The Fundamental Theorem of Algebra

Isaiah Lankham, Bruno Nachtergaele, Anne Schilling

(February 13, 2007)

The set $\mathbb{C}$ of complex numbers can be described as elegant, intriguing, and fun, but why are complex numbers important? One possible answer to this question is the *Fundamental Theorem of Algebra*. It states that every polynomial equation in one variable with complex coefficients has at least one complex solution. In other words, polynomial equations formed over $\mathbb{C}$ can always be solved over $\mathbb{C}$. This result has several equivalent formulations in addition to a myriad of different proofs, one of the first of which was given by the eminent mathematician Carl Gauss in his doctoral thesis [2].

The aim of these notes is to provide a proof of the Fundamental Theorem of Algebra using concepts that should be familiar to you from your study of Calculus, and so we begin by providing an explicit formulation.

Theorem 1 (Fundamental Theorem of Algebra)**.** *Given any positive integer $n \geq 1$ and any choice of complex numbers $a_0, a_1, \ldots, a_n$, such that $a_n \neq 0$, the polynomial equation*

$$a_n z^n + \cdots + a_1 z + a_0 = 0 \tag{1}$$

has at least one solution $z \in \mathbb{C}$.

This is a remarkable statement. No analogous result holds for guaranteeing that a real solution exists to Equation (1) if we restrict the coefficients $a_0, a_1, \ldots, a_n$ to be real numbers. E.g., there does not exist a real number x satisfying an equation as simple as $x^2 + 1 = 0$. Similarly, the consideration of polynomial equations having integer (resp. rational) coefficients quickly forces us to consider solutions that cannot possibly be integers (resp. rational numbers). Thus, the complex numbers are special in this respect.

The statement of the Fundamental Theorem of Algebra can also be read as follows: Any non-constant complex polynomial function defined on the complex plane $\mathbb{C}$ (when thought of as $\mathbb{R}^2$) has at least one root, i.e., vanishes in at least one place. It is in this form that we will provide a proof for Theorem 1.

Given how long the Fundamental Theorem of Algebra has been around, you should not be surprised that there are many proofs of it. There is even an entire book [1] solely devoted to exploring the mathematics behind thirteen distinct proofs. Different proofs arise from attempting to understand the statement of the theorem from the viewpoint of different branches of mathematics. This quickly leads to many non-trivial interactions with such fields of mathematics as Real and Complex Analysis, Topology, and (Modern) Abstract Algebra.

The diversity of proof techniques available is yet another indication of how fundamental and deep the Fundamental Theorem of Algebra really is.

Like many first courses in Linear Algebra, we could easily be content with just accepting the statement of the theorem and deferring a discussion of its proof to a more advanced mathematics course. There is, however, a proof that uses nothing more than ideas that should be familiar to you from the study of Differential Calculus. How could you not want to see such a proof now?

To prove the Fundamental Theorem of Algebra, we will need the *Extreme Value Theorem* for real-valued functions of two real variables, which we state without proof. In particular, we formulate this theorem in the restricted case of functions defined on the *closed disk* D of radius $R > 0$ and centered at the origin, i.e., $D = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq R^2\}$.

Theorem 2 (Extreme Value Theorem)**.** *Let $f : D \to \mathbb{R}$ be a continuous function on the closed disk $D \subset \mathbb{R}^2$. Then f is bounded and attains its minimum and maximum values on D. In other words, there exist points $x_m, x_M \in D$ such that*

$$f(x_m) \leq f(x) \leq f(x_M)$$

for every possible choice of point $x \in D$.

If we define a polynomial function $f : \mathbb{C} \to \mathbb{C}$ by setting $f(z) = a_n z^n + \cdots + a_1 z + a_0$ as in Equation (1), then note that we can regard $(x, y) \mapsto |f(x + iy)|$ as a function $\mathbb{R}^2 \to \mathbb{R}$. By a mild abuse of notation, we denote this function by $|f(\,\cdot\,)|$ or $|f|$. As it is a composition of continuous functions (polynomials and the square root), we see that $|f|$ is continuous.

Lemma 3. *Let $f : \mathbb{C} \to \mathbb{C}$ be any polynomial function. Then there exists a point $z_0 \in \mathbb{C}$ where the function $|f|$ attains its minimum value in $\mathbb{R}$.*

Proof. If f is a constant polynomial function, then the statement of the Lemma is trivially true since $|f|$ attains its minimum value at every point in $\mathbb{C}$. So choose, e.g., $z_0 = 0$.

If f is not constant, then the degree of the polynomial defining f is at least 1. In this case, we can denote f explicitly as in Equation (1). That is, we set

$$f(z) = a_n z^n + \cdots + a_1 z + a_0$$

with $a_n \neq 0$. Now, assume $z \neq 0$, and set $A = \max\{|a_0|, \ldots, |a_{n-1}|\}$. We can obtain a lower bound for $|f(z)|$ as follows:

$$\begin{aligned} |f(z)| &= |a_n|\,|z|^n \left|1 + \frac{a_{n-1}}{a_n}\frac{1}{z} + \cdots + \frac{a_0}{a_n}\frac{1}{z^n}\right| \\ &\geq |a_n|\,|z|^n \left(1 - \frac{A}{|a_n|}\sum_{k=1}^{\infty}\frac{1}{|z|^k}\right) = |a_n|\,|z|^n\Big(1 - \frac{A}{|a_n|}\frac{1}{|z| - 1}\Big). \end{aligned}$$

For all $z \in \mathbb{C}$ such that $|z| \geq 2$, we can further simplify this expression and obtain

$$|f(z)| \geq |a_n|\,|z|^n\Big(1 - \frac{2A}{|a_n||z|}\Big).$$

It follows from this inequality that there is an $R > 0$ such that $|f(z)| > |f(0)|$, for all $z \in \mathbb{C}$ satisfying $|z| > R$. Let $D \subset \mathbb{R}^2$ be the disk of radius R centered at 0, and define a function $g : D \to \mathbb{R}$, by

$$g(x, y) = |f(x + iy)|.$$

Then g is continuous, and so we can apply Theorem 2 in order to obtain a point $(x_0, y_0) \in D$ such that g attains it minimum at (x_0, y_0). By the choice of R we have that for $z \in \mathbb{C} \setminus D$, $|f(z)| > |g(0, 0)| \geq |g(x_0, y_0)|$. Therefore, $|f|$ attains its minimum in $z = x_0 + iy_0$. □

We now prove the Fundamental Theorem of Algebra.

Proof of Theorem 1. For our argument, we rely on the fact that the function $|f|$ attains its minimum value by Lemma 3. Let $z_0 \in \mathbb{C}$ be a point where the minimum is attained. We will show that if $f(z_0) \neq 0$, then z_0 is *not* a minimum, thus proving by contraposition that the minimum value of $|f(z)|$ is zero. Therefore, $f(z_0) = 0$.

If $f(z_0) \neq 0$, then we can define a new function $g : \mathbb{C} \to \mathbb{C}$ by setting

$$g(z) = \frac{f(z + z_0)}{f(z_0)}, \text{ for all } z \in \mathbb{C}.$$

Note that g is a polynomial of degree n, and that the minimum of $|f|$ is attained at z_0 if and only if the minimum of $|g|$ is attained at $z = 0$. Moreover, it is clear that $g(0) = 1$.

More explicitly, g is given by a polynomial of the form

$$g(z) = b_n z^n + \cdots + b_k z^k + 1,$$

with $n \geq 1$ and $b_k \neq 0$, for some $1 \leq k \leq n$. Let $b_k = |b_k| e^{i\theta}$, and consider z of the form

$$z = r|b_k|^{-1/k} e^{i(\pi - \theta)/k}, \tag{2}$$

with $r > 0$. For z of this form we have

$$g(z) = 1 - r^k + r^{k+1} h(r),$$

where h is a polynomial. Then, for $r < 1$, we have by the triangle inequality that

$$|g(z)| \leq 1 - r^k + r^{k+1} |h(r)|.$$

For $r > 0$ sufficiently small we have $r|h(r)| < 1$, by the continuity of the function $rh(r)$ and the fact that it vanishes in $r = 0$. Hence

$$|g(z)| \leq 1 - r^k (1 - r|h(r)|) < 1,$$

for some z having the form in Equation (2) with $r \in (0, r_0)$ and $r_0 > 0$ sufficiently small. But then the minimum of the function $|g| : \mathbb{C} \to \mathbb{R}$ cannot possibly be equal to 1. □

We now conclude these notes with several more fundamental facts about polynomials, including an equivalence form of the Fundamental Theorem of Algebra. While these facts should be familiar to you, they nonetheless require careful formulation and proof.

Theorem 4. *Given a positive integer $n \geq 1$ and any choice of coefficients $a_0, a_1, \ldots, a_n \in \mathbb{C}$, where $a_n \neq 0$, define the function $f : \mathbb{C} \to \mathbb{C}$ by setting*

$$f(z) = a_n z^n + \cdots + a_1 z + a_0, \forall\, z \in \mathbb{C}.$$

In other words, f is a polynomial function of degree n. Then

1. *given any complex number $w \in \mathbb{C}$, we have that $f(w) = 0$ if and only if there exists a polynomial function $g : \mathbb{C} \to \mathbb{C}$ of degree $n-1$ such that*

$$f(z) = (z - w)g(z), \forall\, z \in \mathbb{C}.$$

2. *there are at most n distinct complex numbers w for which $f(w) = 0$. In other words, f has at most n distinct roots.*

3. *(Fundamental Theorem of Algebra, restated) there exist exactly $n+1$ complex numbers $w_0, w_1, \ldots, w_n \in \mathbb{C}$ (not necessarily distinct) such that*

$$f(z) = w_0(z - w_1)(z - w_2)\cdots(z - w_n), \forall\, z \in \mathbb{C}.$$

   In other words, every polynomial function with coefficients over $\mathbb{C}$ can be factored into linear factors over $\mathbb{C}$.

Proof.

1. Let $w \in \mathbb{C}$ be a complex number.

   ("$\Longrightarrow$") Suppose that $f(w) = 0$. Then, in particular, we have that

$$a_n w^n + \cdots + a_1 w + a_0 = 0.$$

   Since this equation is equal to zero, it follows that, given any $z \in \mathbb{C}$,

$$\begin{aligned} f(z) &= a_n z^n + \cdots + a_1 z + a_0 - (a_n w^n + \cdots + a_1 w + a_0) \\ &= a_n(z^n - w^n) + a_{n-1}(z^{n-1} - w^{n-1}) + \cdots + a_1(z - w) \\ &= a_n(z - w)\sum_{k=0}^{n-1} z^k w^{n-1-k} + a_{n-1}(z - w)\sum_{k=0}^{n-2} z^k w^{n-2-k} + \cdots + a_1(z - w) \\ &= (z - w)\sum_{m=1}^{n}\left(a_m \sum_{k=0}^{m-1} z^k w^{m-k}\right). \end{aligned}$$

Thus, upon setting

$$g(z) = \sum_{m=1}^{n} \left(a_m \sum_{k=0}^{m-1} z^k w^{m-k} \right), \forall\, z \in \mathbb{C},$$

we have constructed a degree $n-1$ polynomial function g such that

$$f(z) = (z-w)g(z), \forall\, z \in \mathbb{C}.$$

("$\Longleftarrow$") Suppose that there exists a polynomial function $g : \mathbb{C} \to \mathbb{C}$ of degree $n-1$ such that

$$f(z) = (z-w)g(z), \forall\, z \in \mathbb{C}.$$

Then it follows that $f(w) = (w-w)g(w) = 0$, as desired.

2. We use induction on the degree n of f.

   If $n = 1$, then $f(z) = a_1 z + a_0$ is a linear function, and the equation $a_1 z + a_0 = 0$ has the unique solution $z = -a_0/a_1$. Thus, the result holds for $n = 1$.

   Now, suppose that the result holds for $n-1$. In other words, assume that every polynomial function of degree $n-1$ has at most $n-1$ roots. Using the Fundamental Theorem of Algebra (Theorem 1), we know that there exists a complex number $w \in \mathbb{C}$ such that $f(w) = 0$. Moreover, from Part 1 above, we know that there exists a polynomial function g of degree $n-1$ such that

   $$f(z) = (z-w)g(z), \forall\, z \in \mathbb{C}.$$

   It then follows by the induction hypothesis that g has at most $n-1$ distinct roots, and so f must have at most n distinct roots.

3. This part follows from an induction argument on n that is virtually identical to that of Part 2, and so the proof is left as an exercise to the reader.

□

References

[1] B. Fine and G. Rosenberger. *The Fundamental Theorem of Algebra*. Undergraduate Texts in Mathematics. Springer-Verlag New York, 1997.

[2] C. F. Gauss. *New Proof of the Theorem That Every Algebraic Rational Integral Function In One Variable can be Resolved into Real Factors of the First or the Second Degree*. Ph.D. thesis, University of Helmstedt, 1799. Available online at `http://www.fsc.edu/library/documents/Theorem.pdf`.