Continuity of the Roots of a Polynomial

S. Kumaresan
School of Math. and Stat.
University of Hyderabad
Hyderabad 500046
kumaresa@gmail.com

We give three proofs of the result which says that the roots of a polynomial depend "continuously" on the coefficients of the polynomial. The first proof uses Rouche's theorem. The second proof is quite elementary. The third (if written) will justify what is in the quotes and is a highbrow proof.

Theorem 1. *Let* $f(z) := \sum_{k=0}^{n} a_k z^k = a_n \prod_{j=1}^{p} (z - z_j)^{m_j}$*,* $a_n \neq 0$ *and* $g(z) := (a_0 + \varepsilon_0)z + \cdots + (a_{n-1} + \varepsilon_{n-1})z^{n-1} + a_n z^n$*. Let* $0 < r_k \leq \min_{j \neq k} |z_k - z_j|$*. Then there exists an* $\varepsilon > 0$ *such that if* $|\varepsilon_i| \leq \varepsilon$*, then* g *has precisely* m_k *zeros in* $B(z_k, r_k) =: B_k$*.*

Proof. Note that on B_k, $h := g - f$, $h(z) = \sum_{i=0}^{n-1} \varepsilon_i z^i$ satisfies $|h(z)| \leq \sum \varepsilon_i \left(|z_i| + r_i\right)^i \leq M_k \varepsilon$, where $M_k := \sum [|z_i| + r_i]^i$. But, on ∂B_k, we have

$$|f(z)| = |a_n| \prod_j |z - z_j|^{m_j} = |a_n||z - z_k|^{m_k} \prod_{j \neq k} |z - z_j|^{m_j} \geq |a_n| r_k^{m_k} \prod_{j \neq k} \left(|z_j - z_k| - r_k\right)^{m_j}.$$

Call the right hand side of the last inequality as δ_k. So, if we choose

$$\varepsilon < \min \left\{ \frac{\delta_k}{M_k}, 1 \leq k \leq n \right\},$$

we then have $|h(z)| < |f(z)|$ on ∂B_k. This means by Rouche's theorem that f and $f + h = g$ have the same number of zeroes in B_k. By our choice of r_k, the only zeroes of f in B_k is z_k with multiplicity m_k. Hence the result. □

Theorem 2. *Let*

$$\begin{aligned} f(z) &= z^n + a_{n-1} z^{n-1} + \cdots + a_0 = \prod (z - \lambda_i) \\ g(z) &= z^n + \alpha_{n-1} z^{n-1} + \cdots + \alpha_0 = \prod (z - \mu_i). \end{aligned}$$

Let λ *be a root of* f *with multiplicity* m *and* $\varepsilon > 0$ *be given. Then for* $|a_i - \alpha_i|$ *sufficiently small for each* i*,* g *has at least* m *roots within an* ε*-distance of* λ*.*

Proof. Suppose not. Then there exists a sequence $\{f_k\}$ of polynomials which converge to f such that f_k has fewer than m roots in $B(\lambda, \varepsilon)$. Since the coefficients $\{a_i^k\}_k$ converge, the set $\{(a_0^{(k)}, \dots, a_{n-1}^{(k)} \mid k \in \mathbb{N}\}$ is bounded in $\mathbb{C}^n$.

Let $f_k(z) = (x - \lambda_1^{(k)}) \cdots (x - \lambda_n^{(k)})$. Then $\{(\lambda_1^{(k)}) \dots \lambda_n^{(k)}) \mid k \in \mathbb{N}\}$ is a bounded subset of $\mathbb{C}^n$. (For, if λ is a root of f, then $|\lambda| \leq \max\{1, \sum_{i=1}^n |a_i|\}$.) It, therefore, has a convergent subsequence. Without loss of generality, assume $\lambda_i^k \to \mu_i$ for each i.

Now, since $\lambda_i^k \to \mu_i$, $f_k(z) = (z - \lambda_1^{(k)}) \cdots (z - \lambda_N^{(k)})$ converge to $h(z) = (z - \mu_1) \cdots (z - \mu_n)$. But then, by uniqueness of limits, $h(z) = f(z)$. Hence, m of μ_i's must equal λ, a contradiction. □

Remark 3. By repeated application of the theorem to all distinct roots λ_i with multiplicity m_i, we see that in the statement we can conclude g must have precisely m roots within an ε-distance of λ.

Third Proof

We assume that all polynomials P of degree n are normalized so that $P(z) = z^n + a_1 z^{n-1} + \cdots + a_n$. We identify the polynomial P with the vector $(a_1, \dots, a_n) \in \mathbb{C}^n$. By the fundamental theorem of algebra we can factorize P as follows: $P(z) = \prod_{j=1}^n (z - \xi_j)$ for a finite set of elements $\xi_j \in \mathbb{C}$. We also know that the coefficients a_j are symmetric polynomials of the roots ξ_j:

$$a_k = \sigma_k(\xi_1, \dots, \xi_n) = \sum \xi_{j_1} \cdots \xi_{j_k}$$

where the sum is over all possible k subsets $\{j_1, \dots, j_k\}$ of $\{1, \dots, n\}$. Define $\sigma \colon \mathbb{C} \to \mathbb{C}$ be the map $\sigma(\xi) := (\sigma_1(\xi), \dots, \sigma_n(\xi))$. Then σ is continuous and σ is onto, by the fundamental theorem of algebra. However, σ is not one-one.

Let S_n, the symmetric group on n symbols act on $\mathbb{C}^n$ by $\mu(z_1, \dots, z_n) := (z_{\mu(1)}, \dots, z_{\mu(n)})$. This induces an equivalence relation $\sim$ so that the equivalence classes are the orbits under the group action. Let $\mathbb{C}^n/\sim$ be the quotient space and $\pi \colon \mathbb{C}^n \to \mathbb{C}^n/\sim$ be the quotient map. Now let $\tilde{\sigma}$ be the unique map $\tilde{\sigma} \colon \mathbb{C}^n/\sim \to \mathbb{C}^n$ such that $\tilde{\sigma} \circ \pi = \sigma$. Using the standard results from the theory of quotient spaces, one shows easily that $\tilde{\sigma}$ is a continuous bijection. (At this point, it is very tempting to think of the result which says that a bijective continuous map of a compact space onto a Hausdorff space is a homeomorphism. See Remark 6 at the end.

Proposition 4. *The map $\tilde{\sigma} \colon \mathbb{C}^n/\sim \to \mathbb{C}^n$ is a homeomorphism.*

Proof. We define a metric on $\mathbb{C}^n/\sim$ which induces the quotient topology as follows:

$$d(\pi(z), \pi(w)) := \min\{|z' - w'| : z' \in \pi(z), w' \in \pi(w)\}.$$

Note that

$$d(\pi(z), \pi(w)) := \min\{|z - w'| : w' \in \pi(w)\}.$$

Given $z, v, w \in \mathbb{C}^n$, choose $v' \in \pi(v)$ so that $d(\pi(z), \pi(v)) = |z - v'|$. For each $w' \in \pi(w)$, we then have $|z - w'| \leq |z - v'| + |v' - w'|$. Hence,

$$\begin{aligned} d(\pi(z),\pi(w)) &\leq \min\{|z-v'|+|v'-w'| : w' \in \pi(w)\} \\ &= d(\pi(z),\pi(v)) + d(\pi(v),\pi(w)). \end{aligned}$$

It follows that $\pi \colon \mathbb{C}^n \to (\mathbb{C}^n/\sim, d)$ is continuous.

We now claim that d induces the quotient topology. Since $\pi \colon \mathbb{C} \to (\mathbb{C}^n/\sim, d)$ is a continuous bijection, it is enough to show that π is open. Let $U \subset \mathbb{C}$ be open. Let $[\alpha] \in \pi(U)$. We assume without loss of generality that $\alpha \in U$. Then there exists $r > 0$ such that $B(\alpha, r) \subset U$. We claim that $B([\alpha], r) \subset \pi(U)$. Let $[\beta] \in B([\alpha], r)$. We can find $\beta \in [\beta]$ such that $d([\alpha],[\beta]) = d(\alpha,\beta)$. The result follows form this. (Alternative proof of the Claim: For any set $A \subset \mathbb{C}^n$, we have $\pi^{-1}(\pi(A))$ is open (closed) if A is open (closed). Thus π is an open and closed map. Hence the claim.)

Let $B(0,R)$ denote the open ball in $\mathbb{C}^n/\sim$. We claim that $\tilde{\sigma}$ is a homeomorphism of $\overline{B(0,R)}$ onto its image. It is enough to show that $\tilde{\sigma}$ is closed. Let $K \subset \overline{B(0,R)}$ be closed. Then $\pi^{-1}(K)$ is a closed and bounded subset of $\mathbb{C}^n$ and hence is compact. Thus, $\tilde{\sigma}(K) = \sigma(\pi^{-1}(K))$ is compact and hence closed in $\mathbb{C}^n$.

We complete the proof by showing that $\tilde{\sigma} \colon \mathbb{C}^n/\sim \to \mathbb{C}^n$ is open. Let U be open subset of $\mathbb{C}^n/\sim$ and $x \in U$. Choose $\varepsilon > 0$ and $R > 0$ such that $B(x,\varepsilon) \subset U$ and $\overline{B(x,\varepsilon)} \subset B(0,R)$. Since $\tilde{\sigma}$ is open on $B(0,R)$, it follows that $\tilde{\sigma}(x)$ lies in the interior of $\tilde{\sigma}(B(x,\varepsilon))$. Since x was arbitrary, this completes the proof. $\square$

Theorem 5. *Suppose*

$$P(z) = z^n + a_1 z^{n-1} + \cdots + a_n = \prod_{j=1}^{k} (z-\xi_j)^{m_j}$$

for distinct $\xi_1, \ldots, \xi_k$. Let $\varepsilon > 0$ be given such that for $i \neq j$, we $B(\xi_i,\varepsilon) \cap B(\xi_j,\varepsilon) = \emptyset$. Then there exists $\delta > 0$ so that $b \in B(a,\delta)$ implies that the polynomial

$$Q(z) := z^n + b_1 z^{n-1} + \cdots + b_n$$

has exactly m_j roots (counted with multiplicity) in $B(\xi_j,\varepsilon)$.

Proof. Let $\tau \colon \mathbb{C}^n \to \mathbb{C}^n/\sim$ be inverse of $\tilde{\sigma}$. Given $a \in \mathbb{C}^n$, the coefficients of P, by the last proposition, there exists a $\delta > 0$ such that $\tau(B(a,\delta)) \subset B(\tau(a),\varepsilon)$. Spell this out in terms of the metrics on $\mathbb{C}^n$ and $\mathbb{C}^n/\sim$ to arrive at the theorem. $\square$

Remark 6. It is tempting to introduce some projective notions to "simplify" the proof of the proposition conceptually.