Proof of Cayley-Hamilton theorem using polynomials over the algebra of module endomorphisms

Alexey Muranov

Version of Record: https://www.sciencedirect.com/science/article/pii/S0024379522001136
Manuscript_88a2214b64e67fcae84e7134af6f0a06

PROOF OF CAYLEY-HAMILTON THEOREM USING POLYNOMIALS OVER THE ALGEBRA OF MODULE ENDOMORPHISMS

ALEXEY MURANOV

Abstract. If R is a commutative unital ring and M is a unital R-module, then each element of $\mathrm{End}_R(M)$ determines a left $\mathrm{End}_R(M)[X]$-module structure on $\mathrm{End}_R(M)$, where $\mathrm{End}_R(M)$ is the R-algebra of endomorphisms of M and $\mathrm{End}_R(M)[X] = \mathrm{End}_R(M) \otimes_R R[X]$. These structures provide a very short proof of the Cayley-Hamilton theorem, which may be viewed as a reformulation of the proof in *Algebra* by Serge Lang. Some generalisations of the Cayley-Hamilton theorem can be easily proved using the proposed method.

1. Introduction

Theorem (Cayley-Hamilton theorem)**.** *Let R be a commutative unital ring (i.e., with 1_R) and M be a finite-rank free unital R-module (i.e., which respects 1_R). Let $a\colon M \to M$ be an endomorphism of M and $\chi_a \in R[X]$ be the characteristic polynomial of a. Then $\chi_a(a) = 0$ (in $R[a] \subset \mathrm{End}_R(M)$).*

The goal of this note is to show how basic properties of tensor products provide a very short proof of this theorem and allow to generalise it.

The two main ingredients of the proof presented in this note are:

(1) the canonical isomorphism $\mathrm{End}_R(M)[X] \cong \mathrm{End}_{R[X]}(M[X])$,
(2) certain left actions of $\mathrm{End}_R(M)[X]$ on $\mathrm{End}_R(M)$ associated to elements of $\mathrm{End}_R(M)$.

The presented proof is essentially a reformulation of the one in *Algebra* by Serge Lang[1] [3], eliminating the need to work with bases or with matrices explicitly. However, the author is unaware of the considered actions' of $\mathrm{End}_R(M)[X]$ on $\mathrm{End}_R(M)$ having been used in the literature before to prove the Cayley-Hamilton theorem or generalisations thereof.

The proof by Bourbaki in *Algèbre* [1] is essentially different and more involved. Not only they work with matrices explicitly, but they also need to prove the identity $\tilde{a}a = a\tilde{a}$.

Date: January 16, 2022.
2020 *Mathematics Subject Classification.* Primary 13C10, 15A15; Secondary 13-03.
Key words and phrases. Cayley-Hamilton theorem, determinant, commutative ring, free module, tensor product.

[1]The proof found in the current version of *characteristic polynomial* web page of *nLab* wiki [4] follows the one from the Lang's book.

2. Basic definitions and properties

Necessary definitions and basic properties of modules, their tensor products, and their exterior powers may be found, for example, in expository papers by Keith Conrad [2].

Let R be a commutative unital ring and M be a free unital R-module of finite rank n.

The following usual notation shall be used: $R[X]$ is the ring of polynomials in X over R, $M[X] = M \otimes_R R[X]$, $\mathrm{End}_R(M)[X] = \mathrm{End}_R(M) \otimes_R R[X]$.

Following a common practice, elements of R may be viewed as elements of $R[X]$ or as elements of $\mathrm{End}_R(M)$ (as scalar endomorphisms), elements of M may be viewed as elements of $M[X]$, etc.

Since M is free of finite rank, there is a canonical isomorphism

$$\mathrm{End}_R(M)[X] \cong \mathrm{End}_{R[X]}(M[X]).$$

Using this isomorphism, elements of $\mathrm{End}_R(M)[X]$ may be viewed as elements of $\mathrm{End}_{R[X]}(M[X])$ and vice versa.

For an endomorphism a of M, the *determinant* of a is defined by the identity

$$ax_1 \wedge \cdots \wedge ax_n = (\det a)(x_1 \wedge \cdots \wedge x_n) \qquad (x_1, \ldots, x_n \in M).$$

The *adjugate endomorphism* $\tilde{a}$ of a is defined by the identity

$$ax_1 \wedge \cdots \wedge ax_{n-1} \wedge y = x_1 \wedge \cdots \wedge x_{n-1} \wedge \tilde{a}y \qquad (x_1, \ldots, x_{n-1}, y \in M).$$

Replacing y with ax_n in the last identity, it can be deduced that

$$\tilde{a}a = (\det a)\mathrm{id}_M = \det a$$

(identifying scalar endomorphisms of M with elements of R).

The *characteristic polynomial* of $a \in \mathrm{End}_R(M)$ is the polynomial $\chi_a \in R[X]$ defined as[2]

$$\chi_a = \det(a - X)$$

(where $a - X \in \mathrm{End}_{R[X]}(M[X]) \cong \mathrm{End}_R(M)[X]$).

It is not hard to show that the degree of χ_a is n, and that its leading coefficient is $(-1)^n$. These facts shall not be used in this note however.

Denote

$$t_a = a - X.$$

Then

$$\chi_a = \det t_a = \tilde{t_a} t_a.$$

3. Proof of Cayley-Hamilton theorem through an action of $\mathrm{End}_R(M)[X]$ on $\mathrm{End}_R(M)$

Given $a \in \mathrm{End}_R(M)$, consider the left action of $\mathrm{End}_R(M)[X]$ on the R-module $\mathrm{End}_R(M)$ (forgetting its algebra structure) denoted by the binary operator "$\triangleleft_a$" and defined by the rules:

$$f \triangleleft_a g = fg \quad \text{for} \quad f \in \mathrm{End}_R(M), \quad \text{and} \quad X \triangleleft_a g = ga,$$

[2]The term *characteristic polynomial* is alternatively (and possibly more commonly) used to denote $\det(X - a)$.

where g is an arbitrary element of $\mathrm{End}_R(M)$ acted upon. Thus, if

$$p = f_0 + f_1 X + \cdots + f_k X^k \in \mathrm{End}_R(M)[X] \quad \text{and} \quad g \in \mathrm{End}_R(M),$$

then

$$p \lhd_a g = f_0 g + f_1 g a + \cdots + f_k g a^k \in \mathrm{End}_R(M),$$

and, in particular,[3]

$$p \lhd_a \mathrm{id}_M = f_0 + f_1 a + \cdots + f_k a^k \in \mathrm{End}_R(M).$$

Thus,

$$(a - X) \lhd_a \mathrm{id}_M = a - a = 0 \quad \text{and} \quad \chi_a \lhd_a \mathrm{id}_M = \chi_a(a).$$

Proof of Cayley-Hamilton theorem.

$$\chi_a(a) = \chi_a \lhd_a \mathrm{id}_M = (\tilde{t_a} t_a) \lhd_a \mathrm{id}_M = \tilde{t_a} \lhd_a (t_a \lhd_a \mathrm{id}_M) = \tilde{t_a} \lhd_a 0 = 0. \quad \square$$

4. Generalisation

The method used above to prove the Cayley-Hamilton theorem allows to prove seemingly more general statements, such as the following one.

Proposition. *Let R and M be as before. Let $f_1, \ldots, f_n, a_1, \ldots, a_n$ be endomorphisms of M such that*:

(1) $f_1 a_1 + \cdots + f_n a_n = 0$,
(2) $a_1, \ldots, a_n$ *commute pairwise.*

Let

$$p = f_1 X_1 + \cdots + f_n X_n \in \mathrm{End}_{R[X_1,\ldots,X_n]}(M[X_1, \ldots, X_n])$$

and

$$P = \det p \in R[X_1, \ldots, X_n].$$

Then

$$P(a_1, \ldots, a_n) = 0.$$

Example. Let a and b be two commuting endomorphisms of M, and let

$$p = bX - aY \in \mathrm{End}_{R[X,Y]}(M[X, Y]).$$

Then substituting a for X and b for Y in $P = \det p$ yields 0:

$$P(a, b) = 0.$$

5. Declaration of competing interests

There are none.

6. Acknowledgement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

[3]Since $\mathrm{End}_R(M)$ in general is not commutative, the element $f_0 + f_1 a + \cdots + f_k a^k$ of $\mathrm{End}_R(M)$ should not be viewed as the result of "substitution" of a for X in p. It may be viewed though as the result of *right substitution*, and $f_0 + a f_1 + \cdots + a^k f_k$ may be viewed as the result of *left substitution*.

References

[1] Nicolas Bourbaki. *Algèbre. Chapitres 1 à 3*. French. Reprint of the 1970 original. Berlin: Springer, 2007, pp. xiii + 636. ISBN: 3-540-33849-7/pbk. URL: https://www.springer.com/gp/book/9783540338499.

[2] Keith Conrad. *Expository papers*. URL: https://kconrad.math.uconn.edu/blurbs/ (visited on 04/06/2021).

[3] Serge Lang. *Algebra. Volume 1*. Revised third edition. Vol. 211. Graduate Texts in Mathematics. New York, NY: Springer, 2002, pp. xv + 918. ISBN: 0-387-95385-X/hbk. URL: https://www.springer.com/gp/book/9780387953854.

[4] *nLab*. A wiki devoted to Mathematics, Physics, and Philosophy. URL: https://ncatlab.org/ (visited on 04/06/2021).

I2M, UMR 7373, CNRS, Université d'Aix-Marseille, Marseille, France
Email address: alexey.muranov@univ-amu.fr