On the simultaneous diagonalisability of matrices

1986 J. Phys. A: Math. Gen. 19 3917

(http://iopscience.iop.org/0305-4470/19/18/036)

J. Phys. A: Math. Gen. **19** (1986) 3917-3919. Printed in Great Britain

COMMENT

On the simultaneous diagonalisability of matrices†

W Grimus and G Ecker

Institut für Theoretische Physik, Universität Wien, Boltzmanngasse 5, A-1090 Wien, Austria

Received 12 March 1986

Abstract. We prove two theorems on the simultaneous diagonalisability of a set of complex square matrices by a biunitary transformation.

The problem of simultaneous diagonalisation of a set of square matrices (e.g. Mehta 1977) arises in the discussion of natural flavour conservation in Higgs induced neutral currents (e.g. Sartori 1979, Gatto *et al* 1980, Frère and Yao 1985). In this context the following theorem was formulated.

Theorem 1. Let $\{A_1, \ldots, A_N\}$ be a set of complex $m \times m$ matrices. Then there exist unitary matrices U, V such that $U^\dagger A_i V$ is diagonal for all $i = 1, \ldots, N$ iff the sets $S_1 = \{A_i^\dagger A_j\}_{i,j=1,\ldots,N}$ and $S_2 = \{A_i A_j^\dagger\}_{i,j=1,\ldots,N}$ are Abelian.

This is the original version of Sartori (1979) who, however, only sketched the idea of a proof in his paper. Theorem 1 was discussed again by Gatto *et al* (1980) who presented a proof under the additional assumption of, e.g., $A_1 A_1^\dagger$ being non-degenerate. More recently, Frère and Yao (1985) referred to a similar theorem by Federbush where both the non-degeneracy of $A_1^\dagger A_1$ and the non-singularity of the matrices A_i seem to play an essential role. The purpose of the present comment is twofold: firstly we show that theorem 1 holds in its original form, i.e. no assumptions concerning non-degeneracy or non-singularity of matrices need to be made. Secondly, if at least one of the matrices A_i is non-singular it can be shown that the commutativity of either S_1 or S_2 is actually sufficient for the simultaneous diagonalisation of $A_1, \ldots, A_N$ (theorem 2).

Let us first prove theorem 1. The commutativity of S_1 and S_2 is clearly necessary for the existence of unitary matrices U, V such that $U^\dagger A_i V$ are diagonal for all $i = 1, \ldots, N$. The more interesting part of the theorem concerns the opposite implication. If S_1 is Abelian all elements of S_1 are normal matrices. Consequently, there exists a common orthonormal basis of eigenvectors $\{x_\alpha\}_{\alpha=1,\ldots,m}$ for all matrices of S_1:

$$A_i^\dagger A_j x_\alpha = \lambda_{ij}^\alpha x_\alpha \tag{1}$$

with λ_{ij}^α ($\alpha = 1, \ldots, m$; $i, j = 1, \ldots, N$) the corresponding eigenvalues. Let $\{x_1, \ldots, x_k\}$ be those vectors for which $A_i x_\alpha = 0$ for all $i = 1, \ldots, N$. Then for all $\alpha > k$ (for $k = m$ the theorem holds trivially because $A_i \equiv 0$) there exists an index i_α such that $A_{i_\alpha} x_\alpha \neq 0$. This allows one to define the normalised vectors

$$y_\alpha = A_{i_\alpha} x_\alpha / \| A_{i_\alpha} x_\alpha \| \qquad \alpha = k+1, \ldots, m. \tag{2}$$

† Partially supported by 'Fonds zur Förderung der wissenschaftlichen Forschung in Österreich', Project Nr 5444.

Using

$$\lambda_{ii}^{\alpha} = \|A_i x_\alpha\|^2 \tag{3}$$

one finds that the y_α form an orthonormal system since

$$(y_\alpha | y_\beta) = \lambda_{i_\alpha i_\beta}^{\beta} \delta_{\alpha\beta} / (\|A_{i_\alpha} x_\alpha\| \, \|A_{i_\beta} x_\beta\|) = \delta_{\alpha\beta} \qquad \alpha, \beta > k. \tag{4}$$

Furthermore, we obtain

$$(y_\alpha | A_i x_\beta) = \lambda_{i_\alpha i}^{\beta} \delta_{\alpha\beta} (\lambda_{i_\alpha i_\alpha}^{\alpha})^{-1/2} \tag{5}$$

which implies that each vector $A_i x_\beta$ has only a y_β component and a possible component in the space orthogonal to all y_α $(\alpha = k+1, \ldots, m)$. To show that $A_i x_\beta$ is in fact proportional to y_β it is sufficient to prove

$$|(y_\beta | A_i x_\beta)| = \|A_i x_\beta\|. \tag{6}$$

At this point the commutativity of S_2 enters. From

$$A_i A_i^\dagger A_i A_j^\dagger A_j x_\beta = A_i A_j^\dagger A_i A_i^\dagger A_j x_\beta \tag{7}$$

we obtain the relation

$$\lambda_{ii}^{\beta} \lambda_{jj}^{\beta} = \lambda_{ji}^{\beta} \lambda_{ij}^{\beta} = |\lambda_{ij}^{\beta}|^2 \tag{8}$$

using (1). With the help of (8) we can derive (6):

$$|(y_\beta | A_i x_\beta)| = |\lambda_{i_\beta i}^{\beta}| (\lambda_{i_\beta i_\beta}^{\beta})^{-1/2} = (\lambda_{ii}^{\beta})^{1/2} = \|A_i x_\beta\|.$$

Completing $\{y_{k+1}, \ldots, y_m\}$ to an orthonormal basis $\{y_1, \ldots, y_k, y_{k+1}, \ldots, y_m\}$ we can write†

$$A_i x_\alpha = \rho_i^\alpha y_\alpha \qquad i = 1, \ldots, N; \ \alpha = 1, \ldots, m \tag{9}$$

and

$$A_i = \sum_{\alpha=1}^{m} \rho_i^\alpha y_\alpha x_\alpha^\dagger = (y_1, \ldots, y_m) \begin{pmatrix} \rho_i^1 & & 0 \\ & \ddots & \\ 0 & & \rho_i^m \end{pmatrix} \begin{pmatrix} x_1^\dagger \\ \vdots \\ x_m^\dagger \end{pmatrix}. \tag{10}$$

Thus we have obtained unitary matrices

$$\begin{aligned} U &= (y_1, \ldots, y_m) \\ V &= (x_1, \ldots, x_m) \end{aligned} \tag{11}$$

which diagonalise all A_i, completing the proof of theorem 1.

The commutativity of both S_1 and S_2 was crucial for the proof. If, however, one of the matrices A_i $(i = 1, \ldots, N)$ is non-singular it is already sufficient for the simultaneous diagonalisability of the A_i that either S_1 or S_2 is Abelian. This is the content of the following theorem.

Theorem 2. Let A_1 be non-singular and $S_1 = \{A_i^\dagger A_j\}_{i,j=1,\ldots,N}$ be an Abelian set. Then there exist unitary matrices U, V such that $U^\dagger A_i V$ is diagonal for all $i = 1, \ldots, N$ and thus the set $S_2 = \{A_i A_j^\dagger\}_{i,j=1,\ldots,N}$ is also Abelian.

† Of course, $\rho_i^\alpha = 0$ for $\alpha = 1, \ldots, k$ and all i.

In order to prove theorem 2 we choose vectors $\{x_\alpha\}_{\alpha=1,\ldots,m}$ as before and we define

$$y_\alpha = A_1 x_\alpha / \|A_1 x_\alpha\| \qquad \alpha = 1, \ldots, m. \tag{12}$$

The y_α exist for all α because A_1 is non-singular. As in (5) we obtain

$$(y_\alpha | A_i x_\beta) = \lambda^\beta_{1i} \delta_{\alpha\beta} (\lambda^\alpha_{11})^{-1/2}. \tag{13}$$

Since now $\{y_1, \ldots, y_m\}$ is a complete orthonormal basis of C^m it follows immediately from (13) that $A_i x_\beta$ is proportional to y_β. As for theorem 1 U and V are given by (11).

In conclusion we want to note that the non-singularity of A_1 in theorem 2 is essential. For instance, it is not sufficient to demand only non-degeneracy of $A_1^\dagger A_1$. Finally, we emphasise once again that no assumptions concerning non-degeneracy were needed to prove both theorems, making their actual application much simpler in most cases.

The content of this comment grew out of discussions during an earlier collaboration with our late friend Walter Konetschny.

Note added in proof. After submission of this comment, Professors R Gatto and G Sartori supplied us with an alternative proof of theorem 1.

References

Frère J-M and Yao Y-P 1985 *Phys. Rev. Lett.* **55** 2386
Gatto R, Morchio G, Sartori G and Strocchi F 1980 *Nucl. Phys.* B **163** 221
Mehta M L 1977 *Elements of Matrix Theory* (Delhi: Hindustan Publishing Corporation) ch 4
Sartori G 1979 *Phys. Lett.* **82B** 255