Roots and polynomials as Homeomorphic spaces

Branko Ćurgus[a], Vania Mascioni[b],*

[a]*Department of Mathematics, Western Washington University, Bellingham, WA 98225, USA*
[b]*Department of Mathematical Sciences, Ball State University, Muncie, IN 47306-0490, USA*

Abstract

We provide a unified, elementary, topological approach to the classical results stating the continuity of the complex roots of a polynomial with respect to its coefficients, and the continuity of the coefficients with respect to the roots. In fact, endowing the space of monic polynomials of a fixed degree n and the space of n roots with suitable topologies, we are able to formulate the classical theorems in the form of a homeomorphism. Related topological facts are also considered.
© 2005 Elsevier GmbH. All rights reserved.

1. Introduction

The roots of a polynomial depend continuously on its coefficients. This is probably the best known and most used perturbation theorem and, clearly, it is a continuity statement (see [3] for several historical references; also, see our final remarks in Section 6). Conversely, the coefficients depend continuously on the roots. This is essentially due to Viète's formulas; see Theorem 4.3 below. However, this second result is often formulated separately from the first, and there has been no unanimity as to the topology on the set of roots.

---

* Corresponding author.
  E-mail addresses: curgus@cc.wwu.edu (B. Ćurgus), vdm@cs.bsu.edu (V. Mascioni).

In this note we provide a metric space setting in which both of these classical continuity results can be stated as a unique homeomorphism (our main result will be Theorem 4.4) between the corresponding metric spaces of roots and polynomials. This reveals more than may be widely known about the similar topological structure of these spaces.

We only use very basic background knowledge of the topology of metric spaces for example at the level of Rudin's or Baum's classical books [1,8]. Whenever we refer to a set as a metric space we imply that a specific metric has been earlier defined on it. Each subset of a metric space is considered a metric space with the induced metric. We use the standard notation $\mathbb{N}$ for the set of positive integers, $\mathbb{R}$ for the set of real numbers, $\mathbb{C}$ for the set of complex numbers, and $i = \sqrt{-1}$. Throughout this note $n \geqslant 2$ is a fixed positive integer. We study complex monic polynomials of order n and we consider all their complex roots. Since monic polynomials of degree one are in an obvious one-to-one correspondence with their unique root, the case $n = 1$ is a special, although trivial, case. Note that Theorem 5.2 and Corollary 5.3 are not true in the case $n = 1$.

2. Metric-space preliminaries

Definition 2.1. Let (X, d_X) and (Y, d_Y) be metric spaces, and let $f : X \to Y$ be a bijection. If both f and f^{-1} are continuous then f is called a *homeomorphism* between X and Y.

Definition 2.2. A subset K of a metric space (X, d_X) is said to be *compact* if for each family $\{G_\alpha\}$ of open subsets of X such that $K \subset \bigcup_\alpha G_\alpha$ there exist finitely many indices $\alpha_1, \ldots, \alpha_n$ such that

$$K \subset G_{\alpha_1} \cup \cdots \cup G_{\alpha_n}.$$

The following lemma gives a characterization of compactness in metric spaces in terms of sequences. For its proof see [8, Exercise 26, p. 45].

Lemma 2.3. *A subset K of a metric space (X, d_X) is compact if and only if each sequence in K has a subsequence which converges in K.*

Our first theorem bears a strong resemblance to the classical result that states that *a continuous bijection from a compact space to a Hausdorff space has a continuous inverse* (see [1, Theorem 3.21] or [8, Theorem 4.17], for example).

Theorem 2.4. *Let (X, d_X) and (Y, d_Y) be metric spaces and let $f : X \to Y$ be a bijection. Suppose that the following three conditions are satisfied*:

(a) *Each bounded and closed subset of X is compact.*
(b) *f is continuous.*
(c) *f^{-1} maps each bounded set in Y into a bounded set in X.*
    Then f^{-1} is continuous (and so f is a homeomorphism).

Proof. Let $\{y_k\}$ be a convergent sequence in (Y, d_Y) with limit y. Since $\{y_k\}$ is bounded, assumption (c) implies that the sequence $\{f^{-1}(y_k)\}$ is bounded in X and thus it is contained in a closed and bounded subset of X. Recall that in a metric space if a set is bounded, that is, if it has a finite diameter, then its closure has the same diameter. Therefore, by (a) and Lemma 2.3, $\{f^{-1}(y_k)\}$ has a convergent subsequence. If $\{f^{-1}(y_{m_k})\}$ is an arbitrary convergent subsequence of $\{f^{-1}(y_k)\}$ with, say,

$$\lim_{k\to\infty} f^{-1}(y_{m_k}) = x,$$

the continuity of f (assumption (b)) implies that

$$\lim_{k\to\infty} y_{m_k} = f(x) = \lim_{k\to\infty} y_k = y.$$

Thus, each convergent subsequence of the bounded sequence $\{f^{-1}(y_k)\}$ converges to the same element $f^{-1}(y)$, and this implies that $\{f^{-1}(y_k)\}$ converges to $f^{-1}(y)$. Since the sequence $\{y_k\}$ was an arbitrary convergent sequence in Y, the theorem is proved. $\square$

Proposition 2.5. *If each bounded and closed subset of a metric space (X, d_X) is compact, then (X, d_X) is complete.*

Proof. Each Cauchy sequence in a metric space is bounded and thus contained in a closed ball. Since by assumption a closed ball in (X, d_X) is compact, each Cauchy sequence in (X, d_X) has a convergent subsequence. Consequently, each Cauchy sequence in (X, d_X) converges. $\square$

We denote by $\mathbb{C}^n$ the set of all ordered n-tuples of complex numbers. We equip this space with what is called the "supremum norm"

$$\|\mathbf{v}\|_\infty = \max_{1\leqslant j\leqslant n} |v_j| \quad \text{for } \mathbf{v} = (v_1, \ldots, v_n) \in \mathbb{C}^n$$

and, for $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$, $\mathbf{u} = (u_1, \ldots, u_n)$, $\mathbf{v} = (v_1, \ldots, v_n)$, the corresponding metric

$$d_\infty(\mathbf{u}, \mathbf{v}) = \max_{1\leqslant j\leqslant n} |u_j - v_j| = \|\mathbf{u} - \mathbf{v}\|_\infty.$$

The following proposition is well known and not difficult to prove.

Proposition 2.6 (*Heine–Borel*). *$(\mathbb{C}^n, d_\infty)$ is a metric space. A subset of $(\mathbb{C}^n, d_\infty)$ is compact if and only if it is bounded and closed.*

The metric d_∞ on $\mathbb{C}^n$ is chosen for convenience only. Clearly, it can be replaced with any other equivalent metric.

Next, we prove a topological property of the space $(\mathbb{C}^n, d_\infty)$ which we shall need in Section 5.

Definition 2.7. Let (X, d_X) be a topological space. A subset S of X is *pathwise connected* if for each $u, v \in S$ there exists a continuous function $\Theta : [0, 1] \to X$ such that

$\Theta(0) = u$, $\Theta(1) = v$, and the range of Θ is a subset of S. The range of the function Θ is called a *path* from u to v which is contained in S.

Lemma 2.8. *Let $\mathscr{D}$ be the subset of $\mathbb{C}^n$ consisting of all n-tuples of distinct complex numbers. Then $\mathscr{D}$ is an open pathwise-connected subset of $(\mathbb{C}^n, d_\infty)$.*

Proof. Given the continuous function $f(z_1, \ldots, z_n) = \prod_{i \neq j}(z_i - z_j)$ between $\mathbb{C}^n$ and $\mathbb{C}$, we can write $\mathscr{D} = f^{-1}(\mathbb{C}\backslash\{0\})$, and since $\mathbb{C}\backslash\{0\}$ is open in $\mathbb{C}$, $\mathscr{D}$ must be open in $\mathbb{C}^n$.

To prove that $\mathscr{D}$ is pathwise connected, we let $\mathbf{v} = (v_1, \ldots, v_n)$ and $\mathbf{w} = (w_1, \ldots, w_n)$ be two points in $\mathscr{D}$ and construct a path from $\mathbf{v}$ to $\mathbf{w}$ which is contained in $\mathscr{D}$.

First consider a special case. Assume that there exists $k \in \{1, \ldots, n\}$ such that $v_j = w_j$ for all $j \in \{1, \ldots, n\}\backslash\{k\}$ and $v_k \neq w_k$. Since the numbers $v_k, w_k, v_1, \ldots, v_{k-1}, v_{k+1}, \ldots, v_n$ are mutually distinct, it is not difficult to construct a continuous function $\phi : [0, 1] \to \mathbb{C}$ such that $\phi(0) = v_k$, $\phi(1) = w_k$ and none of the numbers $v_1, \ldots, v_{k-1}, v_{k+1}, \ldots, v_n$ is in the range of ϕ. Consequently, the function

$$\Theta(t) = (v_1, \ldots, v_{k-1}, \phi(t), v_{k+1}, \ldots, v_n), \quad t \in [0, 1],$$

is a path from $\mathbf{v}$ to $\mathbf{w}$ which is contained in $\mathscr{D}$.

Now consider the general case of arbitrary points $\mathbf{v} = (v_1, \ldots, v_n)$ and $\mathbf{w} = (w_1, \ldots, w_n)$ in $\mathscr{D}$. Let $\mathbf{u} = (u_1, \ldots, u_n) \in \mathscr{D}$ be such that

$$\{u_1, \ldots, u_n\} \cap \{v_1, \ldots, v_n, w_1, \ldots, w_n\} = \emptyset.$$

Consider the following sequence of points in $\mathscr{D}$:

$$(v_1, v_2, v_3, \ldots, v_{n-1}, v_n), (u_1, v_2, v_3, \ldots, v_{n-1}, v_n),$$
$$(u_1, u_2, v_3, \ldots, v_{n-1}, v_n), \ldots, (u_1, u_2, u_3, \ldots, u_{n-1}, v_n),$$
$$(u_1, u_2, u_3, \ldots, u_{n-1}, u_n), (w_1, u_2, u_3, \ldots, u_{n-1}, u_n),$$
$$(w_1, w_2, u_3, \ldots, u_{n-1}, u_n), \ldots, (w_1, w_2, w_3, \ldots, u_{n-1}, u_n),$$
$$(w_1, w_2, w_3, \ldots, w_{n-1}, u_n), (w_1, w_2, w_3, \ldots, w_{n-1}, w_n).$$

The special case considered above applies to each of the $2n$ pairs of consecutive points in this sequence. It follows that for each of these pairs there exists a path contained in $\mathscr{D}$ which connects them. Since each two consecutive pairs contains a point in common, these $2n$ paths connect to a path connecting $\mathbf{v}$ and $\mathbf{w}$ which is clearly contained in $\mathscr{D}$. As $\mathbf{v}$ and $\mathbf{w}$ were arbitrary points in $\mathscr{D}$ this proves that $\mathscr{D}$ is pathwise connected. □

We denote by $\mathscr{P}_{n,1}$ the set of all monic complex polynomials of degree n. Let

$$f(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_0, \quad g(z) = z^n + b_{n-1}z^{n-1} + \cdots + b_0, \quad z \in \mathbb{C}$$

be in $\mathscr{P}_{n,1}$. Define a metric on $\mathscr{P}_{n,1}$ by

$$d_{\mathscr{P}}(f, g) := \max\{|a_0 - b_0|, \ldots, |a_{n-1} - b_{n-1}|\}. \tag{2.1}$$

Proposition 2.9. *$(\mathscr{P}_{n,1}, d_{\mathscr{P}})$ is a metric space. A subset of the metric space $(\mathscr{P}_{n,1}, d_{\mathscr{P}})$ is compact if and only if it is bounded and closed.*

Proof. The function

$$(v_1, \ldots, v_n) \longmapsto p, \quad \text{where } p(z) = z^n + v_n z^{n-1} + \cdots + v_1,$$
$$\text{and } (v_1, \ldots, v_n) \in \mathbb{C}^n,$$

is a distance-preserving bijection between the spaces $(\mathbb{C}^n, d_\infty)$ and $(\mathscr{P}_{n,1}, d_\mathscr{P})$. Therefore the proposition follows from Proposition 2.6. $\square$

3. The metric space of roots

At the end of Section 2, we introduced the metric space $(\mathscr{P}_{n,1}, d_\mathscr{P})$ of all monic polynomials of degree n. Now we define the space of sets of roots of these polynomials. Since roots can occur with finite multiplicities, instead of the set of roots of a polynomial we consider the multiset of roots, that is, we allow elements to occur with multiplicities. Denote by $\mathscr{Z}_n$ the family of all multisets of complex numbers with n elements. For multisets $U = \{u_1, \ldots, u_n\}$ and $V = \{v_1, \ldots, v_n\}$ in $\mathscr{Z}_n$, define

$$d_F(U, V) := \min_{\tau \in \Pi_n} \max_{1 \leqslant j \leqslant n} |u_j - v_{\tau(j)}|, \tag{3.1}$$

where Π_n is the set of all permutations of $\{1, \ldots, n\}$. The function d_F, which is a metric by the proposition below, is analogous to the Fréchet metric defined for curves in [2, Chapter 6]. Instead of curves here we have multisets and a function $f : \{1, \ldots, n\} \to \mathbb{C}$ is a parameterization of the multiset $\{f(k) : 1 \leqslant k \leqslant n\}$. If we denote by $\mathscr{U}$ and $\mathscr{V}$ all possible parameterizations of multisets U and V, respectively, then definition (3.1) can be rewritten as

$$d_F(U, V) = \min_{f \in \mathscr{U}, g \in \mathscr{V}} \max_{1 \leqslant k \leqslant n} |f(k) - g(k)|.$$

Proposition 3.1. *The function* $d_F : \mathscr{Z}_n \times \mathscr{Z}_n \to [0, \infty)$ *is a metric on* $\mathscr{Z}_n$.

Proof. Let $U, V, W \in \mathscr{Z}_n$. We need to prove the following three properties of d_F:

$$d_F(U, V) = 0 \Longleftrightarrow U = V, \tag{3.2}$$

$$d_F(U, V) = d_F(V, U), \tag{3.3}$$

$$d_F(U, V) \leqslant d_F(U, W) + d_F(W, V). \tag{3.4}$$

Proving (3.2) is a simple exercise. The definition of d_F can be rewritten as

$$d_F(U, V) = \min_{\sigma, \tau \in \Pi_n} \max_{1 \leqslant j \leqslant n} |u_{\sigma(j)} - v_{\tau(j)}|. \tag{3.5}$$

Since the last expression is symmetric in U and V, this shows that $d_F(U, V) = d_F(V, U)$ and thus (3.3) holds.

To prove (3.4) note that the triangle inequality for complex numbers yields

$$|u_j - v_{\tau(j)}| \leqslant |u_j - w_{\sigma(j)}| + |w_{\sigma(j)} - v_{\tau(j)}| \tag{3.6}$$

for arbitrary $j \in \{1, \ldots, n\}$ and arbitrary $\sigma, \tau \in \Pi_n$. Keeping σ and τ fixed and taking maximums with respect to $j \in \{1, \ldots, n\}$ in (3.6) we obtain

$$\max_{1 \leqslant j \leqslant n} |u_j - v_{\tau(j)}| \leqslant \max_{1 \leqslant l \leqslant n} |u_l - w_{\sigma(l)}| + \max_{1 \leqslant k \leqslant n} |w_{\sigma(k)} - v_{\tau(k)}|. \tag{3.7}$$

Keeping $\sigma \in \Pi_n$ fixed and taking the minimums of both sides in (3.7) with respect to $\tau \in \Pi_n$ we obtained

$$d_F(U, V) \leqslant \max_{1 \leqslant l \leqslant n} |u_l - w_{\sigma(l)}| + d_F(W, V)$$

and so (3.4) follows by taking the minimum of the right-hand side with respect to $\sigma \in \Pi_n$. $\square$

Next, we explore the relationship between the space $(\mathscr{L}_n, d_F)$ and the more familiar space $(\mathbb{C}^n, d_\infty)$. First we define two functions P and K.

Define $P : \mathbb{C}^n \to \mathscr{L}_n$ by

$$P((v_1, \ldots, v_n)) := \{v_1, \ldots, v_n\}, \quad (v_1, \ldots, v_n) \in \mathbb{C}^n. \tag{3.8}$$

Here an n-tuple is simply mapped to the multiset of its elements (once again, with multiplicities preserved). By the definitions of d_F and d_∞ it follows that

$$d_F(P(\mathbf{v}), P(\mathbf{w})) \leqslant \|\mathbf{v} - \mathbf{w}\|_\infty = d_\infty(\mathbf{v}, \mathbf{w}) \quad \text{for all } \mathbf{v}, \mathbf{w} \in \mathbb{C}^n. \tag{3.9}$$

Thus $P : \mathbb{C}^n \to \mathscr{L}_n$ is a contraction (and therefore a continuous function) between $(\mathbb{C}^n, d_\infty)$ and $(\mathscr{L}_n, d_F)$.

Clearly P is onto, but not one-to-one. For each $V \in \mathscr{L}_n$ the set

$$P^{-1}(V) := \{\mathbf{v} \in \mathbb{C}^n : P(\mathbf{v}) = V\}$$

has between 1 and $n!$ elements, depending on the multiplicities of the elements in V. Note that for distinct V and W in $\mathscr{L}_n$ the sets $P^{-1}(V)$ and $P^{-1}(W)$ are disjoint.

To define a partial inverse of P let $\mathscr{K}$ be a subset of $\mathbb{C}^n$ with the property that for each $V \in \mathscr{L}_n$ the set $\mathscr{K} \cap P^{-1}(V)$ has exactly one element. (In Example 3.5 below we give a specific example of a set $\mathscr{K}$ with this property.) This assumption is equivalent to the requirement that the restriction

$$P|_{\mathscr{K}} : \mathscr{K} \to \mathscr{L}_n$$

of P onto $\mathscr{K}$ is a bijection. In this way to each $V = \{v_1, \ldots, v_n\} \in \mathscr{L}_n$ we associate a unique n-tuple $(v_1, \ldots, v_n) \in \mathbb{C}^n$ that has exactly the elements of V as coordinates. Now define the function $K : \mathscr{L}_n \to \mathbb{C}^n$ by

$$K := (P|_{\mathscr{K}})^{-1}. \tag{3.10}$$

As an immediate consequence of the definitions we conclude that $P \circ K$ is the identity on $\mathscr{L}_n$.

Let $O \in \mathscr{L}_n$ be the multiset consisting of n zeros. By the definitions of d_F and K it follows that

$$d_F(V, O) = \|K(V)\|_\infty \quad \text{for all } V \in \mathscr{L}_n. \tag{3.11}$$

Proposition 3.2. *Let $\{V_k\}$ be a sequence in $(\mathscr{L}_n, d_F)$. The following statements are equivalent.*

(a) *The sequence $\{V_k\}$ is bounded in $(\mathscr{L}_n, d_F)$.*
(b) The set $\bigcup_{k=1}^\infty V_k$ *of complex numbers is bounded in* $\mathbb{C}$.
(c) *The sequence $\{K(V_k)\}$ is bounded in $(\mathbb{C}^n, d_\infty)$.*

Proof. Let $\{V_k\}$ be a bounded sequence in $(\mathscr{L}_n, d_F)$. Since $\{V_k\}$ is bounded there exists $M > 0$ such that

$$d_F(O, V_k) < M \quad \text{for all } k \in \mathbb{N}. \tag{3.12}$$

By (3.11) (b) follows trivially, and just as trivially (b) implies (c). If (c) holds, then (3.11) implies that the sequence $\{d_F(V_k, O)\}$ is bounded, and thus (a) follows just as easily. $\square$

In a similar way (3.11) can be used to prove the following proposition.

Proposition 3.3. *The function $K : \mathscr{L}_n \to \mathbb{C}^n$ maps each bounded set in $(\mathscr{L}_n, d_F)$ to a bounded set in $(\mathbb{C}^n, d_\infty)$.*

The continuity of K is discussed in Section 5 (see, in particular, Corollary 5.3).

Theorem 3.4. *A subset of the metric space $(\mathscr{L}_n, d_F)$ is compact if and only if it is bounded and closed.*

Proof. Let $\mathscr{V}$ be an arbitrary bounded and closed subset of $\mathscr{L}_n$. To prove that $\mathscr{V}$ is compact we shall prove that an arbitrary sequence $\{V_k\}$ in $\mathscr{V}$ has a convergent subsequence. By Proposition 3.2 the sequence $\{K(V_k)\}$ is bounded in $(\mathbb{C}^n, d_\infty)$. By the Bolzano–Weierstrass Theorem there exists a subsequence $\{V_{m_k}\}$ of $\{V_k\}$ such that $\{K(V_{m_k})\}$ converges, say, to the n-tuple $\mathbf{w} = (w_1, \ldots, w_n)$, in $(\mathbb{C}^n, d_\infty)$. Since $P : \mathbb{C}^n \to \mathscr{L}_n$ is continuous and $P \circ K$ is the identity on $\mathscr{L}_n$, it follows that $\{V_{m_k}\}$ converges to $P(\mathbf{w})$ in $(\mathscr{L}_n, d_F)$. Since $\mathscr{V}$ is closed $P(\mathbf{w}) \in \mathscr{V}$, and thus $\mathscr{V}$ is compact by Lemma 2.3. Since the converse is true in each metric space the theorem is proved. $\square$

In the next two examples we use the notion of lexicographic ordering in $\mathbb{C}$. Let $a, b, c, d \in \mathbb{R}$. For two complex numbers $a + ib$ and $c + id$ the lexicographic ordering $a + ib \preccurlyeq c + id$ is defined by

$$a + ib \preccurlyeq c + id \iff [(a < c) \vee (a = c \wedge b \leqslant d)].$$

Example 3.5. Let $\mathscr{L}_n$ be the subset of $\mathbb{C}^n$ defined by

$$\mathscr{L}_n := \{(z_1, \ldots, z_n) \in \mathbb{C}^n : z_1 \preccurlyeq z_2 \preccurlyeq \cdots \preccurlyeq z_n\}.$$

Since $\preccurlyeq$ is a total order on $\mathbb{C}$, for each $V \in \mathscr{Z}_n$ the set $P^{-1}(V) \cap \mathscr{L}_n$ has exactly one element. Note that the set $\mathscr{L}_n$ is not closed in $(\mathbb{C}^n, d_\infty)$. To show this consider the sequence $\{(-1/k+i, 1/k-i)\}_{k=1}^\infty$ in $\mathscr{L}_2$ which converges to $(i, -i)$ in $(\mathbb{C}^2, d_\infty)$. Clearly $(i, -i) \notin \mathscr{L}_2$. Thus $\mathscr{L}_2$ is not closed.

Example 3.6. Define the function $L : \mathscr{Z}_n \to \mathbb{C}^n$ by $L := (P|_{\mathscr{L}_n})^{-1}$, where $\mathscr{L}_n$ was defined in Example 3.5. Thus $L(V) = (v_1, \ldots, v_n)$, where $v_1 \preccurlyeq v_2 \preccurlyeq \cdots \preccurlyeq v_n$ and $V = \{v_1, \ldots, v_n\}$. We remark that the operator L is not continuous. To show this we use the same sequence as in Example 3.5 and note that

$$d_F(P((-1/k+i, 1/k-i)), P((-i, i))) = 1/k \to 0 \quad (k \to \infty).$$

Remark 3.7. With a different total order on $\mathbb{C}$, for example,

$$z \preccurlyeq w \iff [(|z| < |w|) \vee (|z| = |w| \wedge \arg(z) \leqslant \arg(w))],$$

the reader can create examples similar to Examples 3.5 and 3.6 (with the same negative conclusions).

The multiplicities of roots play an important role in the classical statement of the continuity of roots of polynomials. The following proposition clarifies the relation between the metric d_F and the multiplicity of the elements in a particular multiset in $\mathscr{Z}_n$.

Proposition 3.8. *Let $V \in \mathscr{Z}_n$ be arbitrary. Let $v_1, \ldots, v_k$ be all the distinct elements of V and let $m_1, \ldots, m_k$ be their respective multiplicities as elements of V, so that $m_1 + \cdots + m_k = n$. Put*

$$\eta(V) := \begin{cases} \frac{1}{2} \min\{|v_j - v_l|, \ j \neq l, \ j, l \in \{1, \ldots, k\}\} & \text{for } k > 1, \\ 1 & \text{for } k = 1. \end{cases}$$

Then for each $U \in \mathscr{Z}_n$, such that $d_F(V, U) < \eta(V)$ we have that each disk $D(v_j, \eta(V))$, $j = 1, \ldots, k$, in the complex plane contains exactly m_j elements of U counted according to their multiplicities in U.

Proof. Let $U \in \mathscr{Z}_n$ be such that $d_F(V, U) < \eta(V)$. Without loss of generality, let us consider the situation around v_1. Let $\sigma \in \Pi_n$ be such that $\sigma(1) = 1$ and $v_{\sigma(j)} = v_{\sigma(1)} = v_1$, $j = 1, \ldots, m_1$. By the definition of $d_F(V, U)$, see also (3.5), there exists a permutation $\tau \in \Pi_n$ such that

$$\max_{1 \leqslant j \leqslant m_1} |v_{\sigma(1)} - u_{\tau(j)}| < \eta(V).$$

Therefore all the elements $u_{\tau(j)}$, $j = 1, \ldots, m_1$, of U lie in the disk $D(v_1, \eta(V))$. Clearly, a similar statement holds for all the other v_j and since the disks $D(v_j, \eta(V))$, $j = 1, \ldots, k$, are disjoint by the definition of $\eta(V)$, the proposition is proved. $\quad\square$

4. Continuity

In this section we prove that the function $Z : \mathscr{P}_{n,1} \rightarrow \mathscr{Z}_n$ which assigns to each polynomial $p \in \mathscr{P}_{n,1}$ the multiset of its roots $Z(p) \in \mathscr{Z}_n$ is a homeomorphism between the corresponding metric spaces.

The next theorem is the classical Cauchy inequality. Cauchy's result is restated in terms of the metrics introduced above to emphasize its topological meaning. We reproduce the simple proof of this fact as found in Marden's book [3, Theorem 27.2].

Theorem 4.1 (*Cauchy's Inequality*). *Define $e_n \in \mathscr{P}_{n,1}$ by $e_n(z) := z^n$, $z \in \mathbb{C}$, and for any $p \in \mathscr{P}_{n,1}$ let $Z(p) \in \mathscr{Z}_n$ be the multiset of the roots of p. Then for an arbitrary polynomial $p \in \mathscr{P}_{n,1}$ we have*

$$d_F(O, Z(p)) < 1 + d_{\mathscr{P}}(e_n, p). \tag{4.1}$$

(Recall that by $O \in \mathscr{Z}_n$ we denote the multiset of n zeros.)

Proof. Let $p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0 \in \mathscr{P}_{n,1}$ and let $Z(p) = \{z_1, \ldots, z_n\}$ be the roots of p. The theorem claims that the following inequality holds:

$$\max_{1 \leqslant j \leqslant n} |z_j| < 1 + \max_{0 \leqslant j \leqslant n-1} |a_j|. \tag{4.2}$$

Let $c := \max\{|a_j| : 0 \leqslant j \leqslant n-1\} = d_{\mathscr{P}}(e_n, p)$. First notice that if any root satisfies $|z_k| \leqslant 1$, then the inequality

$$|z_k| < 1 + \max\{|a_j| : 0 \leqslant j \leqslant n-1\}$$

is trivially satisfied. Now let $z \in \mathbb{C}$, $|z| > 1$. We have

$$|p(z)| \geqslant |z|^n - \sum_{j=0}^{n-1} |a_j||z|^j$$

$$\geqslant |z|^n \left(1 - c \sum_{j=1}^{n} |z|^{-j}\right)$$

$$> |z|^n \left(1 - c \sum_{j=1}^{\infty} |z|^{-j}\right)$$

$$> |z|^n \left(1 - \frac{c}{|z| - 1}\right) = |z|^n \frac{|z| - (1 + c)}{|z| - 1}.$$

Therefore, if we actually have $|z| > 1 + c$, then $|p(z)| > 0$ and z cannot be one of the roots of p. This means that all roots of p must satisfy inequality (4.2).   □

As an immediate consequence we have

Corollary 4.2. *The function $Z : \mathscr{P}_{n,1} \rightarrow \mathscr{Z}_n$ maps each bounded set in $(\mathscr{P}_{n,1}, d_{\mathscr{P}})$ into a bounded set in $(\mathscr{Z}_n, d_F)$.*

As we did above, we denote by Π_n the set of all permutations of $\{1, \ldots, n\}$. In the following theorem and in Section 5 we shall use the notation:

$$\mathbf{u}_\sigma := (u_{\sigma(1)}, \ldots, u_{\sigma(n)}), \quad \text{for } \sigma \in \Pi_n, \quad \mathbf{u} = (u_1, \ldots, u_n) \in \mathbb{C}^n. \tag{4.3}$$

Theorem 4.3. *The function $\Phi : \mathscr{Z}_n \to \mathscr{P}_{n,1}$ defined by*

$$\Phi(\{z_1, \ldots, z_n\}) := \prod_{j=1}^n (z - z_j),$$

is a continuous function between $(\mathscr{Z}_n, d_F)$ and $(\mathscr{P}_{n,1}, d_\mathscr{P})$.

Proof. Let $\{z_1, \ldots, z_n\} \in \mathscr{Z}_n$ be the roots of $p(z) = z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0 \in \mathscr{P}_{n,1}$. By Viète's formulas,

$$
\begin{aligned}
a_0 &= & (-1)^n z_1 z_2 \cdots z_n & &=: & & \psi_1(z_1, \ldots, z_n) \\
a_1 &= & (-1)^{n-1} \sum_{k=1}^n \prod_{j \neq k} z_j & &=: & & \psi_2(z_1, \ldots, z_n) \\
\vdots \ &= & \vdots \quad & &=: & & \vdots \\
a_{n-1} &= & -(z_1 + z_2 + \cdots + z_n) & &=: & & \psi_n(z_1, \ldots, z_n).
\end{aligned}
$$

As a linear combination of products of continuous functions, each function $\psi_k : \mathbb{C}^n \to \mathbb{C}$, $k = 1, \ldots, n$, is continuous. Also note that each function ψ_k is symmetric, that is

$$\psi_k(\mathbf{u}) = \psi_k(\mathbf{u}_\sigma), \quad \text{for all } \mathbf{u} \in \mathbb{C}^n, \ \sigma \in \Pi_n, \ k \in \{1, \ldots, n\}.$$

(In fact each ψ_k is a constant multiple of an elementary symmetric polynomial.)

Consider the function $\Psi : \mathbb{C}^n \to \mathbb{C}^n$ defined by

$$\Psi(\mathbf{v}) = (\psi_1(\mathbf{v}), \ldots, \psi_n(\mathbf{v})), \quad \mathbf{v} \in \mathbb{C}^n.$$

The function $\Psi : \mathbb{C}^n \to \mathbb{C}^n$ is continuous and symmetric, since each of its components ψ_k is continuous and symmetric. Therefore for each $\varepsilon > 0$ and each $\mathbf{v} \in \mathbb{C}^n$ there exists $\delta(\varepsilon, \mathbf{v}) > 0$ such that

$$\mathbf{w} \in \mathbb{C}^n, \quad d_\infty(\mathbf{v}, \mathbf{w}) < \delta(\varepsilon, \mathbf{v}) \implies d_\infty(\Psi(\mathbf{v}), \Psi(\mathbf{w})) < \varepsilon.$$

Also

$$\Psi(\mathbf{u}) = \Psi(\mathbf{u}_\sigma) \quad \text{for all } \mathbf{u} \in \mathbb{C}^n, \ \sigma \in \Pi_n.$$

The last two displayed relations yield

$$\mathbf{w} \in \mathbb{C}^n, \quad \min_{\sigma \in \Pi_n} d_\infty(\mathbf{v}, \mathbf{w}_\sigma) < \delta(\varepsilon, \mathbf{v}) \implies d_\infty(\Psi(\mathbf{v}), \Psi(\mathbf{w})) < \varepsilon. \tag{4.4}$$

Let $K : \mathscr{Z}_n \to \mathbb{C}^n$ be the function defined in (3.10) and let $V, W \in \mathscr{Z}_n$. By the definition of d_F and (4.3) we have

$$d_F(V, W) = \min_{\sigma \in \Pi_n} d_\infty(K(V), K(W)_\sigma).$$

With this observation, (4.4) yields

$$W \in \mathscr{L}_n, \quad d_F(V, W) < \delta(\varepsilon, K(V))$$
$$\implies d_\infty(\Psi(K(V)), \Psi(K(W))) < \varepsilon. \tag{4.5}$$

The definitions of Φ and Ψ and the proof of Proposition 2.9 imply that

$$d_\infty(\Psi(K(V)), \Psi(K(W))) = d_{\mathscr{P}}(\Phi(V), \Phi(W)), \quad V, W \in \mathscr{L}_n. \tag{4.6}$$

Substituting (4.6) into (4.5) we obtain that for each $\varepsilon > 0$ and each $V \in \mathscr{L}_n$ there exists $\delta(\varepsilon, K(V)) > 0$ such that

$$W \in \mathscr{L}_n, \quad d_F(V, W) < \delta(\varepsilon, K(V)) \implies d_{\mathscr{P}}(\Phi(V), \Phi(W)) < \varepsilon.$$

This proves the continuity of Φ. □

Now, we can prove that the space of roots and the space of polynomials are homeomorphic.

Theorem 4.4. *The function* $Z : \mathscr{P}_{n,1} \to \mathscr{L}_n$ *which associates with each polynomial* $p \in \mathscr{P}_{n,1}$ *the multiset of its roots* $Z(p) \in \mathscr{L}_n$ *is a homeomorphism between* $(\mathscr{P}_{n,1}, d_{\mathscr{P}})$ *and* $(\mathscr{L}_n, d_F)$.

Proof. Clearly the functions Z and Φ are each other's inverse, and so $\Phi : \mathscr{L}_n \to \mathscr{P}_{n,1}$ is a bijection. Let us verify the assumptions of Theorem 2.4:

(a) By Theorem 3.4, each bounded and closed subset of the metric space $(\mathscr{L}_n, d_F)$ is compact.
(b) By Theorem 4.3, Φ is continuous.
(c) By Corollary 4.2, the function $\Phi^{-1} = Z$ maps bounded subsets of $\mathscr{P}_{n,1}$ into bounded subsets of $\mathscr{L}_n$.

Thus Theorem 2.4 applies and we conclude that $\Phi^{-1} = Z$ is continuous. Consequently, Z is a homeomorphism and the theorem is proved. □

5. Roots in $\mathbb{C}^n$

In Section 3, we introduced a bijection K between $\mathscr{L}_n$ and a subset $\mathscr{K}$ of $\mathbb{C}^n$ such that for each $V \in \mathscr{L}_n$ the n-tuple $K(V)$ and the multiset V have the same elements, counting multiplicities. Example 3.6 offers a specific bijection L between $\mathscr{L}_n$ and a subset $\mathscr{L}_n$ of $\mathbb{C}^n$. This bijection turns out not to be continuous. Since the space $\mathbb{C}^n$ is more familiar than $\mathscr{L}_n$, it would be desirable to have a bijection $K : \mathscr{L}_n \to \mathscr{K} \subset \mathbb{C}^n$ which is a homeomorphism between $(\mathscr{L}_n, d_F)$ and $(\mathscr{K}, d_\infty)$. In this section, we prove that this is not possible.

Theorem 5.1. *Let* P *be defined by* (3.8). *Let* $\mathscr{K}$ *be a subset of* $\mathbb{C}^n$ *with the property that for each* $V \in \mathscr{L}_n$ *the set* $\mathscr{K} \cap P^{-1}(V)$ *has exactly one element. Let* $K : \mathscr{L}_n \to \mathbb{C}^n$ *be defined by* $K = (P|_{\mathscr{K}})^{-1}$. *Then* K *is continuous if and only if its range* $\mathscr{K}$ *is closed in* $(\mathbb{C}^n, d_\infty)$.

Proof. Assume that K is continuous. Let $\{\mathbf{u}_k\}$ be a Cauchy sequence in $\mathcal{K}$. Since the function P satisfies (3.9), the sequence $\{P(\mathbf{u}_k)\}$ is Cauchy in $\mathcal{L}_n$. As $\mathcal{L}_n$ is complete by Theorem 3.4 and Proposition 2.5, this sequence is convergent, say, to V in $(\mathcal{L}_n, d_F)$. Since K is continuous the sequence $\{\mathbf{u}_k\} = \{K(P(\mathbf{u}_k))\}$ converges to $K(V) \in \mathcal{K}$. Thus, $\mathcal{K}$ is closed in $(\mathbb{C}, d_\infty)$.

To prove the converse assume that $\mathcal{K}$ is closed. Then the function $P|_{\mathcal{K}} : \mathcal{K} \to \mathcal{L}_n$ satisfies all the assumptions of Theorem 2.4 (recall that $\mathcal{K}$ is equipped with the metric d_∞). Assumption (a) in Theorem 2.4 is satisfied since each bounded and closed subset of $\mathcal{K}$ is bounded and closed in $(\mathbb{C}^n, d_\infty)$ and therefore compact in $\mathbb{C}^n$ and consequently compact in $\mathcal{K}$. Assumption (b) in Theorem 2.4 follows from (3.9), and (c) follows from Proposition 3.3. $\square$

Theorem 5.2. *Let $\mathcal{K}$ be as in Theorem* 5.1. *Then $\mathcal{K}$ is not closed in $(\mathbb{C}^n, d_\infty)$.*

Proof. Let $\mathcal{D}$ be the set of all points $\mathbf{u} = (u_1, \ldots, u_n) \in \mathbb{C}^n$ such that $u_k \neq u_j$ whenever $k \neq j$. For a point $\mathbf{w}$ in $\mathbb{C}^n$ and $r > 0$ let

$$B(\mathbf{w}, r) = \{\mathbf{v} \in \mathbb{C}^n : d_\infty(\mathbf{w}, \mathbf{v}) < r\}$$

be the open ball centered at $\mathbf{w}$ and with radius r. Also, define Π_n^* to be the set of all permutations of $\{1, \ldots, n\}$ minus the identity permutation.

By contradiction, suppose that $\mathcal{K}$ is closed in $(\mathbb{C}^n, d_\infty)$. Let $\mathbf{u} \in \mathcal{K} \cap \mathcal{D}$, that is, all the coordinates of $\mathbf{u} \in \mathcal{K}$ are mutually distinct. By the definition of $\mathcal{K}$, for every $\sigma \in \Pi_n^*$ we have that $\mathbf{u}_\sigma \in \mathbb{C}^n \backslash \mathcal{K}$. Since $\mathbb{C}^n \backslash \mathcal{K}$ is open, there exists an $r_\sigma > 0$ such that the entire open ball $B(\mathbf{u}_\sigma, r_\sigma)$ is contained in $\mathbb{C}^n \backslash \mathcal{K}$. Now we put

$$r := \min\{r_\sigma : \sigma \in \Pi_n^*\}$$

and prove that the ball $B(\mathbf{u}, r)$ is entirely contained in $\mathcal{K}$. To observe this, pick a $\mathbf{v} \in B(\mathbf{u}, r)$. Then, by our choice of r it follows that $\mathbf{v}_\sigma$ is contained in $B(\mathbf{u}_\sigma, r_\sigma)$ (and thus $\mathbf{v}_\sigma \notin \mathcal{K}$) for all $\sigma \in \Pi_n^*$. Since our construction of $\mathcal{K}$ requires that some permutation of the coordinates of $\mathbf{v}$ be contained in $\mathcal{K}$, and the only one we have left is $\mathbf{v}$ itself, we conclude that $\mathbf{v} \in \mathcal{K}$. So, $B(\mathbf{u}, r) \subset \mathcal{K}$, as claimed. We have thus proved that all the points in $\mathcal{K} \cap \mathcal{D}$ (i.e., those with n distinct coordinates) are interior points of $\mathcal{K}$.

Now let $\sigma \in \Pi_n^*$. Since $\mathbf{u} \in \mathcal{K} \cap \mathcal{D}$, we have $\mathbf{u}_\sigma \in \mathcal{D} \backslash \mathcal{K}$. By Lemma 2.8, $\mathcal{D}$ is pathwise connected. Therefore, there exists a continuous function $\Theta : [0, 1] \to \mathcal{D}$ such that $\Theta(0) = \mathbf{u}$ and $\Theta(1) = \mathbf{u}_\sigma$. Let

$$a := \sup\{t \in [0, 1] : \Theta(t) \in \mathcal{K}\}. \tag{5.1}$$

This supremum exists since $\Theta(0) = \mathbf{u} \in \mathcal{K}$ so the set on the right-hand side of (5.1) is not empty. As we assume that $\mathcal{K}$ is closed, $\Theta(a) \in \mathcal{K}$. Therefore $a < 1$. The range of Θ is a subset of $\mathcal{D}$, and thus $\Theta(a) \in \mathcal{K} \cap \mathcal{D}$ and consequently $\Theta(a)$ must be an interior point of $\mathcal{K}$. Since Θ is continuous this contradicts the definition of a. Thus $\mathcal{K}$ cannot be closed. $\square$

An immediate consequence of the previous two theorems is as follows:

Corollary 5.3. *The operator K defined in Theorem* 5.1 *is not continuous.*

Example 5.4. Let $\mathscr{L}_n$, L, and Z be as in Examples 3.5, 3.6 and Theorem 4.4. Then the function $L \circ Z : \mathscr{P}_{n,1} \to \mathscr{L}_n \subset \mathbb{C}^n$ is not continuous. For simplicity, we consider $n = 2$. The sequence of polynomials

$$z^2 + 1 + 2i/k - 1/k^2, \quad k \in \mathbb{N},$$

converges to $z^2 + 1$ in $(\mathscr{P}_{2,1}, d_{\mathscr{P}})$, but the sequence of lexicographically ordered pairs of their roots $(-1/k + i, 1/k - i), k \in \mathbb{N}$, does not converge in $(\mathbb{C}^2, d_\infty)$ to the pair of lexicographically ordered roots $(-i, i)$ of $z^2 + 1$.

Remark 5.5. A metric space setting for Theorem 4.4 is also provided in [4] and parts of our proof are similar to the proofs in [4]. In [4], the authors consider two metric spaces: the space of all monic polynomials of degree n and the space of their roots considered as ordered n-tuples of complex numbers (ordered lexicographically as explained in Example 3.5) and equipped with the d_∞ metric. Example 5.4 points out the difficulty with this setting (which invalidates the argument in [4]). Moreover, Corollary 5.3 and Theorem 4.4 imply that it is *not* possible to identify the roots of monic polynomials with unique n-tuples and equip such a set with the d_∞ metric and have a homeomorphism between such a space of roots and the space of polynomials. This indicates that the metric d_F is the natural metric on the roots.

6. Final remarks

We conclude with some historical remarks. In 1939, Ostrowski [5] published his own form of the perturbation theorem for polynomial roots. We quote it from [6, Appendix A].

Theorem 6.1. *Consider two polynomials*:

$$f(x) = a_0 x^n + \cdots + a_n, \quad a_0 = 1,$$
$$g(x) = b_0 x^n + \cdots + b_n, \quad b_0 = 1.$$

Let the n roots of $f(x)$ be $x_1, \ldots, x_n$, and those of $g(x)$, be $y_1, \ldots, y_n$. Put

$$\gamma = 2\Gamma, \quad \Gamma = \max_{v>0} (|a_v|^{1/v}, |b_v|^{1/v}).$$

Introduce the expression

$$\varepsilon = \sqrt[n]{\sum_{v=1}^{n} |b_v - a_v| \gamma^{n-v}}.$$

The roots x_v and y_v can be ordered in such a way that we have

$$|x_v - y_v| < (2n - 1) \quad \varepsilon(v = 1, \ldots, n).$$

We can see that Ostrowski's statement was quite "ready" for the language of the metric d_F, as it essentially contains the definition we give of d_F in Section 3. To show an alternate presentation of the classical perturbation theorem (although this time without the kind of numerical estimate that Ostrowski wished to obtain), here is the one given in [3]:

Theorem 6.2. *Let*

$$f(z) = a_0 + a_1 z + \cdots + a_n z^n = a_n \prod_{j=1}^{p} (z - z_j)^{m_j}, \quad a_n \neq 0,$$

$$F(z) = (a_0 + \varepsilon_0) + (a_1 + \varepsilon_1)z + \cdots + (a_{n-1} + \varepsilon_{n-1})z^{n-1} + a_n z^n$$

and let

$$0 < r_k < \min |z_k - z_j|, \quad j = 1, 2, \ldots, k-1, k+1, \ldots, p.$$

Then there exists a positive number ε such that, if $|\varepsilon_i| \leqslant \varepsilon$ for $i = 0, \ldots, n-1$, then $F(z)$ has precisely m_k zeros in the circle C_k with center z_k and radius r_k.

As a last quote, here is a version of the continuity theorem from the recent major survey of the theory of polynomials by Rahman and Schmeisser [7, Theorem 1.3.1 and Supplement]:

Theorem 6.3. *Let*

$$f(z) = \sum_{\nu=0}^{n} a_\nu z^\nu = \prod_{j=1}^{k} (z - z_j)^{m_j} \quad (m_1 + \cdots + m_k = n)$$

be a monic polynomial of degree n with distinct zeros $z_1, \ldots, z_k$ of multiplicities $m_1, \ldots, m_k$. Then, given a positive $\varepsilon < \min_{1 \leqslant i \leqslant j \leqslant k} |z_i - z_j|/2$, there exists a $\delta > 0$ so that any monic polynomial $g(z) = \sum_{\nu=0}^{n} b_\nu z^\nu$ whose coefficients satisfy $|b_\nu - a_\nu| < \delta$, for $\nu = 1, \ldots, n-1$, has exactly m_j zeros in the disk

$$D(z_j, \varepsilon) \quad (j = 1, \ldots, k).$$

Further, if we let

$$A := \max\{1, 2|a_\nu|^{1/(n-\nu)} : \nu = 0, \ldots, n-1\}$$

and let the zeros of f be denoted by $\zeta_1, \ldots, \zeta_n$, where an m-fold zero is now listed m times, then, for sufficiently small $\delta > 0$, there exists a numbering of the zeros of g as $\omega_1, \ldots, \omega_n$ such that $\max_{1 \leqslant \nu \leqslant n} |\omega_\nu - \zeta_\nu| \leqslant 4A\delta^{1/n}$.

To conclude: in every case known to us, the classical perturbation theorem has been presented as a continuity result (in a more or less convoluted way) and it has been proved by many authors using a variety of techniques (mostly from complex function theory, or trying to obtain useful numerical estimates). We hope that our topological presentation, and the emphasis on the homeomorphic relation between roots and polynomials, may have added to the understanding of this elegant, age-old result.

References

[1] J.D. Baum, Elements of Point Set Topology, Prentice-Hall, Englewood Cliffs, NJ, 1964.
[2] G.M. Ewing, Calculus of Variations with Applications, W.W. Norton & Co., New York, 1969.
[3] M. Marden, Geometry of Polynomials, second ed., reprinted with corrections, American Mathematical Society, Providence, RI, 1985.
[4] R. Naulin, C. Pabst, The roots of a polynomial depend continuously on its coefficients, Rev. Colombiana Mat. 28 (1994) 35–37.
[5] A.M. Ostrowski, Sur la continuité relative des racines d'équations algébriques, C. R. Acad. Sci. Paris 209 (1939) 777–779.
[6] A.M. Ostrowski, Solution of Equations in Euclidean and Banach Spaces, Solution of Equations and Systems of Equations, third ed., Pure and Applied Mathematics, vol. 9, Academic Press, New York, 1973.
[7] Q.I. Rahman, G. Schmeisser, Analytic Theory of Polynomials, Oxford University Press, Oxford, 2002.
[8] W. Rudin, Principles of Mathematical Analysis, third ed., International Series in Pure and Applied Mathematics, McGraw-Hill, New York, 1976.