The Exponential Function, Polar Coordinates, and the Fundamental Theorem of Algebra

Polynomials in the complex plane. Say we have a polynomial in a real variable x:

$$p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \tag{1}$$

To extend it to a polynomial; in the complex variable $z = x + iy$ is is natural to use:

$$p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0 \tag{2}$$

Although the real polynomial $x^2+4 = 0$ has no real roots, the complex polynomial $z^2+4 = 0$ has two roots, $z = 2i$ and $z = -2i$.

Fundamental Theorem of Algebra: *A polynomial of degree n has exactly n complex roots, some of which may be multiple roots.*

For instance, $p(z) = z(z-1)^2$ has a root with multiplicity 2 at $z = 1$.

The essential step to proving the Fundamental Theorem of Algebra is to prove that *A polynomial $p(z)$ with degree $n \ge 1$ has at least one complex root.* Using that it has at least one complex root, say z_1. We now show that it has exactly n of them. Consider the polynomial $p_1(z) := p(z)/(z - z_1)$. It has degree $n - 1$ so it too has at least one complex root, say z_2, which is also a root of $p(z)$. Repeat this using $p_2(z) := p_1(z)/(z - z_2)$ which has degree $n - 2$. After n repe88titions we end up with $p(z)$ having exactly n complex roots.

Exponential Function

It will be useful to introduce the exponential function e^z. The definition (2) of $p(z)$ motivates us for how to define e^z, $\cos z$, and $\sin z$ complex z. For real x we have the familiar power series

$$e^x = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^k}{k!} + \cdots$$

which leads us to *define* e^z for complex $z = x + iy$ by

$$e^z = 1 + z + \frac{z^2}{2!} + \cdots + \frac{z^k}{k!} + \cdots . \tag{3}$$

Using the ratio test, which is valid for complex power series, this series converges for all complex z.

An important observation is that the basic property

$$e^{z+w} = e^z e^w \tag{4}$$

still holds for complex z and w, as one can verify by multiplying the power series for e^z and e^w and collecting the terms – just as in the case when z and w are real.

Similarly we are led to define $\cos z$ and $\sin z$ by the power series

$$\cos z = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \frac{z^6}{6!} + \cdots , \qquad \sin z = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \frac{z^7}{7!} + \cdots .$$

It is clear that adding the power series for $\cos z$ and $\sin z$ you almost get the power series for e^z – *except* for the minus signs. This inspired Euler (1748) to observe that using e^{iz} instead of e^z one obtains the beautiful identity

$$e^{iz} = \cos z + i \sin z. \tag{5}$$

The special case $z = \pi$ is the celebrated formula

$$e^{i\pi} = -1$$

relating the four basic constants e, i, π, and -1. One might never suspect a stunning formula like this. We also see that $e^{2k\pi i} = 1$ for any integer k.

An immediate consequence of (5) is DeMoivre's identity (1707)

$$(\cos\theta + i\sin\theta)^n = \left(e^{i\theta}\right)^n = e^{in\theta} = \cos n\theta + i \sin n\theta, \tag{6}$$

which he perhaps proved by induction.

Polar Coordinates

We will use (5) to introduce *polar coordinates* in the complex plane. Let $z = x + iy$ and let $r = |z| = \sqrt{x^2 + y^2}$ be the distance from z to the origin. Also, let θ, $0 \le \theta < 2\pi$, be the unique angle between z and the x-axis (θ is not defined at $z = 0$). Then $x = r\cos\theta$ and $y = r\sin\theta$ so using (5)

$$z = x + iy = r(\cos\theta + i\sin\theta) = re^{i\theta}. \tag{7}$$

In the previous paragraph we choose $0 \le \theta < 2\pi$. Below it will be useful to observe that by the 2π periodicity of $\cos\theta$ and $\sin\theta$, for any integer k, $e^{2k\pi i} = 1$, we also have $e^{i\theta} = e^{i(\theta + 2k\pi)}$.

Geometric Interpretation of Complex Multiplication. If $w = |w|e^{i\varphi}$, then the product

$$\begin{aligned} zw =& |z|e^{i\theta}|w|e^{i\varphi} = |z||w|e^{i(\theta+\varphi)} \\ =& |z||w|\big(\cos(\theta+\varphi) + i\sin(\theta+\varphi)\big) \end{aligned}$$

so the absolute values are multiplied and the angles are added. This geometric interpretation of multiplication in the complex plane was exploited by Gauss and Wessel around 1799 and used effectively by Argand in 1806, long after the work of DeMovire and Euler (although surely they were smart enough to understand intuitively what this was all about).

Solving $\mathbf{z^n = c}$

As an application we show that the important special polynomial equation $z^n = c$ has exactly n distinct complex roots – by finding the roots explicitly.

As a warm-up, we compute $\sqrt{i}$, that is, to solve $z^2 = i$. Seek z in polar form $z = re^{i\theta}$, where $r > 0$ and $0 \le \theta < 2\pi$ Then we want $i = z^2 = r^2 e^{2i\theta}$ so $r = 1$ and $i = \cos 2\theta + i \sin 2\theta$. Thus $\cos 2\theta = 0$ and $\sin 2\theta = 1$. The only possibilities are $2\theta = \pi/2$ or $2\theta = 3\pi/2$. Consequently $\theta = \pi/4$ or $\theta = 3\pi/4$. The roots are therefore $z = e^{i\pi/4}$ and $z = e^{3\pi i/4}$, that is, $z = \pm(1+i)/\sqrt{2}$.

A more complicated example: Solve $z^3 = 2i$.

STEP 1 Write $c = 2i$ in polar form $c = \rho e^{i\alpha}$. First, $\rho = |c| = 2$. Then note that $i = e^{i\frac{\pi}{2}}$. But since we want *all* of the solutions of $z^3 = 2i$ write $i = e^{i\left(\frac{\pi}{2}+2k\pi\right)}$ where the possible integers k will be chosen later. Seek the solution z as $z = re^{i\theta}$ for all $0 \le \theta < 2\pi$. Then $z^3 = 2i$ is

$$r^3 e^{3i\theta} = 2e^{i\left(\frac{\pi}{2}+2k\pi\right)}.$$

STEP 2 Matching the terms we find

$$r^3 = 2 \quad \text{and} \quad 3\theta = \frac{\pi}{2} + 2k\pi,$$

so $r = 2^{1/3}$ and $\theta = \frac{\pi}{6} + \frac{2k\pi}{3}$. To find the values of θ recall that we want all the θ's in $0 \le \theta < 2\pi$. Therefore $k = 0$, $k = 1$ and $k = 2$ are all possible. This gives the three solutions

$$z_1 = 2^{1/3} e^{\frac{\pi}{6}i}, \qquad z_2 = 2^{1/3} e^{\frac{5\pi}{6}i}, \qquad \text{and} \qquad z_3 = 2^{1/3} e^{\frac{9\pi}{6}i}.$$

If we use integers $k \ge 3$, we just repeat these same roots.

More generally, to solve $z^n = c$ write c in polar form $c = |c|e^{i\alpha}$ and seek $z = re^{i\theta}$. Then, as in the example, the equation is

$$r^n e^{n\theta i} = |c| e^{(\alpha+2k\pi)i}$$

so

$$r = |c|^{1/n}, \qquad \theta = \frac{\alpha + 2k\pi}{n}.$$

The n values of k we use are $k = 0, 1, \ldots, n-1$. This gives n distinct solutions. Notice that $e^{\frac{2k\pi}{n}i}$, $k = 0, 1, \ldots, n-1$, are the n solutions of $z^n = 1$.

REMARK We can use the same procedure to find an infinite number of solutions of $z^{\sqrt{2}} = 1$ (Exercise). Note that $z^{\sqrt{2}} - 1$ is *not* a polynomial.

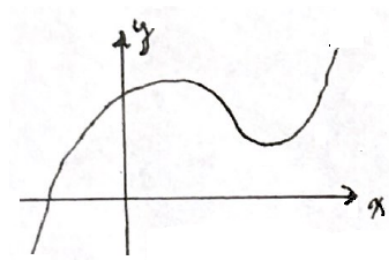
The Fundamental Theorem of Algebra

SOME HISTORY: For quadratic polynomials, in school we learned an explicit formula for the roots. For cubic and quartic polynomials much more complicated formulas were found by Ferro, Tartaglia, Ferrari, and Cardano (1500- 1550). The history is a bit messy. These formulas for the roots involved only sums, products, and roots of the coefficients. Note that the equal sign = was not invented until 1557 (by Robert Recorde).

For several centuries mathematicians sought a similar formula for the roots of polynomials of degree five – but did not succeed. Finally Abel (1824) and in greater depth, Galois (1832), showed there are no such general formulas for polynomials of degree greater than four.

We will prove that roots, possibly involving complex numbers, do exist. The proof does not give a formula for the root.

As a simple example of such an existence theorem with no formula for the solution, we will show that every cubic polynomial $p(x) := x^3 + ax^2 + bx + c$ with real coefficients has at least one real root. Look at a graph of $p(x)$. For large positive x, $p(x) = x^3 +$ lower order terms so $p(+\infty) > 0$. Similarly, $p(-\infty) < 0$. Thus by the intermediate value theorem $p(x_0) = 0$ for at least one real x_0. Similarly, every real polynomial whose degree is odd has at least one real root. This proof by a picture could not have been given before 1637 when Fermat and Descartes gave us analytic geometry to graph curves.



This proof fails for polynomials whose degree is even, $q(x) = x^{2k} +$ (lower order terms) since $q(\pm\infty) = +\infty$. For such polynomials, using $x = \pm\infty$ is crude; for real x infinity does not have much structure. Of course, the above proof only reveals a real root of a polynomial and a polynomial of even degree might not have a real root. The key idea to adapt the above proof to obtain roots, possibly complex, for *all* polynomials is to observe that for $p(z) = z^n$, using polar coordinates lifts a veil revealing that *infinity has a much richer structure* for both even and odd degrees. This is evident using polar form $z = re^{i\theta}$ since then $z^n = r^n e^{in\theta}$. For large $|z| = r$, we know that r is large but the angle θ can be anything (for z real, $\theta = 0$ or π is too constraining). Most proofs of the Fundamental Theorem of Algebra – including the one here – use this observation to adapt the above geometric proof we gave for cubic polynomials.

The Maximum and Minimum Principles

Our proof of the existence of roots of polynomials relies on an important geometric understanding of a polynomial $w = p(z)$ as a map from the complex z-plane to the complex w-plane. Note there are *two* separate pictures, one of the z-plane and one for the w-plane. A few examples show the idea.

Example 0. The simplest example is $w = p(z) = a$, where a is some constant. This maps every point to the same point $w = a$. Completely uninteresting.

Example 1. The next simplest example is $w = p(z) = a + z$. In polar coordinates $z = re^{i\theta} = r\cos\theta + ir\sin\theta$

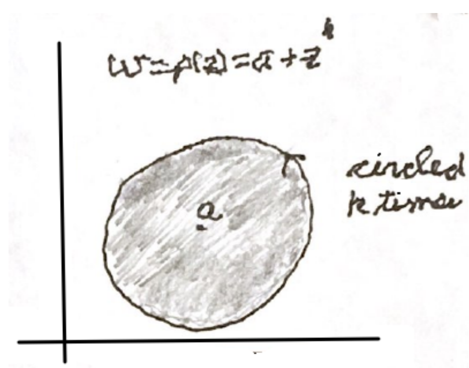
$$w - a = re^{i\theta}.$$

Then for fixed $r > 0$ and $0 \le \theta < 2\pi$, w describes a circle centered at a with radius r that circles a once.

Example 2. Similarly, for any integer $k \ge 1$, say $w = p(z) = a + z^k$, so

$$w - a = r^k e^{ik\theta}.$$

Then for fixed $r > 0$ and $0 \le \theta < 2\pi$, as z circles the origin once in the z-plane then w describes a circle centered at $w = a$ with radius $R = r^k$ that circles the point a k times.



EXAMPLE 3. Slightly more complicated: $w = p(z) = a + bz^k$ where $b \neq 0$ is a constant. Say in polar form, $b = \rho e^{i\beta}$. Then

$$w = a + bz^k, \quad \text{so} \quad w - a = \rho r^k e^{i(\beta + k\theta)}. \tag{8}$$

For fixed $r > 0$ and $0 \leq \theta < 2\pi$, as z circles the origin once in the z-plane then w describes w describes a circle centered at a radius ρr^k that circles the point a k times. Thus if z is in the disk of radius r centered at he origin, then the corresponding points w given by (8) are exactly the points in the disk with center a and radius $\rho r^k = |b| r^k$ (covered k times).

While Examples 1–3 are different, we now give a crude, yet valuable, geometric summary that covers all of them:

LOCAL MAX AND MIN OF $|p(z)|$. There is a point z_1 near $z = 0$ (in fact, many points), in the z-plane whose image, $w_1 = p(z_1)$ is further from the origin than $a = p(0)$: $|p(0)| < |p(z_1)|$.
If $a \neq 0$ the same picture shows there are many points z_2 near $z = 0$ whose image, $p(z_2) \neq 0$ is closer to the origin, $0 < |p(z_2)| < |p(0)| = |a|$.

We want to extend our geometric insight to general polynomials. The result answers the following specific question: Say we are investigating a polynomial $p(z)$ where z is in a disk $D := \{|z| < R\}$. Where in D is $|p(z)|$ largest? Where is is $|p(z)|$ smallest? The (crude) result is surprisingly simple: $|p(z)|$ is largest at some point z on the boundary of D, that is, on the circle $|z| = R$, not inside the disk D. Also, if p is not zero in D then $|p(z)|$ is smallest at some other point z on the boundary of D, not inside D.
To prove this we closely examine a polynomial $p(z)$ locally. near a point, say z_0 (in the above Examples we used $z_0 = 0$). For this we write p as a Taylor polynomial

$$p(z) = b_0 + b_1(z - z_0) + b_2(z - z_0)^2 + \cdots + b_n(z - z_0)^n \tag{9}$$

For z near z_0 as k increases the terms $b_k(z - z_0)^k$ become less important. Thus the most important term (after b_0) will be the smallest $k \geq 1$ for which $b_k \neq 0$. Then, near z_0

$$p(z) = b_0 + b_k(z - z_0)^k + (z - z_0)^{k+1}\left[b_{k+1} + \cdots + b_n(z - z_0)^{n-(k+1)}\right] \tag{10}$$

The polynomial

$$\begin{aligned} p(z) =& z^3 - 3z^2 + 3z - 1 \\ =& (z - 1)^3 \end{aligned}$$

looks quite different near $z = 0$ and near $z = 1$, exhibiting behavior of both Examples 1 and Example 2.

This leads us to the following extension of the Local Max and Min interpretation of Example 3 above[1]. It is the main technical step in our approach. At first glance it is not obvious how useful this result is.

[1] The Examples suggest an even stronger more intuitive result. Let $w = p(z)$ with $w_0 = p(z_0)$. If $p(z)$ is not identically constant, then the image of a small disk $\{|z - z_0| < \epsilon\}$ contains a small disk $\{|w - w_0) < \delta\}$ around w_0. In brief, p maps open open sets in the z-plane to open sets in the w-plane. This is true but more complicated to prove.

Lemma 1. Local Max and Min of $|p(z)|$. Say near $z = z_0$ the polynomial $p(z)$ is

$$w := p(z) = a + b(z - z_0)^k + g(z)(z - z_0)^{k+1}, \tag{11}$$

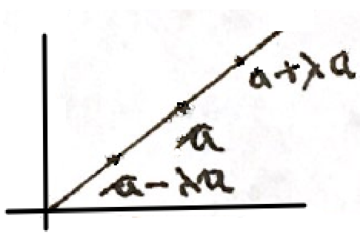
where $b \neq 0$ and $g(z)$ is bounded for $|z - z_0|$ small: $|g(z)| \leq M$ (for equation (10) this clearly holds for $|z - z_0| < 1$).

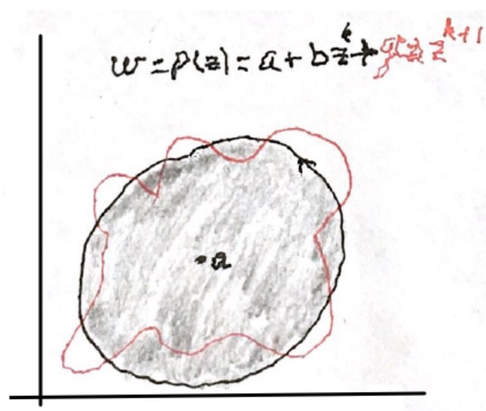
(i). There is a point z_1 (in fact many points) near z_0 whose image $p(z_1)$ is further from the origin than $a = p(z_0)$: $|p(z_1)| > |p(z_0)|$. Thus the real-valued function $|p(z)|$ cannot have a local maximum at z_0.

(ii). If $a \neq 0$ there is a point z_2 near $z = z_0$ such that $p(z_2)$ is closer to the origin than $p(z_0)$: $0 < |p(z_2)| < |p(z_0)| = |a|$. Thus, if $p(z_0) \neq 0$ the real-valued function $|p(z)|$ cannot have a local minimum at z_0.

Proof of (i). By a translation in the z-plane we may assume that $z_0 = 0$. The case $a = 0$ is trivial (pick any small z_1) so assume $a \neq 0$. We will follow the special case $g(z) = 0$ of Example 3. The idea is that if z is small, then the term $|g(z)z^{k+1}|$ will be smaller than $|bz^k|$ so near $z = 0$ the picture will be a small perturbation of Example 3.

To prove this we use a simple explicit choice of the point z_1: choose z_1 so that in the w-plane bz_1^k is on the line from the origin to a, that is, $bz_1^k = \lambda a$ for $\lambda > 0$ to be chosen shortly. Thus, $z_1^k = \lambda a/b$ (we solved equations of the form $z^n = c$ above). Note that $a + bz_1^k = (1 + \lambda)a$ is *further* from the origin than a.





Also, pick z_1 so that $|g(z)z_1^{k+1}| < \frac{1}{2}|bz_1^k|$, that is, $|Mz_1| < \frac{1}{2}|b|$. This will hold by choosing $\lambda > 0$ sufficiently small. Then by the triangle inequality[2]

$$\begin{aligned} |p(z_1)| &\geq |a + bz_1^k| - |g(z)z_1^{k+1}| \\ &\geq (1 + \lambda)|a| - \tfrac{1}{2}|bz_1^k| \\ &= (1 + \lambda)|a| - \tfrac{1}{2}\lambda|a| > |a| = |p(0)| \end{aligned}$$

Thus $p(z_1)$ is further from the origin than a.

Proof of (ii). This is quite similar. Pick z_2 so that bz_2^k is on the line from the origin to a, but so that $a + bz_2^k = (1 - \lambda)a$ is *closer* to the origin than a, that is, $bz_2^k = -\lambda a$ for some $0 < \lambda < 1$. Thus, $z_2^k = -\lambda a/b$. Again pick $\lambda > 0$ so small that $|Mz_2| < \frac{1}{2}|b|$.

[2] Write $|A + B| \leq |A| + |B|$ in the form $|B| \geq |A + B| - |A|$.

By the triangle inequality

$$\begin{aligned}|p(z_2)| \leq& |a + bz_2^k| + |g(z)z_2^{k+1}| \\ \leq& |a - \lambda a| + \tfrac{1}{2}|bz_2^k| \\ =& (1-\lambda)|a| + \tfrac{1}{2}\lambda|a| < |a| = |p(0)|\end{aligned}$$

The following theorem summarizes the local results of the Lemma.

Theorem Let D be any disk, $\{|z| < R\}$ in the complex plane.

(a). [**Maximum Principle**] The function $|p(z)|$ cannot have a local maximum inside D. In particular, a local maximum of $|p(z)|$ in $\{|z| \leq R\}$ can only occur on the boundary $\{|z| = R\}$. In other words, the point z in $\bar{D} = \{|z| \leq R\}$ where $p(z)$ is furthest from the origin is on the boundary of D, that is, on the circle $\{|z| = R\}$.

(b). [**Minimum Principle**] If $p(z)$ is not zero at any point of D, then $|p(z)|$ can not have a local minimum in D. In particular, the minimum of $|p(z)|$ in $\bar{D}$ can only occur on the boundary $\{|z| = R\}$. Consequently, the point z in $\bar{D}$ where $p(z)$ is closest to the origin is on the circle $\{|z| = R\}$.

PROOF. a). Since $\bar{D}$ is a closed and bounded set, the continuous real-valued function $|p(z)|$ has it's maximum value at some point of $\bar{D}$. But by Lemma 1 it cannot have a local maximum at any point inside D. Thus the maximum buts be somewhere on the boundary of D.

b). If $p(z)$ is not zero at any point in D, then by the Lemma $|p(z)|$ cannot have its minimum at any point inside D. Thus the minimum must be at some point on the boundary of D.

What is special about polynomials?

As motivation for the next step, we note that the above Maximum/Minimum principle actually hold for much more general regions D (any bounded open set) and functions than polynomials (any function having a convergent complex power series). An example is the exponential function e^z whose power series converges for all complex z. But since $e^z e^{-z} = 1$, the function e^z is never zero. Thus, to prove that polynomials do have complex zeros we need to use a special property of polynomials whose degree is at least 1: they get large at infinity.

Lemma 2. [$\lim_{|z|\to\infty} |p(z)| = \infty$] [3]. Let $p(z)$ be as in (2) with $n \geq 1$ and $M := |a_{n-1}| + \cdots + +|a_1| + |a_0|$. If $|z| \geq 1$ then

$$|p(z)| \geq |z|^{n-1}(|z| - M). \tag{12}$$

For any $c > 0$, if $|z| \geq 1$ and $|z^{n-1}(z - M)| > c$, then $|p(z)| > c$. In particular, the zeroes of $p(z)$ lie the disk $|z| \leq R$ where $R = \max(1, M)$.

[3] The converse of this is also true: Say $f(z) = \sum_0^\infty a_k z^k$ and the series converges for all complex z. If $|f(z)| \to \infty$ as $|z| \to \infty$, then f is a polynomial.

PROOF: We show that for $|z|$ large, z^n is the dominant term in $p(z)$. We first estimate the lower order terms in $p(z)$. For $|z] \ge 1$.

$$\begin{aligned}|a_{n-1}z^{n-1}| + \cdots + a_1 z + a_0| \le& \big(|a_{n-1}| + \cdots + |a_1| + |a_0|\big)\, |z|^{n-1}\\ =& M|z|^{n-1}.\end{aligned}$$

Therefore

$$\begin{aligned}|z^n| =& \big|p(z) - (a_{n-1}z^{n-1} + \cdots + a_1 z + a_0)\big)\, |\\ \le& |p(z)| + M|z|^{n-1}.\end{aligned}$$

Consequently,

$$|z|^{n-1}(|z| - M) \le |p(z)|.$$

Thus if $|z|$ is large, so is $|p(z)|$. In particular, since this estimate assumed $|z| \ge 1$, all of the zeroes of $p(z)$ lie in the disk of radius $R = \max(1, M)$.

Proof of the Fundamental Theorem of Algebra

PROOF: By contradiction, say $p(z)$ is never zero. We use the Minimum Principle concerning the minimum of $|p(z)|$. Let $\bar{D}$ be the closed disk $\{|z| \le R\}$, where R is a large number to be chosen.

Since $\bar{D}$ is a closed and bounded set, the continuous function $|p(z)|$ attains its minimum at some point z_{min} of $\bar{D}$. Note that $|p(z_{\text{min}})| \le |p(0)|$.

By Lemma 2 we can find R so that if $|z| \ge R$ then $|p)z)| > |p(0)| \ge |p(z_{\text{min}})|$. Thus z_{min} must satisfy $|z_{\text{min}}| < R$. Since we assumed that $p(z)$ has no zeroes, his contradicts the Minimum Principle.

REMARK: As you noticed, I never used the maximum principle here. I included it because it makes the story clearer and because it is so useful in other applications.