

THE FUNDAMENTAL THEOREM OF ALGEBRA, A PROOF BY LINEAR ALGEBRA

LALIT JAIN

1. INTRODUCTION

The fundamental theorem of algebra is one of the most distinguished results of modern mathematics. Simply stated, the theorem guarantees the existence of a root of any complex polynomial. After several attempts to find a proof from Euler, Lagrange and Laplace among others, the theorem was proved in 1799 by Gauss. All of the original methods of proof involved algebraic and topological arguments which encompassed the notion of winding number. The following is a proof using Liouville's Theorem from Complex Analysis.

Theorem 1.1 (The Fundamental Theorem of Algebra). *Any polynomial $P(z) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_0$ where the $a_i \in \mathbb{C}$ has a complex root.*

Proof. Suppose that $P(z)$ is not 0 for any value of z . Let $f(z) = 1/P(z)$. Let

$$w = \frac{a_0}{z^n} + \frac{a_1}{z^{n-1}} + \dots + \frac{a_{n-1}}{z}$$

so that $P(z) = (a_n + w)z^n$. For a sufficiently large R we can show that for $|z| > R$ the absolute value of each term in the sum above can be bounded by $|a_n|/(2n)$, so by the triangle inequality $|w| < |a_n|/2$ and $|a_n + w| > |a_n|/2$. Thus

$$|f(z)| = \frac{1}{|P(z)|} = \frac{1}{|a_n + w||z^n|} < \frac{2}{|a_n|R^n}$$

for $|z| > R$. Also since f is continuous on the disk $|z| \leq R$, it is bounded in this disk so f is bounded on all of \mathbb{C} . Thus by Liouville's theorem, since f is clearly entire, f is constant. Thus $P(z)$ is constant, a contradiction. \square

The goal of this essay is to present a proof of this result motivated by linear algebra. The fundamental theorem of algebra is crucial in most linear algebra courses to demonstrate the existence of an eigenvalue of a complex matrix. Conversely, to prove that every polynomial has a complex root it suffices to show that every matrix has an eigenvalue.

Before we begin the proof, we recall the following necessary definitions and lemmas.

Lemma 1.2. *Every odd degree polynomial with real coefficients has a real root.*

Proof. Let $p(x)$ be such a polynomial and without loss of generality, assume that the coefficient of the highest power of $p(x)$ is positive. As x tends to positive infinity, $p(x)$ will tend to positive infinity, and as x tends to negative infinity, $p(x)$ will tend to negative infinity. Thus for some x_0 , by the intermediate value theorem, $p(x_0)$ will be zero. \square

Note that this result is how we build the necessity of the completeness of \mathbb{R} into our proof.

Lemma 1.3. *Every complex number has a square root.*

Proof. Let $a + bi$ be a complex number, with a and b real. Let $\gamma = \sqrt{a^2 + b^2}$ then we can easily see

$$\left(\sqrt{\frac{\gamma + a}{2}} + i \frac{\gamma - a}{2} \right)^2 = a + bi$$

\square

We also recall the following definitions and lemmas from linear algebra:

Definition 1.4. Let A be a complex square matrix of dimension n . We say that A is *Hermitian* if $A = \overline{A}^t$ and that A is *skew-symmetric* if $A^t = -A$.

Lemma 1.5. *The Hermitian Matrices on an n -dimensional space are a n^2 dimensional real vector space.*

Proof. Note that every Hermitian matrix can be written as $A + iB$ where A and B are n -dimensional matrices with real entries. In particular A must be symmetric and B must have zeros on its diagonal, and the $B_{ij} = -B_{ji}$ for $i \neq j$. Thus the dimension of this vector space must be the dimension of the vector space of symmetric matrices plus the dimension of the real vector space of skew-symmetric matrices. This is just

$$\frac{n(n+1)}{2} + \frac{n(n-1)}{2} = n^2.$$

\square

2. MOTIVATION OF THE PROOF

We will prove that any r commuting endomorphisms of a finite dimensional complex vector space have a common eigenvector. With this result we can prove Theorem 1.

Proof. Let

$$P(x) = x^n + a_1x^{n-1} + \cdots + a_n$$

Now consider the matrix A .

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & -a_n \\ 1 & 0 & \cdots & 0 & -a_{n-1} \\ 0 & 1 & \cdots & 0 & -a_{n-2} \\ \vdots & & \cdots & & \vdots \\ 0 & 0 & \cdots & 1 & -a_1 \end{pmatrix}$$

□

Consider $\det(xI - A)$. From linear algebra we know that this matrix is the companion matrix of $P(x)$ so $P(x) = \det(xI - A)$. Thus from the comment above we know that A has a complex eigenvalue so the polynomial has a root.



3. PRELIMINARY RESULTS

Fix a field k and let r and d be positive integers. In the following we will prove that any r commuting endomorphisms A_1, \dots, A_r of a K -vector space V of dimension n where d does not divide n have a common eigenvector. Let this statement be designated as $P(K, d, r)$. At a glance, the divisibility requirement between d and n is superfluous, however we will demonstrate the role it plays later.

Consider the following lemma.

Lemma 3.1. *If $P(K, d, 1)$ holds then $P(K, d, r)$ holds for $r \geq 1$.*

Proof. We prove the lemma by inducting on r and for each r inducting on n . Assume that $P(K, d, r - 1)$ holds and suppose that A_1, \dots, A_r are commuting endomorphisms of a K -vector space V of dimension n such that d does not divide n . By induction on n we will show the A_i have a common eigenvector. The case of $n = 1$ is clear.

By $P(K, d, 1)$ we know that A_r has an eigenvalue λ in K . Let $T = A_r - \lambda I$. Note that since A_1, \dots, A_{r-1} commute we see that the kernel and image of T are stable under their action.

Note that if $\ker T$ is V then since $P(K, d, r - 1)$ holds then A_1, \dots, A_{r-1} will have a common eigenvector. In particular this will be an eigenvector of A_r .

If the $\ker T$ is not all of V then we know $\dim \ker T + \dim \operatorname{Im} T = \dim V$. So either d does not divide $\dim \ker T$ or $\dim \operatorname{Im} T$. However by induction on n the matrices A_1, \dots, A_r have a common eigenvector in the kernel or image of T . \square

Lemma 3.2. *$P(\mathbb{R}, 2, r)$ holds for all r , i.e. if A_1, \dots, A_r are commuting endomorphisms on an odd dimensional real vector space, then they have a common eigenvector.*

Proof. By the previous lemma, we need to show $P(\mathbb{R}, 2, 1)$ is true. However given A an endomorphism of an odd dimensional vector space, $\det(xI - A)$ is a polynomial of odd degree with real coefficients so it has a real zero. This zero is a real eigenvalue. \square

Lemma 3.3. *$P(\mathbb{C}, 2, 1)$ holds. In other words, every endomorphism of a complex vector space of odd dimension has an eigenvector.*

Proof. Let A be an endomorphism of an odd dimension vector space. Let H denote the set of $n \times n$ Hermitian operators. Define for $B \in H$ the following endomorphisms of V :

$$L_1(B) = \frac{AB + BA^*}{2}$$

and

$$L_2(B) = \frac{AB - BA^*}{2i}.$$

Note that

$$\begin{aligned}
 L_1(L_2(B)) &= L_1\left(\frac{AB - BA^*}{2i}\right) \\
 &= \frac{A\frac{AB-BA^*}{2i} + \frac{AB-BA^*}{2i}A^*}{2} \\
 &= \frac{A^2B - ABA^* + ABA^* - B(A^*)^2}{4i} \\
 &= \frac{A\frac{AB+BA^*}{2} + \frac{AB+BA^*}{2}A^*}{2i}
 \end{aligned}$$

So L_1 and L_2 are commuting operators. Now the dimension of H as a real vector space is n^2 by Lemma 1.5 which is odd. Thus applying the previous lemma, we see that $P(\mathbb{R}, 2, 2)$ implies that L_1 and L_2 have a common eigenvector, B . If $L_1(B) = \lambda B$ and $L_2(B) = \mu B$ where λ, μ are real then $(L_1 + iL_2)B = \frac{AB-BA^*}{2} + \frac{AB+BA^*}{2} = AB = (\lambda + \mu i)B$ so any nonzero column vector of B gives an eigenvector of A . \square

Lemma 3.4. $P(\mathbb{C}, 2^k, r)$ holds for all k and r .

Proof. We prove the lemma by induction on k . The case of $k = 1$ follows from the previous lemmas. Now assume that $P(\mathbb{C}, 2^l, r)$ holds for $l < k$. To establish $P(\mathbb{C}, 2^k, r)$ we need to just establish $P(\mathbb{C}, 2^k, 1)$ by Lemma 3.1. Choose n to be divisible by 2^{k-1} but not by 2^k and take A to be an n dimensional complex endomorphism. Let S denote the complex vector space of skew-symmetric matrices. Define the following endomorphisms of V

$$L_1(B) = AB - BA^t$$

and

$$L_2(B) = ABA^t.$$

Also note that since the dimension of S is $\frac{n(n-1)}{2}$ we can ensure that 2^{k-1} does not divide the dimension of S . Applying our induction hypothesis, we see $P(\mathbb{C}, 2^{k-1}, 2)$ implies L_1 and L_2 have a common eigenvector $B \neq 0$, with $L_1(B) = \lambda B$ and $L_2(B) = \mu B$, for λ, μ some complex numbers.

So we see

$$\mu B = ABA^t = A(AB - \lambda B)$$

and this implies

$$(A^2 - \lambda A - \mu I)B = 0.$$

Fix a nonzero column, c of B .

$$(A^2 - \lambda A - \mu I)c = 0.$$

The discriminant of $x^2 - \lambda x - \mu$ is just $\lambda^2 + 4\mu$. So by Lemma 1.3, we can find a complex square root δ . So $x^2 - \lambda x - \mu = (x - \alpha)(x - \beta)$ where $(\lambda + \delta)/2$ and $\beta = (\lambda - \delta)/2$. So if $w = (A - \beta I)v$, then if $w = 0$, v is an eigenvector of A with eigenvalue β . Otherwise

$$(A - \alpha I)w = 0$$

so w is an eigenvector with eigenvalue α . □

Now can show our main result:

Theorem 3.5. *If A_1, A_2, \dots, A_r are commuting endomorphisms of a finite dimensional nonzero complex vector space then they have a common eigenvector.*

Proof. Let n be the dimension of the vector space. There exists an integer k so that 2^k does not divide n . Since $P(\mathbb{C}, 2^k, r)$ holds by the previous lemma, the theorem follows. □

Note that our divisibility criterion between d and n was crucial to be able to induct on the dimension of the space our endomorphisms were acting.

REFERENCES

- [1] Derksen, H., "The Fundamental Theorem of Algebra and Linear Algebra," *The American Mathematical Monthly*, pp. 620-623, Aug 2003, Vol 110, No 7.
- [2] Freidberg, S., Insel, A., Spence, L., Linear Algebra, Prentice Hall, New Jersey, 2003.