Cayley Hamilton Theorem — A Topological Proof

S. Kumaresan
School of Math. and Stat.
University of Hyderabad
Hyderabad 500046
kumaresa@gmail.com

1. Does there exist a polynomial $P(X)$ such that $P(A) = 0$? Since $\dim M(n, \mathbb{C}) = n^2$, the $(n^2 + 1)$ elements $I, A, \ldots, A^{n^2}$ are linearly dependent. Hence there exists a polynomial $P(X)$ of degree n^2 such that $P(A) = 0$.

2. Since $p(A) = 0$ (0 is the zero matrix here) iff $p(A)v = 0$ (here 0 is the zero vector) for all v, can one find a polynomial $p_v(X)$ such that $p_v(A)v = 0$? Yes, and there exists at least one with degree n.

3. If $\{v_i : 1 \leq i \leq n\}$ is a basis, and if $p_i(X)$ is such that $p_i(A)v_i = 0$, then $P(X) = \prod_{i=1}^{n} p_i(X)$ is of degree n^n and it kills A. This is bad, as $n^n \geq n^2$.

4. Definition of the characteristic polynomial $f_A(X)$ of a matrix. Example of a 2×2 matrix. Relation of the coefficients of the characteristic equation, eigenvalues, trace and determinant.

5. Statement of the theorem. Why is this astonishing? In our attempts, we could only assert the existence of polynomials which kill A, but we do not know how to construct, also their degrees are much higher than n. Cayley-Hamilton theorem produces an explicit polynomial with degree n.

6. If A and B are conjugate, then they have the same characteristic polynomial.

7. What is wrong with the 'proof': put $X = A$ in $\det(XI - A)$?

8. The simplest class of matrices for which the theorem is true: the class of diagonal matrices. For, if $A = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, then the characteristic polynomial $f_A(X) = \prod_{k=1}^{n}(X - \lambda_k)$. In particular, if v_i is a nonzero eigenvector corresponding to λ_i, then $p_i(X) := X - \lambda_i$ is such that $p_i(A)(v_i) = 0$.

9. Since conjugate matrices have the same characteristic polynomial, it follows that the theorem is true for any diagonalizable matrix.

10. Do we have a special class of diagonalizable matrices? Yes, the matrices with distinct eigenvalues.

11. Are there any special class of matrices whose eigenvalues are easily seen by inspection? Yes, the class of upper triangular matrices.

12. The vector space $M(n, \mathbb{C})$ is an inner product space with the inner product $(X, Y) \mapsto \operatorname{Trace} XY^*$. This inner product induces a norm and hence a distance function or metric it.

13. Since trace is conjugate invariant, it folllows that $d(A, B) = d(UAU^*, UBU^*)$.

14. In $\mathbb{C}^n$, the set of points $z = (z_1, \ldots, z_n)$ with distinct components (that is, $z_i \neq z_j$ if $i \neq j$) is an open dense set.

15. Given an upper triangular matrix T, there exists an upper triangular matrix T_ε which has distinct eigenvalues and which is ε-near to A.

16. In view of Item 6, a natural question is: Given a matrix $A \in M(n, \mathbb{C})$, does there exist a upper triangular matrix which is a conjugate of A? Yes, there exists an upper triangular matrix T which is unitarily conjugate to A.

17. It follows from Item 13 and 16 that given any matrix A, there exists an upper triangular matrix $T = UAU^*$ which is conjugate to A, there exists an upper triangular matrix T_ε which is at most ε distance from T and hence $A_\varepsilon := U^* T_\varepsilon U$ is a matrix which has distinct eigenvalues and is at most ε-distance from A.

18. What Item 17 says is this: In the space $M(n, \mathbb{C})$ of matrices endowed with the metric as in Item 15, the set of all diagonalizable matrices is dense.

19. Given two functions $f, g \colon X \to Y$, there are some special situations under which we may conclude that $f = g$ without actually verifying that $f(x) = g(x)$ for **all** $x \in X$.

    (a) If f and g are polynomials, if n is the maximum of their degrees, then if they take the same values at $n + 1$ points, then $f = g$.

    (b) If $f, g \colon V \to W$ are linear maps, if $\{v_i : 1 \leq i \leq n\}$ is a basis of V, and if $f(v_i) = g(v_i)$ for $1 \leq i \leq n$, we conclude that $f = g$.

    (c) If $f, g \colon X \to Y$ are continuous functions on metric spaces and if $f(x) = g(x)$ for all $x \in D$, a dense subset of X, then $f = g$.

    (d) If $f, g \colon U \subset \mathbb{C} \to \mathbb{C}$ are analytic functions on a connected open set and if $f(z_n) = g(z_n)$ for distinct $z_n \in U$ which have an accumulation point $z \in U$, then $f = g$.

20. In view of Items 18 and 19c, it behooves us to wonder whether the map $A \mapsto f_A(A)$ from $M(n, \mathbb{C})$ to itself is continuous. Yes, it is.

21. Thus, the maps $A \mapsto f_A(A)$ and $A \mapsto 0$ are continuous maps that agree on the dense set of diagonalizable matrices. Hence they agree everywhere. That is, the Cayley-Hamilton theorem is proved for complex matrices!

Acknowledgment. This is an outline of the talks given at various places including Hyderabad University, ISI Kolkata, and the Refresher Course in Linear Algebra and Analysis, held at the Department of Mathematics and Statistics, University of Hyderabad during 14-17, September 2006. I thank the enthusiastic audience at all these places. This outline arose out of the request from some members of the audience who asked me to write up the talk. In the near future, I hope to expand this article to give the details and also to show how the proof can be modified to show that the result holds over any algebraically closed field.