Fundamental Theorem of Algebra

Ethan Y. Jaffe

In this note we give a proof of the fundamental theorem of algebra, using only the inverse function theorem and basic point-set topology. The proof is inspired by a proof of Milnor ([1]), which passes from $\mathbf{C}$ to the sphere S^2, but proceeds somewhat differently (in particular does not use regular values, or S^2).

Theorem 1. *Let $p(z)$ be any non-constant polynomial of one complex variable with complex coefficients. Then $p(z)$ has a complex root.*

Proof. Let us identify $\mathbf{R}^2$ with $\mathbf{C}$ via $x+iy \mapsto z$, and identify $p(z)$ as a smooth map $\mathbf{R}^2 \to \mathbf{R}^2$ via $p(x+iy) = \operatorname{Re}(p(x+iy)) + i\operatorname{Im}(p(x+iy))$. We will need to relate the differential dp to the classical (formal) derivative p', computed using the power rule.[1] To do this, we also identify $\mathbf{C}$ with a subset of $M_2(\mathbf{R})$, the space of 2×2 real matrices, via the map

$$\Phi(x+iy) = \begin{pmatrix} x & -y \\ y & x \end{pmatrix}.$$

It is easy to check that Φ is indeed a ring isomorphism onto its image. We have the following easy lemma:

Lemma 2. *Considered as a map $\mathbf{R}^2 \to \mathbf{R}^2$, $dp_z = \Phi(p'(z))$, where $p'(z)$ is the (formal) derivative of p.*

From the lemma, we deduce immediately that p has finitely many critical points. Indeed, dp_z is singular if and only if $\Phi(p'(z))$ is not invertible, which happens if and only if $p'(z) = 0$, since Φ is a ring isomorphism onto its image. Since p' is a nonzero polynomial over a field, it only has finitely many zeroes

To show that p has a root, we prove the stronger claim that p is surjective.

To do so, we need:

Lemma 3. *The image $p(\mathbf{C})$ is closed.*

[1] By formal, we mean for instance that the derivative of az^k is akz^{k-1}, extended by sum rule to all polynomials. We say "formal" since we do not wish to use the derivative of a function of one complex variable. Of course these coincide, but the formal algebraic definition suffices.

Proof. Suppose $x_n \in \mathbf{C}$ and $p(x_n) \to y \in \mathbf{C}$. We need to show that $y = p(x)$ for some x. If the x_n were uniformly bounded, then we may pass to a convergent subsequence $x_{n_k} \to x$, and then $p(x_{n_k}) \to p(x)$ since p is continuous, and $p(x_{n_k}) \to y$ by assumption. Thus $y = p(x)$.

So suppose the sequence x_n were not uniformly bounded. Then we may extract a subsequence x_{n_k} with $|x_{n_k}| \to \infty$. Write

$$p(z) = a_N z^N + a_{N-1} z^{N-1} + \cdots + a_0,$$

with $a_N \neq 0$. Then

$$|p(z)| \geq |a_N||z|^N \left(1 - \frac{|a_{N-1}|}{|a_N|}|z|^{-1} - \cdots - \frac{|a_0|}{|a_N|}|z|^{-k}\right) =: |a_N||z|^N(1 - E(z)).$$

If k is large enough, then $|E(x_{n_k})| \leq 1/2$, and so $|p(x_{n_k})| \geq \frac{1}{2}|a_N||x_{n_k}|^N \to \infty$, a contradiction. Thus the sequence x_n is uniformly bounded. □

Remark 4. This lemma and the proof is a restatement of the following fact: the map $p : \mathbf{C} \to \mathbf{C} \subseteq S^2$ (with S^2 identified as the Riemann sphere, i.e. the one point compactification of S^2), is a proper, and thus p extends to a map $S^2 \to S^2$ mappping ∞ to ∞ and $\mathbf{C} \to \mathbf{C}$. Thus $p(S^2)$ is compact, and hence closed in S^2. Thus

$$p(\mathbf{C}) = p(S^2 \setminus \{\infty\}) = p(S^2) \setminus \{\infty\} = p(S^2) \cap \mathbf{C}$$

is the intersection of two closed sets, and hence is closed.

Denote $A = p(\mathbf{C})$. We have shown already that A is closed. Let $B = \mathbf{C} \setminus p(\mathbf{C})$. Then B is open. Let

$$G = \{y \in p(\mathbf{C}) = A \colon dp_x \text{ is singular whenever } y = p(x)\}.$$

Since p has finitely many singular points, G is finite. We now use the inverse function theorem to show that $A \setminus G$ is open. Indeed, if $y \in C$, then $y = p(x)$ for some x where dp_x is invertible. Thus, by the inverse function theorem, there exist neigbhourhoods $U, V \subseteq \mathbf{C}$, with $x \in U$ and $y \in V$ such that $p : U \to V$ is a homeomorphism. Shrinking U, we may assume that dp is non-singular over U. In particular $V = p(U) \subseteq C$ and thus C contains a neighbourhood of y.

Since $C = A \cup B = (A \setminus G) \cup G \cup B$ is the union of three disjoint sets, $C \setminus G = (A \setminus G) \cup B$ is the union of tow disjoint open sets. Since G is finite, $C \setminus G$ is connected, and so this is possible only if $A \setminus G$ or B is empty. As A is infinite but G is finite, $A \setminus G$ is not empty, and thus $B = \mathbf{C} \setminus p(\mathbf{C})$ is empty. Hence p is surjective. □

References

[1] John W Milnor, *Topology from the Differentiable Viewpoint.* University Press of Virginia, Charlottesville