The Weak Fundamental Theorem of Algebra

Andrew Paul

12/31/2019

1 Introduction

The Fundamental Theorem of Algebra is a powerful result, stating that every non-constant single-variable polynomial in $\mathbb{C}[x]$ has at least one root in $\mathbb{C}$. Interestingly, no purely algebraic proof exists of this fact. I have recently learned of a relatively elementary proof of this fact, and it still requires the notion of continuity (which is topological) and the geometry of complex numbers. Not to mention, the proof is very complicated and unreasonably ingenious.

There is, however, a weaker form of the theorem that provides a maximum for the number of distinct roots of a polynomial. We can state it as follows.

Theorem (Weak Fundamental Theorem of Algebra): Every nonzero polynomial in $\mathbb{C}[x]$ with degree d has at most d distinct roots in $\mathbb{C}$.

This is almost embarrassingly obvious by Factor Theorem. In fact, we can prove it inductively with the Factor Theorem. However, there is in fact a much more interesting proof using linear algebra. This proof showed up as a homework problem in MATH 31AH at UC San Diego.[1]

To establish the result, we require some important lemmas.

Lemma 1: *A linear transformation is injective if and only if its kernel is* $\{\mathbf{0}\}$.

Proof: Suppose T is a linear transformation. Then in one direction, if T is injective, there exists only a single value in the kernel of T. Since every linear transformation must have $\mathbf{0}$ in the kernel, as $T(\mathbf{0}) = 0 \cdot T(\mathbf{v}) = 0$, we are done in this direction.

In the other direction, suppose the kernel of T is $\{\mathbf{0}\}$. Then,

$$T(\mathbf{v}) = T(\mathbf{w}) \Leftrightarrow T(\mathbf{v} - \mathbf{w}) = \mathbf{0} \Rightarrow \mathbf{v} = \mathbf{w},$$

so T is injective. ■

[1] Actually, the problem in homework only applied to polynomials in $\mathbb{R}[x]$, but the method easily generalizes to $\mathbb{C}[x]$.

Lemma 2: *The image of a linear transformation is a subspace of the codomain.*

Proof: Let $T\colon V \to W$ be a linear transformation. Clearly, the image of any linear transformation contains $\mathbf{0}$. Let T map $\mathbf{v}_i \mapsto \mathbf{w}_i$ and let a, b be scalars. Then,

$$\begin{aligned} a\mathbf{w}_1 + b\mathbf{w}_2 &= aT(\mathbf{v}_1) + bT(\mathbf{v}_2) \\ &= T(a\mathbf{v}_1 + b\mathbf{v}_2) \\ &= \mathbf{w}_k, \end{aligned}$$

so the image of T is a subspace of W. ■

Lemma 3: *If the dimension of a vector space and a subspace of it are equal, then the subspace is that vector space.*

Proof: Let the dimension be n. Then, a basis of the subspace has n linearly independent vectors, so it must also be a basis for the original vector space. Since the span of the basis must be both the subspace and the vector space, the subspace must be the vector space. ■

Lemma 4: *Let $T\colon V \to W$ be a linear transformation. If $\dim V = \dim W$, then T is injective if and only if it is surjective.*

Proof: Suppose T is injective, then by the first lemma, its kernel is $\{\mathbf{0}\}$. Then, by the rank-nullity theorem,

$$\dim \operatorname{img} T = \dim V = \dim W.$$

By the second lemma, the image of T is a subspace of W, and then by the third lemma, they must also be the same. Hence, T is also surjective.

In the other direction, suppose T is surjective. Then, by the rank-nullity theorem $\dim \ker T = 0$, and as a simple corollary of the first lemma, this only occurs when $\ker T = \{\mathbf{0}\}$ and T is injective. ■

The fourth lemma is particularly useful because it tells us that if the dimension of the domain and codomain vector spaces are equal, then showing that the linear transformation is either surjective or injective is sufficient to establish that the linear transformation is an isomorphism.

We are now ready to prove the Weak Fundamental Theorem of Algebra.

2 Main Proof

Let $a_1, ..., a_{n+1}$ be distinct complex numbers. Then, we can find a polynomial $p_i(x)$ of degree n such that

$$p_i(a_j) = \begin{cases} 1 & \text{when } j = i \\ 0 & \text{when } j \neq i. \end{cases}$$

This is simply a matter of polynomial interpolation. The solution is guaranteed to exist when our a_i are distinct since the resulting system of equations of coefficients has a Vandermonde matrix with nonzero determinant.[2]

Alternatively, we can explicitly find $p_i(x)$ in factored form to be

$$p_i(x) = \left[\prod_{\substack{0 \leq k \leq n+1 \\ k \neq i}} (a_i - a_k)^{-1} \right] \left[\prod_{\substack{0 \leq j \leq n+1 \\ j \neq i}} x - a_j \right].$$

Let $b_1, ..., b_{n+1}$ be complex numbers, not necessarily distinct. It is clear that we can find $p(x)$ with degree n such that $p(a_i) = b_i$ for $1 \leq i \leq n+1$ by using standard interpolation techniques as discussed before. In particular, we have that

$$p(x) = \sum_{k=1}^{n+1} b_k p_k(x).$$

Consider the set of all polynomials with degree lesser than or equal to n. This set contains 0, and linear combinations of elements of the set remain in the set. Hence, this set forms a vector space. Denote this vector space as P_n.

We define a linear transformation $T \colon P_n \to \mathbb{C}^{n+1}$ by

$$T(p(x)) = [p(a_1), ..., p(a_{n+1})].$$

The set of vectors $\{1, x, ..., x^n\}$ clearly forms a basis for P_n, since every polynomial in P_n can be written as a linear combination of those monomials, so we have $\dim P_n = n + 1 = \dim \mathbb{C}^{n+1}$.

Now, consider an arbitrary vector in $\mathbb{C}^{n+1}$, $\mathbf{b} = [b_1, ..., b_{n+1}]$. From what we have already deduced, we know $\exists p(x) \in P_n$ s.t. $T(p(x)) = \mathbf{b}$, namely $p(x) = \sum_{k=1}^{n+1} b_k p_k(x)$. Therefore, $\forall \mathbf{b} \in \mathbb{R}^{n+1}, \exists p(x) \in P_n$ s.t. $T(p(x)) = \mathbf{b}$. So T is surjective.

Then, by our fourth lemma, T is an isomorphism. Since T is injective, by our first lemma, its kernel is $\{\mathbf{0}\}$. Hence, the only solution to $T(A(x)) = T(B(x))$ for $A, B \in P_n$ must be $A = B$.

So consider $Q(x) \in P_n$ that satisfies $T(Q(x)) = \mathbf{0}$. It must be the only polynomial in P_n that maps to $\mathbf{0}$. Observe that under T, we have

$$0 \mapsto \mathbf{0}.$$

[2]This follows from the Vandermonde determinant formula. We discuss this in the next section.

Hence, $Q(x) = 0$ is the unique polynomial in P_n that is zero at least $n + 1$ times.

That is, no nonzero polynomial with degree less than or equal to n can have any more than n distinct roots. $\square$

3 Doing it Better: Removing the Factor Theorem

Observe that in the proof above, we never required the explicit form of $p(x)$. It was sufficient to know that such a $p(x)$ exists. By directly constructing $p(x)$, we implicitly used the Factor Theorem.

However, polynomial interpolation, at its heart, is a problem of linear algebra, and it is certainly possible to create a nonconstructive proof that a $p(x)$ exists, without explicitly finding it. Essentially, we wish to find a polynomial with degree n that contains the points (a_i, b_i) for $1 \leq i \leq n+1$, where the a_i are distinct. We know that the general form a polynomial is

$$p(x) = c_n x^n + ... + c_1 x + c_0.$$

It suffices to determine the coefficients c_i. Plugging in all known coordinates, we have a system of linear equations that we can put in matrix form as

$$A\mathbf{c} = \mathbf{b},$$

where $\mathbf{c} = [c_0, ..., c_n]^T$, $\mathbf{b} = [b_1, ..., b_{n+1}]^T$, and A is the $(n+1) \times (n+1)$ square matrix whose entry in row i column j is a_i^{j-1}. There is a unique solution to the system if and only if A is invertible. This occurs precisely when the determinant of A is nonzero. In particular, we seek

$$\det \begin{bmatrix} 1 & a_1 & \dots & a_1^{n-1} & a_1^n \\ 1 & a_2 & \dots & a_2^{n-1} & a_2^n \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 1 & a_n & \dots & a_n^{n-1} & a_n^n \\ 1 & a_{n+1} & \dots & a_{n+1}^{n-1} & a_{n+1}^n \end{bmatrix}$$

This determinant can be computed directly using purely linear algebra techniques.[3] It is

$$\prod_{1 \leq i < j \leq n+1} (a_j - a_i).$$

Since each a_i is unique, this product is never zero.

[3] https://en.wikipedia.org/wiki/Vandermonde_matrix#By_Gaussian_elimination,_U-part_of_LU_decomposition

4 Conclusion

This proof is pretty insightful. It implies that the fact that the number of distinct roots of a polynomial cannot exceed the degree of the polynomial is almost purely a consequence of the fact that the polynomials with a capped degree form a vector space. That is, they are closed under addition and scalar multiplication, and they contain a zero element.

The final element that makes the proof possible is the connection between polynomial interpolation and the Vandermonde matrix. Though this is actually a property that is very inherent to the definition of a polynomial. It is simply a finite series of monomials (which each have nonnegative degrees).

Linear algebra looks deep where we cannot and tells us that the maximum number of roots of a polynomial are directly related to their structure.