2016

An analysis of Charles Fefferman's proof of the Fundamental Theorem of Algebra

Kyle O. Linford
Eastern Michigan University

An analysis of Charles Fefferman's proof of the Fundamental Theorem of Algebra

Abstract
Many peoples' first exploration into more rigorous and formalized mathematics is with their early explorations in algebra. Much of their time and effort is dedicated to finding roots of polynomials-a challenge that becomes more increasingly difficult as the degree of the polynomials increases, especially if no real number roots exist. The Fundamental Theorem of Algebra is used to show that there exists a root, particularly a complex root, for any nth degree polynomial. After struggling to prove this statement for over 3 centuries, Carl Friedrich Gauss offered the first fairly complete proof of the theorem in 1799. Further proofs of the theorem were later developed, which included the short proof by contradiction of Charles Fefferman. First published in the American Mathematical Monthly in 1967, this complete proof offers a fairly elementary explanation that only requires an undergraduate understanding of Real Analysis to work through.

This project is a proof analysis of Fefferman's proof for the Fundamental Theorem of Algebra. In this analysis, rigorous detail of the proof is offered as well as an explanation of the purpose behind certain sections and how they help to show the existence of a complex root for nth degree polynomials. It is the goal of this project to work with Fefferman's proof to develop a clearer explanation of the theorem and how it is able to show this property.

Degree Type
Open Access Senior Honors Thesis

Department
Mathematics

First Advisor
Dr. Kenneth Shiskowski

Second Advisor
Dr. Carla Tayeh

Keywords
Fundamental Theorem of Algebra, Fefferman, Algebra, Polynomial, Proof, Complex numbers

Subject Categories
Mathematics

AN ANALYSIS OF CHARLES FEFFERMAN'S PROOF OF THE FUNDAMENTAL THEOREM OF ALGEBRA

By

Kyle O. Linford

A Senior Thesis Submitted to the

Eastern Michigan University

Honors College

in Partial Fulfillment of the Requirements for Graduation

with Honors in Mathematics

Approved at Ypsilanti, Michigan, on this date December 6, 2016

Table of Contents

Abstract

Many peoples' first exploration into more rigorous and formalized mathematics is with their early explorations in algebra. Much of their time and effort is dedicated to finding roots of polynomials—a challenge that becomes more increasingly difficult as the degree of the polynomials increases, especially if no real number roots exist. The Fundamental Theorem of Algebra is used to show that there exists a root, particularly a complex root, for any nth degree polynomial. After struggling to prove this statement for over 3 centuries, Carl Friedrich Gauss offered the first fairly complete proof of the theorem in 1799. Further proofs of the theorem were later developed, which included the short proof by contradiction of Charles Fefferman. First published in the American Mathematical Monthly in 1967, this complete proof offers a fairly elementary explanation that only requires an undergraduate understanding of Real Analysis to work through.

This project is a proof analysis of Fefferman's proof for the Fundamental Theorem of Algebra. In this analysis, rigorous detail of the proof is offered as well as an explanation of the purpose behind certain sections and how they help to show the existence of a complex root for nth degree polynomials. It is the goal of this project to work with Fefferman's proof to develop a clearer explanation of the theorem and how it is able to show this property.

Statement of Theorem: Let $P(z) = a_0 + a_1 z^1 + \cdots + a_n z^n$ be a complex polynomial. Then P has a zero.

Explanation of the Theorem: This version of the theorem suggests that for a polynomial $P(z)$ with variable z and coefficients $(a_0, a_1, \ldots, a_n)$ in the complex number field, there exists a root such that $0 = a_0 + a_1 z^1 + \cdots + a_n z^n$ can be evaluated.

Proof Analysis:

Hypothesis

- Let $P(z) = a_0 + a_1 z^1 + \cdots + a_n z^n$ be a complex polynomial so that $(a_0, a_1, \ldots, a_n) \in \boldsymbol{C}$ the complex field.
- We will show that P has a zero.
- By contradiction, assume that $P(z)$ has no root in $\boldsymbol{C}$.

Part 1: We will show that $|P(z)|$, the modulus of $P(z)$, attains a minimum as z varies over the entire complex plane.

- By our assumption that $P(z)$ has no root in $\boldsymbol{C}$, it can be said that $P(z)$ is a positive degree polynomial in $\boldsymbol{C}[z]$.
- It can be said that $\lim_{|z|\to\infty} |P(z)| = \infty$ for z, a variable with values in $\boldsymbol{C}$.
    - It is known from calculus that $\lim_{x\to\infty} \frac{1}{x^k} = 0$ for any positive integer K.
    - Additionally, $\lim_{|x|\to\infty} |x^k| = \infty$ and $\lim_{|x|\to\infty} \frac{1}{|x^k|} = \lim_{|x|\to\infty} \left|\frac{1}{x^k}\right| = 0$ for any positive integer K.
    - Taking $\lim_{|z|\to\infty} |P(z)|$, substitution can be used to say $\lim_{|z|\to\infty} |P(z)| = \lim_{|z|\to\infty} |a_n z^n + a_{n-1} z^{n-1} + \cdots + a_2 z^2 + a_1 z^1 + a_0|$.

- Factoring out $|a_n z^n|$, the limit can be rewritten as

  $\lim_{|z|\to\infty}|P(z)| = \lim_{|z|\to\infty}|a_n z^n| * \left|1 + \frac{a_{n-1}}{a_n} * \frac{1}{z} + \cdots + \frac{a_2}{a_n} * \frac{1}{z^{n-2}} + \frac{a_1}{a_n} * \frac{1}{z^{n-1}} + \frac{a_0}{a_n} * \frac{1}{z^n}\right|$.

- Rewriting the limit again, it can be said that $\lim_{|z|\to\infty}|P(z)|$

  $= \lim_{|z|\to\infty}|a_n| * |z|^n * \left|1 + \frac{a_{n-1}}{a_n} * \frac{1}{z} + \cdots + \frac{a_2}{a_n} * \frac{1}{z^{n-2}} + \frac{a_1}{a_n} * \frac{1}{z^{n-1}} + \frac{a_0}{a_n} * \frac{1}{z^n}\right|$.

- As previously shown with $\lim_{x\to\infty} \frac{1}{x^k} = 0$, the portion of the limit,

  $1 + \frac{a_{n-1}}{a_n} * \frac{1}{z} + \cdots + \frac{a_2}{a_n} * \frac{1}{z^{n-2}} + \frac{a_1}{a_n} * \frac{1}{z^{n-1}} + \frac{a_0}{a_n} * \frac{1}{z^n}$, can be shown to equal 1 since all other terms involve $\frac{1}{z^i}$ approaching 0 for all $0 \le i \le n$.

- Additionally, $\lim_{|z|\to\infty} |a_n| * |z|^n = \infty$ as shown from , $\lim_{|x|\to\infty}|x^k| = \infty$.

- Therefore, $\lim_{|z|\to\infty}|P(z)|$

  $= \lim_{|z|\to\infty}|a_n| * |z|^n * \left|1 + \frac{a_{n-1}}{a_n} * \frac{1}{z} + \cdots + \frac{a_2}{a_n} * \frac{1}{z^{n-2}} + \frac{a_1}{a_n} * \frac{1}{z^{n-1}} + \frac{a_0}{a_n} * \frac{1}{z^n}\right| = \infty$.

- The limit $\lim_{|z|\to\infty}|P(z)| = \infty$ suggests that for each $r > 0$, where r corresponds to the polynomial output, there exists an $M_r > 0$, corresponding to the z values, such that for all z satisfying $|z| \ge M_r$ we have $|P(z)| \ge r$.

  - Because $P(z)$ is a positive degree polynomial in $\mathbf{C}[z]$ and we assumed it to not have a root, we can say that, by this method, all output values of $P(z)$ satisfy that $P(z) > 0$.

- Let $|P(0)| = r > 0$ because we assumed that $P(z)$ is a positive degree polynomial with no roots.
- We can represent a circle centered at the origin (0,0) with a radius M_r by the term D.
- By the nature of a circle's radius, it can be said that every $z \in D$, which include all of the z values that fall within the radius of the circle, will satisfy the condition that $|z| \leq M_r$.
    - This implies that only the distance of z from the origin is less than or equal to the radius of the circle, M_r. The distance of the values for $P(z)$ they produce can be greater than, equal to, or less than the circle's radius.
- Similarly, each $z \notin D$, which includes all of the z values that fall outside of the radius of the circle, satisfy $|P(z)| \geq r = |P(0)|$ since each $z \notin D$ satisfies $|z| > M_r$.
    - This implies that all values of $P(z)$, which are output values of the polynomial, are more farther from the origin than r because as z increases, $P(z)$ will become farther away from the origin and have a distance greater than M_r.
    - This fact can be seen by noticing that the point $P(z)$ is on the line $x = z$ in the Cartesian plane. Since $|z| > M_r$ and thus outside of the circle, the value of $P(z)$ must also be outside of the circle. This implies that $P(z)$ is farther than M_r distance from the origin.

- Now take $|P|: D \rightarrow R^{\geq 0}$ where D is the circle from before.
- Because $P(z)$ is a positive degree polynomial in $\boldsymbol{C}[z]$ for $\boldsymbol{C}$ the complex numbers, $|P|: D \rightarrow R^{\geq 0}$ is a continuous function from elements inside the region and on any circle D to the non-negative real numbers $R^{\geq 0}$.
  - These elements in or on D can be expressed as $D = \{a + bI \mid |(a + bI) - (H + KI)| \leq r\}$ for r the radius of the disk with center at (H, K).
- Then by the Maximum-Minimum Theorem there are complex numbers $z_0, z_d \in D$ where $|P(z_0)|$ is the minimum value of $|P(z)|$ on the disk D and $|P(z_d)|$ is the maximum value of $|P(z)|$ on the disk D.
- This implies that there exists some complex number z_0 in D so that $|P(z_0)|$ is the minimum value of the function $|P(z)|$ for all z in D.
  - This implies that for all z in D, $|P(z_0)| \leq |P(z)|$.
- Additionally, it can be said that $|P(z_0)| \leq |P(z)|$ for $z = 0$, the center of D, since $|P(0)|$ does not have to be the minimum.
- This statement can be extended beyond all $z \notin D$ to say that $|P(z_0)| \leq |P(z)|$ for all $z \in \boldsymbol{C}$.
  - Because we have $|P(z_0)| \leq |P(0)|$ and for all $\notin D$ $|P(0)| \leq |P(z)|$, the transitive property can be used to say $|P(z_0)| \leq |P(0)| = r \leq |P(z)|$.

Part 2: We will show that if $|P(z_0)|$ is the minimum of $|P(z)|$, then $P(z_0) = 0$ meaning that $P(z_0)$ is a root of the polynomial.

- Let $Q(z)$ be a complex polynomial such that $Q(z) = P(z + z_0)$.
- By definition of a polynomial, we can write this polynomial as $Q(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_l z^l + a_0$ where $Q(z)$ has the same degree as $P(z)$, $a_0 \neq 0$ is the constant term of $Q(z)$, and $a_l z^l$ is the lowest positive order term of $Q(z)$ with $a_l \neq 0$.
- Rewriting $Q(z)$ again, we can say $Q(z) = z^{l+1} * T(z) + a_l x^l + a_0$ where $T(z)$ is another polynomial in $\boldsymbol{C}[z]$ which may be the constant 0 polynomial.
  - The inclusion of z^{l+1} is necessary to increase the terms of $T(z)$ by $l + 1$ degree since they will start lower and increase to a different amount than the degree n.
- Let w be any complex number where, when evaluated in the polynomial, the term $a_l w^l = -a_0$.
  - Solving this expression for w, it can be said that w is the Lth root of $-\frac{a_0}{a_l}$, which can be computed since $a_l \neq 0$ by how we constructed $Q(z)$.
- Define a new continuous function $f(t) = |T(t * w)|$ for all t in [0,1] where $f\colon [0,1] \to \boldsymbol{R}^{\geq 0}$.
- By the Maximum-Minimum Theorem, it can be said that there is a value $d \in [0,1]$ so that $M = f(d)$ if the maximum value of f.
- Let there be some real number ε in the interval [0,1] so that $0 < \varepsilon < 1$.
- Choose this number to be sufficiently small so that $\varepsilon * |w|^{l+1} * M < |a_0|$.

- Take $Q(\varepsilon * w)$ for $Q(z)$.
- By how $Q(z)$ was rewritten, we can evaluate

  $Q(\varepsilon * w) = (\varepsilon^{l+1} * w^{l+1}) * T(\varepsilon * w) + a_l * (\varepsilon^l * w^l) + a_0$.
- Taking the modulus, it can be said that

  $|Q(\varepsilon * w)| = |(\varepsilon^{l+1} * w^{l+1}) * T(\varepsilon * w) + a_l * (\varepsilon^l * w^l) + a_0|$
- By the Triangle Inequality, it can be said that

  $|(\varepsilon^{l+1} * w^{l+1}) * T(\varepsilon * w) + a_l * (\varepsilon^l * w^l) + a_0|$

  $\leq (\varepsilon^{l+1} * |w|^{l+1}) * |T(\varepsilon * w)| + |a_l * (\varepsilon^l * w^l) + a_0|$.
  - The Triangle Inequality shows that $\alpha, \beta \in C, |\alpha + \beta| \leq |\alpha| + |\beta|$.
  - Note that it is not necessary to take the modulus of ε^{l+1} since it is already positive from our assumption that $0 < \varepsilon < 1$.
- We can rewrite ε^{l+1} as $\varepsilon^l * \varepsilon$ since ε^{l+1} just implies that ε is just multiplied by itself l times and then an extra time after that.
- Additionally, because the entries of the polynomial are all elements of the complex field, C, and all fields have commutative and associative multiplication, the terms in $|a_l * (\varepsilon^l * w^l) + a_0|$ can be rearranged such that $|a_l * w^l * \varepsilon^l + a_0|$.
- By our assumption about w where $a_l w^l = -a_0$, we can use substitution to say that

  $|a_l * (\varepsilon^l * w^l) + a_0| = |a_l * w^l * \varepsilon^l + a_0| = |-a_0 * \varepsilon^l + a_0|$.
- Thus, it can be said that $(\varepsilon^{l+1} * |w|^{l+1}) * |T(\varepsilon * w)| + |a_l * (\varepsilon^l * w^l) + a_0|$

  $= (\varepsilon^l * \varepsilon * |w|^{l+1}) * |T(\varepsilon * w)| + |-a_0 * \varepsilon^l + a_0|$.

- Because we defined M to be the maximum value of the function produced by $f(t) = |T(t * w)|$ on the interval $[0,1]$, we can say that $M \geq |T(t * w)|$.
- Using substitution, it can be said that $(\varepsilon^l * \varepsilon * |w|^{l+1}) * |T(\varepsilon * w)| + |-a_0 * \varepsilon^l + a_0| \leq (\varepsilon^l * \varepsilon * |w|^{l+1}) * M + |-a_0 * \varepsilon^l + a_0|$.
- Similarly, by our assumption that $\varepsilon * |w|^{l+1} * M < |a_0|$, we can use substitution to say that $(\varepsilon^l * \varepsilon * |w|^{l+1}) * M + |-a_0 * \varepsilon^l + a_0| < (\varepsilon^l * |a_0|) + |-a_0 * \varepsilon^l + a_0|$
    - Note that we are able to rearrange $(\varepsilon^l * \varepsilon * |w|^{l+1}) * M$ to be $\varepsilon^l * (\varepsilon * |w|^{l+1} * M)$ because these are all entries in the complex field, and multiplication under a field is associative.
- Factoring out the $|a_0|$, it can be said that $(\varepsilon^l * \varepsilon * |w|^{l+1}) * M + |-a_0 * \varepsilon^l + a_0| < (\varepsilon^l * |a_0|) + |a_0| * (1 - \varepsilon^l)$.
- Factoring out $|a_0|$ again, it can be simplified to $|a_0| * [\varepsilon^l + (1 - \varepsilon^l)]$.
- Because the terms are elements of the complex field and addition is associative in the field, the terms can be rearranged to say $|a_0| * [\varepsilon^l + (1 - \varepsilon^l)] = |a_0| * [\varepsilon^l + (-\varepsilon^l) + 1)] = |a_0|$.
- By how we defined the polynomial $Q(z)$, it can be said that $|a_0| = |Q(0)|$.
- Additionally, this implies that $|Q(0)| = |P(0 + z_0)| = |P(z_0)|$, the minimum of the polynomial.
- It can be said that $|P(z_0)| \leq |P(\varepsilon * w + z_0)|$.
- Again, because $Q(z) = P(z + z_0)$, $|P(\varepsilon * w + z_0)| = |Q(\varepsilon * w)|$.

- Combining the expressions, it can be said that

  $|Q(\varepsilon * w)|$

  $= |(\varepsilon^{l+1} * w^{l+1}) * T(\varepsilon * w) + a_l * (\varepsilon^l * w^l) + a_0|$

  $< (\varepsilon^l * |a_0|) + |a_0| * (1 - \varepsilon^l)$

  $= |a_0|$

  $= |Q(0)|$

  $= |P(z_0)|$

  $\leq |Q(\varepsilon + w)|.$

- Thus, $|Q(\varepsilon * w)| < |Q(\varepsilon + w)|$.
- However, this cannot be true since no real number can be less than itself. This is the contradiction that is produced.
- Thus, it cannot be said that $a_0 \neq 0$.
- Therefore, $a_0 = 0$.
- Thus, $Q(0) = 0$.
- Because $Q(z) = P(z + z_0)$, it can be said that $0 = Q(0) = P(0 + z_0) = P(z_0)$.
- Thus, there exists a root for $P(z)$ in the complex field, $\boldsymbol{C}$.

  QED

Bibliography

Fefferman, C. L. (1967). An Easy Proof of the Fundamental Theorem of Algebra. *The American Mathematical Monthly*, 74(7), 854-855. Retrieved from http://www.jstor.org/stable/2315823.