The Cayley-Hamilton Theorem

A short rigidity argument.

Heinrich Hartmann*

doi:10.5281/zenodo.18016109

Abstract

We give a short real-variable proof of the Cayley–Hamilton theorem using a local-to-global rigidity argument: prove the identity on a neighbourhood of a diagonal matrix with simple spectrum and extend to all matrices using polynomiality.

1 Introduction

The Cayley–Hamilton theorem states that every square matrix satisfies its own characteristic polynomial: For $A \in \mathbb{R}^{n \times n}$ the characteristic polynomial is defined as:

$$\chi_A(t) := \det(tI - A) \in \mathbb{R}[t],$$

Inserting the matrix A for t one finds $\chi_A(A) = 0$. Historically, a quaternionic special case was obtained by Hamilton [Ham53]in 1853, Cayley gave the matrix formulation [Cay58] in 1858; a fully general proof was given by Frobenius [Fro78] in 1878.

There are many standard proofs. Over $\mathbb{C}$, one often reduces to Jordan normal form, or proves the statement first for diagonalizable matrices and then extends by continuity; see for example [HJ13]. Textbook treatments also give adjugate-matrix proofs, but these require some care: the theorem asserts a matrix identity $\chi_A(A) = 0$, and it is not legitimate to argue by the bogus substitution $\chi_A(A) = \det(AI - A) = 0$; see [Hig20].

In this note we give a short proof that stays entirely over $\mathbb{R}$ and avoids complex canonical forms and density arguments. The core idea is local-to-global rigidity. We fix a diagonal matrix $D_0 = \mathrm{diag}(1, \ldots, n)$ with simple real spectrum. By an implicit-function-theorem argument, matrices in a neighbourhood of D_0 continue to have n distinct real eigenvalues and are therefore diagonalizable. On this open neighbourhood the identity $\chi_A(A) = 0$ follows immediately by conjugating to a diagonal matrix. Finally, we observe that the entries of the map $A \mapsto \chi_A(A)$ are polynomial functions of the entries of A; hence vanishing on a nonempty open set forces vanishing everywhere on $\mathbb{R}^{n \times n}$.

This framing as rigidity argument has the benefit of keeping the proof "low-tech'': beyond a basic perturbation lemma for simple eigenvalues and an elementary polynomial identity principle, no structure theory for linear maps is required. The result is a proof that is short, intuitive, and pedagogically robust.

2 Main Theorem

(1) Theorem (Cayley-Hamilton). *Let $A \in \mathbb{R}^{n \times n}$ and let $\chi_A(\lambda) = \det(\lambda I - A)$ be its characteristic polynomial. Then*

$$\chi_A(A) = 0,$$

where $\chi_A(A)$ denotes the matrix polynomial obtained by evaluating χ_A at A.

*Hartmann IT GmbH / HeinrichHartmann.com / Heinrich@HeinrichHartmann.com

Proof. The characteristic polynomial is a degree–n polynomial in the entries $A_{i,j}$ of A. The evaluation $\chi_A(A)$ is an $n \times n$ matrix, whose entries $F(A)_{i,j}$ are again polynomials in the $A_{i,j}$. We regard this construction as a polynomial map $F : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n},\ A \mapsto \chi_A(A)$. We want to show that this polynomial map is identically zero: $F \equiv 0$.

To do so it is sufficient to show that there is an open subset in the Euclidean topology where F vanishes identically.

Let $D_0 = \operatorname{diag}(1, 2, \ldots, n)$ be the diagonal matrix with n distinct eigenvalues $1, \ldots, n$. There exists a neighborhood U of D_0 where all matrices $A \in U$ have n distinct real eigenvalues, by Lemma (2) below. Any matrix $A \in U$ with n distinct real eigenvalues is diagonalizable and therefore $F(A) = \chi_A(A) = 0$, by Lemma (3) below.

Thus $\chi_A(A) = 0$ for all $A \in U$, hence by rigidity $A \in \mathbb{R}^{n\times n}$.

□

(2) Lemma. *Let $D_0 \in \mathbb{R}^{n\times n}$ be a matrix with n distinct real eigenvalues. Then there is a neighborhood U of D_0 where all matrices $D \in U$ have n distinct real eigenvalues.*

Proof. Consider the function $G(D, t) = \chi_D(t)$ as a differentiable map $\mathbb{R}^{n\times n} \times \mathbb{R} \to \mathbb{R}$. The condition $G(D, \lambda) = 0$ is equivalent to λ being an eigenvalue of D.

We use the implicit function theorem to show that each real eigenvalue λ_i of D_0 can be continued to a real function $\lambda_i(D)$ in a neighborhood of D_0. For each eigenvalue λ_i of D_0 we have $G(D_0, \lambda_i) = 0$ and

$$\frac{\partial G}{\partial t}(D_0, \lambda_i) = \chi'_{D_0}(\lambda_i) = \prod_{j\neq i}(\lambda_i - \lambda_j) \neq 0,$$

since all eigenvalues of D_0 are distinct.

By the implicit function theorem, there exists a neighborhood U_i of D_0 and a smooth function $\lambda_i : U_i \to \mathbb{R}$ such that $G(D, \lambda_i(D)) = 0$ for all $D \in U_i$, with $\lambda_i(D_0) = \lambda_i$. Taking $U = \bigcap_{i=1}^{n} U_i$, we obtain a neighborhood where all n eigenvalues $\lambda_1(D), \ldots, \lambda_n(D)$ exist as real-valued functions. Since the λ_i are continuous and the values $\lambda_i(D_0)$ are distinct, they remain distinct in a sufficiently small neighborhood.

□

(3) Lemma. *Let $A \in \mathbb{R}^{n\times n}$ have n distinct real eigenvalues. Then $\chi_A(A) = 0$.*

Proof. If A has n distinct real eigenvalues $\lambda_1, \ldots, \lambda_n$ and corresponding eigenvectors $q_1, \ldots, q_n$, then $Q = [q_1, \ldots, q_n]$ is invertible and $Q^{-1}AQ = D$, where $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. The characteristic polynomial satisfies $\chi_A(\lambda) = \chi_D(\lambda) = \prod_{i=1}^{n}(\lambda - \lambda_i)$, and hence

$$\chi_A(A) = Q\, \chi_D(D)\, Q^{-1} = Q \cdot 0 \cdot Q^{-1} = 0.$$

□

References

[Cay58] Arthur Cayley. A memoir on the theory of matrices. *Philosophical Transactions of the Royal Society of London*, 148:17–37, 1858. URL `https://royalsocietypublishing.org/doi/10.1098/rstl.1858.0002`.

[Fro78] Ferdinand Georg Frobenius. Über lineare substitutionen und bilineare formen. *Journal für die reine und angewandte Mathematik*, 84:1–63, 1878. URL `https://www.degruyterbrill.com/document/doi/10.1515/crelle-1878-18788403/html`.

[Ham53] William Rowan Hamilton. *Lectures on Quaternions.* Hodges and Smith, Dublin, 1853. URL `https://archive.org/details/bub_gb_TCwPAAAAIAAJ`.

[Hig20] Nicholas J. Higham. What is the Cayley–Hamilton theorem? Blog post, 2020. URL https://nhigham.com/2020/11/03/what-is-the-cayley-hamilton-theorem/.

[HJ13] Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, 2nd edition, 2013. URL https://www.cambridge.org/highereducation/books/matrix-analysis/3D0C9628C97B0E2E0EBA1C0935C8C34F.