Irish Math. Soc. Bulletin
Number 96, Winter 2025, 81–91
ISSN 0791-5578

The Fundamental Theorem of Algebra

ANTHONY G. O'FARRELL

ABSTRACT. This is an expository note about the Fundamental Theorem of Algebra.

1. INTRODUCTION

Each nonconstant polynomial with complex coefficients has a complex root. In symbols:

Theorem 1.1. *If* $p(z) \in \mathbb{C}[z]$ *has positive degree, then there exists* $a \in \mathbb{C}$ *such that* $p(a) = 0$.

This is one of the foundations on which algebra rests. Burnside and Panton [3][1] state it in article 15, in Chapter II, and use it for most of the rest of Volume I, before giving a proof in article 122. The proof they give is based on the argument principle.

1.1. **Argument principle.** The variation of the argument of the polynomial around a simple closed curve γ on which it does not vanish counts the roots inside:

$$\int_\gamma d\arg(p) = \int_\gamma \frac{p'(z)dz}{p(z)} = 2\pi i n$$

if p has n roots inside γ (where multiple roots are counted a number of times equal to their multiplicity).

Assuming this, Theorem 1.1 follows on applying the principle, taking γ to be a very large circle around 0.

This proof of Theorem 1.1 has the merit of exposing the *real reason* why the theorem is true. The theorem is a consequence of the topological action of polynomials on the plane. More precisely, there are two ingredients: (1) the completeness of the complex plane $\mathbb{C}$, as a metric space; (2) the fact that a polynomial with $b = p(a)$ induces a *positive* map of homotopy groups

$$\pi_1(\{z \in \mathbb{C} : 0 < |z - a| < r\}) \to \pi_1(\mathbb{C} \setminus \{b\})$$

for all sufficiently small positive r. The latter comes down to the fact that the map $z \mapsto z^m$ induces multiplication by m on $\pi_1(\mathbb{C}^\times)$, combined with the remainder theorem.

Usually, people derive the theorem from the argument principle for holomorphic functions, and note that polynomials are entire functions, so that the argument principle applies to them. The argument principle for holomorphic functions depends on Cauchy's Theorem, and hence on the Stokes-Green formula.

2020 *Mathematics Subject Classification.* 12-02.
Key words and phrases. Polynomial, root, fundamental theorem.
Received on 23 October 2025, revised on 27 November 2025.
DOI: 10.33232/BIMS.0096.81.91.

[1]Classic text by two TCD academics, published by Hodges and Figgis, booksellers of happy memory.

1.2. **Issues.** The remainder theorem is elementary algebra, but plane algebraic topology is not. So it is reasonable to ask for proofs of Theorem 1.1 that avoid analysis as much as possible.

It seems obvious to me that you can't avoid analysis altogether, since the completeness of $\mathbb{C}$ is an essential ingredient.

1.3. **From Cauchy-Stokes.** The following proof uses a minimum of complex analysis:

Suppose $p(z) \in \mathbb{C}[z]$ has degree $m > 0$ and has no roots. Assume, as we may, that $p(z)$ is monic. Then $f(z) := z^{m-1}/p(z)$ has $f_{\bar{z}} = 0$ on $\mathbb{C}$. If $D = \mathbb{U}(0, R)$ is the disk of radius R about 0, then by Stokes' Theorem

$$\int_{\partial D} f(z)dz = \int_D df \wedge dz = \int_D (f_z dz + f_{\bar{z}} d\bar{z}) \wedge dz = 0.$$

But, parametrising ∂D by $z = Re^{i\theta}$, we have

$$\begin{aligned}\int_{\partial D} f(z)dz &= \int_{\partial D} \frac{dz}{z(1 + \mathrm{O}(1/R))} = i \int_0^{2\pi} (1 + \mathrm{O}(1/R))d\theta \\ &= 2\pi i(1 + \mathrm{O}(1/R) \to 2\pi i\end{aligned}$$

as $R \uparrow \infty$. This is impossible.

1.4. **Maximum principle.** The maximum principle for polynomials is elementary:

Theorem 1.2. *Let $p(z) \in \mathbb{C}[z]$. Suppose $|p(z)|$ has a local maximum at some point $a \in \mathbb{C}$. Then $p(z) = p(a)$, constant.*

Proof. Suppose p is nonconstant. Composing with translations, we may assume $a = 0$. Applying the remainder theorem, we can factor

$$p(z) - p(0) = z^n g(z),$$

where $n \geq 1$ and $g(z) \in \mathbb{C}[z]$ has $g(0) \neq 0$. Then for small positive r and any $\theta \in \mathbb{R}$, we have

$$p(re^{i\theta}) = p(0) + r^n e^{in\theta} g(0) (1 + \mathrm{o}(1)).$$

Writing $p(0) = \alpha e^{i\beta}$ and $g(0) = \rho e^{i\phi}$ with $\alpha \geq 0$, $\beta \in \mathbb{R}$, $\rho > 0$ and $\phi \in \mathbb{R}$, this gives

$$p(re^{i\theta}) = \alpha e^{i\beta} + r^n e^{i(n\theta+\phi)} \rho (1 + \mathrm{o}(1)). \tag{1}$$

So for $\theta = (\beta - \phi)/n$ and all small positive r we have

$$p(re^{i\theta}) = (\alpha + r^n \rho) e^{i\beta} + \mathrm{o}(r^n),$$

so for arbitrarily small positive r

$$|p(re^{i\theta})| \geq \alpha + r^n \rho - \frac{r^n \rho}{2} > \alpha = |p(0)|,$$

contradicting the assumption that 0 is a local maximum. □

A small twist on the same argument gives the *minimum priciple* away from roots:

Theorem 1.3. *Let $p(z) \in \mathbb{C}[z]$. Suppose $|p(z)|$ has a local minimum at some point $a \in \mathbb{C}$. Then $p(a) = 0$ or $p(z) = p(a)$, constant.*

Proof. Assuming p nonconstant and $p(a) \neq 0$, and proceeding as before, we have Equation (1), where now α is strictly positive. So for $\theta = (\beta - \phi + \pi)/n$ and all small positive r we have

$$p(re^{i\theta}) = (\alpha - r^n \rho) e^{i\beta} + \mathrm{o}(r^n),$$

so for arbitrarily small positive r

$$|p(re^{i\theta})| \leq \alpha - r^n \rho + \frac{r^n \rho}{2} < \alpha = |p(0)|,$$

contradicting the assumption that 0 is a local minimum. □

1.5. **Bolzano-Weierstrass.** The Bolzano-Weierstrass Theorem says that each bounded sequence of real numbers has a convergent subsequence. (See, for instance, [16, Theorem 8.17].) It follows that each bounded sequence of complex numbers has a convergent subsequnce: just apply it to the real parts and then apply it to the imaginary parts of the resulting subsequence. This is enough analysis to give Theorem 1.1.

1.6. **Proof of Theorem 1.1 without winding numbers.** Suppose $p(z) \in \mathbb{C}[z]$ is nonconstant and has no root.

Since $|p(z)| \to +\infty$ as $|z| \to +\infty$, we may choose $R > 0$ such that $|p(z)| \geq 2|p(0)|$ whenever $|z| \geq R$. Let $B := \mathbb{B}(0, R)$.

Let $m := \inf\{|p(z| : |z| \leq R\}$. Then $0 \leq m \leq |p(0)|$.

Suppose $m = 0$. Then we could choose a sequence $(z_n) \subset B$ such that $p(z_n) \to 0$. Passing to a subsequence, we may assume (z_n) converges to some $a \in B$. By continuity of p, $p(a) = 0$, which is impossible. Thus $m > 0$.

Choose a sequence $(z_n) \subset B$ such that $|p(z_n)| \to m$. Passing to a subsequence, we may assume $z_n \to \xi$ for some $\xi \in B$. Then $|p(\xi)| = m$. We cannot have $|\xi| = R$, since otherwise

$$m = |p(\xi)| \geq 2|p(0)| \geq 2m > m.$$

Thus p has a local minimum at ξ, which contradicts Theorem 1.3. □

1.7. **Variation.** A variation on the foregoing proof goes as follows:

Suppose $p(z) \in \mathbb{C}[z]$ is nonconstant and has no root. Then $f := 1/|p(z)|$ is positive and continuous on $\mathbb{C}$, and tends to zero as $|z| \to +\infty$. Thus we may choose $R > 0$ such that $|f(z)| < \frac{1}{2}|f(0)|$ whenever $|z| > R$. Let $B := \mathbb{B}(0, R)$, $D := \mathbb{U}(0, R)$ and $S := B \setminus D$.

There exists some $a \in B$ such that

$$f(a) = \sup_B |f|.$$

By continuity, $|f| \leq \frac{1}{2}|f(0)| \leq \frac{1}{2}|f(a)|$ on S, so $a \in D$. Thus $|p|$ has a local minimum at a, contradicting Theorem 1.3.

1.8. **Proof of Theorem 1.1 using harmonicity.** Harmonic functions may be defined as the twice-differentiable solutions of Laplace's equation, or, equivalently, as the continuous functions having the mean-value property. See [4].

Harmonic functions have a maximum principle.

Theorem 1.4. *Suppose $\Omega \subset \mathbb{C}$ is a connected open set, and $h : \Omega \to \mathbb{R}$ is harmonic on Ω. Then if there is some point $a \in \Omega$ such that*

$$h(a) = \sup_\Omega h,$$

then h is constant on Ω.

This is most conveniently proved by appealing to the mean-value property, and showing that the existence of a global maximum at a implies that the set $h^{-1}(h(a))$ is open-closed relative to Ω.

Now if we had a nonconstant polynomial $p(z)$ having no root, then $u := 2 \log |p|$ would be harmonic on $\mathbb{C}$, because $u = \log(p\bar{p})$ and $p_y = p' z_y = ip_x$, $p_{yy} = -p_{xx}$ so a simple calculation gives

$$u_x = \frac{\bar{p}_x}{\bar{p}} + \frac{p_x}{p},$$

$$u_{xx} = \frac{\bar{p}\bar{p}_{xx} - \bar{p}_x^2}{\bar{p}^2} + \frac{pp_{xx} - p_x^2}{p^2},$$

and similar formulas for the y-derivatives, giving

$$\Delta u = u_{xx} + u_{yy} = \frac{\bar{p}\Delta\bar{p} - \bar{p}_x^2 - \bar{p}_y^2}{\bar{p}^2} + \frac{p\Delta p - p_x^2 - p_y^2}{p^2} = 0.$$

We could then argue much as in Subsection 1.6 that, since it is a continuous real-valued function on $\mathbb{C}$ tending to infinity at infinity, u has a global minimum on $\mathbb{C}$ at some point a, and hence the nonconstant harmonic function $-u$ has a global maximum at a, contradicting from Theorem 1.4.

2. Open Mapping Theorems

The argument of Subsection 1.6 (or Subsection 1.7) can also be used by replacing the minimum principle Theorem 1.3 by the open mapping theorem for polynomials, because an open set that meets a circle must have points inside and outside the circle.

The open mapping theorem is:

Theorem 2.1. *Let $p(z) \in \mathbb{C}[z]$ be nonconstant. Then $p(\Omega)$ is open whenever $\Omega \subset \mathbb{C}$ is open.*

2.1. Holomorphic functions.

The open mapping theorem for holomorphic functions is:

Theorem 2.2. *Let $f : \Omega \to \mathbb{C}$ be holomorphic on the connected open set Ω and nonconstant. Then $p(\Omega)$ is open.*

The usual proof uses Rouchés Theorem. (For an alternative, see Section 4, below.)

Theorem 2.1 is an immediate corollary, and this is the standard way to prove it.

2.2. Smooth functions.

The open mapping theorem for vector-valued differentiable functions is:

Theorem 2.3. *Let $f : \Omega \to \mathbb{R}^d$ be continuously differentiable on the connected open set $\Omega \in \mathbb{R}^d$, with nonsingular Frechet derivative at each point. Then $f(\Omega)$ is open.*

This is a corollary of the inverse function theorem for smooth functions, which can be proved by applying Banach's contraction mapping principle.

Notice that an f satisfying the hypotheses must either preserve or reverse orientation on each connected component of Ω, because the sign of the determinant of its derivative cannot change there.

Theorem 2.3 may be used in a proof of Theorem 2.2, as follows:

Suppose $f : \Omega \to \mathbb{C}$ is holomorphic and nonconstant, and Ω is open. Let $C := \{c \in \Omega : f'(c) = 0\}$ be the set of critical points of f. Then C has no accumulation points in Ω. Theorem 2.3 tells us that $f(\Omega \setminus C)$ is open. So it remains to see that for $c \in C$ and sufficiently-small $r > 0$, the image $f(\mathbb{B}(c, r))$ is a neighbourhood of $a := f(c)$.

The set $P := f^{-1}(f(c))$ of preimages of $f(c)$ has no accumulation points in Ω.

Choose $r > 0$ smaller than half the distance from c to the rest of $C \cup P \cup (\mathbb{C} \setminus \Omega)$. Let B be the closed disc $\mathbb{B}(c, r)$, let U be its interior, and S be its boundary circle. Let A be the annulus $U \setminus \{c\}$.

Let $F := f(B)$. Then F is closed, since f is continuous and B is compact. Let $T := \text{bdy}(F)$, so $T \subset F$. Suppose F is not a neighbourhood of a. Then a must belong to T. Since $f(U \setminus \{a\})$ is open, and is contained in F, it does not meet T. Thus $T \subset \{a\} \cup f(S)$. Since $a \notin f(S)$, this means that T has an isolated point at a.

Now A is dense in B, so $D := f(A)$ is a dense subset of F, thus $T \subset \text{bdy}(D)$. Since D is open, it does not meet $\text{bdy}(D)$, so $\text{bdy}(D) \subset F \setminus D \subset T$. Thus $T = \text{bdy}(D)$, and D is a connected open set having a as an isolated boundary point. This implies that D

contains a deleted neighbourhood of a, and then F contains a full neighbourhood of a, contrary to our assumption. Thus $f(\Omega)$ is open. □

This proof does not simplify materially when f is assumed to be a polynomial, in place of an arbitrary holomorphic function.

3. Roots

The fundamental theorem implies that each nonzero $a \in \mathbb{C}$ has m-th roots of each order $m \in \mathbb{N}$, but this fact is more elementary, and can be proved using De Moivre's formula. Proving De Moivre's formula does require some analysis, of course, since we have to introduce the trigonometric functions first. Look at [16], for instance.

4. Formal power series

4.1. Let $\mathcal{F}$ be the ring of all formal power series over $\mathbb{C}$ in one variable, and $\mathcal{F}^\times$ be the group of invertibles under convolution multiplication. Let $\mathcal{G} \subset \mathcal{F}$ be the group of the series that are invertible under formal composition. Let $\mathfrak{F}$, $\mathfrak{F}^\times$ and $\mathfrak{G}$ be the corresponding subsets of series having positive radius of convergence.

Cartan [5] proves the inverse function theorem for convergent series, using a majorization argument:

Theorem 4.1. *Suppose $f \in \mathcal{G} \cap \mathfrak{F}$. Then the compositional inverse of f belongs to $\mathfrak{G}$.*

This has as a corollary the inverse function theorem for holomorphic functions, already mentioned, and this is an interesting alternative to the use of Rouchés Theorem.

4.2. **Roots.**

Proposition 4.2. *Suppose $f = a_0 + a_1 z + \mathrm{HOT} \in \mathcal{F}$ and $a_0 \neq 0$. Then for each $m \in \mathbb{N}$ there exists $g \in \mathcal{F}$ such that $g(z)^m = f(x)$. Moreover, if $f \in \mathfrak{F}$, then each choice of g also belongs to $\mathfrak{F}$.*

(Here, HOT stands for *higher-order terms.*)

Proof. Since a_0 has m-th roots, it suffices to consider the case $a_0 = 1$. The binomial series for the m-th root:

$$r := (1+x)^{1/m} := \sum_{n=0}^{\infty} \binom{\frac{1}{m}}{n} x^n$$

has radius of convergence $1 > 0$, so the composition $g := r \circ (f-1)$ has positive radius of convergence if f does, and satisfies $g^m = f$. □

This gives us another way to prove the open mapping theorem for holomorphic functions:

Suppose f is holomorphic and nonconstant on a neighbourhood N of a. We want to see that $f(N)$ is a neighbourhood of $b = f(a)$. Translating before and after, we may assume $a = b = 0$. The function f has a convergent power series expansion near 0, so for some $m \in \mathbb{N}$, we have

$$f(z) = z^m(a_0 + a_1 z + \mathrm{HOT}) = z^m h(z),$$

with $a_0 \neq 0$. By Proposition 4.2, there is a convergent series $g = b_0 + b_1 z + \mathrm{HOT})$ such that $g^m = h$. Then $f = (zg(z))^m$ near 0. Now by the inverse function theorem, $zg(z)$ maps N onto a neighbourhood N_1 of 0, and since all complex numbers have mth roots, $z \mapsto z^m$ maps N_1 onto a neighbourhood N_2 of 0, so $f = (z^m) \circ (zg(z))$ maps N onto N_2, and we are done.

5. Connectivity

Recall that a map is *proper* if the preimage of each compact set is compact. Proper maps between metric spaces are continuous. A map $f : \mathbb{C} \to \mathbb{C}$ is proper if it is continuous and $|f(z)| \to +\infty$ as $|z| \to +\infty$.

Theorem 5.1. *Suppose $M \neq \emptyset$ and N are connected manifolds, and $f : M \to N$ is continuous, proper, and open. Then $f(M) = N$.*

Proof. $f(M)$ is nonempty, connected, open and closed in N. Since N is connected, $f(M) = N$. □

This gives Theorem 1.1, once we know that nonconstant polynomials are open. This proof sidesteps the use of maxima and minima.

6. Galois Theory

People who like to use as little analysis as possible are drawn to the following proof of Theorem 1, which uses substantial results from Galois theory and group theory. It is found for instance in van der Waerden [17, Kap 11], or [14]. Lang says it is essentially one of Gauss' proofs, and van der Waerden describes it as the second Gauss proof [17, §81, p.252][2].

The analysis is in the following two lemmas.

Lemma 6.1. *Each odd-degree polynomial over $\mathbb{R}$ has a real root.*

Proof. It suffices to consider monic polynomials. If $p(x) \in \mathbb{R}[x]$ is monic and has odd degree, then for all large enough real $x > 0$, $p(x)$ is positive and $p(-x)$ is negative. By the Axiom of Completeness, there exists a least upper bound λ of the set $\{x \in \mathbb{R} : p(x) < 0\}$. Since $p : \mathbb{R} \to \mathbb{R}$ is a continuous function, it follows readily that $p(\lambda) = 0$. □

Lemma 6.2. *Each positive real number has a positive real square root.*

Proof. If $0 < a \in \mathbb{R}$, then $p(x) := x^2 - a$ is negative at $x = 0$ and positive for all large enough real x, so exactly as in the previous lemma, p has a positive real root. □

Corollary 6.3. *Each nonzero complex number has a complex square root.*

Proof. Indeed, let a, b be arbitrary elements of $\mathbb{R}$. We claim that there are $c, d \in \mathbb{R}$ such that $a + bi = (c + di)^2$. The case $a = 0$ is covered by Lemma 6.2 and the observation that -1 and i have complex square roots. For the case of non-zero a, we may assume that a is positive since -1 has a complex square root, and then by Lemma 6.2 that $a = 1$. We have to solve the system

$$\begin{aligned} 1 &= c^2 - d^2, \\ b &= 2cd. \end{aligned}$$

Squaring the second and multiplying both sides of the first by c^2, we get

$$c^4 - c^2 = \frac{b^2}{4}.$$

Completing the square gives

$$(c^2 - \tfrac{1}{2})^2 = \frac{b^2 + 1}{4}.$$

The right-hand side has a positive real square root, say e. Then $\frac{1}{2} + e$ has a positive real square root, say f. So $c = f$ and $d = b/(2c)$ give us real numbers that solve the system. □

[2]It is worth noting that van der Waerden [17, §80] says that the simplest proof of Theorem 1.1 is one that uses complex analysis: a counterexample $p(z)$ would have $1/p(z)$ nonconstant, entire and bounded, contradicting Liouville's Theorem.

Corollary 6.4. *Each monic quadratic over $\mathbb{C}$ has a complex root.*

Proof. Just use the usual quadratic formula and Corollary 6.3. □

Armed with these, we can prove Theorem 1.1, as follows.

Since -1 is not a square in $\mathbb{R}$, $\mathbb{C} := \mathbb{R}[i]$ is a degree 2 extension of $\mathbb{R}$.

Suppose some monic polynomial $p(z) \in \mathbb{C}[z]$ has no root in $\mathbb{C}$. Then there is a proper finite degree extension K of $\mathbb{C}$, which is Galois over $\mathbb{C}$.

Let S be a Sylow 2-subgroup of $\mathrm{Aut}(K/\mathbb{R})$. The fixed subfield $K^S \subset K$ of S is an odd degree extension of $\mathbb{R}$. Pick $\xi \in K^S$ such that $K^S = \mathbb{R}[\xi]$. Then the minimal polynomial of ξ over $\mathbb{R}$ has odd degree, hence has a root in $\mathbb{R}$, and hence has degree one. Thus $S = \mathrm{Aut}(K/\mathbb{R})$.

Thus $\mathrm{Aut}(K/\mathbb{R})$ is a 2-group, hence so is its subgroup $\mathrm{Aut}(K/\mathbb{C})$.

Every 2-group has a subgroup of index two[3], so choose $H \leq \mathrm{Aut}(K/C)$ of index 2. Then the fixed subfield K^H is a degree 2 extension of $\mathbb{C}$. But we can always solve quadratics over $\mathbb{C}$ in $\mathbb{C}$, so $\mathbb{C}$ does not have a degree 2 extension. This contradiction concludes the proof.

7. Pure Algebra

We can avoid analysis completely by changing the question. As already remarked, $\mathbb{R}$ is a convenient fiction, containing a huge set of 'yellow-pack' numbers which are literally indescribable. One can imagine trying to get along without $\mathbb{R}$. Among reasonable alternatives, three come immediately to mind:

- The field $\mathbb{E}$ of 'Euclidean reals'. These are the numbers corresponding to the points that you can construct on a line using straight-edge and compass and a segment on the line with ends labelled 0 and 1. Algebraically, $\mathbb{E}$ is a quadratic closure of $\mathbb{Q}$, i.e. $\mathbb{E}$ has characteristic zero, each quadratic polynomial over $\mathbb{E}$ has a root in $\mathbb{E}$, and no proper subfield of $\mathbb{E}$ has this property.
- The field $\mathbb{G}$ of 'Gaussian reals', or real algebraic numbers. This can be described without reference to $\mathbb{R}$, as follows. Let $\hat{\mathbb{Q}}$ be an algebraic closure of $\mathbb{Q}$, and let $i \in \hat{\mathbb{Q}}$ denote one of the square roots of -1. The field automorphism of $\mathbb{Q}[i]$ that sends $i \mapsto -i$ extends to an involutive field automorphism of $\hat{\mathbb{Q}}$, which we denote by $\tau : z \mapsto \bar{z}$. Then $\mathbb{G} \subset \hat{\mathbb{Q}}$ is the subfield fixed by τ, and one sees that $\hat{\mathbb{Q}} = \mathbb{G}[i]$.

  If you try, for a moment, to put yourself in Gauss' shoes, at the time before he had found his first proof of Theorem 1.1, you see that the great man had to grapple with the possibility that, big though it may be, $\mathbb{C}$ might not be large enough to embrace $\hat{\mathbb{C}}$ or even $\hat{\mathbb{Q}}$, and $\mathbb{R}$ might not contain a copy of $\mathbb{G}$.
- The field $\mathbb{D}$ of real numbers that have a definite description. Without getting into technicalities, $\mathbb{D}$ contains $\mathbb{G}$ and also numbers such as π and Euler's e and γ, real and imaginary parts of the values of all explicit elementary functions at all rationals, of zeros of Bessel functions, of Riemann's $\zeta(s)$, and so on. But since there are only a countable number of definite descriptions, $\mathbb{D}$ is much smaller than $\mathbb{R}$, even though it contains all the real numbers anyone might ever care about.

[3]Each p-group is nilpotent [8, Theorem 3.3(iii)], so each 2-group G has $[G,G] < G$ so the abelian group $G/[G,G]$ has a subgroup of index 2, hence its preimage under the surjection

$$G \to \frac{G}{[G,G]}$$

also has index 2.

If we ask what might replace Theorem1.1 if we replace $\mathbb{R}$ and $\mathbb{C} = \mathbb{R}[i]$ by one of these alternatives, then we get nowhere with the Euclidean numbers, because there are cubics over $\mathbb{Q}$ with no solution in $\mathbb{E}$. For $\mathbb{G}$ and $\mathbb{D}$, one can formulate reasonable questions.

The fields $\mathbb{R}$, $\mathbb{G}$ and $\mathbb{D}$ are examples of *formally-real fields* in the sense of Artin and Schreier: This just means that -1 is not a sum of squares in the field.

In logical terms, the *real-closed formally-real fields* share the same first-order properties as the ordered field $\mathbb{R}$, and are the subject of the Grand Artin-Schreier Theorem [19, 6]:

Theorem 7.1 (Grand Artin-Schreier Theorem). *Let F be a field. Then the following are equivalent:*
(i) F is formally real and admits no proper formally real algebraic extension.
(ii) F is formally real, every odd degree polynomial over F has a root, and for each $x \in F^\times$, one of $x, -x$ is a square.
(iii) F is formally real and $F(\sqrt{-1})$ is algebraically closed.
(iv) The absolute Galois group of F is finite and nontrivial.

To use this to prove that $\mathbb{G}[i]$ and $\mathbb{D}[i]$ are algebraically-closed, one needs to verify condition (ii) for $F = \mathbb{G}$ and $F = \mathbb{D}$. I don't see any way to do that without applying the fact that it holds for $F = \mathbb{R}$ and hence that (iii) holds for $F = \mathbb{R}$.

8. Wider context

It has been said that Theorem 1.1 is neither fundamental nor algebra. *Algebra* has changed its meaning in the past two centuries, and no longer just means the theory of equations, so the real question is whether the theorem is really fundamental for algebraic equations. The field $\hat{\mathbb{Q}}$ of algebraic numbers is certainly fundamental. It is necessary to deal with all the algebraic numbers, fictions of our imagination though they be. But the field of complex numbers is much larger, even in cardinality, and most complex numbers are even more fictional. In fact, the typical complex number has only generic properties, i.e. it cannot be characterised by a specific finite list of properties. The field $\mathbb{C}$ is convenient, because Theorem 1.1 implies that it contains an isomorphic copy of $\hat{\mathbb{Q}}$ and because we can use the richness of complex analysis on it. It is an interesting consequence of Theorem 1.1 that the field $\mathbb{R}$ has index two in its algebraic closure. However, $\mathbb{R}$ is large and mysterious, and open to the same criticism as $\mathbb{C}$.

8.1. $\mathbb{C}_p$.

Definition 8.1. $|\cdot|$ is a *field norm* on the field F if it satisfies the conditions:
(1) $|x|$ is a nonnegative real number, whenever $x \in F$, and $|x| = 0$ if and only if $x = 0$.
(2) $|x + y| \le |x| + |y|$, whenever $x, y \in F$.
(3) $|xy| = |x| \cdot |y|$, whenever $x, y \in F$.
The norm is *non-archimedean* if it satisfies the stronger condition:
(2') $|x + y| \le \max(|x|, |y|)$, whenever $x, y \in F$.

We remark that one could consider a more general concept, where the values of the norm lie in some totally-ordered abelian group [11].

Only fields of characteristic zero admit a field norm.

Each field of characteristic zero has a subfield isomorphic to $\mathbb{Q}$. Ostrowski [12] proved that the only field norms on $\mathbb{Q}$ are powers of the usual absolute value and powers of the p-adic norms corresponding to primes p.

From the adelic point of view, there is little to choose between $\mathbb{R}$ and any of the p-adic completions $\mathbb{Q}_p$ of the rationals. It is no longer the case that the algebraic closure $\widehat{\mathbb{Q}_p}$ of $\mathbb{Q}_p$ is a finite extension, nor is it complete with respect to the (unique!) extension of the p-adic metric, and we can enlarge it to its metric completion, denoted $\mathbb{C}_p$.

Theorem 8.2. [12, Theorem 13, p72][15, Theorem 4.6] *The field $\mathbb{C}_p$ is algebraically-closed.*

The key step in proving this is Krasner's Lemma:

Lemma 8.3. *Suppose K is a field complete with respect to a non-archimedean field norm $|\cdot|$. Suppose $\alpha, \beta \in \hat{K}$. Let L be the Galois completion of $K(\alpha, \beta)$ and let α_j ($j = 1, \ldots, m$) be the conjugates of α under the group G of automorphisms of L that fix $K(\beta)$. Suppose*

$$|\beta - \alpha| < \min\{|\alpha_i - \alpha_j| : i \neq j\}.$$

Then $\alpha \in K(\beta)$.

Proof. Let $\sigma \in G$. Then, since the norm is invariant under σ, we have

$$|\beta - \sigma(\alpha)| = |\sigma(\beta - \alpha)| = |\beta - \alpha|.$$

Thus

$$\begin{aligned} |\sigma(\alpha) - \alpha| &= |\sigma(\alpha) - \beta + \beta - \alpha| \\ &\leq \max(|\sigma(\alpha) - \beta|, |\beta - \alpha|) \\ &= |\beta - \alpha| \\ &< |\alpha_j - \alpha|, \ \forall \alpha_j \neq \alpha. \end{aligned}$$

Thus $\sigma(\alpha) = \alpha$, and so $\alpha \in K(\beta)$ since it is fixed by G. □

8.2. **Proof of Theorem 8.2.**

Proof. Fix $\alpha \in \widehat{\mathbb{C}_p}$, nonzero. Let $f(x) \in \mathbb{C}_p[x]$ be the (monic) minimal polynomial of α, and α_j ($j = 1, \ldots, n$) be its roots. Let

$$M := \max(1, |\alpha|^n), \text{ and } m := \min_{i \neq j} |\alpha_i - \alpha_j|.$$

Choose a monic polynomial $g(x) \in \widehat{\mathbb{Q}_p}$ of degree n with all coefficients within $(m/2)^n/M$ (with respect to the field norm) of the corresponding coefficients of $f(x)$. This ensures that

$$|g(\alpha) - f(\alpha)| < \left(\frac{m}{2}\right)^n.$$

Let $\beta_1, \ldots, \beta_n$ be the roots of $g(x)$, so that

$$g(x) = \prod_{j=1}^{n} (x - \beta_j).$$

Then

$$\prod_{j=1}^{n} |\alpha - \beta_j)| = |g(\alpha) - f(\alpha)| < \left(\frac{m}{2}\right)^n.$$

It follows that for some j we have $|\alpha - \beta_j| < m/2 < m$, so by Krasner's Lemma it follows that $\alpha \in K(\beta_j)$. Thus $\alpha \in \mathbb{C}_p$. □

8.3. **Spectra.** It is interesting that this proof of Theorem 8.2 is quite different from those we have seen of Theorem 1.1. It uses the extended norm on $\widehat{\mathbb{Q}_p}$. It raises the question whether Theorem 1.1 could be proved in the same way. In fact, if one could prove without assuming the fundamental theorem of algebra that the usual absolute-value norm extends from $\mathbb{C}$ to $\hat{\mathbb{C}}$, then one could deduce the fundamental theorem in various ways. For example, the norm would extend to the metric completion $\overline{\hat{\mathbb{C}}}$, which would then be a Banach algebra (a complete normed complex algebra) and a field, and the Gelfand-Mazur Theorem[1] then yields $\overline{\hat{\mathbb{C}}} = \mathbb{C}$.

Unfortunately, the usual proof of Gelfand-Mazur uses Liouville's Theorem, which may be applied more directly to prove the fundamental theorem. The key ingredient in the proof of the Gelfand-Mazur Theorem is that spectra are always nonempty, for elements of a Banach algebra with unit. The *spectrum* of an element f of a Banach algebra A is defined to be

$$\mathrm{spec}(f) := \{\lambda \in \mathbb{C} : f - \lambda 1 \text{ is noninvertible in } A\}.$$

Theorem 8.4. *Let A be a complete normed algebra with unit over $\mathbb{C}$. If $f \in A$, then $\mathrm{spec}(f) \neq \emptyset$.*

This theorem may be regarded as a *generalisation* of the fundamental theorem of algebra, because each monic complex polynomial of degree n is the characteristic polynomial of a companion $n\times n$ matrix, the set of all $n\times n$ complex matrices forms a Banach algebra, and the spectrum of a matrix in that algebra is the set of its eigenvalues.

Having written the account above, I had a look at Wikipedia [18], and found considerable overlap, along with a good deal of historical information. In particular, it seems that the first correct proof of the full theorem was given by Argand, in 1806, and not by Gauss, as folklore said.

9. Finding the roots

It is one thing to know the number of complex roots of a polynomial, counting multiplicities, but that doesn't butter any parsnips unless you can calculate them all to any desired accuracy. Ever since Galois, we know that this involves more than just the computation of k-th roots. Newton's method, the iteration of

$$q(z) := z - \frac{p(z)}{p'(z)},$$

will approximate any given root of $p(z)$, once you get close enough. For simple roots, it is phenomenally efficient, eventually doubling the number of significant figures at each step. But how can you ensure that you get close enough to each and every root? This problem was solved quite recently by applying twentieth-century advances in the theories of complex dynamical systems, topology, and conformal invariants. The initial breakthrough was made in the doctoral thesis of Scott Sutherland, and a refined and polished account is available in the paper of Hubbard, Schleicher and Sutherland [10]. This is a beautiful exposition of a stunning piece of work. It involves the classical Gauss-Lucas Theorem, a result of F. Riesz about radial limits, the Ahlfors conformal modulus and extremal length. The basic idea is that each root of $p(z)$ lies in a basin of attraction for $q(z)$ that contains a tentacle heading out to infinity, and each sufficiently-large circle meets each of these basins in a set that contains an interval that is not too small. So evenly-spaced points on the circle, provided the spacing is not too large, will do as a set of starting points for iterations of $q(z)$ that will converge to all the roots. In fact, the *same* evenly-spaced points will deliver all the roots of *all possible* polynomials $p(z)$ of a given degree d that have all their roots in the unit disc. By using some fast footwork, the number of starting points can be further reduced by placing them strategically on several circles instead of one. They give an explicit construction of approximately $0.2663 \log d$ circles, each containing $4.1627 d \log d$ points at equal distances.

References

[1] G.R. Allan. Introduction to Banach Spaces and Algebras. OUP. 2011.
[2] Z.I. Borevich and I.R. Shafarevich. Number Theory. Academic Press. 1966.
[3] W.S. Burnside and A.W. Panton. The Theory of Equations with an Introduction to the Theory of Binary Algebraic Forms. Two volumes. 10th ed. Hodges, Figgis. Dublin 1935.

[4] T. Carroll. Geometric Function Theory – A Second Course in Complex Analysis. Springer. 2024.
[5] H. Cartan. Elementary Theory of Analytic Functions of One or Several Complex Variables. Hermann. 1963.
[6] P.L. Clark. Field Theory. PDF. http://alpha.math.uga.edu/~pete/FieldTheory.pdf Accessed 2-5-2025.
[7] H. Cohen. Number Theory. Volume I: Volume II: Analytic and Modern Tools. Springer. 2007.
[8] D. Gorenstein. Finite Groups. Harper and Row. 1968.
[9] E. Hille. Analytic Function Theory. Two volumes. Ginn. Boston. 1962.
[10] J. Hubbard, D. Schleicher and S. Sutherland. How to find all roots of complex polynomials by Newton's method. Invent. Math. 146 (2001) 1–33.
[11] F. Jahnke. An Introduction to valued fields. PDF. https://ivv5hpp.uni-muenster.de/u/fjahn_01/franzi-notes.pdf Accessed 2-5-2025.
[12] N. Koblitz. p-adic Numbers, p-adic Analysis, and Zeta-Functions. 2nd edittion. Springer 1984.
[13] N. Koblitz. p-adic Analysis: a Short Course on Recent Work. LMS Lecture Notes 46. CUP. 1980.
[14] S. Lang. Algebra. Revised third edition. Springer. 2002.
[15] M. Ram Murty. Introduction to p-adic Analytic Number Theory. AMS. 2002.
[16] A.G. O'Farrell. Analysis One. Logic Press. 2015.
[17] B.L. van der Waerden. Algebra I. Achte Auflage der modernen Algebra. Springer. 1971.
[18] Wikipedia. Fundamental theorem of algebra. https://en.wikipedia.org/wiki/Fundamental_theorem_of_algebra. Viewed 25-4-2025.
[19] Wikipedia. Real closed field. https://en.wikipedia.org/wiki/Real_closed_field. Viewed 19-4-2025

Anthony G. O'Farrell studied at UCD and Brown, and worked at UCLA before taking the chair of Mathematics in Maynooth, where he served for 37 years. More at https://www.logicpress.ie/aof.

Department of Mathematics and Statistics, Maynooth University, Co. Kildare W23 HW31

E-mail address: anthony.ofarrell@mu.ie