The Fundamental Theorem of Algebra: A Linear Algebraic Approach

By Olivia C. Riley

* * * * * * * * *

Submitted in partial fulfillment of the requirements for
Honors in the Department of Mathematics

Union College

March 3, 2025

ABSTRACT

OLIVIA C. RILEY The Fundamental Theorem of Algebra: A Linear Algebraic Approach

Department of Mathematics, March 3, 2025.

ADVISOR: HATLEY, JEFFREY The Fundamental Theorem of Algebra guarantees the existence of complex roots for nonconstant polynomials with complex coefficients. Drawing on the work done by Harn Derksen, this paper presents a linear algebraic approach highlighted by Keith Conrad. By reformulating the theorem in terms of eigenvectors and linear operators, the existence of an eigenvector in $\mathbf{C}^n$ for every $n \times n$ square matrix over $\mathbf{C}$ directly implies The Fundamental Theorem of Algebra. Prior to the main work, necessary topics surrounding elementary linear algebra are introduced. Conjoining the work completed by Derksen and Conrad, the inductive method offers an elegant approach to The Fundamental Theorem of Algebra.

ACKNOWLEDGMENT

Thank you to the Union College Mathematics Department for your given support throughout the years! Thank you to Professor Hatley for assisting with this paper and for your helpful insights!

Contents

ABSTRACT ii

ACKNOWLEDGMENT iii

1. INTRODUCTION 1

2. PRELIMINARIES 2

2.1. IMAGE AND KERNEL 2

2.2. EIGENVALUES AND EIGENVECTORS 4

2.3. SYMMETRIC AND HERMITIAN MATRICES 6

2.4. PROPERTIES OF LINEAR OPERATORS AND COMMUTATIVITY 9

3. PROOF OF THEOREM 12

References 16

1. INTRODUCTION

One of the most celebrated results in mathematics, The Fundamental Theorem of Algebra, asserts that any nonconstant polynomial with complex coefficients has a complex root.

Numerous approaches have been developed to prove this theorem, including methods from geometry, complex analysis, topology, and more. However, it's usefulness spans far beyond into various disciplines within mathematics.

Why does this theorem matter? The quadratic formula provides a straightforward method for calculating complex roots of quadratic polynomials. However, for polynomials of degree greater than 5, there is no existent equivalent formula. Through theorems such as the Fundamental Theorem of Algebra, we are able to guarantee the existence of complex roots of high-degree polynomials. The guarantee allows us to understand the structure of high-degree polynomials and their behavior.

Within this paper, we will outline a proof due to Harm Derksen, as highlighted by Keith Conrad. Reformulating the theorem in terms of eigenvectors and linear algebra, Derksen demonstrates that The Fundamental Theorem of Algebra is equivalent to the following. Let

$$f(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$$

have degree $n \geq 1$, with $a_j \in \mathbb{C}$. The matrix

$$A = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & 0 & \dots & 0 & -a_1 \\ 0 & 1 & 0 & \dots & 0 & -a_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & -a_{n-2} \\ 0 & 0 & 0 & \dots & 1 & -a_{n-1} \end{pmatrix}$$

satisfies $\det(\lambda I_n - A) = f(\lambda)$. Therefore, Derksen's reformulation is equivalent to The Fundamental Theorem of Algebra. The theorem is as follows:

Theorem 1. *For each $n \geq 1$, every $n \times n$ square matrix over $\mathbf{C}$ has an eigenvector in $\mathbf{C}^n$. Equivalently, for each $n \geq 1$, every linear operator on an n-dimensional complex space V has an eigenvector in V.*

We will use Derksen's work alongside Conrad's highlights work to prove Theorem 1. The Fundamental Theorem of Algebra will then be a direct consequence.

2. PRELIMINARIES

For the reader's convenience, we recall some basic facts from linear algebra.

2.1. IMAGE AND KERNEL.

Definition 1. Let $T : U \to W$ be a linear transformation. The **kernel** of T is defined by

$$\ker T = \{u \in U \mid T(u) = 0\}.$$

Example 1. Suppose $T : \mathbf{R}^2 \to \mathbf{R}^2$ is a linear transformation defined by the matrix

$$\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}.$$

We can write

$$T \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + 2y \\ 0 \end{bmatrix}.$$

Setting our result equal to the zero vector, we have

$$\begin{bmatrix} x + 2y \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

By equating the entries in the first row, we get $x + 2y = 0$ or $x = -2y$. The kernel of T then consists of all scalar multiples of $\begin{bmatrix} -2 \\ 1 \end{bmatrix}$.

Definition 2. Let $T : U \to W$ be a linear transformation. The **image** of T is defined by

$$\text{im } T = \{u \in U \mid T(u) \in U\}.$$

Example 2. Using the same transformation example 1, suppose $T : \mathbf{R}^2 \to \mathbf{R}^2$ is a linear transformation defined by A. We can write

$$T \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + 2y \\ 0 \end{bmatrix}.$$

We denote $a = x + 2y$, where as x and y range over all real numbers, the same is true of a. However, our second entry will always be equal to 0. Therefore, the image of T is all vectors of the form $\begin{bmatrix} a \\ 0 \end{bmatrix}$.

2.2. **EIGENVALUES AND EIGENVECTORS.** We begin this section with the main definition.

Definition 3. If A is a $n \times n$ matrix, then a nonzero vector x in $\mathbf{C}^n$ is called an **eigenvector** of A if Ax is a scalar multiple of x; that is,

$$Ax = \lambda x$$

for some scalar λ. The scalar λ is called an **eigenvalue** of A, and x is said to be an **eigenvector** corresponding to λ.

From the definition of an eigenvector, the first equation can be written as $Ax = \lambda Ix$. After subtracting the right-hand side of the equation, we get $Ax - \lambda Ix = 0$ or

$$(A - \lambda I)x = 0.$$

Theorem 2. *If A is a $n \times n$ matrix, then λ is an eigenvalue of A if and only if it satisfies the equation*

$$\det(\lambda I - A) = 0.$$

This is called the ***characteristic equation*** *of A.*

Proof. Suppose A is a $n \times n$ matrix.
($\Rightarrow$) Assume λ is an eigenvalue of A. By definition, then $Ax = \lambda x$ for some vector x. We can rewrite this as $(A - \lambda I)x = 0$. However, as x is a nonzero vector, x must be in $\ker A$. The rows of A are then linearly dependent, resulting in a determinant of 0. Hence, if λ is an eigenvalue of A, $\det(A - \lambda I) = 0$ holds.
($\Leftarrow$) Assume $\det(A - \lambda I) = 0$. As a result, there exists a nonzero vector x

such that $(A - \lambda I)x = 0$. We can rewrite this as $Ax = \lambda x$. We have then shown λ is an eigenvalue of A by definition. $\square$

Example 3. Take matrix

$$A = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}.$$

From Theorem 2, it follows that the eigenvalues of A are the solutions to $\det(\lambda I - A) = 0$. We write this as

$$\det\left(\lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}\right).$$

We simplify this as

$$\det \begin{bmatrix} \lambda - 3 & 0 \\ -8 & \lambda + 1 \end{bmatrix} = 0$$

from which we obtain the characteristic polynomial

$$(\lambda - 3)(\lambda + 1) = 0$$

of A. Our characteristic equation results in eigenvalues of $\lambda_1 = 3$ and $\lambda_2 = -1$.

To find the corresponding eigenvectors, we use the definition of an eigenvector alongside Theorem 2. Vector x is an eigenvector of A corresponding to λ if and only if x is a nontrivial solution of $\det(\lambda I - A)x = 0$. We continue using our eigenvalues from Example 3.

Example 4. To find the respective eigenvectors for Example 3, we begin with $\lambda_1 = 3$. We aim to find all nonzero vectors resulting in the solution $(3I - A)x =$

0. This gives

$$\left(3\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} - \begin{bmatrix}3 & 0\\ 8 & -1\end{bmatrix}\right)\begin{bmatrix}x_1\\ x_2\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix}$$

$$\begin{bmatrix}0 & 0\\ -8 & 4\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix}.$$

The corresponding augmented matrix and reduced row-echelon form is given by

$$\left[\begin{array}{cc|c}0 & 0 & 0\\ -8 & 4 & 0\end{array}\right] \to \left[\begin{array}{cc|c}1 & \frac{1}{2} & 0\\ 0 & 0 & 0\end{array}\right].$$

The general solution in matrix form is then any vector of the form

$$\begin{bmatrix}x_1\\ x_2\end{bmatrix} = \begin{bmatrix}\frac{1}{2}t\\ t\end{bmatrix} = t\begin{bmatrix}\frac{1}{2}\\ 1\end{bmatrix}.$$

Thus, our basic eigenvector for $\lambda_1 = 3$ is:

$$\begin{bmatrix}\frac{1}{2}\\ 1\end{bmatrix}.$$

The eigenspace for $\lambda_1 = 3$ is then the kernel of $A - 3I$. Following the same procedure for $\lambda_2 = -1$, our basic eigenvector is:

$$\begin{bmatrix}0\\ 1\end{bmatrix},$$

with the eigenspace being the kernel of $A + 2I$.

2.3. **SYMMETRIC AND HERMITIAN MATRICES.** We now discuss a few special types of matrices which will be crucial in the proof of The Fundamental Theorem of Algebra.

Definition 4. If A is a square matrix, then A is **symmetric** if it is equal to its transpose that is, $A = A^T$.

Example 5. Suppose A is a square matrix defined by

$$A = \begin{bmatrix} 2 & 3 & 6 \\ 3 & 4 & 5 \\ 6 & 5 & 9 \end{bmatrix}.$$

Then

$$A = \begin{bmatrix} 2 & 3 & 6 \\ 3 & 4 & 5 \\ 6 & 5 & 9 \end{bmatrix} = A^T.$$

Hence, A is symmetric.

Definition 5. If A is a $n \times n$ complex matrix, then the **conjugate transpose** of A, denoted by A^*, is defined by

$$A^* = \overline{A}^T.$$

Example 6. Suppose A is a complex matrix defined by:

$$A = \begin{bmatrix} 1-i & i \\ 0 & 2 \end{bmatrix}.$$

We have

$$\overline{A} = \begin{bmatrix} 1+i & -i \\ 0 & 2 \end{bmatrix} \text{ and hence } A^* = \overline{A}^T = \begin{bmatrix} 1+i & 0 \\ -i & 2 \end{bmatrix}.$$

Definition 6. A square matrix A is said to be **Hermitian** if $A^* = A$.

Example 7. Take A to be the complex matrix defined by:

$$A = \begin{bmatrix} 1 & 1-i \\ 1+i & 4 \end{bmatrix}$$

We have

$$\overline{A} = \begin{bmatrix} 1 & 1+i \\ 1-i & 4 \end{bmatrix} \text{ and hence } A^* = \begin{bmatrix} 1 & 1-i \\ 1+i & 4 \end{bmatrix} = A.$$

A is then called Hermitian.

Definition 7. The space of all $n \times n$ Hermitian matrices is defined by

$$H_n = \{M \in M_n(\mathbf{C}) : M^* = M\}.$$

Definition 8. The space of all $n \times n$ symmetric matrices is defined by

$$\text{Sym}(\mathbf{C}) = \{B \in M_n(\mathbf{C}) | B^T = B\}.$$

It is important to discuss the dimension of symmetric matrices as it will aid in our proof of The Fundamental Theorem of Algebra.

Lemma 1. *The space of $n \times n$ symmetric matrices has dimension $\frac{n(n+1)}{2}$.*

Proof. We see there are n diagonal entries. As all values above our diagonal have a corresponding value below the diagonal (by symmetry), then the matrix is determined by the upper triangular. This results in $\frac{n(n-1)}{2}$, or the number of ways to choose 2 different indices. Hence, total dimension = values across diagonal + values above the diagonal, we get dimension $= n + \frac{n(n-1)}{2} = \frac{n(n+1)}{2}$.

□

2.4. **PROPERTIES OF LINEAR OPERATORS AND COMMUTATIVITY.** For the purposes of our proof, we introduce two sets of linear operators and some supporting propositions.

Define $L_1, L_2 : H_n(\mathbf{C}) \to H_n(\mathbf{C})$ as two linear operators defined by

$$L_1(B) = \frac{AB + BA^*}{2}, \quad L_2(B) = \frac{AB - BA^*}{2i}.$$

Proposition 1. *The linear operators stated above, L_1 and L_2, commute.*

Proof. To show $L_1(B)$ and $L_2(B)$ commute, we must show $L_1(L_2(B)) = L_2(L_1(B))$. We compute

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

Hence L_1 and L_2 commute. □

Define $L_1, L_2 : Sym_n(\mathbf{C}) \to Sym_n(\mathbf{C})$ be linear operators defined by

$$L_1(B) = AB + BA^T, L_2(B) = ABA^T.$$

Proposition 2. *The linear operators stated above, L_1 and L_2 commute.*

Proof. To show $L_1(B)$ and $L_2(B)$ commute, we must show $L_1(L_2(B)) = L_2(L_1(B))$. Then

$$\begin{aligned} L_1(L_2(B)) &= L_1(ABA^T) \\ &= A(ABA^T) + (ABA^T)A^T \\ &= A^2BA^T + AB(A^T)^2 \\ &= A(AB + BA^T)A^T \\ &= L_2(AB + BA^T) \\ &= L_2(L_1(B)). \end{aligned}$$

Hence, L_1 and L_2 commute. □

Used as a starting point for the main proof, Conrad uses the hypothesis of the following lemma to support the inductive step.

Lemma 2. *Fix an integer $m > 1$ and a field F. Suppose that, for every F-vector space V whose dimension is not divisible by m, every linear operator on V has an eigenvector in V. Then for every F-vector space V whose dimension is not divisible by m, each pair of commuting linear operators on V has a common eigenvector in V.*

Proof. Conrad [1] proves the case where $m = 2^k$, with $k \geq 0$. We proceed by induction upon $d = \dim_F V$ where d is not divisible by m. Consider the base case where $d = 1$. Trivially, for any linear operator on a one-dimensional vector space V must map a basis vector to a multiple of itself. Therefore, every vector in V is an eigenvector of every linear operator on the space. Hence, any two commuting linear operators on V must share a common eigenvector in V.

Assume now that $d > 1$ and that our lemma holds for all vector spaces with dimensions less than d that are not divisible by m. We proceed with our inductive step for $d > 1$. By our hypothesis, we know L_1 has an eigenvalue in F, λ. Suppose

$$U = \text{im}(L_1 - \lambda I_V), \quad W = \ker(L_1 - \lambda I_V).$$

Both U and W are L_1 stable, meaning that applying L_1 to vectors in U or W results in vectors back in U or W respectively ($\dim_F W \geq 1$ as λ is an eigenvalue of L_1). Using our assumption regarding the commutativity of L_1 and L_2, we can show these subspaces are also L_2-stable. Apply L_2 to $u \in U$ where $u = L_1(v) - \lambda v$. Then

$$\begin{aligned} L_2(u) &= L_2(L_1(v) - \lambda v) \\ &= L_2(L_1(v)) - L_2(\lambda v) \\ &= L_1(L_2(v)) - \lambda L_2(v) \\ &= (L_1 - \lambda I_V)(L_2(v)) \end{aligned}$$

is back in U (a similar process would follow for W). As sum of the dimension of the image and the dimension of the kernel is equal to the dimension of the domain, $\dim_F U + \dim_F W = \dim_F V = d$. Therefore, d is not divisible by m. Therefore, either U or W must have odd dimension. By our inductive hypothesis, L_1 and L_2 must then have a common eigenvector in V. In the case where the dimension of $W = V$ (we already noted $\dim_F W \geq 1$, hence $\dim_F U = 0$ in this case), every vector in V is an eigenvector of L_1. Finally, since L_1 and

L_2 commute, L_2 has at least one eigenvector corresponding to λ as a result of our hypothesis. Therefore L_1 and L_2 have a common eigenvector. □

Corollary 1. *For every real vector space V whose dimension is odd, each pair of commuting linear operators on V has a common eigenvector in V.*

Proof. Let $V = \mathbb{R}$, applying Lemma 2. with $m = 2$. Assume the dimension of V is odd. As the characteristic polynomial will have odd degree, it must then contain at least one real eigenvalue corresponding to a real eigenvector. Then, by Lemma 1, each pair of commuting linear operators on V has a common eigenvector in V. □

3. PROOF OF THEOREM

Theorem 3. *For each $n \geq 1$, every $n \times n$ square matrix over $\mathbf{C}$ has an eigenvector in $\mathbf{C}^n$. Equivalently, for each $n \geq 1$, every linear operator on an n-dimensional complex vector space V has an eigenvector in V.*

Proof. We follow the proof of [2] as explained in [1]. We proceed via induction, inducting on $k \geq 1$ such that $n = 2^k n'$ where n' is odd.

Consider the base case where $k = 0$ (we see that $n = 2^0 n' = n'$). Therefore, n is odd. Take $A \in M_n(\mathbf{C})$ with odd dimension. We aim to show A has an eigenvector in $\mathbf{C}^n$. We will associate to A two linear operators on the space H_n. Any matrix C in $M_n(\mathbf{C})$ can be written as

$$C = \frac{C + C^*}{2} + i\frac{C - C^*}{2i}.$$

We see that each component is Hermitian, as

$$\begin{aligned}\left(\frac{C+C^*}{2}\right)^* &= \frac{C^*+C}{2} \\ &= \frac{C+C^*}{2}\end{aligned}$$

and

$$\begin{aligned}\left(\frac{C-C^*}{2i}\right)^* &= \frac{C^*-C}{-2i} \\ &= \frac{C-C^*}{2i}.\end{aligned}$$

For $B \in M_n(\mathbf{C})$, take $C = AB$. We then write:

$$AB = \frac{AB+(AB)^*}{2} + i\frac{AB-(AB)^*}{2i}.$$

We rewrite as:

$$AB = \frac{AB+B^*A^*}{2} + i\frac{AB-B^*A^*}{2i} = \frac{AB+BA^*}{2} + i\frac{AB-BA^*}{2i}.$$

Consider the following $\mathbf{R}$-linear operators. Recall from Section 2.4, the linear operators $L_1, L_2 : H_n \to H_n$:

$$L_1(B) = \frac{AB+BA^*}{2}, \quad L_2(B) = \frac{AB-BA^*}{2i}.$$

By Proposition 1., we know that L_1 and L_2 commute. As H_n has odd dimension, then by Corollary 3, L_1 and L_2 have a common eigenvector in H_n. Therefore, if $L_1(B) = \lambda_1 B$ and $L_2(B) = \lambda_2 B$ (where our eigenvalues are real due to complexity), then $AB = L_1(B) + iL_2(B) = \lambda_1 B + i\lambda_2 B = (\lambda_1 + i\lambda_2)B$. Therefore, we have shown A has an eigenvector in $\mathbf{C}^n$ as any column vector of

B represents an eigenvector in $\mathbf{C}^n$. This concludes our base case for odd n. After proving the base case for odd n, we will now extend the proof to all n through our inductive step.

Our inductive hypothesis for $k \geq 1$ is the following: Assume for any integer d, $d = 2^j d'$ with d' odd. Then if $j \leq k$, then every linear operator on an odd d-dimensional vector space has an eigenvector in that same vector space. By Lemma 1, each pair of commuting linear operators in that space must have a common eigenvector.

Now, we aim to prove that the result holds for n where the highest power of 2 dividing n is 2^k. We consider the space of all $n \times n$ symmetric matrices. This space has dimension $\frac{n(n+1)}{2}$ by Lemma 1. As $\frac{n(n+1)}{2}$ is not divisible by 2^k, by Lemma 2 every pair of commuting linear operators on $\mathrm{Sym}_n(\mathbf{C})$ has a common eigenvector in $\mathrm{Sym}_n(\mathbf{C})$.

Recall from Section 2.4, the linear operators $L_1, L_2 : \mathrm{Sym}_n(\mathbf{C}) \to \mathrm{Sym}_n(\mathbf{C})$. By Corollary 2, L_1, L_2 commute. The linear operators then have a common eigenvector in $\mathrm{Sym}_n(\mathbf{C})$ by our previous application of Lemma 3. In other words some nonzero $B \in \mathrm{Sym}_n(\mathbf{C})$ satisfies:

$$AB + BA^T = \lambda B, \quad ABA^T = \mu B,$$

(where λ, μ are complex numbers). Applying A to the first equation and then using the second to simplify, we find

$$A^2 B + \mu B = \lambda AB.$$

After subtracting λAB to the left-hand side and factoring out B, we have

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

As every complex number has a complex square root, we can factor the complex polynomial:

$$(A - \alpha I)((A - \beta I)B) = 0.$$

Then, if $(A-\alpha I)B = 0$, then a column vector of B is an eigenvector of A with eigenvalue β. Alternatively, if $(A - \alpha I)B \neq 0$, then a nonzero column vector of $(A - \beta I)B$ is an eigenvector of A, with eigenvalue β same holds but with an eigenvalue of α.

□

References

[1] Keith Conrad, *The fundamental theorem of algebra via linear algebra*, kconrad.math.uconn.edu (2004).

[2] Harm Derksen, *The fundamental theorem of algebra and linear algebra*, American Mathematical Monthly **110** (200308).