The diagonalization of cubic matrices

2000 J. Phys. A: Math. Gen. 33 5669

(http://iopscience.iop.org/0305-4470/33/32/305)

J. Phys. A: Math. Gen. **33** (2000) 5669–5673. Printed in the UK PII: S0305-4470(00)11838-4

The diagonalization of cubic matrices

D Cocolicchio and M Viggiano
Dipartimento di Matematica, Universitá della Basilicata, Via N Sauro 85, Potenza, Italy

Received 9 February 2000

Abstract. This paper is devoted to analysing the problem of the diagonalization of cubic matrices. We extend the familiar algebraic approach which is based on the Cardano formulae. We rewrite the complex roots of the associated resolvent secular equation in terms of transcendental functions and we derive the diagonalizing matrix.

Many important problems in applied mathematics, physics and engineering frequently deal with the question of carrying out a similarity transformation to reduce a matrix to a diagonal form. This represents a powerful mathematical tool whenever the formulation in terms of scalar equations no longer suffices for the description of anisotropic problems and vector equations appear. When the relationship is linear, although anisotropic, it can be expressed in terms of a tensor. Choosing a coordinate system (i.e. an appropriate basis), the tensor can be written as a matrix. One usually assumes that these complex matrices are diagonalizable. However, finite-dimensional complex matrices $\boldsymbol{M}$ cannot always be diagonalized, but can only be brought into a Jordan canonical form $\boldsymbol{E}$ by means of a similarity transformation $\boldsymbol{M} = \boldsymbol{K}\boldsymbol{E}\boldsymbol{K}^{-1}$. In general the matrix $\boldsymbol{K}$ is not unique; also $\boldsymbol{E}\boldsymbol{E}'$ leaves $\boldsymbol{M}$ invariant, with $\boldsymbol{E}$, $\boldsymbol{E}'$ nonsingular commuting matrices. When the n-dimensional matrix $\boldsymbol{M}$ is not symmetric, $\boldsymbol{E}$ has in general n^2 different complex matrix elements. When $\boldsymbol{M}$ is a Hermitian matrix, it can be viewed as an element of the Lie algebra $U(3)$ and decomposed in terms of the (Gell-Mann matrices) generators of $U(3)$. As a consequence of some group theoretical considerations, the parameter space is (a subspace of) the projective manifold $SU(3)/SU(2) = CP^2$. In this paper we examine the case when a matrix

$$\boldsymbol{M} = \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{pmatrix} \tag{1}$$

can be transformed into a diagonal matrix $\boldsymbol{D} = \boldsymbol{V}\boldsymbol{M}\boldsymbol{V}^{\dagger}$ by means of only one diagonalizing matrix $\boldsymbol{V}$. The rotating matrix $\boldsymbol{V}^{\dagger}$ consists of the column eigenvectors $\boldsymbol{v}^t = (\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)$, satisfying

$$\boldsymbol{M}|\boldsymbol{v}_\alpha\rangle = \mu_\alpha|\boldsymbol{v}_\alpha\rangle. \tag{2}$$

The eigenvalues μ_α are solutions of a linear and homogeneous system of the following equations:

$$m_{ij} - \mu\delta_{ij} = 0 \qquad i, j = 1, 2, 3 \tag{3}$$

which have a nontrivial solution if

$$\det(\boldsymbol{M} - \mu\boldsymbol{I}) = 0. \tag{4}$$

The only difficulty consists in attaching the eigenvalue to the right state. This is performed by comparing the solutions of the cubic secular equation with the predicted effective eigenstate of each particle. Once we have identified to which eigenvalue corresponds each state, we can proceed without any further concern about the chosen ordering. The secular equation yields the following cubic equation in μ:

$$\mu^3 + b\mu^2 + c\mu + d = 0 \tag{5}$$

whose invariant coefficients are respectively given by

$$\begin{aligned} -b &= \mathrm{tr}\, \boldsymbol{M} = \sum_{\alpha} \mu_\alpha \\ c &= \tfrac{1}{2}[(\mathrm{tr}\, \boldsymbol{M})^2 - \mathrm{tr}(\boldsymbol{M}^2)] = \sum_{\substack{\alpha,\beta \\ \alpha\neq\beta}} \mu_\alpha \mu_\beta \\ -d &= \det \boldsymbol{M} = \prod_{\alpha} \mu_\alpha. \end{aligned} \tag{6}$$

Of course, the three roots of equation (5) are just $\mu_\alpha, \alpha = 1, 2, 3$. The characteristic polynomial is of order three which makes the direct solution of this eigenvalue problem quite complicated. Substituting one root at a time, we can find the three eigenvectors. There are several methods to determine the roots of a cubic equation. A method to solve many particular cubic equations was first published by Gerolamo Cardano in 1545 [1], though it now seems to have been disclosed to Cardano by Niccolò Fontana, best known as Tartaglia (The Stammerer), who, in turn, was aware of the solution in the context of the mathematical scholars at the University of Bologna. This solution is probably due to Scipione del Ferro, thirty years earlier, and later improved by his team-workers Annibale Della Nave and mainly Antonio Maria Fiore. However, the solution was never published until Cardano [2], who was permitted to be acquainted with it, did so after nagging insistence, and only under the solemn oath '*ad Sacra Dei Evangelia*'; he should never have published their formula. But we often learn of the real value of the promises of career academicians. In the Cardano method, the cubic equation (5) is first transformed in a reduced form in which the quadratic term no longer appears, by means of the following substitution:

$$\mu = \left(z - \frac{b}{3}\right). \tag{7}$$

This transformation leads to a reduced equation of the form

$$z^3 + qz + r = 0 \tag{8}$$

where

$$q = 3u = -\tfrac{1}{3}(b^2 - 3c) \tag{9}$$

$$r = 2v = \tfrac{1}{27}(2b^3 - 9bc + 27d). \tag{10}$$

In the Cardano method, the substitution

$$z = x + y \tag{11}$$

is then made, which further transforms the reduced cubic equation to

$$x^3 + y^3 + 3(x + y)(xy + u) + 2v = 0 \tag{12}$$

which will clearly be satisfied if x and y satisfy the relations

$$x^3 + y^3 + 2v = 0 \tag{13}$$

$$xy + u = 0. \tag{14}$$

Solving for y gives $y = -u/x$; substituting this into equation (13) leads to a single sixth-degree equation for x having the form

$$x^6 + 2vx^3 - u^3 = 0. \tag{15}$$

This equation is simply a quadratic in x^3, which is easily solved to yield

$$x^3 = -v + \sqrt{v^2 + u^3} \tag{16}$$

$$y^3 = -\frac{u^3}{x^3} = -v - \sqrt{v^2 + u^3}. \tag{17}$$

In these expressions, the positive sign of the radical has been chosen. It can be shown that it yields all the roots of the cubic and no loss of generality occurs if the negative sign is rejected. The quantities x and y are the cube roots of the expressions

$$x_0 = \left[-v + \sqrt{v^2 + u^3}\right]^{\frac{1}{3}} \tag{18}$$

$$y_0 = \left[-v - \sqrt{v^2 + u^3}\right]^{\frac{1}{3}} \tag{19}$$

$$x_+ = \mathrm{e}^{+\frac{2\pi}{3}\mathrm{i}} x_0 = -\tfrac{1}{2}(1 - \mathrm{i}\sqrt{3})x_0 \tag{20}$$

$$y_+ = \mathrm{e}^{-\frac{2\pi}{3}\mathrm{i}} x_0 = -\tfrac{1}{2}(1 + \mathrm{i}\sqrt{3})y_0 \tag{21}$$

$$x_- = \mathrm{e}^{-\frac{2\pi}{3}\mathrm{i}} x_0 = -\tfrac{1}{2}(1 + \mathrm{i}\sqrt{3})x_0 \tag{22}$$

$$y_- = \mathrm{e}^{+\frac{2\pi}{3}\mathrm{i}} x_0 = -\tfrac{1}{2}(1 - \mathrm{i}\sqrt{3})y_0 \tag{23}$$

and they lead to the following roots:

$$z_0 = x_0 + y_0 = \left[-v + \sqrt{v^2 + u^3}\right]^{\frac{1}{3}} + \left[-v - \sqrt{v^2 + u^3}\right]^{\frac{1}{3}} \tag{24}$$

$$z_+ = x_+ + y_+ = -\frac{1}{2}(x_0 + y_0) + \frac{\sqrt{3}}{2}\mathrm{i}(x_0 - y_0) \tag{25}$$

$$z_- = x_- + y_- = -\frac{1}{2}(x_0 + y_0) - \frac{\sqrt{3}}{2}\mathrm{i}(x_0 - y_0). \tag{26}$$

These Cardano formulae give the exact roots involving only rational algebra operations [3]. Unfortunately, they present some difficulties due to the presence of the terms containing the cube root of a quadratic expression. Therefore, they are not easy to use as a basis for further applications. For example, it is hard to expand them in series, to manipulate them algebraically or to use them for any purpose other than computing numerical values for the roots. As a result, the general treatment of such problems is usually neglected and one proceeds further only by numerical calculations using appropriately chosen numbers as coefficients. Frequently, we consider the case $u < 0$ and $v^2 + u^3 < 0$ so that there are three real roots and we can write

$$z_0 = -2\sqrt{|u|}\cos\left(\frac{\theta}{3}\right) \tag{27}$$

$$z_\pm = \sqrt{|u|}\left[\cos\left(\frac{\theta}{3}\right) \pm \sqrt{3}\sin\left(\frac{\theta}{3}\right)\right] \qquad \text{with } \cos\theta = -\left(\frac{v}{u\sqrt{|u|}}\right). \tag{28}$$

These expressions induce us to look for a similar form for the general complex roots of the cubic equation, by means of hyperbolic and their inverse functions, in view of the relationships that exist between the trigonometric and hyperbolic functions. They can be based upon the following simple relation:

$$-v \pm \sqrt{v^2 + u^3} = u\sqrt{|u|}\exp\{\pm F(-\beta)\} \tag{29}$$

where

$$F(\beta) = \begin{cases} \cosh^{-1}\beta = \ln\left(\beta + \sqrt{\beta^2 - 1}\right) & \text{if} \quad |\beta| \geqslant 1 \quad \text{and} \quad u < 0 \\ \cos^{-1}\beta = -\mathrm{i}\ln\left(\beta \pm \sqrt{\beta^2 - 1}\right) & \text{if} \quad |\beta| < 1 \quad \text{and} \quad u < 0 \\ \sinh^{-1}\beta = \ln\left(\beta + \sqrt{\beta^2 + 1}\right) & \text{for all } \beta \quad \text{and} \quad u > 0 \end{cases} \tag{30}$$

with $\beta(u, v) = v/|u|\sqrt{|u|}$. The application of these trivial relations may engender the relevant eigenvalues $\mu_\alpha = (z_\alpha - b/3)$ in a more usable form, which can be extracted by one element of the following set:

$$z_0 = -2\sqrt{|u|}\sinh\left(\frac{\omega}{3}\right) \tag{31}$$

$$z_+ = \sqrt{|u|}\left[\sinh\left(\frac{\omega}{3}\right) + \mathrm{i}\sqrt{3}\cosh\left(\frac{\omega}{3}\right)\right] \tag{32}$$

$$z_- = \sqrt{|u|}\left[\sinh\left(\frac{\omega}{3}\right) - \mathrm{i}\sqrt{3}\cosh\left(\frac{\omega}{3}\right)\right] \tag{33}$$

where

$$\sinh\omega \equiv \frac{v}{|u|\sqrt{|u|}}. \tag{34}$$

These solutions have many significant advantages over those given directly by Cardano's formulae. They are more compact and in a sense also much simpler. In fact, they can be manipulated to give closed-form results which would otherwise be practically unobtainable. In addition, they are easily expanded in series using the usual rules for their summation. Finally the hyperbolic/trigonometric forms are better suited for numerical calculations with computers than the usual Cardano expressions. A few sequential keystrokes are all that one needs, for example, to find the roots using a pocket calculator having built-in hyperbolic and trigonometric functions.

From these results, we can derive the transformation matrix $\boldsymbol{V}$. To this end, we can conveniently rewrite the column eigenvectors $(a_I, b_I, c_I)^T$ corresponding to each eigenvalue μ_I, $I = \pm, 0$, in accordance with the following expressions:

$$\begin{aligned} a_I &= (\mu_I - m_{22})m_{13} - m_{12}m_{23} \\ b_I &= -[(\mu_I - m_{11})m_{23} + m_{21}m_{13}] \\ c_I &= (\mu_I - m_{11})(\mu_I - m_{22}) + m_{12}m_{21} \end{aligned} \tag{35}$$

where I labels the subscripts $\pm$, 0, to assign to the eigenstate indices 1, 2, 3. The last step of the problem consists in the identification of each μ_α to one of the μ_I. Starting from the elements of $\boldsymbol{M}$, we can reconstruct the generator $\boldsymbol{A}$ of $\boldsymbol{V} = \mathrm{e}^{\boldsymbol{A}}$ by means of the relation $\mathrm{e}^{\boldsymbol{A}} = \boldsymbol{DVM}^{-1}$, where $\boldsymbol{M}^{-1}$ can be expressed in terms of the components of the eigenvectors of $\boldsymbol{M}$ in the following form:

$$\boldsymbol{M}^{-1} = \frac{1}{\det \boldsymbol{M}} \begin{pmatrix} b_-c_0 - b_0c_- & a_0c_- - a_-c_0 & a_-b_0 - a_0b_- \\ b_0c_+ - b_+c_0 & a_+c_0 - a_0c_+ & a_0b_+ - a_+b_0 \\ b_+c_- - b_-c_+ & a_-c_+ - a_+c_- & a_+b_- - a_-b_+ \end{pmatrix}. \tag{36}$$

In particular, assuming that $\boldsymbol{V}$ is unitary ($\boldsymbol{A}^\dagger = -\boldsymbol{A}$) and using the Cayley–Hamilton decomposition, we can write

$$\boldsymbol{V} = \mathrm{e}^{\boldsymbol{A}} = \boldsymbol{I} + \left(\frac{\sin\mathcal{X}}{\mathcal{X}}\right)\boldsymbol{A} + \frac{1}{2}\left[\frac{\sin(\mathcal{X}/2)}{\mathcal{X}/2}\right]^2\boldsymbol{A}^2 \tag{37}$$

in terms of the elements of the matrix $\boldsymbol{A}$. If $\boldsymbol{A}$ is real, it must be of the form

$$\boldsymbol{A} = \begin{pmatrix} 0 & \chi_1 & \chi_3 \\ -\chi_1 & 0 & -\chi_2 \\ -\chi_3 & \chi_2 & 0 \end{pmatrix} \tag{38}$$

and, consequently, $\boldsymbol{V}$ will be given by

$$\boldsymbol{V} = \begin{pmatrix} 1 - \frac{\mathcal{X}^2 - \chi_2^2}{2} & \chi_1 + \frac{1}{2}\chi_2\chi_3 & \chi_3 - \frac{1}{2}\chi_1\chi_2 \\ -\chi_1 + \frac{1}{2}\chi_2\chi_3 & 1 - \frac{\mathcal{X}^2 - \chi_3^2}{2} & -\chi_2 - \frac{1}{2}\chi_1\chi_3 \\ -\chi_3 - \frac{1}{2}\chi_1\chi_2 & \chi_2 - \frac{1}{2}\chi_1\chi_3 & 1 - \frac{\mathcal{X}^2 - \chi_1^2}{2} \end{pmatrix} \tag{39}$$

with $\mathcal{X} = \sqrt{\chi_1^2 + \chi_2^2 + \chi_3^2}$.

Let us end by remarking that this diagonalizing approach may be considered as a suitable tool for finding the solution of many different fundamental physical problems which involve three-dimensional Hilbert spaces. Recently, there has been a renewed interest in the study of quantum three-level systems and adiabatic geometric phases [4]. The occurrence of the (non-Abelian) geometric phases can be described by solving the difficulties involving the solution of the eigenvalue problem for the ensuing Hamiltonian. When the Hamiltonian is a Hermitian matrix, it can be viewed as an element of the Lie algebra $U(3)$. In general, our approach could provide a basic tool to fully analyse the problem. A further application is represented by the form of quark and neutrino mass matrices and eventually different degeneracy structures, which correspond to distinct ranges of the relevant mixing parameters [5]. Although there are already hundreds of publications on mass matrices, this strategy has not been worked out in detail, simply due to a lack of suitable formulae of the type we introduce in this paper.

In conclusion, in this paper we address the question of the diagonalization of a cubic matrix with the introduction of transcendental functions and their inverse. Then, we derive the diagonalizing matrix, which could be useful in many physical problems which involve three-dimensional Hilbert spaces.

References

[1] Hieronimi Cardani 1545 Praestantissimi mathematici philosophi, ac medici *Artis Magnae sive de Regulis Algebraicis Liber Unus* per Joh. Petreium (Norinbergae)

[2] Tartaglia N 1546 *Quesiti et Inventioni Diverse* (Venezia)

[3] Uspensky J V 1948 *Theory of Equations* 1st edn (New York: McGraw-Hill) p 84ff

[4] Dattoli G and Torre A 1990 Cayley–Klein parameters and evolution of two- and three-level systems and squeezed states *J. Math. Phys.* **31** 236

Arvind P K, Mallesh K S and Mukunda N 1997 A generalized Pancharatnam geometric phase formula for three level quantum systems *J. Phys. A: Math. Gen.* **30** 2417

Khanna G, Mukhopadhyay S, Simon R and Mukunda N 1997 *Ann. Phys., NY* **253** 55

Ali Mostafazadeh 1996 Geometric phase, bundle classification, and group representation *J. Math. Phys.* **37** 1218

Ali Mostafazadeh 1999 *J. Math. Phys.* **40** 3697

Ali Mostafazadeh 1997 Nonabelian geometrical phase for general three-dimensional quantum systems *J. Phys. A: Math. Gen.* **30** 7525

Byrd M 1999 Geometric phases for three state systems *Preprint* quant-ph/9902061

[5] Cocolicchio D and Viggiano M 1999 The exact parametrization of the neutrino mixing *Preprint* hep-ph/9906228