2 Roots of Polynomials

By passing to a suitable extension of the coefficient field, one can ensure that a univariate polynomial has roots. For the field $\mathbb{C}$ of complex numbers, we also show that no extension is required.

2.0 Existence of Roots

How can we enlarge a field to guarantee that a polynomial has a root? We start with two perspectives on the complex numbers.

2.0.0 Remark (Hamilton 1835)**.** The field $\mathbb{C}$ of complex numbers is the set $\mathbb{R}^2$ equipped with addition and multiplication defined by

$$\begin{aligned}(a,b)+(c,d) &= (a+c, b+d)\\ (a,b)(c,d) &= (ac-bd, ad+bc)\,.\end{aligned}$$

One verifies that these operations make $\mathbb{R}^2$ into a field with $(1,0)$ as the multiplicative identity. Since $(0,1)(0,1) = (-1,0) = -(1,0)$, we set $\mathrm{i} := (0,1)$. We identify $\mathbb{R}$ with the subset $\{(a,0) \in \mathbb{R}^2 \mid a \in \mathbb{R}\}$.

This algebraic description of the complex numbers appears in W.R. Hamilton, *Theory of Conjugate Functions, or Algebraic Couples; with a Preliminary Essay on Algebra as the Science of Pure Time*, Trans. R. Irish Acad., **17** (1837) 293–422.

2.0.1 Remark (Cauchy 1847)**.** Consider the quotient $\mathbb{R}[x] \big/ \langle x^2+1 \rangle$. Applying the Euclidean algorithm, we see that the remainder of any polynomial in $\mathbb{R}[x]$ modulo x^2+1 has the form $a+bx$ where $a, b \in \mathbb{R}$. Hence, the set $\{a+bx \mid a, b \in \mathbb{R}\}$ of all polynomials in $\mathbb{R}[x]$ having degree at most 1 is a complete system of distinct representatives for the cosets in the quotient ring $\mathbb{R}[x] \big/ \langle x^2+1 \rangle$. Moreover, we have

$$\begin{aligned}(a+bx)+(c+dx) &\equiv (a+c)+(b+d)x \mod x^2+1\\ (a+bx)(c+dx) &\equiv (ac)+(ad+bc)x+(bd)x^2\\ &\equiv (ac-bd)+(ad+bc)x \mod x^2+1\,.\end{aligned}$$

Writing $\pi \colon \mathbb{R}[x] \to \mathbb{R}[x] \big/ \langle x^2+1 \rangle$ for the quotient map, it follows that $\mathrm{i} := \pi(x) = x + \langle x^2+1 \rangle$. It remains to show that this quotient ring is a field; see the subsequent proposition. We identify $\mathbb{R}$ with the subset $\{a+0x \in \mathbb{R}[x] \mid a \in \mathbb{R}\}$.

This alternative description of the complex numbers appears in A.-L. Cauchy, *Mémoire sur une nouvelle théorie des imaginaires, et sur les racines symboliques des équations et des équivalences*, C. R. Acad. Sci. Paris, **24** (1847) 1120–1130.

Fortuitously, the quotients of a univariate polynomial ring that are fields have already been characterized.

2.0.2 Proposition. *Let K be a field. For all polynomials $f \in K[x]$, the following are equivalent.*
a. *The polynomial f is irreducible in $K[x]$.*
b. *The ideal $\langle f \rangle := \{fg \mid g \in K[x]\}$ is a maximal ideal.*
c. *The quotient ring $K[x] \big/ \langle f \rangle$ is a field.*

Comment on proof. See MATH 210. □

This discussion motivates the following definition.

2.0.3 Definition. Given a ring homomorphism $\varphi \colon K \to L$ between fields, we say that L *is a field extension of* K. We identify K with its image $\varphi(K) := \{\varphi(\alpha) \in L \mid \alpha \in K\}$ and write $K \subseteq L$.

The ring homomorphism $\varphi \colon K \to L$ satisfies $\varphi(1_K) = 1_L$. Since the only ideals in the field K are $\langle 0 \rangle = \{0_K\}$ and $\langle 1 \rangle = K$, it follows that $\mathrm{Ker}(\varphi) = \langle 0 \rangle$, so the map φ is injective.

Armed with this notion, we demonstrate that every irreducible polynomial has a root in a field extension.

2.0.4 Proposition. *Let K be a field. For all irreducible $f \in K[x]$, there exists a field extension $K \subseteq L$ and $\alpha \in L$ such that $f(\alpha) = 0$.*

Proof. Consider the principal ideal $I := \langle f \rangle$ in $K[x]$ and the quotient ring $L := K[x] \,/\, I$. Proposition 2.0.2 shows that the quotient L is a field. The composition of the canonical inclusion $\eta \colon K \to K[x]$ with the canonical surjection $\pi \colon K[x] \to K[x] \,/\, I$ produces a ring homomorphism from K to L. Thus, we have a field extension $K \subseteq L$.

It remains to show that there exists $\alpha \in L$ such that $f(\alpha) = 0$. Set $\alpha := x + I$. Suppose that $f = a_0 x^n + a_1 x^{n-1} + \cdots + a_{n-1} x + a_n$ where $a_0, a_1, \dots, a_n \in K$. It follows that

The addition and multiplication operations in a quotient ring are inherited from the ambient ring.

$$\begin{aligned} f(\alpha) &= (a_0 + I)\alpha^n + (a_1 + I)\alpha^{n-1} + \cdots + (a_n + I)\alpha^0 \\ &= (a_0 + I)(x+I)^n + (a_1 + I)(x+I)^{n-1} + \cdots + (a_n + I)(x+I)^0 \\ &= (a_0 x^n + a_1 x^{n-1} + \cdots + a_n) + I \\ &= f + I = 0 + I \,, \end{aligned}$$

Since $0 + I$ is the additive identity, we deduce that $f(\alpha) = 0$. □

Division with remainder implies that field element $\alpha \in L$ is a root of a polynomial $f \in L[x]$ if and only if $x - \alpha$ is a factor of f in $L[x]$. Extending this idea leads to the following notion.

2.0.5 Definition. The polynomial $f \in K[x]$ *splits completely over L* if there exists a field extension $K \subseteq L$ and elements $\alpha_1, \alpha_2, \dots, \alpha_n \in L$ such that $f = a_0(x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n)$ where $a_0 \in K$.

The existence of these larger fields is essentially a consequence of the existence of a single root.

2.0.6 Theorem. *Let K be a field. For any nonconstant $f \in K[x]$, there is a field extension $K \subseteq L$ such that f splits completely over L.*

Proof. We proceed by induction on $n := \deg(f)$. When $n = 1$, it follows that $f = a_0 x + a_1$ where $a_0 \neq 0$ and $a_0, a_1 \in K$. Setting $L = K$ and $\alpha_1 = -a_1 / a_0$ implies that $f = a_0(x - \alpha_1)$, which shows that the base case holds.

Suppose that $\deg(f) = n > 1$. Since K is a field, the polynomial ring $K[x]$ is a unique factorization domain; see MATH 210. Hence, f has an irreducible factor g. Applying Proposition 2.0.4 to $g \in K[x]$, there exists a field extension $K \subseteq K_1$ and an element $\alpha_1 \in K_1$ such that $g(\alpha_1) = 0$. Since g is a factor of f, we also have $f(\alpha_1) = 0$, which implies that $x - \alpha_1$ is a factor of f in $K_1[x]$. In other words, there exists a polynomial $h \in K_1[x]$ such that $f = (x - \alpha_1)h$. Notice that $\deg(h) = \deg(f) - 1 = n - 1$. The induction hypothesis applied to h ensures that there exists a field extension $K_1 \subseteq L$ and elements $\alpha_2, \alpha_3, \dots, \alpha_n \in L$ such that $h = a_0(x - \alpha_2)(x - \alpha_3) \cdots (x - \alpha_n)$. We see that $f = a_0(x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n)$, so f splits over L. □

2.1 Fundamental Theorem of Algebra

How do we know that the field $\mathbb{C}$ of complex numbers is a splitting field for every nonconstant polynomial $f \in \mathbb{C}[x]$?

2.1.0 Proposition. *The following are equivalent:*
a. *Every nonconstant $f \in \mathbb{C}[x]$ has at least one root in $\mathbb{C}$.*
b. *Every nonconstant $f \in \mathbb{C}[x]$ splits completely over $\mathbb{C}$.*
c. *Every nonconstant $f \in \mathbb{R}[x]$ has at least one root in $\mathbb{C}$.*

Proof.
a ⇒ **b**: We proceed by induction on $n := \deg(f)$. When $n = 1$, we have $f = ax + b = a\big(x - (-b/a)\big)$, so f splits completely over $\mathbb{C}$.
Suppose that $n > 1$. When $f \in \mathbb{C}[x]$ has degree n, part **a** implies that $f(\alpha) = 0$ for some $\alpha \in \mathbb{C}$. It follows that there exists $g \in \mathbb{C}[x]$ such that $f = (x-\alpha)g$ and $\deg(g) = n-1$. Hence, the induction hypothesis implies that g splits completely over $\mathbb{C}$ and $f = (x - \alpha)g$ shows that the same is true for f.
b ⇒ **c**: Since $\mathbb{R} \subset \mathbb{C}$ is implication is trivial.
c ⇒ **a**: Let $f = a_0 x^n + a_1 x^{n-1} + \cdots a_n \in \mathbb{C}[x]$ with $a_0 \neq 0$. It suffices to show that f has a root in $\mathbb{C}$. Setting $h := f\overline{f}$, observe that

$$\overline{h} = \overline{f\overline{f}} = \overline{f}f = f\overline{f} = h\,,$$

so $h \in \mathbb{R}[x]$. By Part **c**, there exists $\alpha \in \mathbb{C}$ such that $h(\alpha) = 0$. It follows that $f(\alpha)\overline{f}(\alpha) = 0$. Since $\mathbb{C}$ is a domain, we deduce that either $f(\alpha) = 0$ or $\overline{f}(\alpha) = 0$. In the first case, α is a root of f and, in the second,

$$\overline{\overline{f}(\alpha)} = f(\overline{\alpha}) = 0$$

and $\overline{\alpha}$ is a root of f. □

2.1.1 Proposition. *Every polynomial $f \in \mathbb{R}[x]$ of odd degree has at least one root in $\mathbb{R}$.*

Proof. Let $f \in \mathbb{R}[x]$ be a polynomial of odd degree. We can assume that f is monic by multiplying f by a suitable nonzero constant. Hence, we have $f = x^n + a_1 x^{n-1} + a_2 x^{n-2} + \cdots + a_{n-1}x + a_n$ where the integer n is odd and $a_1, a_2, \dots, a_n \in \mathbb{R}$. Set $M := |a_1| + |a_2| + \cdots |a_n| + 1$. It follow that

$$\begin{aligned}|a_1 M^{n-1} &+ a_2 M^{n-2} + \cdots + a_{n-1}M + a_n| \\ &\leqslant |a_1|\,M^{n-1} + |a_2|\,M^{n-2} + \cdots + |a_{n-1}|\,M + |a_n| \\ &\leqslant \big(\,|a_1| + |a_2| + \cdots + |a_n|\big)M^{n-1} < M^n\,.\end{aligned}$$

Hence, we obtain

$$f(M) = M^n + \big(a_1 M^{n-1} + a_2 M^{n-2} + \cdots + a_{n-1}M + a_n\big) > 0\,,$$

because the expression in parentheses has absolute value less than M^n. We also see that

$$f(-M) = -M^n + \big(a_1(-M)^{n-1} + a_2(-M)^{n-2} + \cdots + a_n\big) < 0\,,$$

because n is odd and the expression in parenthesis has absolute value less than M^n. In summary, we have $f(-M) < 0 < f(M)$.

Since $f \in \mathbb{R}[x]$ is continuous, the Intermediate Value Theorem guarantees that there exists $c \in (-M, M)$ such that $f(c) = 0$. In other words, f has a real root. □

The Intermediate Value Theorem depends on the completeness of the real numbers, so one could argue that the Fundamental Theorem of Algebra is really a theorem in analysis or topology.

2.1.2 Lemma. *Every quadratic in* $\mathbb{C}[x]$ *splits completely over* $\mathbb{C}$.

Proof. Given $f = ax^2 + bx + c \in \mathbb{C}[x]$ with $a \neq 0$, the roots of f are $\frac{1}{2a}\big(-b \pm \sqrt{b^2 - 4ac}\big)$. Since every complex number has a square root in $\mathbb{C}$, we deduce that f splits completely over $\mathbb{C}$. □

2.1.3 Fundamental Theorem of Algebra (Girard 1649, Argand 1813)**.** *Every nonconstant polynomial* $f \in \mathbb{C}[x]$ *splits completely over* $\mathbb{C}$.

Proof. By Proposition 2.1.0, it suffices to prove that every $f \in \mathbb{R}[x]$ of positive degree n has at least one root in $\mathbb{C}$. Let $n = 2^m k$, where k is odd and $m \in \mathbb{N}$. We proceed by induction on m. Proposition 2.1.1 shows that a polynomial of odd degree in $\mathbb{R}[x]$ has a root in $\mathbb{C}$, so the base case holds.

This strategy appears in L. Euler, *Recherches sur les racines imaginaires des équations*, Mém. Acad. Roy. Sci. Berlin, **5** (1749) 222–288.

Suppose that $m > 0$. Regarding f as a polynomial in $\mathbb{C}[x]$, there exists a field extension $\mathbb{C} \subseteq L$ such that f splits completely over L. Let $\alpha_1, \alpha_2, \dots, \alpha_n \in L$ denote the roots of f. For all $\lambda \in \mathbb{R}$, consider

$$g_\lambda(x) := \prod_{1 \leqslant j < k \leqslant n} \big(x - (\alpha_j + \alpha_k) + \lambda \alpha_j \alpha_k\big) .$$

which has degree $\binom{n}{2} = \frac{1}{2} n(n-1)$. We first claim that $g \in \mathbb{R}[x]$. By construction, g_λ is invariant under any transposition of the roots, so its coefficients are symmetric polynomials in the roots. Since $\lambda \in \mathbb{R}$, Corollary 1.1.5 establishes that $g_\lambda \in \mathbb{R}[x]$.

Since $n = 2^m k$, the degree of g_λ is

$$\tfrac{1}{2} n(n-1) = \tfrac{1}{2}(2^m k)(2^m k - 1) = 2^{m-1} k (2^m k - 1) .$$

Since k is odd and $m > 0$, the integer $k(2^m k - 1)$ is also odd. Even though g_λ has larger degree than f, the exponent of 2 has been reduced by one. It follow that, for all $\lambda \in \mathbb{R}$, the induction hypothesis ensures that g_λ has a root in $\mathbb{C}$. By construction, the roots of g_λ are $\alpha_j + \alpha_k - \lambda \alpha_j \alpha_k$. In other words, for all $\lambda \in \mathbb{R}$, we can find a pair (j,k) such that $1 \leqslant j < k \leqslant n$ and $\alpha_j + \alpha_k - \lambda \alpha_j \alpha_k \in \mathbb{C}$. Although the pair (j,k) might depend on λ, as we range over the infinitely many possibilities of λ, there are only finitely many possibilities for the corresponding pair (j,k). Hence, there must exist $\lambda \neq \mu$ in $\mathbb{R}$ that use the same pair (j,k), so $\alpha_j + \alpha_k - \lambda \alpha_j \alpha_k \in \mathbb{C}$ and $\alpha_j + \alpha_k - \mu \alpha_j \alpha_k \in \mathbb{C}$. Subtraction gives

$$(\alpha_j + \alpha_k - \lambda \alpha_j \alpha_k) - (\alpha_j + \alpha_k - \mu \alpha_j \alpha_k) = (\mu - \lambda) \alpha_j \alpha_k \in \mathbb{C} ,$$

which implies that $\alpha_j \alpha_k \in \mathbb{C}$. The equation $\alpha_j + \alpha_k - \lambda \alpha_j \alpha_k \in \mathbb{C}$ thereby implies that $\alpha_j + \alpha_k \in \mathbb{C}$.

Finally, consider the quadratic polynomial

$$(x - \alpha_j)(x - \alpha_k) = x^2 - (\alpha_j + \alpha_k)x + \alpha_j \alpha_k \in \mathbb{C}[x] .$$

Lemma 2.1.2 shows that the roots of this quadratic polynomial lie in $\mathbb{C}$. However, the roots are clearly α_j and α_k. Therefore, f has a complex root. □

2.2 Minimal Polynomials

How are elements in a field extension related to a subfield? There is a basic dichotomy.

2.2.0 Definition. Let L be a field extension of a field K. An element $\alpha \in L$ is *algebraic* over K if there exists a nonconstant polynomial $f \in K[x]$ such that $f(\alpha) = 0$. Otherwise the element $\alpha \in L$ is *transcendental* over K.

2.2.1 Problem. Show that $\sqrt{2} + \sqrt{3}$ is algebraic over $\mathbb{Q}$.

Solution. Consider the polynomial

$$\begin{aligned}(x - \sqrt{2} - \sqrt{3})(x - \sqrt{2} + \sqrt{3})(x + \sqrt{2} - \sqrt{3})(x + \sqrt{2} + \sqrt{3})& \\ = (x^2 - 2\sqrt{2}x - 1)(x^2 + 2\sqrt{2}x - 1)& \\ = x^4 - 10x^2 + 1\,.&\end{aligned}$$

Since $\sqrt{2} + \sqrt{3}$ is a root of a nonconstant polynomial in $\mathbb{Q}[x]$, it is algebraic over $\mathbb{Q}$. □

2.2.2 Lemma/Definition. *When $\alpha \in L$ is algebraic over K, there exists a unique nonconstant monic polynomial $p \in K[x]$ such that*

(root) *The element α is a root of p.*

(minimal) *For all $f \in K[x]$ having α as a root, p divides f.*

The polynomial p is called the minimal polynomial *of α over K.*

Proof. Among all nonconstant polynomials in $K[x]$ having α as a root, there is one of smallest degree, say p. Rescaling if necessary, we may assume that p is monic.

We are using the well-ordering of the set $\mathbb{N}$. The hypothesis that $\alpha \in L$ is algebraic means that there is a nonconstant polynomial in $K[x]$ having α as a root.

Suppose that $f \in K[x]$ with $f(\alpha) = 0$. Division with remainder produces $q, r \in K[x]$ such that $f = qp + r$ and either $r = 0$ or $\deg(r) < \deg(p)$. Evaluating at α gives

$$0 = f(\alpha) = q(\alpha)p(\alpha) + r(\alpha) = r(\alpha)\,.$$

If r were nonzero, then it would be a polynomial of degree less that p having α as a root, which would contradict the choice of p. We conclude that $r = 0$ and p divides f. □

2.2.3 Proposition. *Let $\alpha \in L$ be an algebraic element over K with minimal polynomial $p \in K[x]$. For every nonconstant monic polynomial $f \in K[x]$, the following are equivalent:*

a. *$f = p$,*

b. *f is a polynomial of minimal degree such that $f(\alpha) = 0$,*

c. *f is irreducible over K and $f(\alpha) = 0$.*

Proof.

a ⇔ **b**: This follows from the proof of Lemma 2.2.2.

b ⇔ **c**: We claim that the minimal polynomial f is irreducible over K. Suppose that $f = gh$ where $g, h \in K[x]$ have smaller degree that p. It would follows that $0 = f(\alpha) = g(\alpha)h(\alpha)$ which would imply that $g(\alpha) = 0$ or $h(\alpha) = 0$. Since this would contradict **b**, the polynomial f must be irreducible.

c ⇔ b: Suppose that $f(\alpha) = 0$ and f is irreducible. Lemma 2.2.2 shows that p divides f, so $f = ph$ for some $h \in K[x]$. Since f is irreducible and p is nonconstant, h must be a constant. Thus, we deduce that $f = p$ because both f and p are monic. □

2.2.4 Remark. The irrationality of $\sqrt{2}$ implies that $x^2 - 2 \in \mathbb{Q}[x]$ is the minimal polynomial of $\sqrt{2}$ over $\mathbb{Q}$.

2.2.5 Problem. Demonstrate that $x^4 - 10x^2 + 1$ is the the minimal polynomial of $\sqrt{2} + \sqrt{3}$ over $\mathbb{Q}$.

Solution. By Problem 2.2.1 and Proposition 2.2.3, it suffices to show that $f := x^4 - 10x^2 + 1$ is irreducible over $\mathbb{Q}$. By the Gauss Lemma, this is equivalent to proving that f is irreducible over $\mathbb{Z}$. Reducing modulo 3, the polynomial f becomes $x^4 + 2x^2 + 1 \equiv (x^2+1)^2 \in \mathbb{F}_3[x]$. Observe that $x^2 + 1$ is irreducible over $\mathbb{F}_3$, because

$$0^2 + 1 \equiv 1 \mod 3\,, \quad 1^2 + 1 \equiv 2 \mod 3\,, \quad 2^2 + 1 \equiv 5 \equiv 2 \mod 3\,.$$

The image of f in $\mathbb{F}_3[x]$ is the square of an irreducible polynomial, so any quadratic factor must be an associate of $x^2 + 1$. Lifting such a factorization back to $\mathbb{Z}[x]$ would force f to be the square of a quadratic polynomial of $\mathbb{Q}$. Comparing coefficients, the equation

$$x^4 - 10x^2 + 1 = (x^2 + ax + b)^2 = x^4 + 2ax^3 + (a^2 + 2b)x^2 + 2abx + b^2\,,$$

gives $2a = 0$, $a^2 + 2b = -10$, $2ab = 0$, and $b^2 = 1$, so $a = 0$, $b = -5$, and $5^2 = 1$ which is absurd. We see that f is irreducible over $\mathbb{Q}$, and $x^4 - 10x^2 + 1$ is the the minimal polynomial of $\sqrt{2} + \sqrt{3}$ over $\mathbb{Q}$. □

2.2.6 Notation. Let $K \subseteq L$ be a field extension. For all field elements $\alpha_1, \alpha_2, \dots, \alpha_n \in L$, we define

$$K[\alpha_1, \alpha_2, \dots, \alpha_n] := \{h(\alpha_1, \alpha_2, \dots, \alpha_n) \mid h \in K[x_1, x_2, \dots, x_n]\}\,.$$

2.2.7 Lemma. *Let $K \subset L$ be a field extension. For every $\alpha \in L$ that is algebraic over K with minimal polynomial $p \in K[x]$, there exists a unique K-algebra isomorphism $K[x] \,/\, \langle p \rangle \cong K[\alpha]$ that sends the coset $x + \langle p \rangle$ to α.*

Proof. Consider the evaluation map $\varphi \colon K[x] \to L$ that sends x to α. By construction, the image of φ is $K[\alpha]$.

We claim that $\operatorname{Ker}(\varphi) = \langle p \rangle$. For all $g \in K[x]$, we have

$$\varphi(gp) = \varphi(g)\varphi(p) = g(\alpha)p(\alpha) = g(\alpha)0 = 0\,,$$

so $\langle p \rangle \subseteq \operatorname{Ker}(\varphi)$. Conversely, suppose that $f \in \operatorname{Ker}(\varphi)$. It follows that $f(\alpha) = 0$, so Lemma 2.2.2 implies that f is a multiple of p. We deduce that $\operatorname{Ker}(\varphi) \subseteq \langle p \rangle$, so $\operatorname{Ker}(\varphi) = \langle p \rangle$.

Given the image and kernel of φ, the First Isomorphism Theorem shows that the K-algebra homomorphism $\varphi \colon K[x] \to L[\alpha]$ induces an the K-algebra isomorphism

$$\widetilde{\varphi} \colon \frac{K[x]}{\langle p \rangle} \to K[\alpha]$$

where $\widetilde{\varphi}(x) = \varphi(x) = \alpha$. □