs in practice and because the single unitary diagonalization to be performed in the new procedure concerns a hermitian matrix of order m (cf. the order $2m$ of the matrices in sections 2–4).

Starting from a complex positive-definite hermitian $2m$-square matrix $\mathfrak{D}$ of the form

$$\mathfrak{D} \equiv \begin{pmatrix} A & B \\ B & A \end{pmatrix}, \tag{5.1}$$

we describe a practical procedure to find a para-unitary matrix $\mathcal{T}^{-1}$ of the form [cf. eq. (3.7)]:

$$\mathcal{T}^{-1} \equiv \begin{pmatrix} U & V \\ V & U \end{pmatrix}, \qquad \mathcal{T} = \begin{pmatrix} U^\dagger & -V^\dagger \\ -V^\dagger & U^\dagger \end{pmatrix}, \tag{5.2}$$

which diagonalizes $\mathfrak{D}$ according to (in this section $\lambda_1, \lambda_2, \dots, \lambda_m$ are always positive):

$$(\mathcal{T}^\dagger)^{-1}\mathfrak{D}\mathcal{T}^{-1} = \mathscr{E} \equiv \mathrm{diag}(\lambda_1, \lambda_2, \dots, \lambda_m, \lambda_1, \lambda_2, \dots, \lambda_m). \tag{5.3}$$

By definition [see the second eq. (2.10b)] the para-unitarity of $\mathcal{T}^{-1}$ means that [u_r and v_r denote the columns of the m-square matrices U and V in eq. (5.2)]

$$\begin{aligned} u_{r'}^\dagger u_r - v_{r'}^\dagger v_r &= \delta_{r'r}, \\ u_{r'}^\dagger v_r - v_{r'}^\dagger u_r &= 0, \end{aligned} \qquad r', r = 1, 2, \dots, m. \tag{5.4}$$

From eqs. (4.4), (4.9) and (4.11) we see that the procedure to be described can be used for treating the general problem of section 4 in case of a real grand-dynamical matrix $\mathfrak{D}$‡ and according to eqs. (2.3) and (2.7) also in the special cases of the problem of section 2 in which $\mathfrak{D}$, though possibly complex, has the form (5.1). In the problem of section 4 the diagonalization procedure follows the scheme (4.10) with $\boldsymbol{a}$ and $\boldsymbol{c}$ defined by eqs. (4.5) and (4.8), respectively; in the problem of section 2 the scheme (2.6) with $\boldsymbol{a}$ and $\boldsymbol{c}$ defined by eqs. (2.4) and (2.5), respectively.

‡ Lieb et al. restricted their considerations to this case (for fermions). Like all the considerations in this paper also this section allows a self-evident translation into the real variant. Thus to a *real* $\mathfrak{D}$ of the form (5.1) one can find, along the lines to be described in this section, a *real* $\mathcal{T}^{-1}$ of the form (5.2) and consequently of the required form (4.9).

By comparing eqs. (5.3) and (2.11), we see from eq. (2.13) that the first m columns $(u_r\, v_r)'$ and the last m columns $(v_r\, u_r)'$ of $\mathcal{T}^{-1}$ should satisfy:

$$\begin{pmatrix} A & B \\ B & A \end{pmatrix}\begin{pmatrix} u_r \\ v_r \end{pmatrix} = \lambda_r \begin{pmatrix} u_r \\ -v_r \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} A & B \\ B & A \end{pmatrix}\begin{pmatrix} v_r \\ u_r \end{pmatrix} = -\lambda_r \begin{pmatrix} v_r \\ -u_r \end{pmatrix}, \tag{5.5a}$$

respectively. Each of the two equalities in eq. (5.5a) can be written more explicitly as

$$\begin{aligned} Au_r + Bv_r &= \lambda_r u_r, \\ Bu_r + Av_r &= -\lambda_r v_r. \end{aligned} \tag{5.5b}$$

Adding and subtracting the two equalities in eq. (5.5b) give the relations:

$$(A+B)\phi_r = \lambda_r \psi_r, \tag{5.6a}$$

$$(A-B)\psi_r = \lambda_r \phi_r, \qquad \phi_r \equiv u_r + v_r, \quad \psi_r \equiv u_r - v_r. \tag{5.6b}$$

Multiplying eqs. (5.6) on the left by $A - B$ and $A + B$, respectively, we obtain:

$$(A-B)(A+B)\phi_r = \lambda_r (A-B)\psi_r = \lambda_r^2 \phi_r, \tag{5.7a}$$

$$(A+B)(A-B)\psi_r = \lambda_r (A+B)\phi_r = \lambda_r^2 \psi_r. \tag{5.7b}$$

Now the matrices $(A-B)(A+B)$ and $(A+B)(A-B)$ in general are not hermitian (if so, it is possible and often preferable to apply the theorem in appendix A, see remark 2 at the end of appendix A; both matrices then equal $A^2 - B^2$, which matrix then is called the dynamical matrix of the corresponding hamiltonian). Therefore we proceed as follows. The hermiticity of the matrix $\mathcal{D}$ implies that the matrices A and B and consequently also $A+B$ and $A-B$ are hermitian. Furthermore, together with $\mathcal{D}$ also the matrices $A+B$ and $A-B$ are positive definite. This is easily recognized if one makes for an arbitrary column m-vector u the evaluation: $(u^\dagger \pm u^\dagger)\mathcal{D}(u \pm u)' = 2u^\dagger(A \pm B)u$, and realizes that the single element in this 1-square matrix is positive (see also remark 2 below). We now choose an m-square matrix K with $A - B = K^\dagger K$. Then the eigenvalue problem (5.7a) can be written as a unitary-diagonalization problem of a hermitian matrix:

$$\begin{aligned} &K^\dagger[K(A+B)K^\dagger - \lambda_r^2 I](K^\dagger)^{-1}\phi_r = 0 \\ &[K(A+B)K^\dagger - \lambda_r^2 I]\chi_r, \qquad \chi_r \equiv (K^\dagger)^{-1}\phi_r. \end{aligned} \tag{5.8}$$

Starting from an arbitrarily chosen orthogonal set of m eigenvectors χ_r of the (hermitian) matrix $K(A+B)K^\dagger$ which are normalized in such a way that $\chi_r^\dagger \chi_r = \lambda_r^{-1}$ [take the λ_r positive; together with $A+B$ also $K(A+B)K^\dagger$ is positive definite], we determine the corresponding ϕ_r from eq. (5.8) and ψ_r from eq. (5.6a) (always taking the λ_r positive); eq. (5.6b) then holds automatically. Now we can show that the u_r and v_r as derived from these ϕ_r and ψ_r according to eq. (5.6), for which obviously eqs. (5.5b) and, equivalently, (5.5a) hold‡, represent the columns of the matrices U and V in a para-unitary matrix $\mathcal{T}^{-1}$ of the form (5.2), which diagonalizes $\mathfrak{D}$ according to eq. (5.3). In fact, eq. (5.4) is satisfied because from eqs. (5.7a), (5.6b) and (5.8) and in view of our normalization of the χ_r we have:

$$\begin{aligned}\lambda_r^2\psi_{r'}^\dagger\phi_r &= \psi_{r'}^\dagger(A-B)(A+B)\phi_r = \lambda_{r'}\phi_{r'}^\dagger(A+B)\phi_r\\ &= \lambda_{r'}\chi_{r'}^\dagger K(A+B)K^\dagger\chi_r = \lambda_r^2\delta_{r'r},\end{aligned} \tag{5.9}$$

so that

$$\begin{aligned}\psi_{r'}^\dagger\phi_r &= \delta_{r'r} \quad \text{or rather} \quad (u_{r'}^\dagger - v_{r'}^\dagger)(u_r + v_r) = \delta_{r'r},\\ \phi_{r'}^\dagger\psi_r &= \delta_{r'r} \quad \text{or rather} \quad (u_{r'}^\dagger + v_{r'}^\dagger)(u_r - v_r) = \delta_{r'r}.\end{aligned} \tag{5.10}$$

The two equalities in eq. (5.4) follow from eq. (5.10) by adding and subtracting. Our problem has been solved.

Remark 1. Like in remark 1 of section 3, in computer programs it is recommended to carry out a Cholesky decomposition to find a matrix K with $A - B = K^\dagger K$. The procedure is very efficient and simultaneously the positive definiteness of $A - B$ can be checked without performing a unitary diagonalization, which would take a lot more computer time.

Remark 2. In case one succeeds in the Cholesky or another decomposition of $A - B$ (i.e. $A - B$ is positive definite) and if moreover all eigenvalues of the matrix $K(A+B)K^\dagger$ are positive (i.e. $A + B$ is positive definite), it is ensured that $\mathfrak{D}$ is positive definite also: no additional check is necessary to verify that the considerations of this section are applicable. To recognize the positive definiteness of $\mathfrak{D}$ we only have to consider for an arbitrary non-zero column $2m$-vector $w \equiv (u\ v)'$ the single element in the 1-square matrix

$$w^\dagger\mathfrak{D}\,w = \tfrac{1}{2}[(u^\dagger + v^\dagger)(A+B)(u+v) + (u^\dagger - v^\dagger)(A-B)(u-v)], \tag{5.11}$$

which evidently is positive since the column m-vectors $u + v$ and $u - v$ cannot vanish simultaneously.

6. Concluding remarks

6.1. *Connection between the treatments of the Bogoliubov and the general quadratic hamiltonian*

The form (4.6) of the Bogoliubov hamiltonian (2.1) contains a $4m$-square matrix (including the factor $\frac{1}{2}$), which has the form (4.4) required to call it the

‡ In the terminology of section 3, remark 3, the vectors $(u_r\, v_r)'$ and $(v_r\, u_r)'$ are paravectors of $\mathfrak{D}$ at the paravalues λ_r and $-\lambda_r$, respectively.

grand-dynamical matrix of a general quadratic hamiltonian. Thus one may in theory diagonalize a hamiltonian of the form (2.1) by the treatment of section 4 rather than by the treatment of sections 2 and 3. That eq. (4.6) never helps to solve the diagonalization problem in the case that the treatment of sections 2 and 3 does not apply, follows from the fact that the two grand-dynamical matrices – the $2m$-square matrix (2.3) and the $4m$-square matrix occurring in eq. (4.6) – are both positive definite as soon as one of them is.

6.2. *Algorithm to be used*

The question which of the algorithms presented in this paper is to be used when a homogeneous quadratic boson hamiltonian is actually to be diagonalized, is easily answered. First of all one tries to write the hamiltonian, possibly by adding a real constant, into the standard form (2.1) [the grand-dynamical matrix (2.3) is then hermitian]; if this is not possible, into the standard form (4.1) with a hermitian grand-dynamical matrix of the form (4.4). If the resulting grand-dynamical matrix $\mathcal{D}$ has the form (5.1), the considerations of section 5 can be used for the diagonalization. Otherwise one has to use an algorithm which involves a unitary diagonalization of a hermitian matrix of two times larger order: the algorithm described in section 3, remark 1 for the standard form (2.1) [scheme according to eq. (2.6)] and the algorithm described in section 4 for the standard form (4.1) [scheme (4.10)]. The algorithm of section 4 follows completely the one in section 3, remark 1, but only requires the calculation of the first m columns of $\mathcal{T}^{-1}$ whereas the last m columns must be obtained from the first according to eq. (4.9). A simplification can be achieved if one can use the theorem in appendix A (dynamical matrix, see subsection 6.3); therefore it is recommended always to check whether that theorem or a theorem of that kind is applicable. All considerations of this paper can obviously be translated into their real variants, which implies another simplification for hamiltonians with all coefficients real. Complications can arise in the case of soft modes (zero mode energies, positive-*semi*definite $\mathcal{D}$), see subsection 6.5.

6.3. *Dynamical matrices*

The approach in this paper differs widely from those used in almost all special cases treated up to now in the literature. In solid-state physics, e.g. one usually finds positive-definite *hermitian* matrices (called dynamical matrices, see also section 1), like the grand-dynamical matrices directly constructed from the coefficients in the hamiltonian, the *eigenvalues* of which yield the *squares* of the mode energies. In appendix A we discuss a connection between our results and those in the literature: it is made plausible that rather than being a general result the occurrence of dynamical matrices is specific for a very limited class of quadratic boson hamiltonians.

6.4. *The fermion counterpart*

All considerations in this paper, including those in appendix A, can in a self-evident way and without any difficulties be translated into the fermion case. The "para" language of this paper then passes into the familiar matrix-algebraic language: the diagonalization of a homogeneous quadratic fermion hamiltonian amounts to the *unitary* diagonalization of a grand-dynamical matrix $\mathcal{D}$ which is constructed from the coefficients in exactly the same way as in the boson case. Also in the fermion case $\mathcal{D}$ is hermitian; no additional properties like positive definiteness are required, however, for the diagonalizability of the hamiltonian. A decomposition like Cholesky's (cf. section 3, remark 1) is not involved. Since results of this kind for the fermion case are well known (see also section 1), we do not consider this case further. The discussion in appendix B on some more details of the "para" terminology of this paper gives further help to see the diagonalization problem in the boson and fermion case as counterpart problems.

6.5. *Soft modes (zero mode energies)*

The most unsatisfactory aspect of this paper is that we did not include in our treatment the case of a quadratic boson hamiltonian with a positive-*semi*definite grand-dynamical matrix $\mathcal{D}$, which often occurs in practice. In the last two examples in remark 4 at the end of section 3 we saw that not all positive-semidefinite matrices can be para-unitarily diagonalized. In accordance with this it is essential that the matrix $\mathcal{D}$ is positive definite in order that the algorithm in section 3, remark 1 for finding a diagonalizing para-unitary matrix $\mathcal{T}^{-1}$ can be applied. However, we want to point out that there exists a simple calculation in connection with a positive-semidefinite matrix which often is physically interesting. To be specific we turn our attention to solid-state physics where a positive-semidefinite grand-dynamical matrix $\mathcal{D}$ often occurs for certain isolated values $\boldsymbol{k}_0$ of the wave vector $\boldsymbol{k}$ (often $\boldsymbol{k}_0 = \boldsymbol{0}$). If one does not bother about the possible non-existence of a diagonalizing para-unitary matrix $\mathcal{T}^{-1}$ whose elements may tend to infinity, the algorithm in section 3, remark 1 can be used to calculate from the positive-semidefinite $\mathcal{D}$ the (limit of the) mode energies for $\boldsymbol{k}$ tending to such a $\boldsymbol{k}_0$. To see this we recognize – e.g. by considering a positive-semidefinite matrix as a limiting case of positive-definite matrices – from section 3, remark 1 that the unitary diagonalization of the hermitian matrix $\mathcal{K}\hat{\mathcal{I}}\mathcal{K}^{\dagger}$ gives the desired mode energies. Rather than trying to generalize the procedure of the Cholesky decomposition to the case of positive-*semi*definite matrices $\mathcal{D}$, one can apply for this single case an additional unitary diagonalization to find a decomposition $\mathcal{D}=\mathcal{K}^{\dagger}\mathcal{K}$: if $\mathcal{S}$ is a unitary matrix with $\mathcal{S}^{\dagger}\mathcal{D}\mathcal{S}=\mathcal{F}$ diagonal, one easily verifies that we can take $\mathcal{K}=\mathcal{K}^{\dagger}=\mathcal{S}\mathcal{F}^{1/2}\mathcal{S}^{\dagger}$, where the meaning of the positive-semidefinite diagonal matrix $\mathcal{F}^{1/2}$ is self-evident. Analogous arguments can be used in section 5 to find the λ_r (not the $\mathcal{T}^{-1}$) also in case of positive-*semi*definite matrix (5.1).

6.6. *Comparison with other approaches*

In section 1 we already indicated in what aspects our treatment differs from those in the literature. In most publications we referred to, special (solid-state) problems were solved. It may be interesting to compare our approach with a purely theoretical treatment. As such the book of Berezin[10]) seems to be most suitable: theorem 8.1 of ref. 10, to which we referred already in section 1 and in the last footnote of section 4, is concerned with almost the same problem as section 4 of this paper. However, the contents of both approaches are by no means identical.

On the one hand Berezin includes the case of an infinite number of construction operators. This means a much more difficult problem to solve and simultaneously makes the mathematics less accessible. On the other hand Berezin treats only the case of real coefficients and does not diagonalize completely: in his final expressions terms with $\gamma_{r'}^{\dagger}\gamma_r$ occur where $r'\neq r$ so that an additional (unitary, see below) transformation is necessary for a complete diagonalization. Also no explicit algorithm is given in ref. 10. Expressed in the notation of section 4 with real matrices A and B the treatment of ref. 10 just results in the real matrix:

$$\mathcal{T}^{-1}=\tfrac{1}{2}\begin{pmatrix} R^{1/2}+R^{-1/2} & R^{1/2}-R^{-1/2} \\ R^{1/2}-R^{-1/2} & R^{1/2}+R^{-1/2}\end{pmatrix}, \qquad \begin{array}{l} R\equiv P^{-1/2}(P^{1/2}QP^{1/2})^{1/2}P^{-1/2}, \\ P\equiv A+B,\quad Q\equiv A-B. \end{array} \tag{6.1}$$

This matrix has the required partitioned form (4.9) and one indeed verifies directly (see below) that this $\mathcal{T}^{-1}$ transforms the matrix $\mathcal{D}$ [eq. (4.4)] para-unitarily into a $2m$-square matrix $(\mathcal{T}^{\dagger})^{-1}\mathcal{D}\mathcal{T}^{-1}$ with the two off-diagonal m-square blocks zero. Hereafter the diagonalization of the diagonal blocks can easily be performed by an additional unitary diagonalization, that is by a para-unitary and at the same time unitary matrix which has the form of the right-hand side of eq. (3.6) in which T is an m-square orthogonal (i.e. real and unitary) matrix and $c=1$, $s=0$. The "square roots" in eq. (6.1) are defined according to the example of P: $P^{1/2}$ is the real positive-definite symmetric matrix for which $P^{1/2}P^{1/2}=P$, and $P^{-1/2}\equiv(P^{1/2})^{-1}$ (see section 5 for the proof that $\mathcal{D}$ is positive definite if and only if the same holds for $P \equiv A+B$ and $Q \equiv A-B$). Indeed, the matrix (6.1) is para-unitary since it satisfies the first eq. (2.10b). Moreover the off-diagonal blocks mentioned vanish according to the simple reduction

$$\tfrac{1}{2}[R^{1/2}PR^{1/2} - R^{-1/2}QR^{-1/2}] = \tfrac{1}{2}R^{-1/2}[RPR - Q]R^{-1/2} = 0. \tag{6.2}$$

Since in general one needs to carry out a unitary diagonalization for each "square root" occurring in eq. (6.1) (see the end of subsection 6.5), one recognizes from eq. (6.1) that Berezin did not intend to give a practical algorithm for the calculation of a para-unitary matrix with the properties mentioned, but just wanted to present a proof for its *existence*: our treatment (cf. section 5 which is applicable in the problem as considered by Berezin) requires only a single unitary diagonalization of a hermitian matrix of the same order as the matrices which are to be unitarily diagonalized in the procedure suggested by ref. 10.

Appendix A

The dynamical matrix

In section 1 and subsection 6.3 we announced a discussion on the frequently encountered dynamical matrices. By a dynamical matrix we mean a *hermitian* matrix D, the *eigenvalues* of which give the *squares* of the mode energies. According to this paper these mode energies can also be obtained from the paravalues of a grand-dynamical matrix $\mathcal{D}$; in general the order of D is half of the order of $\mathcal{D}$. In this sense the matrices $(A-B)(A+B)$ and $(A+B)(A-B)$ occurring in eq. (5.7) are obviously only dynamical matrices if they are hermitian (see also remark 2 at the end of this appendix). In this appendix we prove an almost trivial theorem which we believe to be a typical example of the kind of theorems which lie behind the existence of dynamical matrices. Application of the theorem yields a matrix D and allows to find the coefficients in a corresponding diagonalizing Bogoliubov transformation [i.e. the elements of the matrix $\mathcal{T}^{-1}$ occurring in eqs. (2.7) and (4.11)] from the eigenvectors of D. Since a matrix decomposition like Cholesky's (cf. section 3, remark 1 and section 5, remark 1) is not involved, the relations between the elements of $\mathcal{T}^{-1}$ and the coefficients in the original hamiltonian are more direct than if for example the $\mathcal{T}^{-1}$ is found with the help of the considerations in section 5; this may have advantages in actual calculations. Although the requirements which the theorem imposes on the elements of $\mathcal{D}$ are rather strong, a check on a great many investigations in the solid-state literature shows that, though not in all cases where a dynamical matrix exists, in the vast majority of such cases one could have derived the dynamical-matrix results also by application of just this theorem. This makes the theorem an almost systematic key for finding dynamical matrices D (in case of their existence) but simultaneously makes it plausible that the existence of the dynamical matrix in many theories is indeed only due to the peculiar relations which hold between the coefficients in the hamiltonian and that the dynamical matrix in general is not naturally connected with the boson diagonalization problem as the grand-dynamical matrix is.

Theorem. Let A and B be commuting (complex) m-square matrices; A is hermitian, B is normal (i.e. $BB^\dagger = B^\dagger B$, e.g. B is hermitian). Let furthermore the m-square hermitian matrices A and $D \equiv A^2 - B^\dagger B$ be positive definite. Then the $2m$-square hermitian matrix

$$\mathscr{D} \equiv \begin{pmatrix} A & B \\ B^\dagger & A \end{pmatrix} \tag{A.1}$$

is positive definite also and its paravalues and paravectors (and also the diagonalizing para-unitary matrix) can be constructed from the eigenvalues and those sets of m orthogonal eigenvectors of D where these eigenvectors are simultaneous eigenvectors of the matrices A and B. Each p-fold (positive) eigenvalue of D is the square of p positive and p negative paravalues of the matrix $\mathscr{D}$.

Proof. On the conditions of the theorem the (hermitian) matrices in the left-hand side of the equation

$$\begin{pmatrix} A & B \\ B^\dagger & A \end{pmatrix}\begin{pmatrix} A & -B \\ -B^\dagger & A \end{pmatrix} = \begin{pmatrix} A^2 - B^\dagger B & 0 \\ 0 & A^2 - B^\dagger B \end{pmatrix} \tag{A.2}$$

commute. Consequently there exists an orthonormal set of $2m$ simultaneous eigenvectors of these two matrices. Let $(u\, v)'$ be an eigenvector of this set with eigenvalues x_1 and x_2, respectively. We have

$$\begin{aligned} Au + Bv &= x_1 u, & Au - Bv &= x_2 u, \\ B^\dagger u + Av &= x_1 v, & -B^\dagger u + Av &= x_2 v, \end{aligned} \tag{A.3}$$

and by adding the upper and the lower two equations we obtain:

$$2Au = (x_1 + x_2)u \quad \text{and} \quad 2Av = (x_1 + x_2)v, \tag{A.4}$$

respectively. Since A is positive definite and since the vectors u and v do not simultaneously vanish, $x_1 + x_2$ is positive. On the other hand $(u\, v)'$ is also an eigenvector of the matrix in the right-hand side of eq. (A.2) at the product $x_1 x_2$ as eigenvalue which is positive since $A^2 - B^\dagger B$ is positive definite. So both x_1

and x_2 are positive and we see that all eigenvalues x_1 of $\mathcal{D}$ [$\mathcal{D}$ is the first matrix in the left-hand side of eq. (A.2)] are positive: $\mathcal{D}$ is positive definite.

Since B is normal, B is unitarily diagonalizable. Because A and B commute and A is hermitian one can even find a unitary matrix which diagonalizes A and B simultaneously. Since B is normal such a unitary matrix also diagonalizes the matrix $B^\dagger$ and consequently the matrix $D \equiv A^2 - B^\dagger B$. In what follows we start from the columns ξ_r $(r = 1, 2, \dots, m)$ of such a unitary matrix or from another orthogonal set of m simultaneous eigenvectors of A and B (and automatically also of $B^\dagger$ and D). The corresponding eigenvalues of D are denoted by λ_r^2 (in what follows we take always $\lambda_r > 0$).

We presently prove that the $2m$ $2m$-vectors constructed from the m-vectors ξ_r according to:

$$\mathscr{W}_r \equiv C_r \begin{pmatrix} (A + \lambda_r I)\xi_r \\ -B^\dagger \xi_r \end{pmatrix} \equiv \begin{pmatrix} u_r \\ v_r \end{pmatrix}, \qquad \mathscr{W}_{m+r} \equiv C_{m+r} \begin{pmatrix} -B\xi_r \\ (A + \lambda_r I)\xi_r \end{pmatrix} \equiv \begin{pmatrix} u_{m+r} \\ v_{m+r} \end{pmatrix} \quad \text{(A.5)}$$

constitute a set of paraperpendicular paravectors of $\mathcal{D}$ at paravalues λ_r and $-\lambda_r$, respectively [in eq. (A.5) the scalars C_r, C_{m+r} are determined such that the vectors are ("para")normalized: $u_r^\dagger u_r - v_r^\dagger v_r = 1$ and $u_{m+r}^\dagger u_{m+r} - v_{m+r}^\dagger v_{m+r} = -1$]. According to section 3, remark 3, these paravectors consequently can serve as the columns of a para-unitary matrix

$$\mathscr{T}^{-1} \equiv (\mathscr{W}_1 \mathscr{W}_2 \dots \mathscr{W}_m \mathscr{W}_{m+1} \mathscr{W}_{m+2} \dots \mathscr{W}_{2m}), \quad \text{(A.6)}$$

which diagonalizes $\mathcal{D}$, i.e.

$$(\mathscr{T}^\dagger)^{-1} \mathcal{D} \mathscr{T}^{-1} = \mathscr{E} \equiv \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_m, \lambda_1, \lambda_2, \dots, \lambda_m) \quad \text{(A.7)}$$

[cf. eqs. (2.7) and (4.11)].

It is easily checked that the vectors (A.5) are paravectors of $\mathcal{D}$ at the paravalues mentioned, indeed: we have

$$\begin{pmatrix} A & B \\ B^\dagger & A \end{pmatrix} \begin{pmatrix} (A + \lambda_r I)\xi_r \\ -B^\dagger \xi_r \end{pmatrix} = \lambda_r \begin{pmatrix} (A + \lambda_r I)\xi_r \\ B^\dagger \xi_r \end{pmatrix}$$

and

$$\begin{pmatrix} A & B \\ B^\dagger & A \end{pmatrix} \begin{pmatrix} -B\xi_r \\ (A + \lambda_r I)\xi_r \end{pmatrix} = -\lambda_r \begin{pmatrix} -B\xi_r \\ -(A + \lambda_r I)\xi_r \end{pmatrix}. \quad \text{(A.8)}$$

The first m vectors (A.5) and also the last m are paraperpendicular: $u_{r'}^\dagger u_r - v_{r'}^\dagger v_r = 0$ and $u_{m+r'}^\dagger u_{m+r} - v_{m+r'}^\dagger v_{m+r} = 0$ $(r' \neq r)$, because u_r, v_r, u_{m+r} and v_{m+r} all equal ξ_r multiplied by a scalar. Moreover the first m vectors (A.5) are paraperpendicular to the last m since they correspond to different (because of the different sign) paravalues [see the discussion around eq. (3.11)]. We conclude that the $2m$ vectors (A.5) are paraperpendicular. With this we have proved the theorem completely.

Remark 1. We constructed the $2m$ paravectors (A.5) of $\mathcal{D}$ from a particular orthogonal set of m eigenvectors of $D \equiv A^2 - B^\dagger B$, viz. from an arbitrary set of simultaneous eigenvectors of the matrices A and B. That such a particular choice is necessary can be seen from the following counterexample ($m = 2$):

$$A \equiv \begin{pmatrix} 4 & 1 \\ 1 & 4 \end{pmatrix}, \quad B \equiv \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}, \quad D \equiv A^2 - B^2 = \begin{pmatrix} 9 & 0 \\ 0 & 9 \end{pmatrix}. \tag{A.9}$$

We have

$$A\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 5\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad A\begin{pmatrix} 1 \\ -1 \end{pmatrix} = 3\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad B\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 4\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad B\begin{pmatrix} 1 \\ -1 \end{pmatrix} = 0\begin{pmatrix} 1 \\ -1 \end{pmatrix}. \tag{A.10}$$

According to our construction we should start from eigenvectors ξ_1 and ξ_2 proportional to $(1\ 1)'$ and $(1\ -1)'$, respectively, and then obtain after a simple calculation which follows eqs. (A.5) and (A.6) the para-unitary matrix

$$\mathcal{T}^{-1} \equiv \frac{1}{\sqrt{6}} \begin{pmatrix} 2 & \sqrt{3} & -1 & 0 \\ 2 & -\sqrt{3} & -1 & 0 \\ -1 & 0 & 2 & \sqrt{3} \\ -1 & 0 & 2 & -\sqrt{3} \end{pmatrix}, \tag{A.11}$$

which diagonalizes $\mathcal{D}$ according to eq. (A.7) with $\mathscr{E} \equiv \mathrm{diag}(3, 3, 3, 3)$. However, if we would take the set of orthogonal eigenvectors of $A^2 - B^2$, where $\xi_1 \equiv (1\ 0)'$ and $\xi_2 \equiv (0, 1)'$, we would get according to eq. (A.5) ($\lambda_1 = \lambda_2 = \sqrt{9} = 3$) the column vectors (not normalized):

$$w_1 \equiv \begin{pmatrix} 7 \\ 1 \\ -2 \\ -2 \end{pmatrix}, \quad w_2 \equiv \begin{pmatrix} 1 \\ 7 \\ -2 \\ -2 \end{pmatrix}, \quad w_3 \equiv \begin{pmatrix} -2 \\ -2 \\ 7 \\ 1 \end{pmatrix}, \quad w_4 \equiv \begin{pmatrix} -2 \\ -2 \\ 1 \\ 7 \end{pmatrix}. \tag{A.12}$$

Obviously these vectors do not form a paraperependicular system.

Remark 2. It is clear that the theorem is applicable to the problem of section 5 if there the matrices $(A - B)(A + B)$ and $(A + B)(A - B)$ are hermitian. The $\mathcal{D}$ in this appendix has then the form (5.1) so that the matrices A and B are hermitian. $\mathcal{D}$ is positive definite, so also A. The hermiticity of (one of) the products $(A - B)(A + B)$ and $(A + B)(A - B)$ implies that the matrices A and B commute and that both products equal $D \equiv A^2 - B^2$. From eqs. (5.7) we see that the matrix D is positive definite.

Appendix B

On the "para" terminology of this paper

In several places in this paper we introduced "para" concepts (cf. section 2 and section 3, remark 3), the properties of which strongly suggest that the diagonalization problem of homogeneous quadratic hamiltonians in the boson and fermion case is governed by counterpart matrix-algebraic formalisms (cf. also subsection 6.4). The fermion case need not be considered any further since the connected concepts are well known. Therefore we only give a few details concerning the boson case: some points of view which were not used in this paper, may contribute to a better understanding of the "para" concepts used.

The central "para" concept is the para unit matrix $\hat{\mathcal{I}}$ defined by:

$$\hat{\mathcal{I}} \equiv \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \equiv \operatorname{diag}(1, 1, \ldots, 1, -1, -1, \ldots, -1), \quad \text{I unit matrix,} \tag{B.1}$$

where here and in what follows it is understood that the order of script-letter matrices and vectors is $2m$ and the order of Latin-letter matrices and vectors m. Next we define the para-inverse $\bar{\mathcal{T}}$ of an arbitrary non-singular matrix $\mathcal{T}$ by

$$\mathcal{T} \equiv \begin{pmatrix} P & Q \\ R & S \end{pmatrix}, \qquad \mathcal{T}^{-1} = \begin{pmatrix} \tilde{P} & \tilde{R} \\ \tilde{Q} & \tilde{S} \end{pmatrix}, \qquad \bar{\mathcal{T}} \equiv \begin{pmatrix} \tilde{P} & -\tilde{R} \\ -\tilde{Q} & \tilde{S} \end{pmatrix} \tag{B.2a}$$

(the second equation defines the m-square matrices $\tilde{P}$, $\tilde{Q}$, $\tilde{R}$ and $\tilde{S}$), or, equivalently, by one of the following equivalent three definitions:

$$\bar{\mathcal{T}}\hat{\mathcal{I}}\mathcal{T} = \hat{\mathcal{I}} \quad \text{or} \quad \mathcal{T}\hat{\mathcal{I}}\bar{\mathcal{T}} = \hat{\mathcal{I}} \quad \text{or} \quad \bar{\mathcal{T}} \equiv \hat{\mathcal{I}}\mathcal{T}^{-1}\hat{\mathcal{I}}. \tag{B.2b}$$

From the first two of these equations we see that each non-singular matrix is the para-inverse of its para-inverse: $\mathcal{T} = \bar{\bar{\mathcal{T}}}$. Furthermore, since $(\mathcal{T}^{-1})^{-1} = \mathcal{T}$, the last definition equation (B.2b) implies that

$$\overline{\mathcal{T}^{-1}} = \hat{\mathcal{I}}\mathcal{T}\hat{\mathcal{I}} = \hat{\mathcal{I}}^{-1}\mathcal{T}\hat{\mathcal{I}}^{-1} = (\hat{\mathcal{I}}\mathcal{T}^{-1}\hat{\mathcal{I}})^{-1} = \bar{\mathcal{T}}^{-1}, \tag{B.3}$$

i.e. the para-inverse of the inverse of a (non-singular) matrix is the inverse of its para-inverse.

Now we can place some considerations of this paper in a new light. From eqs. (2.10b) and (B.2b) we see that a $2m$-square matrix $\mathcal{T}$ is para-unitary if and only if

$$\mathcal{T}^{\dagger} = \bar{\mathcal{T}}. \tag{B.4}$$

[Using that one can express the para-inverse of a product of matrices in a reversed-order product of para-inverses of the individual matrices (this is seen from the reduction $\bar{\mathcal{T}}\bar{\mathcal{S}}\hat{\mathcal{I}}\mathcal{S}\mathcal{T} = \bar{\mathcal{T}}\hat{\mathcal{I}}\mathcal{T} = \hat{\mathcal{I}}$), one easily verifies from eq. (B.4) that the product of a number of para-unitary matrices is again para-unitary.] Consequently eqs. (2.7) and (4.11) can be written in the form [cf. also eq. (2.11)]:

$$\begin{aligned}\bar{\mathcal{S}}\mathcal{D}\mathcal{S} &\equiv \bar{\mathcal{S}}\begin{pmatrix} A & B \\ C & D\end{pmatrix}\mathcal{S} = \mathcal{E} \\ &\equiv \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_m, -\lambda_{m+1}, -\lambda_{m+2}, \ldots, -\lambda_{2m}) \equiv \hat{\mathcal{I}}\mathcal{L},\end{aligned} \tag{B.5}$$

where we replaced $\mathcal{T}^{-1}$ by $\mathcal{S}$ and where $\mathcal{D}$ also is given in partitioned form. We recognize that this equation has not the form of a similarity transformation (i.e. the form $S^{-1}PS = Q$; as is well known such a transformation does not affect the *eigen*values) we are accustomed to. Instead, it has the form of a parasimilarity transformation (by which we mean the form $\bar{\mathcal{S}}\mathcal{P}\mathcal{S} = \mathcal{Q}$), which type of transformation does not affect the *para*values as we will show now. In agreement with section 3, remark 3 we call a non-zero column vector $w \equiv (u\ v)'$ a paravector of an arbitrary matrix $\mathcal{D}$ (in what follows we do not require the hermiticity of $\mathcal{D}$) at the paravalue λ if

$$\mathcal{D}w = \lambda\hat{\mathcal{I}} \quad \text{or} \quad (\mathcal{D} - \lambda\hat{\mathcal{I}})w = 0 \quad \text{or} \quad \begin{pmatrix} A & B \\ C & D\end{pmatrix}\begin{pmatrix} u \\ v\end{pmatrix} = \lambda\begin{pmatrix} u \\ -v\end{pmatrix} \tag{B.6}$$

[cf. eq. (2.13)]. From eq. (B.6) we see immediately that the paravalues λ of a matrix $\mathcal{D}$ are given by the roots of the "parasecular equation":

$$\det(\mathcal{D} - \lambda\hat{\mathcal{I}}) = 0 \quad \text{or} \quad \begin{vmatrix} A - \lambda I & B \\ C & D + \lambda I\end{vmatrix} = 0 \tag{B.7}$$

[cf. eq. (2.14)]. By considering the determinant of the matrix $\bar{\mathcal{S}}\mathcal{D}\mathcal{S} - \lambda\hat{\mathcal{I}} = \bar{\mathcal{S}}(\mathcal{D} - \lambda\hat{\mathcal{I}})\mathcal{S}$ [see the first eq. (B.2b); we do not require here the para-unitary of $\mathcal{S}$: $\mathcal{S}$ should only be non-singular in order that it has a para-inverse], we see that a parasimilarity transformation does not affect the paravalues, indeed.

Acknowledgements

We are very much indebted to Dr. A.H. 't Hooft, Dr. P.E. Brommer and Dr. H.G.J. Pijls for several valuable discussions and suggestions which appreciably contributed to this paper. We are also grateful to the referee for useful remarks on a previous version of this paper. We thank Dr. Brommer for critically reading the manuscript.

Note added in proof

After submission of this paper to Physica another paper on the same subject, by C. Tsallis, appeared in the Journal of Mathematical Physics **19**

(1978) 277. It may be interesting to compare the two treatments which differ in many respects.

References

1) N. Bogoliubov, J. Phys. (USSR) **11** (1947) 23, also reprinted in D. Pines, The Many-Body Problem (W.A. Benjamin, Inc., New York, 1961) p. 292.
2) T. Holstein and H. Primakoff, Phys. Rev. **58** (1940) 1098.
3) E. Lieb, T. Schultz and D. Mattis, Ann. Phys. **16** (1961) 407, appendix A.
4) C.F. Coll, III and A.B. Harris, Phys. Rev. **B4** (1971) 2781.
5) P.A. Lindgård, A. Kowalska and P. Laut, J. Phys. & Chem. Solids **28** (1967) 1357.
6) A.A. Maradudin, E.W. Montroll, G.H. Weiss and I.P. Ipatova, Theory of Lattice Dynamics in the Harmonic Approximation, 2nd edn. (Academic Press, New York, London, 1971). Solid State Physics, H. Ehrenreich, F. Seitz and D. Turnbull, eds., Suppl. 3, section II.3.
7) R.D. Mattuck, A guide to Feynman Diagrams in the Many-Body Problem (McGraw-Hill Publ. Co. Ltd., London etc., 1967) section 16.2.
8) C.F. Coll, III and A.B. Harris, Phys. Rev. **B2** (1970) 1176.
9) R.J. Elliott, J.A. Krumhansl and P.L. Leath, Rev. Mod. Phys. **46** (1974) 465, section II.A.2.
10) F.A. Berezin, The Method of Second Quantization (Academic Press, New York, London, 1966) theorem 8.1.
11) J.H. Wilkinson, The Algebraic Eigenvalue Problem (Clarendon Press, Oxford, 1965) p. 229.