Atti Accademia Nazionale dei Lincei

Classe Scienze Fisiche Matematiche Naturali

Rendiconti

Luiz C. Martins

An analytical approach to Cayley-Hamilton theorem

Atti della Accademia Nazionale dei Lincei. Classe di Scienze Fisiche, Matematiche e Naturali. Rendiconti, Serie 8, Vol. **81** (1987), n.3, p. 279–282.

Accademia Nazionale dei Lincei

<http://www.bdim.eu/item?id=RLINA_1987_8_81_3_279_0>

Atti della Accademia Nazionale dei Lincei. Classe di Scienze Fisiche, Matematiche e Naturali. Rendiconti, Accademia Nazionale dei Lincei, 1987.

Atti Acc. Lincei Rend. fis.
(8), LXXXI (1987), pp. 279-282

Analisi matematica. — *An analytical approach to Cayley-Hamilton theorem.* Nota di LUIZ C. MARTINS, presentata (*) dal Socio G. GRIOLI.

ABSTRACT. — Cayley-Hamilton theorem is proved by an analytical approach by recalling certain interesting properties of density. In this process, the classical expressions of the principal invariants follow immediately from the proposed proof's scheme.

KEY WORDS: Linear Algebra; Principal invariants; Characteristic polynomial.

RIASSUNTO. — *Un approccio analitico al teorema di Cayley-Hamilton.* Il teorema di Cayley-Hamilton viene dimostrato usando un approccio analitico che fa uso di certe interessanti proprietà di densità. Nel corso della dimostrazione si ottengono in modo immediato le espressioni classiche per gli invarianti principali.

1. INTRODUCTION

There are instances in Continuum Physics where Linear Algebra and Real Analysis blend nicely and the Cayley-Hamilton Theorem is a key tool. For this well-known theorem we offer here a new proof which compares with the usual algebraic proof given, e.g., in Gel'fand [1]; as a bonus, we get the explicit expressions of the principal invariants. In this note, the case of a three-dimensional real vector space is considered; a generalized version of the proof is easy to arrive at along the same line of reasoning.

2. NOTATION

Let V be a three-dimensional vector space over the reals, and let the inner product of $\boldsymbol{u}, \boldsymbol{v} \in V$ be denoted by $\boldsymbol{u} \cdot \boldsymbol{v}$. Moreover, let $\mathbf{A}$ be a linear transformation of V into itself, and let $\det \mathbf{A}$ denote the determinant of $\mathbf{A}$ and $\operatorname{tr} \mathbf{A}$ its trace. Recall that the set $\operatorname{Lin}(V)$ of all linear transformations can be made into an inner product space by setting $\mathbf{A} \cdot \mathbf{B} := \operatorname{tr}(\mathbf{A}\mathbf{B}^{\mathrm{T}})$, where $\mathbf{B}^{\mathrm{T}}$ is the transpose of $\mathbf{B}$ (so that, in particular, $\operatorname{tr} \mathbf{A} = \mathbf{A} \cdot \mathbf{I}$, with $\mathbf{I}$ the identity transformation). Recall also that, if $t \mapsto \mathbf{F}(t)$ is a smooth curve in $\operatorname{Lin}(V)$ with

(*) Nella seduta del 24 aprile 1987.

invertible values, then

$$(\det \mathbf{F})' = (\det \mathbf{F}) \operatorname{tr} (\mathbf{F}^{-1} \mathbf{F}'), \tag{2.1}$$

where a superscript dot denotes differentiation with respect to the parameter t. Finally, let $P(\omega)$ denote the characteristic polynomial of $\mathbf{A}$:

$$P(\omega) = \det (\mathbf{A} - \omega \mathbf{I}) = - \omega^3 + i_1 \omega^2 - i_2 \omega + i_3, \tag{2.2}$$

where i_1, i_2 and i_3 are the principal invariants of $\mathbf{A}$.

3. Cayley-Hamilton Theorem

The theorem states that each $\mathbf{A} \in \operatorname{Lin}(V)$ satisfies its own characteristic equation, i.e.,

$$P(\mathbf{A}) = - \mathbf{A}^3 + i_1 \mathbf{A}^2 - i_2 \mathbf{A} + i_3 \mathbf{I} = 0. \tag{3.1}$$

Proof. Let $\mathbf{A} \in \operatorname{Lin}(V)$ be chosen, and let ω be sufficiently large for $(\mathbf{A} - \omega \mathrm{I})$ to be invertible. First, differentiate the characteristic polynomial (2.2) with respect to ω, taking into account also (2.1), to get:

$$- 3 \omega^2 + 2 i_1 \omega - i_2 = - (- \omega^3 + i_1 \omega^2 - i_2 \omega + i_3) \operatorname{tr} (\mathbf{A} - \omega \mathbf{I})^{-1}. \tag{3.2}$$

Then notice that, as

$$(\mathbf{A} - \omega \mathbf{I})^{-1} = - \omega^{-1} (\mathbf{I} - \omega^{-1} \mathbf{A})^{-1}, \tag{3.3}$$

equation (3.2) can be given the form

$$- 3 \omega^3 + 2 i_1 \omega^2 - i_2 \omega = (- \omega^3 + i_1 \omega^2 - i_2 \omega + i_3) \operatorname{tr} (\mathbf{I} - \omega^{-1} \mathbf{A})^{-1}. \tag{3.4}$$

Next, on recalling von Neumann series expansion

$$(\mathbf{I} - \mathbf{B})^{-1} = \mathbf{I} + \mathbf{B} + \mathbf{B}^2 + \mathbf{B}^3 + \ldots \tag{3.5}$$

for $\mathbf{B} \in \operatorname{Lin}(V)$ of sufficiently small norm, proceed to collect equal powers of ω in (3.4) so as to get:

$$\begin{aligned}
2 i_1 &= 3 i_1 - \operatorname{tr} \mathbf{A}, \\
- i_2 &= - 3 i_2 + i_1 \operatorname{tr} \mathbf{A} - \operatorname{tr} \mathbf{A}^2, \\
0 &= - \operatorname{tr} \mathbf{A}^3 + i_1 \operatorname{tr} \mathbf{A}^2 - i_2 \operatorname{tr} \mathbf{A} + 3 i_3, \\
0 &= - \operatorname{tr} \mathbf{A}^4 + i_1 \operatorname{tr} \mathbf{A}^3 - i_2 \operatorname{tr} \mathbf{A}^2 + i_3 \operatorname{tr} \mathbf{A}, \\
0 &= - \operatorname{tr} \mathbf{A}^5 + i_1 \operatorname{tr} \mathbf{A}^4 - i_2 \operatorname{tr} \mathbf{A}^3 + i_3 \operatorname{tr} \mathbf{A}^2, \\
& \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots , \\
& \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots .
\end{aligned} \tag{3.6}$$

The first two equations (3.6) yield well-known formulas for the invariants i_1 and i_2 of $\mathbf{A}$ in terms of the traces of $\mathbf{A}$ and $\mathbf{A}^2$:

$$i_1 = \operatorname{tr} \mathbf{A} \quad , \quad i_2 = \frac{1}{2}\left((\operatorname{tr} \mathbf{A})^2 - \operatorname{tr} \mathbf{A}^2\right) \tag{3.7}$$

the remaining ones can be written as

$$\begin{aligned} 0 &= \mathbf{P}(\mathbf{A}) \cdot \mathbf{I}\,, \\ 0 &= \mathbf{P}(\mathbf{A}) \cdot \mathbf{A}^{\mathrm{T}}\,, \\ 0 &= \mathbf{P}(\mathbf{A}) \cdot (\mathbf{A}^{\mathrm{T}})^2\,, \\ &\ldots\ldots\ldots\ldots\ldots, \\ &\ldots\ldots\ldots\ldots\ldots \end{aligned} \tag{3.8}$$

The two lemmata proved in the next section, which are interesting for their own sake, imply that, for all $\mathbf{A}$ of a dense, open subset of Lin (V), if $\mathbf{P}(\mathbf{A})$ is orthogonal to the span of $\mathbf{I}$, $\mathbf{A}^{\mathrm{T}}$, $(\mathbf{A}^{\mathrm{T}})^2$, ... then $\mathbf{P}(\mathbf{A}) = 0$ [1]. On extending by continuity the validity of this result to all of Lin (V), the proof of Cayley-Hamilton theorem is completed. □

4. Two Lemmata

The three first equations (3.8) can be given the form:

$$\begin{vmatrix} \mathbf{I}\cdot\mathbf{I} & \mathbf{A}\cdot\mathbf{I} & \mathbf{A}^2\cdot\mathbf{I} \\ \mathbf{I}\cdot\mathbf{A}^{\mathrm{T}} & \mathbf{A}\cdot\mathbf{A}^{\mathrm{T}} & \mathbf{A}^2\cdot\mathbf{A}^{\mathrm{T}} \\ \mathbf{I}\cdot(\mathbf{A}^{\mathrm{T}})^2 & \mathbf{A}\cdot(\mathbf{A}^{\mathrm{T}})^2 & \mathbf{A}^2\cdot(\mathbf{A}^{\mathrm{T}})^2 \end{vmatrix} \begin{vmatrix} i_3 \\ -i_2 \\ i_1 \end{vmatrix} = \begin{vmatrix} \operatorname{tr}\mathbf{A}^3 \\ \operatorname{tr}\mathbf{A}^4 \\ \operatorname{tr}\mathbf{A}^5 \end{vmatrix}, \tag{4.1}$$

where the symmetric square matrix on the left-hand side, which shall be denoted by $\mathbf{H}(\mathbf{A})$, reminds us of Gram matrix. We shall now show that, for all $\mathbf{A}$ in a dense open set of Lin (V), (i) $\mathbf{H}(\mathbf{A})$ is invertible; (ii) there exists a unique triplet (α, β, γ) of real numbers such that $\mathbf{A}^3$ admits the representation

$$\mathbf{A}^3 = \alpha\,\mathbf{A}^2 + \beta\,\mathbf{A} + \gamma\,\mathbf{I}; \tag{4.2}$$

(1) Notice that, if the polynomial $\mathbf{Q}(\mathbf{A})$ is not the characteristic polynomial, this result does not hold true in general. Indeed, for V two-dimensional, picking

$$|\mathbf{A}| = \begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix} \quad \text{and} \quad \mathbf{Q}(\mathbf{A}) = \mathbf{A} - \mathbf{I}$$

we have a counter example.

consequently, (iii) $(\gamma, \beta, \alpha) = (i_3, -i_2, i_1)$, i.e., $\mathbf{A}$ satisfies its characteristic equation in a dense open set of $\mathrm{Lin}(V)$.

LEMMA 1. *For all* $\mathbf{A}$ *in a dense open set of* $\mathrm{Lin}(V)$, $\det \mathrm{H}(\mathbf{A}) \neq 0$.

Proof. Had the smooth mapping $\mathbf{A} \mapsto \det \mathbf{H}(\mathbf{A})$ identically value zero over a non-empty open subset of $\mathrm{Lin}(V)$, it would have value zero over the whole of $\mathrm{Lin}(V)$. Therefore, in order to reach the desired conclusion, it is enough to exhibit an element $\mathbf{A}$ of $\mathrm{Lin}(V)$ for which $\det \mathbf{H}(\mathbf{A}) \neq 0$; but this is the case for any symmetric $\mathbf{A}$ having distinct proper numbers. □

LEMMA 2. *For all* $\mathbf{A}$ *in a dense open set of* $\mathrm{Lin}(V)$, $\mathbf{A}^3$ *can be uniquely written as a linear combination of* $\mathbf{I}$, $\mathbf{A}$ *and* $\mathbf{A}^2$.

Proof. For $\boldsymbol{v} \in \mathrm{V}$ a non-null vector, and for $\mathbf{A} \in \mathrm{Lin}(\mathrm{V})$, consider the triplet $(\boldsymbol{v}, \mathbf{A}\boldsymbol{v}, \mathbf{A}^2\boldsymbol{v})$. Were this triplet linearly dependent over a non-empty open subset of $\mathrm{Lin}(V)$, the matrix whose columns are $\boldsymbol{v}$, $\mathbf{A}\boldsymbol{v}$ and $\mathbf{A}^2\boldsymbol{v}$ would have zero determinant for such $\mathbf{A}$'s. Moreover, this would imply that the triplet $(\boldsymbol{v}, \mathbf{A}\boldsymbol{v}, \mathbf{A}^2\boldsymbol{v})$ is linearly dependent over the whole of $\mathrm{Lin}(V)$, a contradiction. Let now $(\boldsymbol{v}, \mathbf{A}\boldsymbol{v}, \mathbf{A}^2\boldsymbol{v})$ be linearly independent. Then,

$$\mathbf{A}^3\boldsymbol{v} = \alpha\,\mathbf{A}^2\boldsymbol{v} + \beta\,\mathbf{A}\boldsymbol{v} + \gamma\boldsymbol{v}, \tag{4.3}$$

or rather,

$$\tilde{\mathbf{P}}(\mathbf{A})\,\boldsymbol{v} = (\mathbf{A}^3 - \alpha\,\mathbf{A}^2 - \beta\,\mathbf{A} - \gamma\,\mathbf{I})\,\boldsymbol{v} = \mathbf{0}. \tag{4.4}$$

It is the matter of a straightforward computation to check that

$$\tilde{\mathbf{P}}(\mathbf{A})\,\mathbf{A}\boldsymbol{v} = \mathbf{A}\tilde{\mathbf{P}}(\mathbf{A})\,\boldsymbol{v} = \mathbf{0}, \quad \tilde{\mathbf{P}}(\mathbf{A})\,\mathbf{A}^2\boldsymbol{v} = \mathbf{A}^2\tilde{\mathbf{P}}(\mathbf{A})\,\boldsymbol{v} = \mathbf{0}. \tag{4.5}$$

(4.4) and (4.5) together show that $\tilde{\mathbf{P}}(\mathbf{A})$ maps the basis $(\boldsymbol{v}, \mathbf{A}\boldsymbol{v}, \mathbf{A}^2\boldsymbol{v})$ of V into the null element of V. Thus, $\tilde{\mathbf{P}}(\mathbf{A}) = \mathbf{0}$. □

REFERENCES

[1] I.M. GEL'FAND (1961) - *Lectures in Linear Algebra*, John Wiley, New York.