On Hamilton's Contribution to the Cayley-Hamilton Theorem

Nicholas J. Rose

Abstract

In 1853 Hamilton showed that a general linear vector transformation in three dimensions satisfied a third-order equation. In this paper we indicate how Hamilton came to consider this transformation and what he did with it. Hamilton's work was written in the language of quaternions, which he invented. We describe this work using vectors instead of the quaternions. In addition, we express Hamilton's transformation as a matrix and show that the matrix and the transformation satisfy the same equation.

1 INTRODUCTION

The concept of a matrix as a square array associated with a linear transformation, together with the algebraic properties of matrices, was first put forth by Cayley in 1858 in his "A Memoir on the Theory of Matrices." In this Memoir Cayley stated and proved what we know as the Cayley-Hamilton theorem for 2×2 matrices ([1], pp. 24-25). Cayley mentioned that he had proved it for 3×3 matrices, but did not write out the proof. As for $n \times n$ matrices, Cayley stated "but I have not thought it necessary to undertake the labour of a formal proof of the theorem in the general case of a matrix of any degree." Cayley evidently seriously underestimated the difficulty of a proof for $n \times n$ matrices; it was not until 1878 that a proof was given by Frobenius.

A little earlier, in 1853, Hamilton, in his "Lectures on Quaternions" published a three-dimensional version of the Cayley-Hamilton Theorem. Hamilton did not use matrices nor have a characteristic equation to start with. What Hamilton did show ([2], p. 256) was that a general linear transformation $\phi : \mathcal{R}^3 \to \mathcal{R}^3$, expressed in terms of what we now call dot and cross products, satisfies a third order equation.

The details of Hamilton's work relating the Cayley-Hamilton theorem seems not to have been discussed in the literature. This may be due to the fact that the "Lectures" are written in a terse style in the relatively unfamiliar language of quaternions. Also using 'points' instead of parenthesis for grouping of terms makes for difficult reading.

To appreciate Hamilton's treatment one should see how he arrived at this transformation and what he did with it. The purpose of this paper is to describe the gist of Hamilton's work in this regard ([2] pp. 559-569) using vectors instead of quaternions. Finally we express Hamilton's linear transformation as a matrix and show that the Cayley-Hamilton theorem for the matrix jibes with the linear transformation version.

Some notation and facts about three dimensional vectors are needed. Let $\mathbf{a} = (a_1, a_2, a_3)$, $\mathbf{b} = (b_1, b_2, b_3)$, $\mathbf{c} = (c_1, c_2, c_3)$, $\mathbf{d} = (d_1, d_2, d_3)$ and $\mathbf{p} = (p_1, p_2, p_3)$ be three-dimensional real vectors. We denote the scalar product of two vectors by $\mathbf{a}\cdot\mathbf{b}$, the vector product by $\mathbf{a}\times\mathbf{b}$ and the triple scalar product (or box product) by

$$[\mathbf{a},\mathbf{b},\mathbf{c}] = \mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}.$$

The following well-known identities will be used

$$[\mathbf{a},\mathbf{b},\mathbf{c}] = [\mathbf{b},\mathbf{c},\mathbf{a}] = [\mathbf{c},\mathbf{a},\mathbf{b}] = -[\mathbf{b},\mathbf{a},\mathbf{c}] = -[\mathbf{a},\mathbf{c},\mathbf{b}] = -[\mathbf{c},\mathbf{b},\mathbf{a}],$$

the 'bac-cab' rule

$$\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = \mathbf{b}(\mathbf{a}\cdot\mathbf{c}) - \mathbf{c}(\mathbf{a}\cdot\mathbf{b}), \tag{1}$$

the Lagrange identity

$$(\mathbf{a}\times\mathbf{b})\cdot(\mathbf{c}\times\mathbf{d}) = (\mathbf{a}\cdot\mathbf{c})(\mathbf{b}\cdot\mathbf{d}) - (\mathbf{a}\cdot\mathbf{d})(\mathbf{b}\cdot\mathbf{c}), \tag{2}$$

and the expansions

$$[\mathbf{a},\mathbf{b},\mathbf{c}]\mathbf{p} = (\mathbf{b}\times\mathbf{c})(\mathbf{a}\cdot\mathbf{p}) + (\mathbf{c}\times\mathbf{a})(\mathbf{b}\cdot\mathbf{p}) + (\mathbf{a}\times\mathbf{b})(\mathbf{c}\cdot\mathbf{p}), \tag{3}$$

$$[\mathbf{a},\mathbf{b},\mathbf{c}]\mathbf{p} = [\mathbf{b},\mathbf{c},\mathbf{p}]\mathbf{a} - [\mathbf{c},\mathbf{p},\mathbf{a}]\mathbf{b} + [\mathbf{p},\mathbf{a},\mathbf{b}]\mathbf{c}. \tag{4}$$

These identities were all derived by Hamilton but stated in terms of quaternions. He also introduced 'scalar' and 'vector' (see appendix 1) into the mathematical vocabulary. For these reasons Hamilton should be considered as one of the founders of vector analysis.

2 A LINEAR VECTOR EQUATION

After Hamilton developed quaternions and investigated their fundamental properties he considers how to solve a general linear equation in quaternions. Separating out the vector part of the quaternion equation he arrives at a linear vector equation of the form

$$\sum_{i=1}^{N} \mathbf{a}_i(\mathbf{b}_i\cdot\mathbf{p}) + \beta\mathbf{p} + \mathbf{c}\times\mathbf{p} = \mathbf{f}, \tag{5}$$

where $\mathbf{a}_i, \mathbf{b}_i, \mathbf{c}, \mathbf{f}$ are given vectors, β is a real number, $\mathbf{p}$ is the unknown vector and N is a positive integer. The details of how this equation is obtained from the linear equation in quaternions we have put in appendix 1, for it is not essential to the understanding of Hamilton's work relating to the Cayley-Hamilton theorem. We may begin our discussion with the Hamilton's solution of this linear vector equation. Throughout this paper, the ingenuity and algebraic agility of Hamilton is evident.

We shall make one simplifcation, namely, we assume $\mathbf{c} = \mathbf{0}$. This reduces the algebra considerably while keeping the essence of Hamilton's work. Moreover, there is no loss of generality since the cross product term may be written as a sum of terms of the form $\mathbf{a}(\mathbf{b}\cdot\mathbf{p})$. The equation we shall consider is

$$\sum_{i=1}^{N} \mathbf{a}_i(\mathbf{b}_i\cdot\mathbf{p}) + \beta\mathbf{p} = \mathbf{f}. \tag{6}$$

Hamilton's uses a clever method of solving the linear vector equation (6). Let $\mathbf{r}, \mathbf{s}$ be vectors such that

$$\mathbf{r}\times\mathbf{s} = \mathbf{f}.$$

Taking the dot product of both sides of (6) with $\mathbf{r}$ and $\mathbf{s}$ yields

$$\sum_{i}(\mathbf{a}_i\cdot\mathbf{r})(\mathbf{b}_i\cdot\mathbf{p}) + \beta(\mathbf{p}\cdot\mathbf{r}) = \mathbf{f}\cdot\mathbf{r} = \mathbf{0}, \tag{7}$$

$$\sum_{i}(\mathbf{a}_i\cdot\mathbf{s})(\mathbf{b}_i\cdot\mathbf{p}) + \beta(\mathbf{p}\cdot\mathbf{s}) = \mathbf{f}\cdot\mathbf{s} = \mathbf{0}. \tag{8}$$

Here and in the sequel, all summations range from 1 to N, unless otherwise indicated. If we define the vectors $\mathbf{r}'$ and $\mathbf{s}'$ by

$$\mathbf{r}' = \sum_{i}(\mathbf{a}_i\cdot\mathbf{r})\mathbf{b}_i + \beta\mathbf{r}, \quad \mathbf{s}' = \sum_{i}(\mathbf{a}_i\cdot\mathbf{s})\mathbf{b}_i + \beta\mathbf{s}, \tag{9}$$

equations (7) and(8) may be written as

$$\mathbf{r}'\cdot\mathbf{p} = \mathbf{0}, \quad \mathbf{s}'\cdot\mathbf{p} = \mathbf{0}.$$

Since $\mathbf{p}$ is perpendicular to both $\mathbf{r}'$ and $\mathbf{s}'$ we must have

$$m\mathbf{p} = \mathbf{r}'\times\mathbf{s}' \tag{10}$$

for some scalar m. Using equations (9) produces

$$m\mathbf{p} = \Big(\sum_i (\mathbf{a}_i\cdot\mathbf{r})\mathbf{b}_i + \beta\mathbf{r}\Big) \times \Big(\sum_i (\mathbf{a}_i\cdot\mathbf{s})\mathbf{b}_i + \beta\mathbf{s}\Big) \qquad (11)$$

We expand this cross product and eliminate $\mathbf{r}, \mathbf{s}$ by using the fact that $\mathbf{r}\times\mathbf{s} = \mathbf{f}$. The product is the sum of three terms

$$\begin{aligned}
\text{Term } 1 &= \beta^2(\mathbf{r}\times\mathbf{s}) = \beta^2\mathbf{f}. && (12)\\
\text{Term } 2 &= \beta\sum_i \left((\mathbf{a}_i\cdot\mathbf{r})(\mathbf{b}_i\times\mathbf{s}) + (\mathbf{a}_i\cdot\mathbf{s})(\mathbf{r}\times\mathbf{b}_i)\right) \\
&= \beta\sum_i \mathbf{b}_i\times(\mathbf{s}(\mathbf{a}_i\cdot\mathbf{r}) - \mathbf{r}(\mathbf{a}_i\cdot\mathbf{s})) \\
&= -\beta\sum_i \mathbf{b}_i\times(\mathbf{a}_i\times\mathbf{f}). && (13)\\
\text{Term } 3 &= \sum_{i,j}(\mathbf{a}_i\cdot\mathbf{r})(\mathbf{a}_j\cdot\mathbf{s})(\mathbf{b}_i\times\mathbf{b}_j) \\
&= \frac{1}{2}\sum_{i,j}\Big((\mathbf{a}_i\cdot\mathbf{r})(\mathbf{a}_j\cdot\mathbf{s}) - (\mathbf{a}_j\cdot\mathbf{r})(\mathbf{a}_i\cdot\mathbf{s})\Big)(\mathbf{b}_i\times\mathbf{b}_j).
\end{aligned}$$

Using the Lagrange identity (2) yields

$$\begin{aligned}
\text{Term } 3 &= \frac{1}{2}\sum_{i,j}\big((\mathbf{a}_i\times\mathbf{a}_j)\cdot(\mathbf{r}\times\mathbf{s})\big)(\mathbf{b}_i\times\mathbf{b}_j). \\
&= \frac{1}{2}\sum_{i,j}((\mathbf{a}_i\times\mathbf{a}_j)\cdot\mathbf{f}))(\mathbf{b}_i\times\mathbf{b}_j). && (14)
\end{aligned}$$

Summing (12), (13), (14) we obtain

$$m\mathbf{p} = \beta^2\mathbf{f} - \beta\sum_i \mathbf{b}_i\times(\mathbf{a}_i\times\mathbf{f}) + \frac{1}{2}\sum_{i,j}((\mathbf{a}_i\times\mathbf{a}_j)\cdot\mathbf{f}))(\mathbf{b}_i\times\mathbf{b}_j). \qquad (15)$$

The value of m may be found by substituting the $m\mathbf{p}$ from (15) into the linear vector equation (6) (see appendix 2 for the details). The result is

$$\begin{aligned}
m =& \beta^3 + \beta^2\sum_i(\mathbf{a}_i\cdot\mathbf{b}_i) + \frac{\beta}{2}\sum_{i,j}(\mathbf{a}_i\times\mathbf{a}_j)\cdot(\mathbf{b}_i\times\mathbf{b}_j) + \\
&\sum_{i<j<k}[\mathbf{a_i},\mathbf{a_j},\mathbf{a_k}][\mathbf{b_i},\mathbf{b_j},\mathbf{b_k}]. && (16)
\end{aligned}$$

If $m \neq 0$, the solution of the linear vector equation (6) is given by (15), with the value of m given in (16).

3 THE 'CAYLEY-HAMILTON' THEOREM'

Having completed the solution of the linear vector equation, Hamilton states, almost as an afterthought, "the general linear and vector equation may also be treated as follows." Hamilton defines the linear transformation ψ by

$$\psi(\mathbf{p}) = \sum_i \mathbf{a}_i(\mathbf{b}_i\cdot\mathbf{p}) + \beta\mathbf{p}, \tag{17}$$

and reasons that the equation $\psi(\mathbf{p}) = \mathbf{f}$ can be solved if the inverse transformation ψ^{-1} can be found. Instead of dealing with the transformation ψ directly Hamilton writes

$$\psi(\mathbf{p}) = \phi(\mathbf{p}) + \beta\mathbf{p} \tag{18}$$

where

$$\phi(\mathbf{p}) = \sum_i \mathbf{a}_i(\mathbf{b}_i\cdot\mathbf{p}). \tag{19}$$

He then shows that ϕ satisfies a third order equation. Since $\psi = \phi + \beta I$, where I is the identity transformation, it follows that ψ also satisfies a third order equation. Thus ψ^{-1} can be expressed in terms of powers of ψ.

Hamilton proceeds to find a third order equation satisfied by ϕ. His method is to find suitable scalars n'', n' so that, with $\mathbf{p}', \mathbf{p}''$ defined below we have:

$$\phi(\mathbf{p}) = \mathbf{p}' - n''\mathbf{p},\ \phi(\mathbf{p}') = \mathbf{p}'' - n'\mathbf{p},\ \phi(\mathbf{p}'') = -n\mathbf{p} \tag{20}$$

The tricky part is to define n'', n' properly so that $\phi(\mathbf{p}'')$ comes out a multiple of $\mathbf{p}$. If the equations in (20) are satisfied we would have

$$\mathbf{p}' = \phi(\mathbf{p}) + n''\mathbf{p}, \tag{21}$$

$$\mathbf{p}'' = \phi^2(\mathbf{p}) + n''\phi(\mathbf{p}) + n'\mathbf{p}, \tag{22}$$

and, since $\phi(\mathbf{p}'') + n\mathbf{p} = \mathbf{0}$,

$$\phi^3(\mathbf{p}) + n''\phi^2(\mathbf{p}) + n'\phi(\mathbf{p}) + n\mathbf{p} = \mathbf{0}.$$

This means that the transformation ϕ satisfies the cubic equation

$$\phi^3 + n''\phi^2 + n'\phi + nI = O.$$

To find n'', Hamilton writes the linear vector equation (19) as

$$\phi(\mathbf{p}) = \sum_i \mathbf{b}_i \times (\mathbf{a}_i \times \mathbf{p}) + \mathbf{p}\sum_i \mathbf{b}_i\cdot\mathbf{a}_i. \tag{23}$$

This follows immediately from the bac-cab rule (1). Letting

$$\mathbf{p}' = \sum_i \mathbf{b}_i \times (\mathbf{a}_i \times \mathbf{p}), \tag{24}$$

$$n'' = -\sum_i \mathbf{b}_i \cdot \mathbf{a}_i, \tag{25}$$

(19) becomes

$$\phi(\mathbf{p}) = \mathbf{p}' - n''\mathbf{p}. \tag{26}$$

Computing $\phi(\mathbf{p}')$ from (24) yields

$$\phi(\mathbf{p}') = \sum_{i,j} (\mathbf{b}_i \times (\mathbf{a}_i \times \mathbf{a}_j))\,(\mathbf{b}_j \cdot \mathbf{p}). \tag{27}$$

Since the sum in (27) is unchanged if the indices i, j in the summand are interchanged, we find

$$\begin{aligned}
\phi(\mathbf{p}') &= \frac{1}{2}\sum_{i,j} \Big(\big(\mathbf{b}_i \times (\mathbf{a}_i \times \mathbf{a}_j)\big)(\mathbf{b}_j \cdot \mathbf{p}) - \big(\mathbf{b}_j \times (\mathbf{a}_i \times \mathbf{a}_j)\big)\,(\mathbf{b}_i \cdot \mathbf{p})\Big) \\
&= \frac{1}{2}\sum_{i,j} \big(\mathbf{b}_i(\mathbf{b}_j \cdot \mathbf{p}) - \mathbf{b}_j(\mathbf{b}_i \cdot \mathbf{p})\big) \times (\mathbf{a}_i \times \mathbf{a}_j) \\
&= \frac{1}{2}\sum_{i,j} \big(\mathbf{p} \times (\mathbf{b}_i \times \mathbf{b}_j)\big) \times (\mathbf{a}_i \times \mathbf{a}_j) \\
&= \frac{1}{2}\sum_{i,j} (\mathbf{b}_i \times \mathbf{b}_j)(\mathbf{p} \cdot (\mathbf{a}_i \times \mathbf{a}_j)) - \frac{\mathbf{p}}{2}\sum_{i,j} (\mathbf{b}_i \times \mathbf{b}_j) \cdot (\mathbf{a}_i \times \mathbf{a}_j) \\
&= \mathbf{p}'' - n'\mathbf{p},
\end{aligned} \tag{28}$$

where

$$\mathbf{p}'' = \frac{1}{2}\sum_{i,j} (\mathbf{b}_i \times \mathbf{b}_j)(\mathbf{p} \cdot (\mathbf{a}_i \times \mathbf{a}_j)), \tag{29}$$

$$\begin{aligned}
n' &= \frac{1}{2}\sum_{i,j} (\mathbf{b}_i \times \mathbf{b}_j) \cdot (\mathbf{a}_i \times \mathbf{a}_j) \\
&= \frac{1}{2}\sum_{i,j} \Big((\mathbf{b}_i \cdot \mathbf{a}_i)(\mathbf{b}_j \cdot \mathbf{a}_j) - (\mathbf{b}_i \cdot \mathbf{a}_j)(\mathbf{b}_j \cdot \mathbf{a}_j) \Big) \\
&= \sum_{i<j} \Big((\mathbf{b}_i \cdot \mathbf{a}_i)(\mathbf{b}_j \cdot \mathbf{a}_j) - (\mathbf{b}_j \cdot \mathbf{a}_i)(\mathbf{b}_i \cdot \mathbf{a}_j) \Big).
\end{aligned} \tag{30}$$

The Lagrange identity (2) was used in the penultimate step. Note that $n' = 0$ if $N = 1$.

Computing $\phi(\mathbf{p}'')$ produces

$$\phi(\mathbf{p}'') = \frac{1}{2}\sum_{i,j,k} (\mathbf{b}_i \times \mathbf{b}_j)[\mathbf{a_i}, \mathbf{a_j}, \mathbf{a_k}](\mathbf{b}_k \cdot \mathbf{p}). \tag{31}$$

In appendix 3 we show that (31) may be written as

$$\phi(\mathbf{p}'') = -n\mathbf{p}, \tag{32}$$

where

$$n = -\sum_{i<j<k} [\mathbf{a_j},\mathbf{a_j},\mathbf{a_k}][\mathbf{b_i},\mathbf{b_j},\mathbf{b_k}]. \tag{33}$$

Note that $n = 0$, if $N = 1$ or 2.

From equations (26, 28, 32) we conclude that ϕ satisfies the cubic

$$\phi^3 + n''\phi^2 + n'\phi + nI = O, \tag{34}$$

where n'', n', n are given in equations (25, 30, 33) respectively. This is Hamilton's version of the Cayley-Hamilton theorem for the linear transformation ϕ.

Before proceeding we note that the coefficient m given in (16) of section 2 may be written in terms of the coefficients n'', n', n.

$$\begin{aligned} m =&\beta^3 + \beta^2 \sum_i (\mathbf{a}_i \cdot \mathbf{b}_i) + \frac{\beta}{2} \sum_{i,j} (\mathbf{a}_i \times \mathbf{a}_j) \cdot (\mathbf{b}_i \times \mathbf{b}_j) + \\ &\sum_{i<j<k} [\mathbf{a_i},\mathbf{a_j},\mathbf{a_k}][\mathbf{b_i},\mathbf{b_j},\mathbf{b_k}] \\ =&\beta^3 - \beta^2 n'' + \beta n' - n. \end{aligned} \tag{35}$$

To find the inverse ψ^{-1}, we need the cubic equation satisfied by ψ. Suppose that

$$\psi^3 - m''\psi^2 + m'\psi - mI = \mathbf{0}. \tag{36}$$

Using the fact the $\phi = \psi - \beta I$, we find from (34) that the coefficients are

$$m = \beta^3 - n''\beta^2 + n'\beta - n, \tag{37}$$

$$m' = 3\beta^2 - 2n''\beta + n', \tag{38}$$

$$m'' = 3\beta - n''. \tag{39}$$

Observe that the constant term m is exactly the coefficient m given in (35). Thus

$$\begin{aligned} m\mathbf{p} &= m\psi^{-1}\mathbf{f} = (\psi^2 - m''\psi + m')\mathbf{f} \\ &= \Big(\phi^2 + n''\phi + n' - \beta(\phi + n'') + \beta^2\Big)\mathbf{f} \\ &= \mathbf{f}'' - \beta\mathbf{f}' + \beta^2\mathbf{f}, \end{aligned} \tag{40}$$

where

$$\mathbf{f}'' = (\phi^2 + n''\phi + n')\mathbf{f}, \quad \mathbf{f}' = (\phi + n'')\mathbf{f}. \tag{41}$$

Replacing **p** with **f** in equations (22, 29) and equations (21, 24) we obtain

$$\mathbf{f}'' = \frac{1}{2}\sum_{i,j}(\mathbf{b}_i \times \mathbf{b}_j)(\mathbf{f}\cdot(\mathbf{a}_i \times \mathbf{a}_j)), \quad \mathbf{f}' = \sum_i \mathbf{b}_i \times (\mathbf{a}_i \times \mathbf{f}). \tag{42}$$

Substituting $\mathbf{f}''$ and $\mathbf{f}'$ from (42) into (40) reproduces the solution given in (15):

$$m\mathbf{p} = \beta^2\mathbf{f} - \beta\left(\sum_i \mathbf{b}_i \times (\mathbf{a}_i \times \mathbf{f})\right) + \frac{1}{2}\sum_{i,j}(\mathbf{b}_i \times \mathbf{b}_j)(\mathbf{f}\cdot(\mathbf{a}_i \times \mathbf{a}_j)).$$

4 THE 'CHARACTERISTIC EQUATION'

Hamilton next looks at the homogeneous equation

$$\psi(\mathbf{p}) = \mathbf{0}, \text{ or } \phi(\mathbf{p}) + \beta\mathbf{p} = \mathbf{0}. \tag{43}$$

If $m \neq 0$, then ψ^{-1} exists and the solution is

$$\mathbf{p} = m^{-1}\psi^{-1}(\mathbf{0}) = \mathbf{0}.$$

However if $m = 0$, that is if β is a solution of

$$\beta^3 - n''\beta^2 + n'\beta - n = 0, \tag{44}$$

a nontrivial solution of $\phi(\mathbf{p})+\beta\mathbf{p} = \mathbf{0}$ may exist. If we call a root of (44) a 'characteristic value', we observe that the cubic equation satisfied by ϕ has the same coefficients, except for sign, as the equation satisfied by the characteristic values. Hamilton did not explicitly point out this fact. Of course, if β is a characteristic value, $-\beta$ would be an eigenvalue of ϕ and then the eigenvalues would satisfy the same equation as ϕ.

Hamilton then finds a nontrivial solution of the homogenous equation when β is a characteristic value. Note that if $m = 0$,

$$\psi\left(\psi^2 - m''\psi + m'I\right)\mathbf{g} = \mathbf{0},$$

where **g** is an arbitrary vector. Thus

$$\mathbf{p} = \left(\psi^2 - m''\psi + m'I\right)\mathbf{g}, \ \mathbf{g} \neq \mathbf{0} \tag{45}$$

furnishes a nontrivial solution of

$$\psi(\mathbf{p}) = \mathbf{0} \text{ or, } \phi(\mathbf{p}) + \beta\mathbf{p} = \mathbf{0}.$$

Using $\psi = \phi + \beta I$, and equations (37, 38, 39), we may write (45) as

$$\begin{aligned}\mathbf{p} &= \Big(\phi^2 + n''\phi + n'I - \beta(\phi + n''I) + \beta^2 I\Big)\mathbf{g} \\ &= \mathbf{g}'' - \beta\mathbf{g}' + \beta^2\mathbf{g},\end{aligned} \tag{46}$$

where $\mathbf{g}''$ and $\mathbf{g}'$ are given by equations (22, 29) and equations (21, 24) with $\mathbf{p}$ replaced by $\mathbf{g}$.

Hamilton considers the case when $\beta_1, \beta_2, \beta_3$ are distinct characteristic values. Defining

$$\mathbf{p}_i = \mathbf{g}'' - \beta_i\mathbf{g}' + \beta_i^2\mathbf{g}, \quad i = 1, 2, 3 \quad (\mathbf{g} \neq \mathbf{0}),$$

we obtain nontrivial solutions of

$$\phi(\mathbf{p}_i) = -\beta_i\mathbf{p}_i, \quad i = 1, 2, 3.$$

Hamilton clearly recognizes the importance of these equations. He notes in ([2], p. 569) "This opens up a very interesting train of research,..., respecting the principal axes of a surface of the second order, and the axis of inertia of a body, on which I cannot enter here."

We should point out that Hamilton's method of finding a nontrivial solution doesn't always work. The method fails if, in (45), we have,

$$\psi^2 - m''\psi + m'I = O,$$

or equivalently

$$\phi^2 + n''\phi + n'I - \beta(\phi + n''I) + \beta^2 I = O, \tag{47}$$

for a given characteristic value β. This occurs, for instance, if $\phi = I$. However in the case of distinct roots of the characteristic equation that Hamilton considers, it does work. Using present day knowledge we know that, in the case of distinct roots, the minimal polynomial for the transformation ϕ is

$$(\phi + \beta_1 I)(\phi + \beta_2 I)(\phi + \beta_3),$$

and is of the third degree. But if equation (47) were satisfied for, say, β_1, this would mean that the minimal polynomial would be of degree 2.

5 MATRIX FORMULATION

In this section we express the linear function ϕ as a matrix, and show that the matrix satisfies the same equation as ϕ. We write the vectors $\mathbf{a}_j$, $\mathbf{b}_j$, $\mathbf{p}$ as column vectors:

$$\mathbf{a}_j = \begin{bmatrix} a_{1,j} \\ a_{2,j} \\ a_{3,j} \end{bmatrix}, \quad \mathbf{b}_j = \begin{bmatrix} b_{1,j} \\ b_{2,j} \\ b_{3,j} \end{bmatrix}, \quad \mathbf{p} = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix}.$$

Since

$$\mathbf{b}_j \cdot \mathbf{p} = \mathbf{b}_j^T \mathbf{p},$$

we have

$$\phi(\mathbf{p}) = \sum_{j=1}^{N} \mathbf{a}_j (\mathbf{b}_j \cdot \mathbf{p}) = \Big(\sum_{j=1}^{N} \mathbf{a}_j \mathbf{b}_j^T \Big) \mathbf{p}.$$

Therefore the matrix representation of ϕ is the 3×3 matrix

$$C = \sum_{j=1}^{N} \mathbf{a}_j \mathbf{b}_j^T .$$

If we let

$$A = [\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_N] \quad \text{and} \quad B = [\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_N],$$

we find that

$$C = AB^T.$$

Let the characteristic polynomial of C be

$$c(\lambda) = \lambda^3 + c_1 \lambda^2 + c_2 \lambda + c_3.$$

By the Cayley-Hamilton theorem,

$$C^3 + c_1 C^2 + c_2 C + c_3 I = O, \tag{48}$$

where

$c_1 = -\mathrm{Trace}\, C, \quad c_2 =$ sum of the 2 by 2 principal minors of C,
$c_3 = -\det C$.

To verify that equation (48) is this is same as equation (19) satisfied by ϕ, we need to show that

$$c_1 = n'', \quad c_2 = n', \quad c_3 = n,$$

where the values of $n", n', n$ are given in equations (25, 30, 33) respectively. It is convenient to also consider the matrix $D = B^T A$, an $N \times N$ matrix whose i, j-th element is

$$d_{i,j} = \mathbf{b}_i^T \mathbf{a}_j = \mathbf{b}_i \cdot \mathbf{a}_j.$$

Since $\mathrm{Trace}\,(AB^T) = \mathrm{Trace}\,(B^T A)$ we find

$$c_1 = -\mathrm{Trace}\,(C) = -\mathrm{Trace}\, D = -\sum_i \mathbf{b}_i \cdot \mathbf{a}_i = n'',$$

using (25).

The characteristic equations of C and D are closely connected, in fact (see [4]), if the characteristic equation of D is

$$d(\lambda) = \lambda^N + d_1\lambda^{N-1} + d_2\lambda^{N-2} + \cdots + d_N,$$

then

$$\lambda^N c(\lambda) = \lambda^3 d(\lambda). \tag{49}$$

From (49) it follows that, if $N = 1$, $c_2 = d_2 = 0$, while for $N \geq 2$, $c_2 = d_2$. Thus

$$\begin{aligned} c_2 &= \text{ sum of the 2 by 2 principal minors of } C \\ &= \text{ sum of the 2 by 2 principal minors of } D \\ &= \sum_{1\leq i<j\leq N} \Big((\mathbf{b}_i\cdot\mathbf{a}_i)(\mathbf{b}_j\cdot\mathbf{a}_j) - (\mathbf{b}_j\cdot\mathbf{a}_i)(\mathbf{b}_i\cdot\mathbf{a}_j) \Big) \\ &= n', \end{aligned}$$

using (30).

Since $c_3 - \det C = -\det(AB^T)$, the Binet Cauchy Theorem ([3], p. 39) states that $c_3 = 0$, for $N = 1, 2$ while for $N \geq 3$ we have

$$\begin{aligned} c_3 &= -\det C = -\det(AB^T) \\ &= -\sum_{1\leq i<j<k\leq N} A\binom{1 \quad 2 \quad 3}{i \quad j \quad k} B^T \binom{i \quad j \quad k}{1 \quad 2 \quad 3} \\ &= -\sum_{1\leq i<j<k\leq N} A\binom{1 \quad 2 \quad 3}{i \quad j \quad k} B \binom{1 \quad 2 \quad 3}{i \quad j \quad k} \\ &= -\sum_{1\leq i<j<k\leq N} [\mathbf{a_i},\mathbf{a_j},\mathbf{a_k}]\,[\mathbf{b_i},\mathbf{b_j},\mathbf{b_k}] \\ &= n, \end{aligned}$$

where we have used (33). In this derivation the notation

$$A\binom{i \quad j \quad k}{t \quad u \quad v},$$

for example, stands for the determinant of the minor of the matrix A formed from the elements in rows $1 \leq i < j < k \leq N$ and columns $1 \leq t < u < v \leq N$ of A.

1 APPENDIX-LINEAR QUATERNION EQUATIONS

In this appendix, after a brief review of quaternions, we will outline how Hamilton arrived at the general linear vector equation (5).

Hamilton defines quaternions as a set of "hyper-complex numbers" of the form

$$\widehat{\mathbf{p}} = p + p_1 i + p_2 j + p_3 k, \tag{50}$$

where p, p_1, p_2, p_3 are real numbers. These numbers are added in the usual manner, while for multiplication the i, j, k are considered as "imaginary units", i.e.

$$i^2 = j^2 = k^2 = -1, \tag{51}$$

and multiplication of different units satisfy

$$ij = -ji = k,\ jk = -kj = i,\ ki = -ik = j. \tag{52}$$

With these definitions, the quaternions form a division algebra, that is, they obey the usual laws of algebra for addition and multiplication except for the commuative law of multiplication. Each non zero element has an inverse.

Looking at only the coefficients in (50), a quaternion may also be considered as an ordered pair

$$\widehat{\mathbf{p}} =< p,\, \mathbf{p} >,$$

where p is a real number called the 'scalar part of $\widehat{\mathbf{p}}$' and $\mathbf{p}$ is a real three dimensional vector called the 'vector part of $\widehat{\mathbf{p}}$'. Let $\widehat{\mathbf{q}} =< q,\, \mathbf{q} >$ be another quaternion, then the sum is given by

$$\widehat{\mathbf{p}} + \widehat{\mathbf{q}} =< p + q,\, \mathbf{p} + \mathbf{q} > .$$

We shall use the abbreviations

$$p =< p,\, 0 >, \quad \mathbf{p} =< 0,\, \mathbf{p} >,$$

so that we may write the quaternion as a "scalar plus a vector":

$$\widehat{\mathbf{p}} = p + \mathbf{p}.$$

Computing the product of two quaternions in the hyper-complex number form (50) using the rules (51, 52), and translating the results into component form results in

$$\widehat{\mathbf{p}}\widehat{\mathbf{q}} =< p,\, \mathbf{p} >< q,\, \mathbf{q} >=< pq - \mathbf{p}\cdot\mathbf{q},\, p\mathbf{q} + q\mathbf{p} + \mathbf{p}\times\mathbf{q} > . \tag{53}$$

It is clear that $\widehat{\mathbf{p}}\widehat{\mathbf{q}} \neq \widehat{\mathbf{q}}\widehat{\mathbf{p}}$ since the cross product does not commute. (At first glance, it might appear that the associative law of multiplication does not hold, since it doesn't hold for cross products. However, using the bac-cab rule, it can be shown that associativity holds.)

The (quaternion) product of two vectors is

$$\begin{aligned}\mathbf{pq} &= <0, \mathbf{p}><0, \mathbf{q}> = <-\mathbf{p}\cdot\mathbf{q}, \mathbf{p}\times\mathbf{q}> . \\ &= -\mathbf{p}\cdot\mathbf{q} + \mathbf{p}\times\mathbf{q}\end{aligned} \tag{54}$$

Note that the product of two vectors is not a vector but a quaternion. In computing the product of quaternions, formula (54) and the distributive law are all that is needed:

$$\begin{aligned}\widehat{\mathbf{p}}\widehat{\mathbf{q}} &= (p+\mathbf{p})(q+\mathbf{q}) = pq + q\mathbf{p} + p\mathbf{q} + \mathbf{pq} \\ &= pq - \mathbf{p}\cdot\mathbf{q} + q\mathbf{p} + p\mathbf{q} + \mathbf{p}\times\mathbf{q},\end{aligned}$$

which is equation (53).

It was certainly natural for Hamilton to consider the solution of a general linear equation in quaternions ([2], p. 559) :

$$\sum_{i=1}^{M} \widehat{\mathbf{r}}_i \widehat{\mathbf{p}}\, \widehat{\mathbf{q}}_i = \widehat{\mathbf{s}}, \tag{55}$$

where $\widehat{\mathbf{r}}_i, \widehat{\mathbf{q}}_i, \widehat{\mathbf{s}}$ are given quaternions, $\widehat{\mathbf{p}}$ is the unknown quaternion and M is a positive integer. Since quaternions do not commute, it is necessary to both pre-multiply and post-multiply the unknown quaternion $\widehat{\mathbf{p}}$. To solve this equation, Hamilton breaks it up into its scalar and vector parts. First substitute $p+\mathbf{p}$ for $\widehat{\mathbf{p}}$ to get

$$p\sum_{i=1}^{M} \widehat{\mathbf{r}}_i \widehat{\mathbf{q}}_i + \sum_{i=1}^{M} \widehat{\mathbf{r}}_i \mathbf{p}\, \widehat{\mathbf{q}}_i = \widehat{\mathbf{s}}. \tag{56}$$

Let

$$\begin{aligned}\widehat{\mathbf{d}} &= \sum_{i=1}^{M} \widehat{\mathbf{r}}_i \widehat{\mathbf{q}}_i = d + \mathbf{d}, \\ \widehat{\mathbf{e}}(\mathbf{p}) &= \sum_{i=1}^{M} \widehat{\mathbf{r}}_i \mathbf{p}\, \widehat{\mathbf{q}}_i = e(\mathbf{p}) + \mathbf{e}(\mathbf{p}), \\ \widehat{\mathbf{s}} &= s + \mathbf{s},\end{aligned}$$

where $\widehat{\mathbf{d}}$ and $\widehat{\mathbf{s}}$ are known while $\widehat{\mathbf{e}}(\mathbf{p})$ contains the unknown vector $\mathbf{p}$. Equating the scalar parts and vector parts of (56) yields

$$\begin{aligned}pd + e(\mathbf{p}) &= s, \\ p\mathbf{d} + \mathbf{e}(\mathbf{p}) &= \mathbf{s}.\end{aligned}$$

Eliminating p gives

$$d\mathbf{e}(\mathbf{p}) - e(\mathbf{p})\mathbf{d} = d\mathbf{s} - s\mathbf{d} = \mathbf{f}, \tag{57}$$

where $\mathbf{f}$ is a known vector. Let us look at

$$\begin{aligned}\widehat{\mathbf{e}}(\mathbf{p}) &= \sum_{i=1}^{M} \widehat{\mathbf{r}}_i \mathbf{p}\, \widehat{\mathbf{q}}_i. \\ &= \sum_{i=1}^{M} (r_i + \mathbf{r}_i)\mathbf{p}(q_i + \mathbf{q}_i) \qquad (58)\end{aligned}$$

Expanding (58) using the product of vectors formula (54) and the bac-cab rule (1) , we find the scalar and vector parts of $\widehat{\mathbf{e}}(\mathbf{p})$ to be

$$e(\mathbf{p}) = \Big(-\sum_{i=1}^{M} \big(r_i \mathbf{q}_i + q_i \mathbf{r}_i + \mathbf{r}_i \times \mathbf{q}_i\big)\Big)\cdot\mathbf{p} = \mathbf{g}\cdot\mathbf{p}, \qquad (59)$$

$$\begin{aligned}\mathbf{e}(\mathbf{p}) &= \Big(\sum_{i=1}^{M} \big(r_i q_i + \mathbf{q}_i\cdot\mathbf{r}_i\big)\Big)\mathbf{p} + \Big(\sum_{i=1}^{M} \big(q_i \mathbf{r}_i - r_i \mathbf{q}_i\big)\Big) \times \mathbf{p} \\ &\quad - \sum_{i=1}^{M} \Big(\mathbf{r}_i(\mathbf{q}_i\cdot\mathbf{p}) + \mathbf{q}_i(\mathbf{r}_i\cdot\mathbf{p})\Big) \\ &= \alpha\mathbf{p} + \mathbf{h}\times\mathbf{p} - \sum_{i=1}^{M} \big(\mathbf{r}_i(\mathbf{q}_i\cdot\mathbf{p}) + \mathbf{q}_i(\mathbf{r}_i\cdot\mathbf{p})\big). \qquad (60)\end{aligned}$$

where $\mathbf{g}, \mathbf{h}$ are known vectors and α is a known scalar. Subsituting into (57) gives

$$\mathbf{d}(\mathbf{g}\cdot\mathbf{p}) + d\alpha\mathbf{p} + d(\mathbf{h}\times\mathbf{p}) - \sum_{i=1}^{M} \Big(\mathbf{r}_i(\mathbf{q}_i\cdot\mathbf{p}) + \mathbf{q}_i(\mathbf{r}_i\cdot\mathbf{p})\Big) = \mathbf{f}. \qquad (61)$$

This equation is of the form

$$\sum_{i=1}^{N} \mathbf{a}_i(\mathbf{b}_i\cdot\mathbf{p}) + \beta\mathbf{p} + \mathbf{c}\times\mathbf{p} = \mathbf{f}, \qquad (62)$$

for some appropriate vectors $\mathbf{a}_i$, $\mathbf{b}_i$, $\mathbf{c}$, scalar β, and positive integer N. The quaternion $\widehat{\mathbf{p}}$ satisfies the linear equation (55), if and only if the vector part $\mathbf{p}$ satisfies (62). Hamilton ([2], p. 559) states that (62) "appears to be the most general possible form for a linear and vector equation."

2 APPENDIX:DETAILS OF SECTION 2

To evalutate m, we substitute $m\mathbf{p}$ given in (15) into the linear vector equation (6) to obtain

$$\begin{aligned} m\mathbf{f} =& \beta^3\mathbf{f} + \beta^2\left(\sum_i \mathbf{a}_i(\mathbf{b}_i\cdot\mathbf{f}) - \sum_i \mathbf{b}_i\times(\mathbf{a}_i\times\mathbf{f})\right) + \\ & \beta\left(\frac{1}{2}\sum_{i,j}[\mathbf{a_i},\mathbf{a_j},\mathbf{f}](\mathbf{b}_i\times\mathbf{b}_j) + \sum_i \mathbf{a}_i[\mathbf{b_i},\mathbf{a_j},\mathbf{f}\times\mathbf{b_j}]\right) + \\ & \frac{1}{2}\sum_{i,j,k}\mathbf{a}_i[\mathbf{b_i},\mathbf{b_j},\mathbf{b_k}][\mathbf{a_j},\mathbf{a_k},\mathbf{f}]. \end{aligned} \tag{63}$$

The coefficient of β^2 is

$$\left(\sum_i \mathbf{a}_i(\mathbf{b}_i\cdot\mathbf{f}) - \sum_i \mathbf{b}_i\times(\mathbf{a}_i\times\mathbf{f})\right) = \mathbf{f}\sum_i(\mathbf{a}_i\cdot\mathbf{b}_i). \tag{64}$$

The coefficeint of β is

$$\begin{aligned} & \frac{1}{2}\sum_{i,j}[\mathbf{a_i},\mathbf{a_j},\mathbf{f}](\mathbf{b}_i\times\mathbf{b}_j) + \sum_{i,j}\mathbf{a}_i[\mathbf{b_i},\mathbf{b_j},\mathbf{f}\times\mathbf{a_j}] \\ &= \frac{1}{2}\sum_{i,j}\Big([\mathbf{a_i},\mathbf{a_j},\mathbf{f}](\mathbf{b}_i\times\mathbf{b}_j) + \mathbf{a}_i[\mathbf{b_i},\mathbf{b_j},\mathbf{f}\times\mathbf{a_j}]\Big) + \frac{1}{2}\sum_{i,j}\mathbf{a}_i[\mathbf{b_i},\mathbf{b_j},\mathbf{f}\times\mathbf{a_j}] \\ &= \frac{1}{2}\sum_{i,j} -\mathbf{a}_i\cdot(\mathbf{f}\times\mathbf{a}_j)(\mathbf{b}_i\times\mathbf{b}_j) + \mathbf{a}_i\big(\mathbf{b}_i\times\mathbf{b}_j)\cdot(\mathbf{f}\times\mathbf{a}_j)\Big) + \frac{1}{2}\sum_{i,j}\mathbf{a}_i[\mathbf{b_i},\mathbf{b_j},\mathbf{f}\times\mathbf{a_j}] \\ &= \frac{1}{2}\sum_{i,j}(\mathbf{f}\times\mathbf{a}_j)\times\big(\mathbf{a}_i\times(\mathbf{b}_i\times\mathbf{b}_j)\big) + \frac{1}{2}\sum_{i,j}\mathbf{a}_i[\mathbf{b_i},\mathbf{b_j},\mathbf{f}\times\mathbf{a_j}] \\ &= -\frac{1}{2}\mathbf{f}\sum_{i,j}[\mathbf{a_j},\mathbf{a_i},\mathbf{b_i}\times\mathbf{b_j}] + \frac{1}{2}\sum_{i,j}\mathbf{a}_j[\mathbf{f},\mathbf{a_i},\mathbf{b_i}\times\mathbf{b_j}] + \frac{1}{2}\sum_{i,j}\mathbf{a}_i[\mathbf{b_i},\mathbf{b_j},\mathbf{f}\times\mathbf{a_j}] \\ &= +\frac{1}{2}\mathbf{f}\sum_{i,j}(\mathbf{a}_i\times\mathbf{a}_j)\cdot(\mathbf{b}_i\times\mathbf{b}_j). \end{aligned} \tag{65}$$

The constant term is

$$\frac{1}{2}\sum_{i,j,k}\mathbf{a}_i[\mathbf{b_i},\mathbf{b_j},\mathbf{b_k}][\mathbf{a_j},\mathbf{a_k},\mathbf{f}].$$

Since the sum is not changed by a cyclic permutation of the indices in the summand, we may write

$$\frac{1}{2}\sum_{i,j,k}\mathbf{a}_i[\mathbf{b_i},\mathbf{b_j},\mathbf{b_k}][\mathbf{a_j},\mathbf{a_k},\mathbf{f}]$$
$$=\frac{1}{6}\sum_{i,j,k}\Big(\mathbf{a}_i[\mathbf{b_i},\mathbf{b_j},\mathbf{b_k}][\mathbf{a_j},\mathbf{a_k},\mathbf{f}]+\mathbf{a}_j[\mathbf{b_j},\mathbf{b_k},\mathbf{b_i}][\mathbf{a_k},\mathbf{a_i},\mathbf{f}]+\mathbf{a}_k[\mathbf{b_k},\mathbf{b_i},\mathbf{b_j}][\mathbf{a_i},\mathbf{a_j},\mathbf{f}]$$
$$=\frac{1}{6}\sum_{i,j,k}[\mathbf{b_i},\mathbf{b_j},\mathbf{b_k}]\Big(\mathbf{a}_i[\mathbf{a_j},\mathbf{a_k},\mathbf{f}]+\mathbf{a}_j[\mathbf{a_k},\mathbf{a_i},\mathbf{f}]+\mathbf{a}_k[\mathbf{a_j},\mathbf{a_k},\mathbf{f}]\Big)$$
$$=\frac{\mathbf{f}}{6}\sum_{i,j,k}[\mathbf{b_i},\mathbf{b_j},\mathbf{b_k}][\mathbf{a_i},\mathbf{a_j},\mathbf{a_k}]$$
$$=\mathbf{f}\sum_{i<j<k}[\mathbf{b_i},\mathbf{b_j},\mathbf{b_k}][\mathbf{a_i},\mathbf{a_j},\mathbf{a_k}], \qquad (66)$$

where we have used expansion (4) in the penultimate step. Substituting equations (64), (65) and (66) into (63) we find the value of m to be

$$m=\beta^3+\beta^2\sum_i(\mathbf{a}_i\cdot\mathbf{b}_i)+\frac{\beta}{2}\sum_{i,j}(\mathbf{a}_i\times\mathbf{a}_j)\cdot(\mathbf{b}_i\times\mathbf{b}_j)+$$
$$\sum_{i<j<k}[\mathbf{a_i},\mathbf{a_j},\mathbf{a_k}][\mathbf{b_i},\mathbf{b_j},\mathbf{b_k}].$$

as given in (16).

3 APPENDIX-DETAILS OF SECTION3

We start with the expression for $\phi(\mathbf{p}'')$ given in (31)

$$\phi(\mathbf{p}'')=\frac{1}{2}\sum_{i,j,k}(\mathbf{b}_i\times\mathbf{b}_j)[\mathbf{a_i},\mathbf{a_j},\mathbf{a_k}](\mathbf{b}_k\cdot\mathbf{p}).$$

Using arguments similar to the previous appendix and the expansion (3) we find

$$\begin{aligned}
\phi(\mathbf{p}'') =& \frac{1}{6}\sum_{i,j,k}(\mathbf{b}_i \times \mathbf{b}_j)[\mathbf{a_j},\mathbf{a_j},\mathbf{a_k}](\mathbf{b}_k\cdot\mathbf{p})+ \\
& \frac{1}{6}\sum_{i,j,k}(\mathbf{b}_k \times \mathbf{b}_i)[\mathbf{a_k},\mathbf{a_j},\mathbf{a_i}](\mathbf{b}_j\cdot\mathbf{p}) + \frac{1}{6}\sum_{i,j,k}(\mathbf{b}_j \times \mathbf{b}_k)[\mathbf{a_j},\mathbf{a_k},\mathbf{a_i}](\mathbf{b}_i\cdot\mathbf{p}) \\
& = \frac{1}{6}\sum_{i,j,k}[\mathbf{a_i},\mathbf{a_j},\mathbf{a_k}]\left((\mathbf{b}_i \times \mathbf{b}_j)(\mathbf{b}_k\cdot\mathbf{p}) + (\mathbf{b}_k \times \mathbf{b}_i)(\mathbf{b}_j\cdot\mathbf{p}) + (\mathbf{b}_j \times \mathbf{b}_k)(\mathbf{a}_i\cdot\mathbf{p})\right) \\
& = \frac{\mathbf{p}}{6}\sum_{i,j,k}[\mathbf{a_i},\mathbf{a_j},\mathbf{a_k}][\mathbf{b_i},\mathbf{b_j},\mathbf{b_k}] \\
& = \mathbf{p}\sum_{i<j<k}[\mathbf{a_i},\mathbf{a_j},\mathbf{a_k}][\mathbf{b_i},\mathbf{b_j},\mathbf{b_k}].
\end{aligned}$$

Therefore

$$n = -\sum_{i<j<k}[\mathbf{a_i},\mathbf{a_j},\mathbf{a_k}][\mathbf{b_i},\mathbf{b_j},\mathbf{b_k}],$$

as claimed in (33).

References

[1] Arthur Cayley. A memoir on the theory of matrices. *Philosophical Transactions of the Royal Society of London*, 148:17–37, 1858.

[2] William Rowan Hamilton. *Lectures on Quaternions.* Addison-Wesley, Cambridge, MA, 1853.

[3] Peter Lancaster and Miron Tismenetsky. *The Theory of Matrices: with Applications.* Academic Press, San Diego, CA, 1985.

[4] J. Schmid. A remark on characteristic polynomials. *Am. Math. Monthly*, 77:998–999, 1970.