Real analysis proof of Fundamental Theorem of Algebra using polynomial interlacing

Soham Basu

January 15, 2018

Abstract

The existence of a real quadratic polynomial factor, given any polynomial with real coefficients, is proven using only elementary real analysis. The aim is to provide an approachable proof to anyone familiar with the least upper bound property for real numbers, continuity and growth property of polynomials. Complex numbers naturally arise as solutions to the general real quadratic divisors of any polynomial.

1 Introduction

Polynomials are fundamental objects appearing in presence of an addition operator, multiplication operator and distributive property. Given a set equipped with these relations, any finite combination of constant numbers from the set and an independent variable using these operations gives rise to a polynomial. The fundamental question about finding the roots of these generic structures over the set of real numbers was answered by the Fundamental Theorem of Algebra (FTA). The first proofs of this theorem (Argand [1], Gauss [2]) were the very first existence proofs, which showed existence of a root by topological considerations, without specifying a closed form formula for the root or providing an algorithm to find those.

Other considerations about the closed-form solutions of polynomials led to the birth of group theory via Galois theory, which established that a general closed-form solution exists only up to the fourth degree polynomial [3].

FTA says that every polynomial with complex coefficients has a complex root. The first solutions of this theorem which are acceptable by modern standards were given by Argand [1] and Gauss [2]. However, both Argand and Gauss presumed topological arguments which were rigorously proven decades later and still needs a lot of background material to understand. Historically, FTA has generated the most number of proofs beside the Pythagorean theorem. However, most proofs of FTA start with assuming the existence of complex numbers. The complex number based proofs use lot of background theory e.g. complex analysis (Liouville [5]), compactness (Argand [1]), Jordan’s curve theorem (Gauss [2]), along with
the assumption of existence of complex numbers themselves. It was only recently showed that using Jordan’s curve theorem in Gauss’s first proof can be completely avoided by using elementary real analysis, although use of complex numbers is still necessary [4]. There are proofs which avoid the existence of complex numbers but use the concept of field extension (Gauss’s second proof [6]), which is not an elementary approach either and is intuitively almost equivalent to assuming the existence of complex numbers. There are only two existing completely real analysis based proofs, by Pukhlikov and Pushkar [7]. Although they establish that the real version of FTA can be solved without assuming complex numbers or using field extension, non-elementary topology background is necessary to follow these proofs. The present proof offers an elementary, albeit involved approach to the real version of FTA. Using only least upper bound principle, concepts of continuity, supremum and infimum, and growth properties of polynomials, it is shown that the two remainders by division of a given real polynomial by \( x^2 - ax - b \) attain zero for some \( a \in \mathbb{R} \) and \( b \in \mathbb{R} \). The elementary methodology makes this proof approachable to anyone familiar with basic real analysis. This is also a more natural way to arrive at complex numbers, by trying to factorize the constituent irreducible quadratic divisors of any real polynomial.

Symbolically, the complex version of FTA states:

Given \( c_n \in \mathbb{C} \), \( 0 \leq n \leq N \) for any natural number \( N \), \( \exists z \in \mathbb{C} \) such that

\[
\sum_{n=0}^{N} c_n z^n = 0 \tag{1}
\]

The real analysis counterpart of FTA, which is how FTA was stated originally, says that every polynomial with real coefficients is divisible by some real quadratic polynomial. This is because the complex conjugate \( \bar{z} \) of a root \( z \) of \( \sum_{n=0}^{N} r_n x^n = 0 \), \( r_n \in \mathbb{R} \) also solves the real polynomial

\[
\sum_{n=0}^{N} r_n z^n = \sum_{n=0}^{N} r_n \bar{z}^n = \sum_{n=0}^{N} r_n z^n = 0 = 0 \tag{2}
\]

Therefore \( (x - z)(x - \bar{z}) \) divides \( f(x) = \sum_{n=0}^{N} r_n x^n \). However, \( (x - z)(x - \bar{z}) = x^2 - (z + \bar{z})x + z\bar{z} \). Since \( (z + \bar{z}) \) and \( z\bar{z} \) both are real numbers, the quadratic polynomial \( x^2 - (z + \bar{z})x + z\bar{z} \) has real coefficients. This shows that the complex number version of FTA is assumed to be true, the real number version of FTA follows. Now we argue that going the other way is also possible, e.g. proving the real number version of FTA is independently possible in the context of real analysis only. Then the complex number version of FTA follows as a corollary.

2 Theorem

Every real polynomial \( f(x) = \sum_{n=0}^{N} r_n x^n \), \( r_n \in \mathbb{R} \) of degree \( N \geq 2 \) has a real quadratic divisor \( x^2 - Ax - B \), \( A \in \mathbb{R}, B \in \mathbb{R} \).
First of all, we consider monic polynomials which are irreducible. Irreducibility forces $c_0 > 0$ and $N$ to be even, since otherwise there is some real root $x_0$ and a corresponding real factor $(x - x_0)$ of the polynomial. For $N=2$ the result is obvious, $A = -r_{N-1}$ and $B = -r_{N-2}$. For $N \geq 3$, if the polynomial has two such real factors $(x - t_1)$ and $(x - t_2)$, their product $(x^2 - (t_1 + t_2)x + t_1t_2)$ gives the required real quadratic divisor. Otherwise, either the polynomial is irreducible or whatever is left by division by a single real factor is an irreducible polynomial. The main idea is to show that the remainders of an irreducible polynomial on division by $x^2 - ax - b$ are identically zero for some $a = A, b = B \in \mathbb{R}$. For $N > 2$, the division of the polynomial by $x^2 - ax - b$ leaves a linear polynomial as a remainder. The coefficients of the remainder are themselves polynomials of both $a$ and $b$. Referring the coefficient of the linear term as $P(\sum_{n=0}^N r_n x^n, a, b)$ and that of the constant term as $Q(\sum_{n=0}^N r_n x^n, a, b)$, the polynomial division by $x^2 - ax - b$ can be expressed as follows

$$\sum_{n=0}^N r_n x^n = (x^2 - ax - b) \sum_{n=0}^{N-2} r_n' x^n + xP(\sum_{n=0}^N r_n x^n, a, b) + Q(\sum_{n=0}^N r_n x^n, a, b)$$

(3)

In this format, the real number version of FTA symbolically translates into

$$\exists A, B \in \mathbb{R} \text{ such that } P(\sum_{n=0}^N r_n x^n, A, B) = 0 \text{ and } Q(\sum_{n=0}^N r_n x^n, A, B) = 0$$

(4)

### 3 Background of the proof

Equation 4 will be proven by comparing the structures of the bivariate polynomials $P(\sum_{n=0}^N r_n x^n, a, b)$ and $Q(\sum_{n=0}^N r_n x^n, a, b)$. Let us build up some background to elucidate the structure between $P(\sum_{n=0}^N r_n x^n, a, b)$ and $Q(\sum_{n=0}^N r_n x^n, a, b)$ for different degrees of $n$. $P$ and $Q$ both turn out to be linear in powers of $x$, since polynomial division is a linear operation over linear combination of polynomials.

$$P(\sum_{n=0}^N r_n x^n, a, b) = \sum_{n=0}^N r_n P(x^n, a, b)$$

$$Q(\sum_{n=0}^N r_n x^n, a, b) = \sum_{n=0}^N r_n Q(x^n, a, b)$$

(5)

Since division by $x^2 - ax - b$ is equivalent to replacing a $x^2$ by $ax + b$, we can directly obtain recursion relation of different orders of $P$ or different orders of $Q$.

$$P(x^{n+1}) = P(x^{n-1}x^2) = P(x^{n-1}(ax + b))$$

$$= P(ax^n + bx^{n-1}) = aP(x^n) + bP(x^{n-1})$$

(6)
The same relation holds for $Q$. Thus

$$
P(x^{n+1}, a, b) = aP(x^n, a, b) + bP(x^{n-1}, a, b)$$

$$
Q(x^{n+1}, a, b) = aQ(x^n, a, b) + bQ(x^{n-1}, a, b)
$$

(7)

Are $P(x^n, a, b)$ and $Q(x^n, a, b)$ interconnected? Let us see for the first few $n$ and make a guess for general $n$.

Table 1: Formula for $P(x^n, a, b)$ and $Q(x^n, a, b)$ for first few $n$

| $n$ | $P(x^n, a, b)$ | $Q(x^n, a, b)$ |
|-----|----------------|----------------|
| 0   | 0              | 1              |
| 1   | 1              | 0              |
| 2   | $a$           | $b$            |
| 3   | $a^2 + b$     | $ab$           |

Consequently we guess that $Q(x^{n+1}, a, b) = bP(x^n, a, b)$. This definitely holds for $n = 0, 1, 2$. We assume that this hold up to some $\bar{n}$. Then by equation 7,

$$
Q(x^{\bar{n}+1}, a, b) = aQ(x^{\bar{n}}, a, b) + bQ(x^{\bar{n}-1}, a, b) = abP(x^{\bar{n}-1}, a, b) + b^2P(x^{\bar{n}-2}, a, b)
$$

Thus by induction we prove that $Q(x^{n+1}, a, b) = bP(x^n, a, b)$ for all $n$.

Since $Q(x^n, a, b)$ can be written in terms of $P(x^n, a, b)$, let us rewrite the remainders $P(\sum_{n=0}^{\infty} r_n x^n, a, b)$ and $Q(\sum_{n=0}^{\infty} r_n x^n, a, b)$ in terms of $P(x^n, a, b)$

$$
P(\sum_{n=0}^{N} r_n x^n, a, b) = \sum_{n=0}^{N} r_n P(x^n, a, b)
$$

$$
Q(\sum_{n=0}^{N} r_n x^n, a, b) = b(\sum_{n=1}^{N} r_n P(x^{n-1}, a, b)) + r_0
$$

(8)

In this structure, $P(\sum_{n=0}^{\infty} r_n x^n, a, b)$ and $Q(\sum_{n=0}^{\infty} r_n x^n, a, b)$ show a striking similarity. The structure in equation 8 gives us a rich amount of information, sufficient to build a proof for the real number version of FTA. Let us outline the proof. Figure 1 provides an illustration to aid in the outline, using an example polynomial $x^6 + 2.222378523653520 x^5 + 2.33242796386145 x^4 + 4.82229783120557 x^3 + 4.07089739997468 x^2 + 5.69672043788538 x + 6.38111082658457$ which has no real roots. Here

$$
r_6 = 1.000000000000000
$$

$$
r_5 = 2.222378523653520
$$

$$
r_4 = 2.33242796386145
$$

$$
r_3 = 4.82229783120557
$$

$$
r_2 = 4.07089739997468
$$

$$
r_1 = 5.69672043788538
$$

$$
r_0 = 6.38111082658457
$$
Figure 1: Plots of $P(\sum_{n=0}^{N} r_n x^n, a, b) = 0$ (red) and $Q(\sum_{n=0}^{N} r_n x^n, a, b) = 0$ (green), for the particular example polynomial mentioned in the text.

**Claim 1** There exists a $b_0 < 0$ (for example $b_0 = -4$ in figure 1), such that for any fixed $b < b_0$, two points are satisfied

- $P(\sum_{n=0}^{N} r_n x^n, a, b)$ and $Q(\sum_{n=0}^{N} r_n x^n, a, b)$ have $N - 1$ and $N - 2$ number of roots respectively in the variable $a$.
- The roots of $P$ and $Q$ alternate each other.

Let us call the simultaneous condition, firstly the existence of maximum number of roots in variable $a$ for a polynomial pair differing by degree 1 in $a$, and secondly alternating of the roots, as the interleaving condition.

Thus the non-interleaving set of $P(\sum_{n=0}^{N} r_n x^n, a, b)$ and $Q(\sum_{n=0}^{N} r_n x^n, a, b)$ is bounded below by all $b < b_0$.

We are considering $Q(\sum_{n=0}^{N} r_n x^n, a, b)$ instead of $Q(\sum_{n=0}^{N} r_n x^n, a, b)$ since it is a monic polynomial in $a$, which makes the analysis easier.

**Claim 2** Interleaving of $P(\sum_{n=0}^{N} r_n x^n, a, b)$ and $Q(\sum_{n=0}^{N} r_n x^n, a, b)$ can fail in two ways, either by decrease in number of roots of any of these or by failure in alternating of the roots. For $N = 4$, if $P(\sum_{n=0}^{N} r_n x^n, a, 0)$ has $N - 1 = 3$ real roots, actually $\sum_{n=0}^{N} r_n x^n$ is enforced to have four real roots by polynomial division. So $P(\sum_{n=0}^{N} r_n x^n, a, 0)$ for an irreducible polynomial must have degree less than 3. For $N > 4$, due to the term $\frac{a}{b}$ which blows up for small $b$, it turns out that $Q(\sum_{n=0}^{N} r_n x^n, a, b)$ can have only two roots for small $b$, which is less than $N - 2$ for any $N > 4$ (For example, $Q$ has only two roots at $b = -1$ in figure 1). Putting these together, we can see that interleaving fails close to $b = 0$ for irreducible polynomials. So the non-interleaving set is non-empty.
Claim 3 Putting together Claim 1 and Claim 2, and using the least upper bound principle for \( b \), we can conclude that the non-interleaving set has an infimum \( b_I \). Only a few more points are then needed to finalize the proof–

- The roots in \( a \) of \( P(\sum_{n=0}^{N} r_n x^n, a, b) \) and \( Q(\sum_{n=0}^{N} r_n x^n, a, b) \) vary continuously for \( b \in (-\infty, b_I) \). This comes from the existence of maximum number of roots and alternating of the roots.
- At \( b = b_I \), the roots in \( a \) of \( P(\sum_{n=0}^{N} r_n x^n, a, b) \) and \( Q(\sum_{n=0}^{N} r_n x^n, a, b) \) have limits. This will be proven by eliminating different cases for the limsup and liminf of the individual roots as \( b \to b_I^- \) and showing that only limsup=liminf can hold for the roots at that limit.
- Interleaving can only then fail for \( b > b_I \) by existence of a common root (by failure of alternating of roots) or by abrupt halting of a root. Arguments over continuity and sign changes will show that abrupt halting is not possible and the only way a root can halt is by merging with another one of same kind (which means a root of \( P \) must merge with another root of \( P \) if it halts at \( b = B_I \)). However, in such a case, considering continuity of the roots again leads to existence of a common root of \( P \) and \( Q \).

4 Technical proofs of claims 1,2 and 3

Let us get into the technicalities, which are quite intricate but based on elementary real analysis theory only.

4.1 Proof of claim 1

There exists a \( b_0 < 0 \), such that for any fixed \( b < b_0 \), \( P(\sum_{n=0}^{N} r_n x^n, a, b) \) and \( Q(\sum_{n=0}^{N} r_n x^n, a, b) \) have \( N - 1 \) and \( N - 2 \) number of roots respectively in the variable \( a \), which alternate.

Proof We define chain of function \( h_m \), which are bivariate polynomials of \( a \) and \( b \). This is necessary because a direct proof of interleaving of \( P \) and \( Q \), by comparing roots of the equations seems too complicated. On the other hand, as we will shortly see, it is easier to compare roots of the triplet \( h_m, h_{m+1} \) and \( h_{m+2} \). We will derive a recurrence relation between such a triplet, which provides sufficient material to compare roots.

\[
h_m(a, b) = \sum_{n=m+1}^{N} r_n P(x^{n-m}, a, b) + r_m, m = 0, 1, ..., N \quad (9)
\]

\( P \) and \( Q \) are related to \( h_m \) in the following manner

\[
P(a, b) = h_0(a, b)
\]

\[
Q(a, b) = bh_1(a, b) + c_0 = b\left(h_1(a, b) + \frac{c_0}{b}\right) \quad (10)
\]
The useful recurrence property, which is inherited from the recurrence relation $P(x^{n+1}, a, b) = aP(x^n, a, b) + bP(x^{n-1}, a, b)$ is as follows

$$h_m(a, b) = ah_{m+1}(a, b) + bh_{m+2}(a, b) + r_m$$  

(11)

This is interesting because it will be shown that interleaving of the pair $(h_{m+1}(a, b), h_{m+2}(a, b))$ for fixed $b$ leads to interleaving of $(h_m(a, b), h_{m+1}(a, b))$ for all $b < b_0$. In the end of this analysis, it will be shown that $\exists b_0 \leq \min\{b_m\}$ such that $(P(a, b), Q(a, b))$ also interleave $\forall b < b_0$.

If some $r_m = 0$, it is particularly easy to show the interleaving of $(h_m(a, b), h_{m+1}(a, b))$ for fixed $b$, given that $(h_{m+1}(a, b), h_{m+2}(a, b))$ interleave [3]. To show this, let us call the $k^{th}$ root of $h_{m+1}$ in $a$ as $\beta_{m+1, k}$, in ascending order. Thus for $1 \leq k \leq N - m - 2$, $h_{m+1}(\beta_{m+1, k}(b), b) = 0$ and $\beta_{m+1, k}(b) < \beta_{m+1, k+1}(b)$. The recurrence relation in equation 10 gives

$$h_m(\beta_{m+1, k}(b), b) = bh_{m+2}(\beta_{m+1, k}(b), b)$$  

(12)

Due to interleaving of $(h_{m+1}(a, b), h_{m+2}(a, b))$(13)

$$h_{m+2}(\beta_{m+1, k}(b), b)h_{m+2}(\beta_{m+1, k+1}(b), b) < 0$$  

(13)

Equation [12] and [13] imply that $h_m(\beta_{m+1, k}(b), b)h_m(\beta_{m+1, k+1}(b), b) < 0$. Using intermediate value theorem, which is a consequence of the least upper bound principle for real numbers applied to continuous functions, it can be seen that $h_m(a, b)$ has a root in $a \in (\beta_{m+1, k}(b), \beta_{m+1, k+1}(b))$. We show further that $h_m(a, b)$ has two more roots in $a \in (-\infty, \beta_{m+1, 1}(b))$ and $(\beta_{m+1, 1}(b), \infty)$, using growth property of polynomials. That will complete the proof of interleaving of $(h_m(a, b), h_{m+1}(a, b))$ for $r_m = 0$ in the variable $a$.

All $h_m$ are monic in $a$, which combined with the interleaving condition imply the following for $b < 0$

$$h_m(\beta_{m+1, N-m-2}(b), b) = bh_{m+2}(\beta_{m+1, N-m-2}(b), b) < 0$$

$$h_m(\beta_{m+1, 1}(b), b) = bh_{m+2}(\beta_{m+1, 1}(b), b) < 0$$ if $m$ is even

$$h_m(\beta_{m+1, 1}(b), b) = bh_{m+2}(\beta_{m+1, 1}(b), b) > 0$$ if $m$ is odd

(14)

However,

$$\lim_{a \to \infty} ah_{m+1}(a, b) \to \infty$$

$$\lim_{a \to -\infty} ah_{m+1}(a, b) \to \infty$$ if $m$ is even

$$\lim_{a \to -\infty} ah_{m+1}(a, b) \to -\infty$$ if $m$ is odd

(15)

Since $ah_{m+1}$ dominates $ah_{m+2}$ as $|a| \to \infty$, application of intermediate value theorem on equations [11], [14] and [15] shows that $h_{m+2}(a, b)$ has two further roots in $a \in (-\infty, \beta_{m+1, 1}(b))$ and $a \in (\beta_{m+1, 1}(b), \infty)$. Thus $\forall b < 0$ interleaving of $(h_{m+1}(a, b), h_{m+2}(a, b))$ in the variable $a$ gets satisfied if $r_m = 0$. 

7
The slightly more complicated case of \( r_m \neq 0 \) is solved by showing that \( \lim_{b \to -\infty} |h_{m+2}(\beta_{m+1,k}(b), b)| \to \infty \). Thus for sufficiently negative \( b \), \( |h_{m+2}(\beta_{m+1,k}(b), b)| > |r_m| \) and the argument of change of signs can be applied again, similar to the case \( r_m = 0 \).

\[
\lim_{b \to -\infty} h_{m+2}(\beta_{m+1,k}(b), b) \to \infty \text{ is shown by contradiction. Assuming this is not the case implies } \exists M > 0 \text{ and a sequence } b_i < 0 \text{ with } \lim_{i \to \infty} b_i = -\infty \text{ such that } |h_m(\beta_{m+1,k}(b_i), b_i)| < M. \text{ The limits } \lim_{b \to -\infty} h_s(\beta_{m+1,k}(b_i), b_i) \text{ can be checked } \forall s > m + 2 \text{ by using the following formula}
\]

\[
h_{s+2}(\beta_{m+1,k}(b_i), b_i) = \frac{h_s(\beta_{m+1,k}(b_i), b_i)}{b_i} + \frac{\beta_{m+1,k}(b_i)}{b_i} h_{s+1}(\beta_{m+1,k}(b_i), b_i) + \frac{r_s}{b_i}
\]

(16)

This gives rise to the inequality

\[
|h_{s+2}(\beta_{m+1,k}(b_i), b_i)| \leq \frac{h_s(\beta_{m+1,k}(b_i), b_i)}{b_i} + \frac{\beta_{m+1,k}(b_i)}{b_i} |h_{s+1}(\beta_{m+1,k}(b_i), b_i)| + \frac{r_s}{b_i}
\]

(17)

Provided \( \lim_{s \to \infty} h_s(\beta_{m+1,k}(b_i), b_i) \to 0 \) and \( h_{s+1}(\beta_{m+1,k}(b_i), b_i) \) is bounded, the quantity in the left hand of equation (17) will also have limit of zero given \( \lim_{i \to \infty} \frac{\beta_{m+1,k}(b_i)}{b_i} \to 0 \). If this holds, then symbolically \( \lim_{s \to \infty} h_{s+2}(\beta_{m+1,k}, b_i) \to 0 \).

\[
\lim_{i \to \infty} \frac{\beta_{m+1,k}(b_i)}{b_i} \to 0 \text{ can be shown by combining the following facts}
\]

- Writing \( h_{m+1}(a, b_i) = \sum_{n=0}^{N-m} c_n(b_i)a^n \), the real roots \( \beta_{m+1,k}(b_i) \) turn out to be bounded by the Fujiwara bound \( \max(|c_n(b_i)|^{-\frac{1}{n-m-n}}) \).
- However, \( c_n(b_i) \) are polynomials of \( b_i \) which has a maximum degree of \( |N-n| \). So the ratio \( \frac{\beta_{m+1,k}(b_i)}{b_i} \) can be bounded by

\[
\left| \frac{\beta_{m+1,k}(b_i)}{b_i} \right| < K \max(|b_i|^{(\frac{N-n}{2}-n-m-n)-1}) \text{ for some } K > 0
\]

\[
\implies \left| \frac{\beta_{m+1,k}(b_i)}{b_i} \right| < K|b_i|^{\alpha} \text{ for some } \alpha < 0
\]

\[
\implies \lim_{i \to \infty} \left| \frac{\beta_{m+1,k}(b_i)}{b_i} \right| \to 0
\]

(18)

So the only point left to complete the proof of \( \lim_{b \to -\infty} h_s(\beta_{m+1,k}, b_i) \to 0 \) \( \forall s > m + 2 \) is to show the validity for \( s = m + 3 \). This is clearly true according to equation (17) since \( h_{m+3}(\beta_{m+1,k}, b_i) = 0 \) and \( h_{m+2}(\beta_{m+1,k}, b_i) \) is bounded. However, putting \( s = m + 2 \) in equation (17) then leads to \( \lim_{b \to -\infty} h_{m+4}(\beta_{m+1,k}, b_i) \to 0 \). This provides sufficient starting material to apply induction and show that \( \lim_{b \to -\infty} h_s(\beta_{m+1,k}, b_i) \to 0 \) \( \forall s > m + 2 \).

Continuing this way, we arrive at a contradiction at the end of our defined chain in equation (11)

\[
h_{N-3}(\beta_{m+1,k}(b_i), b_i)) = \beta_{m+1,k}(b_i)^2 + r_{N-1}\beta_{m+1,k}(b_i) + (b + r_n)
\]

\[
h_{N-2}(\beta_{m+1,k}(b_i), b_i) = \beta_{m+1,k}(b_i) + r_{N-1}.
\]

(19)
Since \( \lim_{b \to \infty} h_{N-2}(\beta_{m+1,k}(b), b) \to 0 \) only if \( \lim_{b \to \infty} \beta_{m+1,k}(b) \to -r_{N-1} \), \( \lim_{b \to \infty} h_{N-3}(\beta_{m+1,k}(b), b) \neq 0 \). Thus the only assumption in the previous analysis, \( |h_m(\beta_{m+1,k}(b), b)| < M \) is not true. This means that for every \( m, \exists b_m < 0 \) such that \( |h_{m+2}(\beta_{m+1,k}, b)| > r_m \) \( \forall b < b_m \).

Thus we have reached a point apply intermediate value theorem to show interleaving of \( (h_m(a, b), h_{m+1}(a, b)) \) \( \forall m \), for sufficiently large negative \( b \). For example, the interleaving will hold \( \forall b < \min\{b_m\} \). Since \( \lim_{b \to -\infty} \frac{a}{b} = 0 \), the previous arguments hold for \( (P(a, b), Q(a, b)) \), \( \forall b < b_0 \), for some \( b_0 \leq \min\{b_m\} \). This proves Claim 1.

### 4.2 Proof of claim 2

The non-interleaving set in \( b \) is non-empty for \( P(a, b) \) and \( Q(a, b) \).

**Proof** We need to show that \( Q \) has less than \( N - 2 \) number of roots near \( b = 0 \), for any even \( N > 4 \). \( Q \) only has two roots for small \( b < 0 \), since \( h_1(a, b) = \frac{a}{b} \) can hold for \( a \) with very large magnitude, in two specific regions where \( h_1(a, b) \) is monotonic in \( a \). This is because for any interval of \( b \leq 0 \), let us say \( b \in [0, 1] \), the interval of \( a \) in which \( h_1(a, b) \) can be non-monotonic is bounded. This is evident from the following derivation for \( |a_1| > |a_2| \).

\[
\begin{align*}
    h_1(a_2, b_i) - h_1(a_1, b_i) &= (a_2^{N-m} - a_1^{N-m}) + \sum_{n=0}^{N-1} c_n(b_i)(a_2^n - a_1^n) \\
    \implies \left| \frac{h_1(a_2, b_i) - h_1(a_1, b_i)}{a_2 - a_1} \right| &> \sum_{j=0}^{N-1} |a_1|^j |a_2|^{N-m-1-j} \\
    &- \sum_{n=0}^{N-1} |c_n(b_i)||\sum_{j=0}^{n-1} |a_1|^j |a_2|^{n-1-j} | \\
    &= |a_2|^{N-m} \left( 1 - \sum_{n=0}^{N-1} |c_n(b_i)||a_2|^{n-N-m} \sum_{j=0}^{n-1} |a_1|^j |a_2|^{n-j} \right) \\
    &> |a_2|^{N-m} \left( 1 - \sum_{n=0}^{N-1} |c_n(b_i)||a_2|^{n-N+m} \right) \\
    &> 0 \quad \forall |a_2| > 2\max\{|nc_n(b_i)|^{\frac{1}{-m-n}}\}
\end{align*}
\]

by using the optimal Fujimura bound. \( 2\max\{|nc_n(b_i)|^{\frac{1}{-m-n}}\} \) is bounded for \( b \in [0, 1] \). The value of \( h_1(a, b) \) is also bounded by \( 2\max\{|c_n(b_i)|^{\frac{1}{-m-n}}\} \). Thus for small \( b \), \( h_1(a, b) = \frac{a}{b} \) can only hold in two monotonic regions of \( h_1(a, b) \) for sufficiently large \( |a| \). Since \( Q(a, b) \) is an even polynomial in \( a \) for \( b < 0 \) and \( \lim_{b \to 0} Q(0, b) \to -\infty \), it must have at least one negative and one positive root for small \( b < 0 \), by application of intermediate value theorem. Thus \( Q(a, b) \) must have two roots for sufficiently small \( b < 0 \).

From this analysis it is inferred that the non-interleaving set of \( b \) for \( (P, Q) \) is non-empty. This proves Claim 2.
4.3 Proof of claim 3

Consequently the non-interleaving set of $P$ and $Q$ set has a greatest lower bound, say $b_I < 0$. Let us show that the roots of $P(a,t)$ and $Q(a,t)$ vary continuously with $t < b_I$. It suffices to check the roots of $P(a,t < b_I)$, which are abbreviated $\beta_{P,k}(t)$ in increasing order with $1 \leq k \leq N - 1$ for fixed $t$. For any $0 < \epsilon < \min\{\beta_{P,k+1}(t) - \beta_{P,k}(t)/2\}$, we have $P(\beta_{P,k}(t) - \epsilon, t)P(\beta_{P,k}(t) + \epsilon, t) < 0$. By continuity of $P$, $\exists \delta > 0$, such that $P \neq 0$ inside the circles of radius $\delta$, centred around the points $(\beta_{P,k}(t) \pm \epsilon, t)$. Thus $\forall |t - t'| < \delta$ we have $P(\beta_{P,k}(t) - \epsilon, t')P(\beta_{P,k}(t) + \epsilon, t') < 0$. Immediate application of intermediate value theorem shows that $\exists \beta_{P,k}(t')$ such that $P(\beta_{P,k}(t'), t') = 0$ and $|\beta_{P,k}(t') - \beta_{P,k}(t)| < \epsilon$. This shows continuity of the roots $\forall t < b_I$.

Thus $\forall b < b_I$, $(P, Q)$ are interleaving and their roots vary continuously.

Let us now show that roots also have limits as $b \to b_I$ from the negative side. This is because there are only three possible fates of the liminf and limsup of a particular root (identified by a particular index) as $b \to b_I$, due to law of trichotomy for real numbers. The liminf and limsup exist for $b \to b_I$ by continuity.

- Either the limsup or liminf is unbounded. This cannot happen since all roots of $P(a,b \leq b_I < 0)$ and $Q(a,b \leq b_I < 0)$ are bounded (by Fujisawa bound, for example).
- limsup and liminf exists but are unequal. This cannot occur as well, since fixing $a = \gamma$ makes both $P$ and $Q$ polynomials in $b$. Then for any $\gamma$ between the limsup and liminf of a root, $P$ or $Q$ cross $\gamma$ infinite times. Otherwise $\gamma$ itself is either smaller than the liminf or bigger than the limsup, which is false by the definition of liminf and limsup. This is clearly impossible since a polynomial can only oscillate around any fixed value only a finite number of times.
- liminf=limsup and the roots have limits as $b \to b_I$ from the negative side. Since this is the only case left, it must hold.

Finally we have that the roots of $P$ and $Q$ exist $\forall b \leq b_I$ and vary continuously $\forall b < b_I$. If interleaving at $b_I$ fails due to a root of $P$ and $Q$ being equal, there is nothing more to prove. Otherwise, the interleaving fails by decrease in number of roots of $P$ or $Q$. Supposing that a root $\alpha$ does not continue beyond $b = b_I$, say for $P(\alpha, b_I) = 0$, the only possibility that exists is that it is a multiple root of even degree. Otherwise if the isolated root is of odd degree at $b = b_I$, checking the sign of $P(a,b)$ on a semicircle $\sqrt{(a - \alpha) + (b - b_I)} = \epsilon$, $b \geq b_I$ with sufficiently small $\epsilon > 0$ leads to a contradiction due to incongruity in the sign of $P(a,b)$. Let us consider the case where at least two roots of $P$ converge to the same limit at $b = b_I$, $\forall b < b_I$, there is a root of $Q$ between two such roots of $P$. By application of continuity and least upper bound principle, for some $b \leq b_I$ the specific root of $Q$ must equal one of the roots of $P$. $\square$
References

[1] Aigner, M, Ziegler, G. M, Proofs from THE BOOK, fourth Edition, Springer-Verlag, Berlin and Heidelberg, 2010

[2] Carl Friedrich Gauss. Demonstratio nova theorematis omnem functionem algebraicam rationalem integrum unius variabilis in factores reales primi vel secundi gradus resolvi posse. PhD thesis, Universität Helmstedt, 1799. In Werke III, 1–30.

[3] B. Fine and G. Rosenberger, The Fundamental Theorem of Algebra. Springer-Verlag, New York, 1997.

[4] Basu, S, Velleman, D.J. On Gauss's First Proof of the Fundamental Theorem of Algebra, to appear in American Mathematical Monthly

[5] Boas, RP, Invitation to Complex Analysis, Random House, New York, 1987

[6] Gauss, Carl Friedrich Demonstratio nova altera theorematis omnem functionem algebraicam rationalem integrum unius variabilis in factores reales primi vel secundi gradus resolvi posse. Comm. Recentiores (Gottingae), 3:107–142, 1816. In Werke III, 31–56. English translation available in [http://www.paultaylor.eu/misc/gauss.html](http://www.paultaylor.eu/misc/gauss.html)

[7] Konrad, K THE FUNDAMENTAL THEOREM OF ALGEBRA VIA PROPER MAPS url: [http://www.math.uconn.edu/~kconrad/blurbs/fundthmalg/propermaps.pdf](http://www.math.uconn.edu/~kconrad/blurbs/fundthmalg/propermaps.pdf)

[8] Sturm, Jacques Charles François (1829). Mémoire sur la résolution des équations numériques. Bulletin des Sciences de Férussac. 11: 419–425.

[9] Fisk, Steve (1829). Polynomials, roots, and interlacing. Lemma 1.87, [arXiv:math/0612833](http://arxiv.org/abs/math/0612833)