A matrix diagonalisation problem in quantum mechanics

1987 J. Phys. A: Math. Gen. 20 6281

(http://iopscience.iop.org/0305-4470/20/18/028)

J. Phys. A: Math. Gen. 20 (1987) 6281-6291. Printed in the UK

A matrix diagonalisation problem in quantum mechanics

Joseph G Conlon

Department of Mathematics, University of Missouri, Columbia, Missouri 65211, USA

Received 12 March 1987, in final form 7 August 1987

Abstract. In this paper we are concerned with discussing a matrix diagonalisation problem which occurs when one attempts to estimate the ground-state energy of a Bose fluid. The matrix problem comes about when one makes Bogoliubov's approximation to the Hamiltonian of an interacting system. The Hamiltonian is then quadratic in creation and annihilation operators. The ground-state energy can be computed if one solves an appropriate matrix diagonalisation problem. Here the roots of the characteristic polynomial for the matrix problem turn out to be the turning points of the roots of some simpler polynomials which depend on a parameter. Using this fact one can estimate them.

1. Introduction

In this paper we discuss a matrix diagonalisation problem which is related to the problem of estimating the ground-state energy of a Bose fluid. In 1947 Bogoliubov, in an attempt to understand the phenomenon of superfluidity, introduced a method for estimating the ground-state energy of a Bose fluid. Bogoliubov's basic ansatz is that 'most particles are in a zero-momentum state'. If the Hamiltonian H of the system is written in second quantised form then the kinetic energy operator is quadratic in creation and annihilation operators whereas the potential energy operator is quartic. From Bogoliubov's ansatz one can argue that the potential energy operator may be well approximated by an expression which is quadratic in creation and annihilation operators. This yields a new Hamiltonian H_{B} which is quadratic in creation and annihilation operators. Furthermore this Hamiltonian can be written as a sum

$$H_{\mathrm{B}} = \sum_k A_k \tag{1.1}$$

where the k vary in momentum space and the A_k commute with each other for different values of k. To calculate the ground-state energy E_{B} of H_{B} one can then calculate the ground-state energy of each A_k and sum. The problem of calculating the ground-state energy of A_k is just a 2×2 matrix diagonalisation problem which can easily be solved. Hence one obtains a formula for E_{B} and also for the ground-state wavefunction ψ_{B}. The energy E_{B} may bear little relationship to the ground-state energy E of H. To determine whether it does, one computes the number of particles in the state ψ_{B} with non-zero momentum. If the ratio of the number of these particles to the total number of particles is small then one claims that Bogoliubov's ansatz is 'consistent' and therefore E_{B} is a good approximation to E.

There have been many calculations of the ground-state energy of Bose systems based on Bogoliubov's method (see Lieb (1965) for a review). In a calculation with a Bogoliubov Hamiltonian, Foldy (1961) observed that the energy per particle of a high-density Coulomb fluid should be proportional to the fourth root of the density. This suggests that the ground-state energy E_N of a neutral system of N charged bosons with Coulomb interaction should be of order $-N^{7/5}$ for large N. By constructing a Bogoliubov-type wavefunction Dyson (1967) proved that $E_N \leqslant -C_1 N^{7/5}$, for some constant $C_1 > 0$. Recently (Conlon 1985, 1987, Conlon *et al* 1987) it has been shown that $E_N \geqslant -C_2 N^{7/5}$, for a constant $C_2 > C_1$. This result then proves that in a particular case Bogoliubov's calculation yields the correct ground-state energy at least to order of magnitude.

Here we are concerned with the problem of calculating the ground-state energy of the operators A_k of (1.1) and its generalisations. Let us consider annihilation operators a_i, $1 \leqslant i \leqslant n$, on Fock space with adjoints a_i^*, $1 \leqslant i \leqslant n$. We require the a_i to satisfy canonical commutation relations

$$[a_i, a_j] = 0 \qquad [a_i, a_j^*] = \delta_{ij} \tag{1.2}$$

where [ ] is the commutator and δ_{ij} is the Dirac δ function $\delta_{ij} = 0$ for $i \neq j$, $\delta_{ii} = 1$. Let us consider an operator

$$A = 2 \sum_{i=1}^{n} \varepsilon_i a_i^* a_i + \left(\sum_{i=1}^{n} \lambda_i a_i + \mu_i a_i^* \right)^* \left(\sum_{i=1}^{n} \lambda_i a_i + \mu_i a_i^* \right). \tag{1.3}$$

The numbers λ_i, μ_i, $1 \leqslant i \leqslant n$, are arbitrary complex numbers but the ε_i, $1 \leqslant i \leqslant n$, are all assumed to be positive. We shall be concerned with putting A in the diagonal form

$$A = 2 \sum_{i=1}^{n} \alpha_i \eta_i^* \eta_i + \beta \tag{1.4}$$

where the η_i, $1 \leqslant i \leqslant n$, also satisfy canonical commutation relations. The α_i, $1 \leqslant i \leqslant n$, are positive numbers and the α_i, β are to be determined in terms of the λ_i, μ_i, ε_i. Clearly the ground-state energy of A is β.

We relate the problem of diagonalising the operator A in (1.3) to the problem of estimating the ground-state energy of the Bose fluid. For the Bogoliubov Hamiltonian (1.1) A_k has the form (1.3) with $n = 2$, $\varepsilon_1 = \varepsilon_2$, $\lambda_1 = \mu_2$ real, $\lambda_2 = \mu_1 = 0$. In Conlon (1985, 1987) and Conlon *et al* (1987) where we show that Bogoliubov's method is in some sense valid we approximate the Hamiltonian H of the system by a modified Bogoliubov Hamiltonian $H_{\mathrm{B,M}}$. This Hamiltonian is, like H_{B}, quadratic in creation and annihilation operators, and may be written as in (1.1):

$$H_{\mathrm{B,M}} = \sum_k A_{k,\mathrm{M}}. \tag{1.5}$$

The operators $A_{k,\mathrm{M}}$ are of the form (1.3) but now n can be an arbitrary even integer. This corresponds to the fact that $H_{\mathrm{B,M}}$ is derived from H by making the modified Bogoliubov ansatz 'most particles are in low momentum states', the number of such states being $n/2$. This generalisation of Bogoliubov's ansatz appears to be necessary if one wishes to understand how Bogoliubov's ideas yield a good approximation to the ground-state energy of the real Hamiltonian.

The modified Bogoliubov ansatz then leads in a natural way to the diagonalisation problem (1.3) and (1.4). There is a particular case of (1.3) and (1.4) in which the diagonalisation problem has a simple form. It occurs when the λ_i, μ_i, $1 \leqslant i \leqslant n$ are real and $\lambda_i = \mu_i$, $1 \leqslant i \leqslant n$. In that case the a_i of (1.4) are the positive roots of the polynomial equation

$$1 + \sum_{i=1}^{n} \lambda_i^2 \left(\frac{1}{\varepsilon_i - \alpha} + \frac{1}{\varepsilon_i + \alpha} \right) = 0 \tag{1.6}$$

and β is given by the formula

$$\beta=\sum_{i=1}^{n}(\alpha_i-\varepsilon_i). \tag{1.7}$$

This problem has been studied in detail by Conlon (1985). In particular let us suppose the ε_i, $1\leqslant i\leqslant n$, are all distinct and can be labelled as

$$0<\varepsilon_1<\varepsilon_2<\varepsilon_3<\cdots<\varepsilon_n. \tag{1.8}$$

Then it is straightforward to see from (1.6) that the positive roots α_i, $1\leqslant i\leqslant n$ can be labelled in such a way that

$$\varepsilon_i<\alpha_i<\varepsilon_{i+1} \qquad 1\leqslant i\leqslant n \qquad (\varepsilon_{n+1}=\infty). \tag{1.9}$$

In this paper we are concerned with studying the diagonalisation problem (1.3) and (1.4) without restrictions on λ_i, μ_i, $1\leqslant i\leqslant n$. In this situation the polynomial equation determining the α_i, $1\leqslant i\leqslant n$, is considerably more complicated than (1.6). However it turns out that the α_i are the turning points of roots of equations like (1.6) which depend on a parameter, the parameter varying on the unit sphere. We have no adequate explanation of why this strange phenomenon occurs. By applying this fact we are able to obtain estimates on the α_i, β which are analogous to those which can be obtained from the simple polynomial equation (1.6). In particular we have the following theorem.

Theorem 1.1. The number β in (1.4) satisfies the inequality

$$0\leqslant\beta\leqslant\sum_{i=1}^{n}|\mu_i|^2 \tag{1.10}$$

and is given in terms of the α_i, $1\leqslant i\leqslant n$, by

$$\beta=\sum_{i=1}^{n}(\alpha_i-\varepsilon_i+\tfrac{1}{2}|\mu_i|^2-\tfrac{1}{2}|\lambda_i|^2). \tag{1.11}$$

Suppose the ε_i, $1\leqslant i\leqslant n$, are all distinct and satisfy (1.8). Then in any interval $(\varepsilon_i, \varepsilon_{i+r})$ the number of roots α_j is at least $r-1$ and at most $r+1$.

The inequality (1.10) follows easily from the representations (1.3) and (1.4). In fact from (1.4) it follows that β is the ground-state energy of A. From (1.3) we see that A is a positive operator and hence $\beta\geqslant 0$. On the other hand the vacuum expectation of A is the right-hand side of (1.10) and hence we obtain the second inequality of (1.10).

In the following our main concern will be to prove the result of theorem 1.1 on the α_j, $1\leqslant j\leqslant n$. We shall also show how (1.10) follows from (1.11) and some appropriate estimates on the α_i obtained from the polynomial equation determining the α_i, without reference to the formulation (1.3) and (1.4).

One should note here an invariance of the operator A in (1.3) up to a constant under the operation K defined by

$$Ka_i=a_i^* \qquad Ka_i^*=a_i \qquad 1\leqslant i\leqslant n. \tag{1.12}$$

From (1.12) it follows that $K^2=1$. This invariance is then an invariance of Kramers type (Kramers 1930, 1956, Messiah 1958). We shall see that it gives rise to the fact that the roots of the polynomial determining the α_i come in pairs α, $-\alpha$.

2. The characteristic polynomial

We now proceed to construct the canonical form (1.4) for A. Let $\boldsymbol{a}$ be the vector

$$\boldsymbol{a} = (a_1, \ldots, a_n) \tag{2.1}$$

where the a_i satisfy canonical commutation relations (1.2), and let $\boldsymbol{a}^*$ be the corresponding vector with a_i replaced by a_i^*. We wish to make transformations

$$\boldsymbol{a} = V\boldsymbol{\eta} + W\boldsymbol{\eta}^* \qquad \text{on vectors} \tag{2.2}$$

$$\boldsymbol{\eta} = (\eta_1, \ldots, \eta_n) \tag{2.3}$$

in such a way that the η_i, $1 \leq i \leq n$, also satisfy canonical commutation relations. For this to be the case the $n \times n$ matrices V, W, must satisfy the identities

$$VW' - WV' = 0 \qquad VV^* - WW^* = I \tag{2.4}$$

where I is the identity $n \times n$ matrix. Here prime denotes adjoint matrix and star Hermitian conjugate matrix. The identities (2.4) can be summarised by defining a matrix M by

$$M = \begin{pmatrix} V & W \\ \bar{W} & \bar{V} \end{pmatrix} \tag{2.5}$$

where the bar denotes taking the complex conjugate of the matrix elements. Then (2.4) is equivalent to the identity

$$MJM^* = J \tag{2.6}$$

with J the $2n \times 2n$ matrix given by

$$J = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}. \tag{2.7}$$

Next we write (1.3) in matrix form as

$$A = [\boldsymbol{a}^*, \boldsymbol{a}]\begin{pmatrix} C & D \\ \bar{D} & \bar{C} \end{pmatrix}\begin{bmatrix} \boldsymbol{a} \\ \boldsymbol{a}^* \end{bmatrix} + \sum_{i=1}^{n} (\tfrac{1}{2}|\mu_i|^2 - \tfrac{1}{2}|\lambda_i|^2 - \varepsilon_i) \tag{2.8}$$

where the matrices C and D are defined by

$$C_{ij} = \varepsilon_i \delta_{ij} + \tfrac{1}{2}[\bar{\lambda}_i \lambda_j + \bar{\mu}_j \mu_i] \tag{2.9}$$

$$D_{ij} = \tfrac{1}{2}[\bar{\lambda}_i \mu_j + \bar{\lambda}_j \mu_i].$$

It is easy to see that if the ε_i are all positive then the matrix B occuring in (2.8),

$$B = \begin{pmatrix} C & D \\ \bar{D} & \bar{C} \end{pmatrix} \tag{2.10}$$

is positive definite Hermitian. Furthermore, in the $\boldsymbol{\eta}$ representation, (2.8) becomes

$$A = [\boldsymbol{\eta}^*, \boldsymbol{\eta}]M^*BM\begin{bmatrix} \boldsymbol{\eta} \\ \boldsymbol{\eta}^* \end{bmatrix} + \sum_{i=1}^{n} (\tfrac{1}{2}|\mu_i|^2 - \tfrac{1}{2}|\lambda_i|^2 - \varepsilon_i). \tag{2.11}$$

We introduce the operator K of Kramers type defined in (1.12) into the matrix representation (2.8) of A. Let K be the operator defined by

$$K = \mathscr{C}\begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \tag{2.12}$$

where $\mathscr{C}$ denotes the operation of complex conjugation. The adjoint K' of K is given by

$$K' = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \mathscr{C} \tag{2.13}$$

whence

$$K'K = KK' = I. \tag{2.14}$$

One can easily see that

$$K'BK = B \tag{2.15}$$

and that the identity (2.15) corresponds to the invariance of (1.3) under the operator K of (1.12). It is also easy to see that a $2n \times 2n$ matrix M has the form (2.5) if and only if

$$K'MK = M. \tag{2.16}$$

To put (1.3) in the form (1.4) we need to put the matrix B in diagonal form. Hence from (2.11) we need to have a matrix M to satisfy

$$M^*BM = \text{diagonal matrix}. \tag{2.17}$$

In order that the η_i satisfy canonical commutation relations we also require that (2.6) and (2.16) hold. It is easily seen that (2.6) is equivalent to the identity

$$M^*JM = J. \tag{2.18}$$

In fact from (2.6) we see that

$$M^* = JM^{-1}J \tag{2.19}$$

and (2.18) follows immediately from this fact.

We consider first the problem of finding M such that (2.17) and (2.18) hold. Since B is positive definite and J is Hermitian it is a well known fact of matrix theory (Mehta 1967, 1977) that such a matrix M exists and that the columns of M are obtained as the eigenvectors $\boldsymbol{x}$ of the equation

$$B\boldsymbol{x} = \alpha J\boldsymbol{x}. \tag{2.20}$$

The additional requirement that M must also satisfy (2.16) follows from the invariance (2.15) of B. In fact for any matrix M let M_K be the matrix given by

$$M_K = K'MK. \tag{2.21}$$

Then from (2.15) it follows that if M satisfies

$$M^*BM = \begin{pmatrix} D_1 & 0 \\ 0 & D_2 \end{pmatrix} \tag{2.22}$$

where D_1 and D_2 are real $n \times n$ diagonal matrices, then

$$M_K^*BM_K = \begin{pmatrix} D_2 & 0 \\ 0 & D_1 \end{pmatrix}. \tag{2.23}$$

Further, in view of the fact that

$$K'JK = -J \tag{2.24}$$

it follows that if (2.18) holds then one also has that

$$M_K^* J M_K = J. \tag{2.25}$$

Thus if M solves the diagonalisation problem (2.17) and (2.18) then so also does the matrix M_K. Since in the generic case M is unique we have generically that $M = M_K$. In the case when M is not unique one sees by continuity considerations that it is possible to choose M in such a way that $M = M_K$. Hence by virtue of the invariance (2.15) we are always able to solve the diagonalisation problem (2.6), (2.16) and (2.17).

We proceed now to find the canonical form (1.4) of A. First we observe from (2.22) and (2.23) that the diagonal matrices D_1, D_2 are identical. Let us write

$$D_1 = D_2 = \begin{bmatrix} \alpha_1 & & & \\ & \alpha_2 & & \\ & & \ddots & \\ & & & \alpha_n \end{bmatrix}. \tag{2.26}$$

From (2.11) it follows that

$$\begin{aligned} A &= \sum_{i=1}^{n} \alpha_i[\eta_i^* \eta_i + \eta_i \eta_i^*] + \sum_{i=1}^{n} (\tfrac{1}{2}|\mu_i|^2 - \tfrac{1}{2}|\lambda_i|^2 - \varepsilon_i) \\ &= 2\sum_{i=1}^{n} \alpha_i \eta_i^* \eta_i + \sum_{i=1}^{n} (\alpha_i + \tfrac{1}{2}|\mu_i|^2 - \tfrac{1}{2}|\lambda_i|^2 - \varepsilon_i) \end{aligned} \tag{2.27}$$

and this latter expression is exactly the canonical form (1.4).

Next we find the polynomial equation determining the α_i, $1 \leqslant i \leqslant n$. From (2.20) one sees that this polynomial is the characteristic polynomial for the $2n \times 2n$ matrix $B^{-1/2} J B^{-1/2}$. This latter polynomial has degree $2n$ and therefore has $2n$ roots, n of which are α_i, $1 \leqslant i \leqslant n$. The remaining n roots are given by $-\alpha_i$, $1 \leqslant i \leqslant n$. This follows from the invariances (2.15) and (2.24). It is clear from these invariances that if $\boldsymbol{x}$ is an eigenfunction of (2.20) with eigenvalue α, then $K\boldsymbol{x}$ is also an eigenfunction with eigenvalue $-\alpha$.

To obtain an explicit formula for the characteristic polynomial for (2.20) we write $\boldsymbol{x} = [\boldsymbol{v}, \boldsymbol{w}]$ where $\boldsymbol{v}, \boldsymbol{w}$ are column vectors of the $n \times n$ matrices V, W. Equation (2.20) then becomes

$$\begin{aligned} (C - \alpha)\boldsymbol{v} + D\boldsymbol{w} &= 0 \\ \bar{D}\boldsymbol{v} + (\bar{C} + \alpha)\bar{\boldsymbol{w}} &= 0. \end{aligned} \tag{2.28}$$

Let $\boldsymbol{\lambda}, \boldsymbol{\mu}$ be vectors

$$\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n) \qquad \boldsymbol{\mu} = (\mu_1, \ldots, \mu_n) \tag{2.29}$$

and for any two vectors $\boldsymbol{\rho}, \boldsymbol{\nu}$ with real or complex entries ρ_i, ν_i, $1 \leqslant i \leqslant n$, let $\langle \boldsymbol{\rho}, \boldsymbol{\nu} \rangle$ denote the sum

$$\langle \boldsymbol{\rho}, \boldsymbol{\nu} \rangle = \sum_{i=1}^{n} \rho_i \nu_i. \tag{2.30}$$

Let x, y be defined by

$$\begin{aligned} x &= \langle \boldsymbol{\lambda}, \boldsymbol{v} \rangle + \langle \boldsymbol{\mu}, \bar{\boldsymbol{w}} \rangle \\ y &= \langle \bar{\boldsymbol{\mu}}, \boldsymbol{v} \rangle + \langle \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{w}} \rangle \end{aligned} \tag{2.31}$$

and E be the $n\times n$ diagonal matrix with diagonal entries ε_i, $1\leqslant i\leqslant n$. Then the equations (2.28) are the same as

$$\begin{aligned} &x\bar{\boldsymbol{\lambda}}+y\boldsymbol{\mu}+2(E-\alpha)\boldsymbol{v}=0\\ &y\boldsymbol{\lambda}+x\bar{\boldsymbol{\mu}}+2(E+\alpha)\tilde{\boldsymbol{w}}=0. \end{aligned} \tag{2.32}$$

It follows easily then from (2.31) and (2.32) that we have two simultaneous equations for x, y, namely

$$\begin{aligned} &x[2+\langle\boldsymbol{\lambda},(E-\alpha)^{-1}\bar{\boldsymbol{\lambda}}\rangle+\langle\boldsymbol{\mu},(E+\alpha)^{-1}\bar{\boldsymbol{\mu}}\rangle\\ &\qquad+y[\langle\boldsymbol{\lambda},(E-\alpha)^{-1}\boldsymbol{\mu}\rangle+\langle\boldsymbol{\lambda},(E+\alpha)^{-1}\boldsymbol{\mu}\rangle]=0\\ &x[\langle\bar{\boldsymbol{\lambda}},(E-\alpha)^{-1}\bar{\boldsymbol{\mu}}\rangle+\langle\bar{\boldsymbol{\lambda}},(E+\alpha)^{-1}\bar{\boldsymbol{\mu}}\rangle]\\ &\qquad+y[2+\langle\boldsymbol{\mu},(E-\alpha)^{-1}\bar{\boldsymbol{\mu}}\rangle+\langle\boldsymbol{\lambda},(E+\alpha)^{-1}\bar{\boldsymbol{\lambda}}\rangle]=0. \end{aligned} \tag{2.33}$$

Since (2.33) must have non-trivial solutions x, y, we find that the characteristic polynomial for α is given by the determinant of the 2×2 matrix defined by (2.33) being zero. The polynomial is therefore

$$\begin{aligned} &[2+\langle\boldsymbol{\lambda},(E-\alpha)^{-1}\bar{\boldsymbol{\lambda}}\rangle+\langle\boldsymbol{\mu},(E+\alpha)^{-1}\bar{\boldsymbol{\mu}}\rangle]\cdot[2+\langle\boldsymbol{\mu},(E-\alpha)^{-1}\bar{\boldsymbol{\mu}}\rangle+\langle\boldsymbol{\lambda},(E+\alpha)^{-1}\bar{\boldsymbol{\lambda}}\rangle]\\ &\qquad=|\langle\boldsymbol{\lambda},(E-\alpha)^{-1}\boldsymbol{\mu}\rangle+\langle\boldsymbol{\lambda},(E+\alpha)^{-1}\boldsymbol{\mu}\rangle|^2 \end{aligned} \tag{2.34}$$

where we are assuming here that α is real.

It is clear that the polynomial (2.34) is of degree $2n$ and to every positive root $\boldsymbol{\alpha}$ there corresponds a negative root $-\alpha$. We know from the matrix formulation that the roots of (2.34) must be real and there are n positive roots which we have labelled $\alpha_1,\ldots,\alpha_n$. It is not, however, obvious from the form of equation (2.34) that there are any real roots. In the following we shall show this and also where these roots must be located.

3. Finding the roots of the characteristic polynomial

We approach this problem by considering certain auxiliary equations. We can assume without loss of generality that

$$|\lambda_i|^2+|\mu_i|^2\neq 0 \qquad 1\leqslant i\leqslant n \tag{3.1}$$

for if one of the inequalities (3.1) does not hold it merely reduces the dimension of the problem. Now we define three-dimensional vectors v_i and w_i by

$$\begin{aligned} v_i&=\tfrac{1}{2}(|\lambda_i|^2-|\mu_i|^2)\boldsymbol{i}+\mathrm{Re}(\lambda_i\mu_i)\boldsymbol{j}+\mathrm{Im}(\lambda_i\mu_i)\boldsymbol{k}\\ w_i&=\tfrac{1}{2}(|\mu_i|^2-|\lambda_i|^2)\boldsymbol{i}+\mathrm{Re}(\lambda_i\mu_i)\boldsymbol{j}+\mathrm{Im}(\lambda_i\mu_i)\boldsymbol{k} \end{aligned} \tag{3.2}$$

for $1\leqslant i\leqslant n$, where $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ are assumed to be the unit vectors along the coordinate axes in $\boldsymbol{R}^3$. From (3.1) it follows that the vectors (3.2) are non-trivial since

$$|v_i|=|w_i|=\tfrac{1}{2}(|\lambda_i|^2+|\mu_i|^2). \tag{3.3}$$

Next we parametrise unit vectors u in $\boldsymbol{R}^3$ differentiably as $u(\Omega)$ where Ω varies on the unit sphere in $\boldsymbol{R}^3$. We wish to consider the polynomial equation in α,

$$\sum_{i=1}^{n}\frac{|v_i|-v_i\cdot u(\Omega)}{\varepsilon_i-\alpha}+\sum_{i=1}^{n}\frac{|w_i|-w_i\cdot u(\Omega)}{\varepsilon_i+\alpha}+2=0. \tag{3.4}$$

Lemma 3.1. Let $\alpha(\Omega)$ be a real root of (3.4) which is differentiable at $\Omega=\Omega_0$ and $\alpha(\Omega_0)\neq\varepsilon_i$, $1\leq i\leq n$. Then if grad $\alpha(\Omega_0)=0$, $\alpha(\Omega_0)$ is also a root of equation (2.34).

Proof. Let $r(\Omega_0)$ and $s(\Omega_0)$ be two unit vectors such that $(u(\Omega_0), r(\Omega_0), s(\Omega_0))$ form an orthonormal basis for $\boldsymbol{R}^3$. It is easy to see that by differentiating $u(\Omega)$ at Ω_0 in appropriate directions one obtains vectors parallel to $r(\Omega_0)$ and $s(\Omega_0)$. Now let

$$\begin{aligned} A&=\sum_{i=1}^{n}\frac{|v_i|}{\varepsilon_i-\alpha}+\sum_{i=1}^{n}\frac{|w_i|}{\varepsilon_i+\alpha}+2 \\ \boldsymbol{B}&=\sum_{i=1}^{n}\frac{v_i}{\varepsilon_i-\alpha}+\sum_{i=1}^{n}\frac{w_i}{\varepsilon_i+\alpha}\end{aligned}\tag{3.5}$$

so A is a scalar and $\boldsymbol{B}$ is a vector. Then equation (3.4) reads

$$A(\alpha)=\boldsymbol{B}(\alpha)\cdot u(\Omega).\tag{3.6}$$

If we differentiate (3.4) with respect to Ω at $\Omega=\Omega_0$ and use the fact that grad $\alpha(\Omega_0)=0$ we also obtain the equations

$$\boldsymbol{B}(\alpha)\cdot r(\Omega_0)=0 \qquad \boldsymbol{B}(\alpha)\cdot s(\Omega_0)=0.\tag{3.7}$$

The equations (3.6) and (3.7) then yield

$$A(\alpha)^2=|\boldsymbol{B}(\alpha)|^2\tag{3.8}$$

and this is the same as

$$A(\alpha)^2-[\boldsymbol{B}(\alpha)\cdot \boldsymbol{i}]^2=[\boldsymbol{B}(\alpha)\cdot \boldsymbol{j}]^2+[\boldsymbol{B}(\alpha)\cdot \boldsymbol{k}]^2\tag{3.9}$$

$$\{A(\alpha)+[\boldsymbol{B}(\alpha)\cdot \boldsymbol{i}]\}\{A(\alpha)-[\boldsymbol{B}(\alpha)\cdot \boldsymbol{i}]\}=[\boldsymbol{B}(\alpha)\cdot \boldsymbol{j}]^2+[\boldsymbol{B}(\alpha)\cdot \boldsymbol{k}]^2.\tag{3.10}$$

It is easy to see that the left-hand side of (3.10) is identical to the left-hand side of (2.34) and the right-hand sides are also identical. It follows then that $\alpha(\Omega_0)$ is a root of (2.34).

Lemma 3.2. In the region Ω in which $u(\Omega)$ is not parallel to any v_i, $1\leq i\leq n$, equation (3.4) has n positive roots $\alpha_i(\Omega)$, $1\leq i\leq n$. The $\alpha_i(\Omega)$ may be chosen locally to be C^∞ functions satisfying the inequality

$$\varepsilon_i<\alpha_i(\Omega)<\varepsilon_{i+1} \qquad 1\leq i\leq n.\tag{3.11}$$

Here the ε_i are assumed to satisfy the condition (1.8) and $\varepsilon_{n+1}=+\infty$.

Proof. This is a simple consequence of the implicit function theorem and the fact that the function of α on the left-hand side in (3.4) goes from $-\infty$ to $+\infty$ as α increases from ε_i to ε_{i+1}.

Lemma 3.3. Suppose at $\Omega=\Omega_i$, $u(\Omega_i)$ is parallel to v_i. Then $\alpha=\varepsilon_i$ is a zero of (3.4) at $\Omega=\Omega_i$ if and only if the coefficient of $(\varepsilon_i-\alpha)^{-1}$ in (2.34) is zero.

Proof. We take

$$u(\Omega_i)=\frac{v_i}{\frac{1}{2}(|\lambda_i|^2+|\mu_i|^2)}\tag{3.12}$$

so the coefficient of $(\varepsilon_i - \alpha)^{-1}$ in (3.4) is zero. Then ε_i is a root of the equation for $\Omega = \Omega_i$ only if

$$\sum_{j\neq i}\left(|v_j| - \frac{v_j \cdot v_i}{\frac{1}{2}(|\lambda_i|^2 + \mu_i|^2)}\right)(\varepsilon_j - \varepsilon_i)^{-1} + \sum_{j=1}^{n}\left(|w_j| - \frac{w_j \cdot v_i}{\frac{1}{2}(|\lambda_i|^2 + |\mu_i|^2)}\right)(\varepsilon_j + \varepsilon_i)^{-1} + 2 = 0. \quad (3.13)$$

This is the same as

$$\sum_{j\neq i} \frac{|\lambda_i\bar{\mu}_j - \lambda_j\bar{\mu}_i|^2}{\varepsilon_j - \varepsilon_i} + \sum_{j\neq i}^{n} \frac{|\lambda_i\bar{\lambda}_j - \bar{\mu}_i\mu_j|^2}{\varepsilon_j - \varepsilon_i} + 2(|\lambda_i|^2 + |\mu_i|^2) = 0. \quad (3.14)$$

It is not difficult now to verify that the left-hand side of (3.14) is exactly the coefficient of $(\varepsilon_i - \alpha)^{-1}$ in equation (2.34).

Lemma 3.4. Suppose the coefficients of $(\varepsilon_i - \alpha)^{-1}$ in (2.34) are all non-zero, $1 \leqslant i \leqslant n$. Then the functions $\alpha_i(\Omega)$, $1 \leqslant i \leqslant n$, defined in lemma 3.2 are infinitely differentiable on the unit sphere.

Proof. The function $\alpha_i(\Omega)$ can only fail to be C^∞ if $\Omega = \Omega_i$ or Ω_{i+1}, and $\lim_{\Omega\to\Omega_i} \alpha_i(\Omega) = \varepsilon_i$ or $\lim_{\Omega\to\Omega_{i+1}} \alpha_i(\Omega) = \varepsilon_{i+1}$ respectively. We consider the former case. We write (3.4) as

$$\frac{|v_i| - v_i \cdot u(\Omega)}{\varepsilon_i - \alpha} + g(\alpha) = 0 \quad (3.15)$$

where by assumption we have $g(\varepsilon_i) \neq 0$. Equation (3.15) is the same as

$$\frac{|v_i| - v_i \cdot u(\Omega)}{g(\alpha)} + \varepsilon_i - \alpha = 0 \quad (3.16)$$

and this clearly has a solution at $\Omega = \Omega_i$ given by $\alpha(\Omega_i) = \varepsilon_i$. The implicit function theorem then guarantees a C^∞ solution $\alpha(\Omega)$ in a neighbourhood of $\Omega = \Omega_i$ satisfying $\alpha(\Omega_i) = \varepsilon_i$, and it is easy to see that

$$\text{grad}\, \alpha(\Omega_i) = 0. \quad (3.17)$$

Let us assume that for some Ω close to Ω_i we have $\alpha(\Omega) > \varepsilon_i$. If follows from (3.16) then that $g(\alpha(\Omega)) > 0$. Thus $g(\alpha) > 0$ for α close to ε_i. It is therefore clear that for *all* Ω close to Ω_i, $\alpha(\Omega) > \varepsilon_i$. Hence $\alpha(\Omega)$ is identical to the function $\alpha_i(\Omega)$ when $\Omega \neq \Omega_i$. We have shown then that $\alpha_i(\Omega)$ is C^∞ at $\Omega = \Omega_i$. Similar arguments apply to the case $\Omega = \Omega_{i+1}$.

Lemma 3.5. Assume the conditions of lemma 3.4. Then the $2n$ numbers $\max_\Omega \alpha_i(\Omega)$, $\min_\Omega \alpha_i(\Omega)$, $1 \leqslant i \leqslant n$, are identical to the $2n$ numbers $\alpha_1, \ldots, \alpha_n, \varepsilon_1, \ldots, \varepsilon_n$, where the α_i are defined by (1.4).

Proof. From lemma 3.1 one sees that an extremum of $\alpha_i(\Omega)$ is one of the numbers α_i in (1.4) provided it is not an ε_j, $1 \leqslant j \leqslant n$.

Proof of theorem 1.1. We assume again the conditions of lemma 3.4. The functions $\alpha_j(\Omega)$, $1 \leqslant j < i + r$, take their values in the interval $[\varepsilon_i, \varepsilon_{i+r}]$. The number of maxima and minima of these functions is $2r$. If $\varepsilon_i < \min \alpha_i(\Omega) < \max \alpha_{i+r-1}(\Omega) < \varepsilon_{i+r}$ then there are $2r - (r-1) = r + 1$ roots α_i in the interval $(\varepsilon_i, \varepsilon_{i+r})$. On the other hand if the equalities $\varepsilon_i = \min \alpha_i(\Omega) < \max \alpha_{i+r-1}(\Omega) = \varepsilon_{i+r}$ hold there are $2r - (r+1) = r - 1$ roots α_i in $(\varepsilon_i, \varepsilon_{i+r})$. The other parts of theorem 1.1 follow from the matrix formulation.

We conclude this paper by showing how to obtain the inequality (1.10) from the properties of the polynomial (3.4). This method has the advantage of showing how one can improve on the inequality (1.10). An improvement of (1.10) was needed in the simple case discussed by Conlon (1985).

Theorem 3.6. Let β be defined by (1.11). Then β satisfies the inequality $\beta \geqslant 0$.

Proof. Define a transformation $\Omega \to \Omega'$ such that $u(\Omega')$ is the reflection of $u(\Omega)$ in the j, k plane. It is easy to see then that $-\alpha_i(\Omega')$, $1 \leqslant i \leqslant n$, are the negative roots of (3.4). It follows then, on calculating the coefficient of α^{2n-1} in (3.4) that

$$\sum_{i=1}^{n} [\alpha_i(\Omega) - \alpha_i(\Omega')] = \tfrac{1}{2} \sum_{i=1}^{n} (w_i - v_i) \cdot u(\Omega). \tag{3.18}$$

Thus from (3.2) we have that

$$\sum_{i=1}^{n} [\max \alpha_i(\Omega) - \min \alpha_i(\Omega)] \geqslant \frac{1}{2} \left| \sum_{i=1}^{n} |\mu_i|^2 - |\lambda_i|^2 \right|. \tag{3.19}$$

From lemma 3.5 we have

$$\begin{aligned} \sum_{i=1}^{n} (\alpha_i + \varepsilon_i) &= \sum_{i=1}^{n} [\max \alpha_i(\Omega) + \min \alpha_i(\Omega)] \\ &= \sum_{i=1}^{n} [\max \alpha_i(\Omega) - \min \alpha_i(\Omega)] + 2 \sum_{i=1}^{n} \min \alpha_i(\Omega) \\ &\geqslant \frac{1}{2} \left| \sum_{i=1}^{n} |\mu_i|^2 - |\lambda_i|^2 \right| + 2 \sum_{i=1}^{n} \varepsilon_i. \end{aligned} \tag{3.20}$$

The result follows clearly from (3.20).

Theorem 3.7. With β defined as in (1.11) there is the inequality

$$\beta \leqslant \sum_{i=1}^{n} |\mu_i|^2. \tag{3.21}$$

Proof. Consider the polynomial equation

$$\sum_{i=1}^{n} \frac{|v_i| - v_i \cdot u(\Omega)}{\varepsilon_i - \alpha'} + 2 = 0. \tag{3.22}$$

It is clear that (3.22) has n positive roots $\alpha_i'(\Omega)$, $1 \leqslant i \leqslant n$, where

$$\varepsilon_i < \alpha_i'(\Omega) < \varepsilon_{i+1} \qquad 1 \leqslant i \leqslant n. \tag{3.23}$$

Furthermore by comparing (3.4) and (3.22) one also has that

$$\alpha_i(\Omega) \leqslant \alpha_i'(\Omega) \qquad 1 \leqslant i \leqslant n. \tag{3.24}$$

Next let $Q(\alpha) = 0$ be the polynomial equation obtained from (2.34) by setting all the terms in $(\varepsilon_i + \alpha)^{-1}$ to zero, $1 \leqslant i \leqslant n$. This gives an equation with n roots $\alpha_1', \ldots, \alpha_n'$. One can show by mimicking the previous arguments that the $2n$ numbers $\alpha_1', \ldots, \alpha_n'$, $\varepsilon_1, \ldots, \varepsilon_n$ are identical to the numbers $\max \alpha_i'(\Omega)$, $\min \alpha_i'(\Omega)$, $1 \leqslant i \leqslant n$. Now by computing the coefficient of α^{n-1} in the equation $Q(\alpha) = 0$ one sees that

$$\sum_{i=1}^{n} \alpha_i' = \tfrac{1}{2} \sum_{i=1}^{n} [|\mu_i|^2 + |\lambda_i|^2 + 2\varepsilon_i]. \tag{3.25}$$

We have then from (3.24) the inequalities

$$\sum_{i=1}^{n} (\alpha_i + \varepsilon_i) = \sum_{i=1}^{n} [\min \alpha_i(\Omega) + \max \alpha_i(\Omega)]$$

$$\leqslant \sum_{i=1}^{n} [\min \alpha_i'(\Omega) + \max \alpha_i'(\Omega)]$$

$$= \sum_{i=1}^{n} (\alpha_i' + \varepsilon_i). \tag{3.26}$$

Inequality (3.21) then follows from (3.25) and (3.26).

Acknowledgments

I should like to thank Mark Ashbaugh for helpful conversations, and also the referee for bringing my attention to the Kramers invariance. This research was supported by US National Science Foundation grant DMS 8600748.

References

Bogoliubov N N 1947 *J. Phys. (USSR)* **11** 23–32
Conlon J 1985 *Commun. Math. Phys.* **100** 355–97
—— 1987 *Commun. Math. Phys.* **108** 363–74
Conlon J, Lieb E and Yau H-T 1987 in preparation
Dyson F 1967 *J. Math. Phys.* **8** 1538–45
Foldy L 1961 *Phys. Rev.* **124** 649–51
Kramers H A 1930 *Proc. Acad. Amst.* **33** 959–72
—— 1956 *Collected Scientific Papers* (Amsterdam: North-Holland)
Lieb E 1965 *The Bose Fluid, Lectures in Theoretical Physics* vol VII C (Boulder, CO: University of Colorado Press) pp 175–224
Mehta M L 1967 *Random Matrices* (New York: Academic)
—— 1977 *Elements of Matrix Theory* (Delhi: Hindustan)
Messiah A 1958 *Quantum Mechanics* vol II (New York: Wiley)