ANOTHER PROOF OF THE THEOREMS ON THE EIGENVALUES OF A SQUARE QUATERNION MATRIX

by YIK-HOI AU YEUNG

(Received 27 September, 1963)

1. Introduction. The nature of the eigenvalues of a square quaternion matrix had been considered by Lee [1] and Brenner [2]. In this paper the author gives another elementary proof of the theorems on the eigenvalues of a square quaternion matrix by considering the equation $Gy = \mu\bar{y}$, where G is an $n \times n$ complex matrix, y is a non-zero vector in C^n, μ is a complex number, and $\bar{y}$ is the conjugate of y. The author wishes to thank Professor Y. C. Wong for his supervision during the preparation of this paper.

2. Notations. Let R and C be the field of real numbers and the field of complex numbers respectively, and Q be the algebra of real quaternions. Then Q has a base composed of four elements e_0, e_1, e_2, e_3 whose multiplication table is given by the following formulae:

$$e_0 e_\alpha = e_\alpha e_0 = e_\alpha, \quad e_0^2 = e_0,$$

$$e_\alpha^2 = -e_0, \quad e_\alpha e_\beta = -e_\beta e_\alpha = e_\gamma,$$

where $1 \leqq \alpha, \beta, \gamma \leqq 3$, and (α, β, γ) is a cyclic permutation of (1, 2, 3). If $q \in Q$, then

$$q = a_0 e_0 + a_1 e_1 + a_2 e_2 + a_3 e_3,$$

where $a_i \in R$ $(i = 0, 1, 2, 3)$. We shall identify e_0 and e_1 with 1 and $i\ (=\sqrt{-1})$ respectively, so that we can write $q = a_0 + ia_1 + e_2(a_2 - ia_3) = \lambda + e_2\mu$, where $\lambda, \mu \in C$ (see Chevalley [3, pp. 16–17]). We define the norm of q as the real number $\sum_{i=0}^{3} a_i^2$, and the trace of q as a_0.

We regard R^n and C^n as vector spaces over R and C, respectively, and Q^n as a right vector space over Q.

3. The nature of the eigenvalues of a square quaternion matrix.

THEOREM 1. *Let $F = G_1 + e_2 G_2$ be an $n \times n$ quaternion matrix, where G_1 and G_2 are complex matrices, and let*

$$G(\lambda) \equiv \begin{pmatrix} \bar{G}_2 & G_1 - \lambda I_n \\ -\bar{G}_1 + \lambda I_n & G_2 \end{pmatrix}, \quad g(\lambda) \equiv |\, G(\lambda) \,|,$$

where the bar denotes the complex conjugate, $|\, G(\lambda) \,|$ the determinant of the matrix $G(\lambda)$, I_n the $n \times n$ identity matrix, and λ a complex variable.

(*a*) *If $\alpha + i\beta + e_2(\gamma + i\delta)$ is any eigenvalue of F, then $\alpha + ik$, where $k^2 = \beta^2 + \gamma^2 + \delta^2$, is a zero point of $g(\lambda)$.*

(*b*) *Conversely, if $\alpha + ik$ is any zero point of $g(\lambda)$, then $\alpha + i\beta + e_2(\gamma + i\delta)$, for any real numbers, β, γ and δ such that $\beta^2 + \gamma^2 + \delta^2 = k^2$, is an eigenvalue of F.*

Proof. As a first step in the proof of Theorem 1, we consider the following equation

$$Fx = xq, \tag{1}$$

where $x = x_1 + e_2 x_2 \neq 0$ with $x_1, x_2 \in C^n$ and $q = \lambda + e_2\mu$ with $\lambda, \mu \in C$. Since

$$Fx = G_1 x_1 - \bar{G}_2 x_2 + e_2(G_2 x_1 + \bar{G}_1 x_2),$$
$$xq = x_1\lambda - \bar{x}_2\mu + e_2(x_2\lambda + \bar{x}_1\mu),$$

equation (1) is equivalent to

$$\left.\begin{aligned} G_1 x_1 - \bar{G}_2 x_2 &= x_1\lambda - \bar{x}_2\mu, \\ G_2 x_1 + \bar{G}_1 x_2 &= x_2\lambda + \bar{x}_1\mu, \end{aligned}\right\} \tag{2}$$

which we can write as

$$\begin{pmatrix} G_1 - \lambda I_n & -\bar{G}_2 \\ G_2 & \bar{G}_1 - \lambda I_n \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \mu\begin{pmatrix} -\bar{x}_2 \\ \bar{x}_1 \end{pmatrix}. \tag{3}$$

But

$$\begin{pmatrix} G_1 - \lambda I_n & -\bar{G}_2 \\ G_2 & \bar{G}_1 - \lambda I_n \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} G_1 - \lambda I_n & -\bar{G}_2 \\ G_2 & \bar{G}_1 - \lambda I_n \end{pmatrix}\begin{pmatrix} O & I_n \\ -I_n & O \end{pmatrix}\begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$$
$$= \begin{pmatrix} \bar{G}_2 & G_1 - \lambda I_n \\ -\bar{G}_1 + \lambda I_n & G_2 \end{pmatrix}\begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}.$$

Therefore, equation (1) is equivalent to

$$G(\lambda)y = \mu\bar{y}, \tag{4}$$

where
$$G(\lambda) = \begin{pmatrix} \bar{G}_2 & G_1 - \lambda I_n \\ -\bar{G}_1 + \lambda I_n & G_2 \end{pmatrix}, \quad y = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix} \in C^{2n}.$$

Several lemmas are required to complete the proof of Theorem 1.

LEMMA 1. *Let U, V, W be $n \times n$ complex matrices and μ be a complex number; then*

$$\begin{vmatrix} \mu I_n & U \\ V & W \end{vmatrix} = |\, \mu W - VU \,| .$$

Proof. If $\mu = 0$, the result follows from Laplace's expansion.
If $\mu \neq 0$, then

$$\begin{vmatrix} \mu I_n & U \\ V & W \end{vmatrix} = \begin{vmatrix} I_n & O \\ -\frac{1}{\mu}V & I_n \end{vmatrix}\begin{vmatrix} \mu I_n & U \\ V & W \end{vmatrix} = \begin{vmatrix} \mu I_n & U \\ O & W - \frac{1}{\mu}VU \end{vmatrix},$$

and again Laplace's expansion yields the result.

LEMMA 2. *Let $G = H_1 + iH_2$, where H_1 and H_2 are $n \times n$ real matrices, and let*

$$h(\gamma, \delta) \equiv \begin{vmatrix} H_1 - \gamma I_n & -H_2 - \delta I_n \\ H_2 - \delta I_n & H_1 + \gamma I_n \end{vmatrix}, \quad p(t) \equiv |\, G\bar{G} - tI_n \,| ,$$

where γ, δ and t are real variables; then (γ_1, δ_1) is a zero point of $h(\gamma, \delta)$ if and only if $\gamma_1^2+\delta_1^2$ is a zero point of $p(t)$.

Proof.

$$h(\gamma, \delta) = (-1)^n \begin{vmatrix} iI_n & I_n \\ O & I_n \end{vmatrix} \begin{vmatrix} H_1-\gamma I_n & -H_2-\delta I_n \\ H_2-\delta I_n & H_1+\gamma I_n \end{vmatrix} \begin{vmatrix} iI_n & O \\ I_n & I_n \end{vmatrix}$$
$$= (-1)^n \begin{vmatrix} 2\bar{\mu} I_n & \bar{G}+\bar{\mu} I_n \\ G+\bar{\mu} I_n & H_1+\gamma I_n \end{vmatrix},$$

where $\mu = \gamma + i\delta$. By Lemma 1, we have

$$\begin{aligned} h(\gamma, \delta) &= (-1)^n \,|\, 2\bar{\mu}(H_1+\gamma I_n)-(G+\bar{\mu} I_n)(\bar{G}+\bar{\mu} I_n) \,| \\ &= (-1)^n \,|\, 2\bar{\mu} H_1 + 2\bar{\mu}\gamma I_n - G\bar{G} - 2\bar{\mu} H_1 - \bar{\mu}^2 I_n \,| \\ &= (-1)^n \,|\, \bar{\mu}(2\gamma-\bar{\mu}) I_n - G\bar{G} \,| \\ &= (-1)^n \,|\, (\gamma^2+\delta^2) I_n - G\bar{G} \,| \\ &= (-1)^{2n} \,|\, G\bar{G} - (\gamma^2+\delta^2) I_n \,| \\ &= p(\gamma^2+\delta^2). \end{aligned}$$

Thus Lemma 2 is proved.

LEMMA 3. *Let $G = H_1 + iH_2$, where H_1 and H_2 are $n \times n$ real matrices, and let $h(\gamma, \delta)$ and $p(t)$ be defined as in Lemma 2. Then the equation*

$$Gy = \mu \bar{y}, \tag{5}$$

where $y = y_1 + iy_2 \neq 0$ with $y_1, y_2 \in R^n$ and $\mu = \gamma + i\delta$ with $\gamma, \delta \in R$, is consistent if and only if $p(\gamma^2+\delta^2) = 0$.

Proof. Since

$$Gy = H_1 y_1 - H_2 y_2 + i(H_2 y_1 + H_1 y_2),$$
$$\mu \bar{y} = \gamma y_1 + \delta y_2 + i(\delta y_1 - \gamma y_2),$$

equation (5) is equivalent to

$$\left.\begin{aligned} (H_1-\gamma I_n) y_1 + (-H_2-\delta I_n) y_2 = O \\ (H_2-\delta I_n) y_1 + (H_1+\gamma I_n) y_2 = O \end{aligned}\right\}, \tag{6}$$

where y_1, y_2 are not both zero. It follows from our definition of $h(\gamma, \delta)$ that equations (6) are consistent if and only if (γ, δ) is a zero point of $h(\gamma, \delta)$. Therefore, by Lemma 2, equations (6), and hence also equation (5), are consistent if and only if $p(\gamma^2+\delta^2) = 0$. Thus Lemma 3 is proved.

LEMMA 4. *Let $G(\lambda)$ be defined as in Theorem 1 and let $p(\lambda, t) = |\, G(\lambda)\overline{G(\lambda)} - tI_{2n} \,|$, where t is a real variable. Then*

(a) Equation (4), and hence also equation (1), and $\lambda = \alpha + i\beta$, $\mu = \gamma + i\delta$ are consistent if and only if $p(\alpha+i\beta, \gamma^2+\delta^2) = 0$.

(b) $p(\alpha+i\beta, \gamma^2+\delta^2) = p(\alpha+i\beta_1, \gamma_1^2+\delta_1^2)$ for all real numbers β_1, γ_1 and δ_1 such that $\beta_1^2+\gamma_1^2+\delta_1^2 = \beta^2+\gamma^2+\delta^2$.

O

Proof. (*a*) follows directly from Lemma 3. To prove (*b*), we note that for all real α, β, γ and δ,

$$p(\alpha+i\beta, \gamma^2+\delta^2) = | G(\alpha+i\beta)\overline{G(\alpha+i\beta)}-(\gamma^2+\delta^2)I_{2n} |$$

$$= \begin{vmatrix} \bar{G}_2G_2-G_1G_1+2\alpha G_1-(a^2+\beta^2)I_n -(\gamma^2+\delta^2)I_n & \bar{G}_2\bar{G}_1+G_1\bar{G}_2-2\alpha\bar{G}_2 \\ -\bar{G}_1G_2-G_2G_1+2\alpha G_2 & -\bar{G}_1\bar{G}_1+G_2\bar{G}_2+2\alpha\bar{G}_1-(\alpha^2+\beta^2)I_n -(\gamma^2+\delta^2)I_n \end{vmatrix}$$

$$= \begin{vmatrix} \bar{G}_2G_2-G_1G_1+2\alpha G_1-\alpha^2I_n -(\beta^2+\gamma^2+\delta^2)I_n & \bar{G}_2\bar{G}_1+G_1\bar{G}_2-2\alpha\bar{G}_2 \\ -\bar{G}_1G_2-G_2G_1+2\alpha G_2 & -\bar{G}_1\bar{G}_1+G_2\bar{G}_2+2\alpha\bar{G}_1-\alpha^2I_n -(\beta^2+\gamma^2+\delta^2)I_n \end{vmatrix}$$

$$= | G(\alpha)\overline{G(\alpha)}-(\beta^2+\gamma^2+\delta^2)I_{2n} | = p(\alpha, \beta^2+\gamma^2+\delta^2).$$

From this it follows that

$$p(\alpha+i\beta, \gamma^2+\delta^2) = p(\alpha+i\beta_1, \gamma_1^2+\delta_1^2) \tag{7}$$

for all real β_1, γ_1 and δ_1 such that $\beta_1^2+\gamma_1^2+\delta_1^2 = \beta^2+\gamma^2+\delta^2$.

Thus assertion (*b*) of Lemma 4 is proved.

The proof of Theorem 1 is now immediate. It follows from the definition that

$$g(\lambda)\overline{g(\lambda)} = | G(\lambda)\overline{G(\lambda)} | = p(\lambda, 0).$$

Therefore we have, by Lemma 4,

$$g(\alpha+ik)\overline{g(\alpha+ik)} = p(\alpha+ik, 0) = p(\alpha+i\beta, \gamma^2+\delta^2), \tag{8}$$

where β, γ and δ are any real numbers such that $k^2 = \beta^2+\gamma^2+\delta^2$. If $q = \alpha+i\beta+e_2(\gamma+i\delta)$ is any eigenvalue of F, then, by Lemma 4, $p(\alpha+i\beta, \gamma^2+\delta^2) = 0$. Therefore it follows from (8) that $\alpha+ik$ is a zero point of $g(\lambda)$. Thus assertion (*a*) of Theorem 1 is proved. Conversely, if $\alpha+ik$ is any zero point of $g(\lambda)$, then it follows from (8) that $p(\alpha+i\beta, \gamma^2+\delta^2) = 0$ for any real β, γ and δ such that $\beta^2+\gamma^2+\delta^2 = k^2$. Therefore, by Lemma 4, $\alpha+i\beta+e_2(\gamma+i\delta)$ is an eigenvalue of F. Thus assertion (*b*) of Theorem 1 is proved.

COROLLARY 1. *If τ is an eigenvalue of F and q is a quaternion such that τ and q have equal norms and traces, then q is an eigenvalue of F.*

Proof. This is an immediate consequence of Theorem 1.

COROLLARY 2. *If q_1 and q_2 are two quaternions having equal norms and traces, then there exists a quaternion $\sigma \neq 0$ such that $q_2 = \sigma^{-1}q_1\sigma$.*

Proof. Take $F = q_1$; then, since $q_1 1 = 1q_1$, Corollary 2 follows from Corollary 1.

THEOREM 2. *Let F and* $g(\lambda)$ *be defined as in Theorem* 1; *then a complex number* λ *is an eigenvalue of F if and only if* λ *is a zero point of* $g(\lambda)$. *And if* τ *is an eigenvalue of F, then* $\sigma^{-1}\tau\sigma$ *is also an eigenvalue of F for all* $\sigma \neq 0$ *in Q. The class* $\sigma^{-1}\tau\sigma$ *contains just two complex numbers* (λ *and* $\bar{\lambda}$).

Proof. Since τ and $\sigma^{-1}\tau\sigma$ have equal norms and traces, by Theorem 1, Corollaries 1 and 2, Theorem 2 follows.

4. Remark. The polynomial $g(\lambda)$ defined in Theorem 1 has real coefficients. In fact, we have

$$g(\lambda) = \begin{vmatrix} \bar{G}_2 & G_1 - \lambda I_n \\ -\bar{G}_1 + \lambda I_n & G_2 \end{vmatrix} = \begin{vmatrix} O & I_n \\ -I_n & O \end{vmatrix} \begin{vmatrix} \bar{G}_2 & G_1 - \lambda I_n \\ -\bar{G}_1 + \lambda I_n & G_2 \end{vmatrix} \begin{vmatrix} O & -I_n \\ I_n & O \end{vmatrix}$$

$$= \begin{vmatrix} G_2 & \bar{G}_1 - \lambda I_n \\ -G_1 + \lambda I_n & \bar{G}_2 \end{vmatrix} = \bar{g}(\lambda).$$

REFERENCES

1. H. C. Lee, Eigenvalues and canonical forms of matrices with quaternion coefficients, *Proc. Roy. Irish Acad. Sect. A*, **52** (1949), 253–260.
2. J. L. Brenner, Matrices of quaternions, *Pacific J. Math.* **1** (1951), 329–335.
3. C. Chevalley, *Theory of Lie groups* (Princeton, 1946).

UNIVERSITY OF HONG KONG