5 Fundamental theorem of algebra and other facts about polynomials

5.1 Polynomials and their roots. Basic facts

5.2 Cardano's formula and complex numbers

5.3 Geometric interpretation of complex numbers

5.4 The fundamental theorem of algebra

Now that we got some experience working with complex numbers, I can give maybe not one hundred percent rigorous but hopefully very convincing proof of one of the most important facts in basic mathematics: the fundamental theorem of algebra. To prepare for the proof, let me start with the definition of an auxiliary function, which will be the key tool I will use in the proof below.

Consider the complex variable $z = \rho(\cos\theta + \mathrm{i}\sin\theta)$ for some $\rho > 0$ and $\theta \in [0, 2\pi]$. When the variable θ changes from 0 to 2π, the corresponding point z will travel along the full circle of radius ρ on the complex plane $\mathbf{C}$, let me call this circle C. Now consider any continuous function f of complex argument z. When z travels along C complex number $w = f(z)$ travels along some closed curve Γ on another complex plane (u, v).

Definition 5.1. *Consider the vector that connects the origin O with the point $f(z)$ on the plane (u, v). By definition, the order of the point O for function f with respect to the curve C is the number of full turns this vector completes when z travels along the curve C.*

To get a feeling of this definition, explain carefully why the order of $f(z) = const$ is zero, the order of $f(z) = z$ is one, the order of $f(z) = z^2$ is two, and generally the order of $f(z) = z^n$ is equal to n.

Now, finally, let me define the function $\varphi_f(\rho)$, which is the order of O for function f when the circle C has the radius ρ. From the given definition φ_f is well defined at any point when $f(z) \neq 0$, and accepts only integer values 0,1,2,.... Moreover, it is almost obvious that if f is continuous and φ_f is defined for all ρ then it also must be continuous. This implies immediately that if $f(z) \neq 0$ for all z then $\varphi_f(\rho)$ must be constant because the only continuous integer valued function is a constant. Ok, everything is ready for the proof.

Theorem 5.2 (Fundamental theorem of algebra)**.** *If polynomial $p \in \mathbf{C}[z]$ is such that*

$$p(z) = c_0 + c_1 z + c_2 z^2 + \ldots + c_{n-1} z^{n-1} + z^n, \quad n \geq 1, \tag{5.1}$$

then there exists a point $\xi \in \mathbf{C}$ such that $p(\xi) = 0$.

Proof. By contradiction. Assume that $p(z) \neq 0$ for any $z \in \mathbf{C}$. It implies that $\varphi_p(\rho)$ is defined for all $\rho > 0$. Since p is a continuous function, hence φ_p is also continuous and hence must be constant. We have that $c_0 \neq 0$ (otherwise zero would be the root and hence theorem would be proved), therefore, when $\rho \to 0$ the variable z approaches zero, and hence $p(z) \approx c_0$, which yields that $\varphi_p(0^+) = 0$.

On the other hand, if we take ρ sufficiently large, then $p(z) \approx z^n$ (the highest power of our polynomial), and therefore $\varphi_p(\rho)$ must be n, which supplies the required contradiction.

To fully justify the last sentence let me take $\rho > 1$ and $\rho > |c_0| + |c_1| + \ldots + |c_{n-1}|$. In this case

$$\begin{aligned}
|p(z) - z^n| &= |c_0 + c_1 z + \ldots c_{n-1} z^{n-1}| \leq \\
&\leq |c_0| + |c_1||z| + \ldots + |c_{n-1}||z|^{n-1} = \\
&= |c_0| + |c_1|\rho + \ldots + |c_{n-1}|\rho^{n-1} = \\
&= \rho^{n-1}\left(\frac{|c_0|}{\rho^{n-1}} + \frac{|c_1|}{\rho^{n-2}} + \ldots + |c_{n-1}|\right) \leq \\
&\leq \rho^{n-1}(|c_0| + |c_1| + \ldots + |c_{n-1}|) < \\
&< \rho^n = |z|^n.
\end{aligned}$$

In words: for sufficiently large ρ the distance from the point $p(z)$ to the point z^n is strictly less than the distance from O to the point z^n, and hence the line segment connecting $p(z)$ and z^n cannot pass through the origin. Therefore we can always transform the point $p(z)$ into the point z^n for all $p(z) \in \Gamma$ without changing the value of $\varphi_p(\rho)$, which finishes the proof of this theorem. ■

Remark 5.3. I copied this proof from Courant, R., & Robbins, H. (1996). What is Mathematics?: An elementary approach to ideas and methods. Oxford University Press, USA, whose first edition was published in 1941. It is known that in the Soviet Union exactly the same proof was given by Andrey Nikolaevich Kolmogorov around 1937 in his lectures on the fundamental theorem of algebra. In the Russian school this proof got a name "The lady with the dog " (after one of the most famous short stories by Anton Chekhov) meaning that if a lady makes n laps around a house having a dog on a leash then the dog will make exactly the same n turns around the house no matter how much it runs around the owner.

Remark 5.4. To make a proof completely rigorous one must be more careful with the definition of the order of O for given f, and it must be proved that φ_f stays constant with continuous transformations of f that do not pass through $f(z) = 0$. This all can be done but requires a somewhat more advanced mathematics, hence will be omitted in these notes.

Putting together the fundamental theorem of algebra and the preliminary facts we proved about polynomials, I immediately obtain

Corollary 5.5. *Consider $p \in \mathbf{C}[z]$ of the form*

$$p(z) = c_0 + c_1 z + \ldots + c_{n-1} z^{n-1} + c_n z^n, \quad c_n \neq 0, \quad n \geq 1.$$

Then there exist $\xi_1, \ldots, \xi_k \in \mathbf{C}$ such that

$$p(z) = c_n(z - \xi_1)^{\alpha_1}(z - \xi_2)^{\alpha_2} \ldots (z - \xi_k)^{\alpha_k},$$

where $\xi_l \neq \xi_j$ and $\alpha_j \in \mathbf{N}$ and $\alpha_1 + \ldots + \alpha_k = n$.

Remark 5.6. The constants α_j are called the multiplicities of the roots ξ_j. Therefore the corollary can be restated as "Any nonconstant polynomial of degree n has exactly n complex roots if they are counted according to their multiplicities."

Proof. Since $c_n \neq 0$ I can consider $q(z) = \frac{p(z)}{c_n}$, which has the same form as in Theorem 5.2 and hence must have complex root $\xi_1 \in \mathbf{C}$. We know that it implies that $q(z) = (z - \xi_1)g(z)$, where g has the degree $n-1$ and moreover the coefficient at z^{n-1} is 1, hence I can apply Theorem 5.2 again obtaining $q(z) = (z-\xi_1)(z-\xi_2)h(z)$, where h has degree $n-2$ and so on. Note that here ξ_1 and ξ_2 can coincide. Returning to p and grouping together the identical roots finishes the proof. ■

Now I can say even more.

Corollary 5.7 (Vieta's theorem)**.** *Let*

$$p(z) = c_0 + c_1 z + \ldots + c_{n-1} z^{n-1} + c_n z^n, \quad c_n \neq 0, \quad n \geq 1.$$

and let $\xi_1, \ldots, \xi_n$ be its roots, which can be identical in this list. Then

$$\begin{aligned} \xi_1 + \ldots + \xi_n &= -\frac{c_{n-1}}{c_n}\,, \\ (\xi_1\xi_2 + \ldots \xi_1\xi_n) + (\xi_2\xi_3 + \ldots + \xi_2\xi_n) + \ldots + \xi_{n-1}\xi_n &= \frac{c_{n-2}}{c_n}\,, \\ &\vdots \\ \xi_1 \ldots \xi_n &= (-1)^n \frac{c_0}{c_n}\,. \end{aligned}$$

Proof. These formulas follow from the last corollary by expanding the product $c_n(z-\xi_1)^{\alpha_1}(z-\xi_2)^{\alpha_2}\ldots(z-\xi_k)^{\alpha_k}$ and using the fact that two polynomials are equal if and only if the coefficients at equal powers are equal. ■

Remark 5.8. It may be useful to derive the formulas in Vieta's theorem for the cases $n = 2$ and $n = 3$.

Now consider a polynomial

$$p(x) = a_0 + a_1 x + \ldots + a_{n-1} x^{n-1} + a_n x^n, \quad a_n \neq 0, \quad n \geq 1, \quad a_j \in \mathbf{R}.$$

I claim that

Corollary 5.9. *Polynomial p with real coefficients can be written as the product of linear and quadratic polynomials with real coefficients.*

Proof. To prove this corollary I need two more facts. First I claim that if $\xi \in C$ is a root of polynomial p with real coefficients then its conjugate $\bar{\xi}$ is also a root. It follows from the fact that conjugate of a sum is a sum of conjugates, conjugate of a product is a product of conjugates, and conjugate of a real number is real (fill in all the details). And second, if I have two conjugate roots $\xi_1 = \bar{\xi_2} = \alpha + \mathrm{i}\beta$ then the product

$$(x - \xi_1)(x - \xi_2) = x^2 - (\xi_1 + \xi_2)x + \xi_1\xi_2 = x^2 - 2\,\mathrm{Re}(\xi_1)x + |\xi_1|^2$$

is a quadratic polynomial with real coefficients. Now the application of Corollary 5.5 finishes the proof. ■

5.5 Another proof of the fundamental theorem of algebra

There are a lot of various proofs of the fundamental theorem of algebra. Here I would like to present a proof[1] that relies on the theory of ordinary differential equations — the mathematical topic, which is very close to my own research. I will certainly need a few facts from the classical ODE theory, but nothing beyond first proof based ODE course.

Theorem 5.10 (Fundamental theorem of algebra)**.** *Any nonconstant polynomial p has at least one root $\xi \in \mathbf{C}$.*

Proof. Let $p(z)$ be nonconstant polynomial. Consider the ordinary differential equation (ODE)

$$\frac{\mathrm{d}z}{\mathrm{d}t} = -\frac{p(z)}{p'(z)}, \quad z(0) = z_0 \in \mathbf{C}. \tag{5.2}$$

(If you never dealt with complex differential equations, think of a system of two real differential equations, since t here is a real variable.) This ODE is defined everywhere except at the points where $p'(z) = 0$, but since p' is a polynomial as well, there are only finitely many such points. In the following we assume that for any such point z that $p'(z) = 0$, $p(z) \neq 0$ otherwise the theorem is proved and we can stop here.

The nicest thing about the initial value problem (5.2) is that it has the solution $z(t; z_0)$ (which exists by the existence and uniqueness theorem at least locally for all such z_0 for which $p'(z_0) \neq 0$) that satisfies

$$p(z(t; z_0)) = e^{-t} p(z_0). \tag{5.3}$$

Indeed, the initial condition is satisfied, and the differentiation yields that

$$p'(z(t; z_0))z'(t; z_0) = -e^{-t} p(z_0),$$

hence, after plugging z' into (5.3), $z(t; z_0)$ satisfies the equation. Equation (5.3) implies that

$$|p(z)| = e^{-t}|p(z_0)|,$$

that is, $|p(z)|$ monotonically (exponentially) decreasing along the solutions $z(t; z_0)$. This means, first of all, that there is sufficiently large disk such that the solutions to (5.2) cannot leave this disk (since for large z $|p(z)| \to \infty$). Inside this disk the only problematic points are those that satisfy $p'(z) = 0$, but we already know that there are only finitely many of them. Moreover, solution (5.3) means that the movement in the plane of the variable $p(z)$ is *isogonal*, since $\arg p(z) = \arg p(z_0)$ for all t, hence it is always possible to choose such z_0 that $z(t; z_0)$ will miss all the roots of p' together with small neighborhoods of these points. But this implies that there are always solutions to (5.2) that do not approach a boundary of a compact set on the complex plane $\mathbf{C}$ and therefore, by another important classical theorem of ODE theory, are defined for all times $t \in (0, +\infty)$, and hence I can pass to the limit $t \to \infty$ in (5.3) yielding that there must be a point $\xi \in \mathbf{C}$ such that $p(\xi) = 0$. ■

[1] I am copying this proof from Anton, R., Mihalache, N., & Vigneron, F. (2023). A short ODE proof of the Fundamental Theorem of Algebra. The Mathematical Intelligencer, 1-2.

5.6 Descartes' rule of signs

Here I will give a full proof of the so-called Descartes' rule of signs that allows to get an upper estimate on the number of real roots of the given polynomial. Let me start with introducing the notation. In the following I consider the polynomial

$$p(x) = a_0 x^{b_0} + a_1 x^{b_1} + \ldots + a_n x^{b_n}, \tag{5.4}$$

where $0 \leq b_0 < b_1 < \ldots < b_n$ are integers and $a_0, a_1, \ldots, a_n$ are nonzero real numbers. Consider the sequence $(a_0, a_1, \ldots, a_n)$ and call *a variation in sign* if $a_{j-1}a_j < 0$ in this sequence. Let $v(p)$ be the total number of variations in signs of the polynomial p, and $r(p)$ be the number of positive roots of p counting multiplicities.

Theorem 5.11 (Descartes' rule of signs)**.**

$$\begin{aligned} &r(p) \leq v(p), \\ &r(p) \equiv v(p) \pmod 2. \end{aligned}$$

Example 5.12.

The proof will rely on two lemmas.

Lemma 5.13. *Consider polynomial* (5.4)*. If* $a_0 a_n > 0$ *then* $r(p) \equiv 0 \pmod 2$*. If* $a_0 a_n < 0$ *then* $r(p) \equiv 1 \pmod 2$*.*

Proof. Consider the case $a_0 > 0, a_n > 0$. This means that $p(x) \to \infty$ if $x \to +\infty$ and that $p(x) > 0$ if $x \in (0, \varepsilon)$ therefore the number of intersections of the graph of p with the x-axis must be even. If a root has even multiplicity then the graph only touches the x-axis, if the multiplicity is odd then the graph still intersects the axis, therefore the total number of positive roots must be even as the sum of even number of intersection and even number that comes from root multiplicities. Other cases are treated similarly. ■

The main idea of the proof of Theorem 5.11 is to use the induction, therefore we must have a connection how a polynomial of degree $b_n - 1$ is connected with a polynomial of degree b_n. The former can be obtained by differentiating, hence we will need

Lemma 5.14. *Consider polynomial* (5.4) *and let* p' *be its derivative. Then*

$$r(p') \geq r(p) - 1. \tag{5.5}$$

Exercise 1. Recalling the facts that if p has a root of multiplicity k then p' has the same root of multiplicity $k-1$, and Rolle's theorem that states that if for differentiable f it is true that $f(a) = f(b)$ then there is $c \in (a, b)$ such that $f'(c) = 0$, prove this lemma.

Proof of Theorem 5.11. Assume without loss of generality that $b_0 = 0$ in (5.4).

We will prove the theorem by induction.

For the base case, take the polynomial $p(x) = a_0 + a_1 x$. By considering all possible cases of signs of a_0, a_1, we get that it is always true that $r(p) = v(p)$.

For the induction step suppose that the theorem is true for any polynomial of degree $b_n - 1$ and again consider (5.4) with $b_0 = 0$. Therefore we have

$$p'(x) = a_1 b_1 x^{b_1 - 1} + \ldots + a_n b_n x^{b_n - 1},$$

and the assumption is that

$$r(p') \leq v(p'), \quad r(p') \equiv v(p') \pmod 2. \tag{5.6}$$

Two cases are possible. Case 1: $a_0 a_1 > 0$. Hence the total number of variations in sings for both p and p' is the same, and moreover, $r(p) \equiv r(p') \pmod 2$ since if $a_0 a_n > 0$ then $a_1 a_n > 0$ or if $a_0 a_n < 0$ then $a_1 a_n < 0$ and invoking Lemma 5.13. In short,

$$v(p) = v(p'), \quad r(p) \equiv r(p') \pmod 2. \tag{5.7}$$

Now, (below all the congruences are taken $\pmod 2$)

$$v(p) = [\text{due to (5.7)}] = v(p') \equiv [\text{due to (5.6)}] \equiv r(p') \equiv [\text{due to (5.7)}] \equiv r(p),$$

hence the second statement in the theorem in this case has been proven.

Similarly,

$$v(p) = [\text{due to (5.7)}] = v(p') \geq [\text{due to (5.6)}] \geq r(p') \geq [\text{due to (5.5)}] \geq r(p) - 1.$$

But we already proved that $v(p)$ and $r(p)$ are either even or odd simultaneously, hence it is impossible to have $v(p) = r(p) - 1$ therefore, $v(p) > r(p) - 1$ or, equivalently, $v(p) \geq r(p)$ as required.

Case 2: $a_0 a_1 < 0$. Hence,

$$v(p) = v(p') + 1, \quad r(p) \equiv r(p') + 1 \pmod 2. \tag{5.8}$$

Therefore,

$$v(p) = [\text{due to (5.8)}] = v(p') + 1 \equiv [\text{due to (5.6)}] \equiv r(p') + 1 \equiv [\text{due to (5.8)}] \equiv r(p),$$

and the second statement has been proven.

Similarly,

$$v(p) = [\text{due to (5.8)}] = v(p') + 1 \geq [\text{due to (5.6)}] \geq r(p') + 1 \geq [\text{due to (5.5)}] \geq r(p)$$

as required. The proof is finished. ■