Complex numbers, the exponential function, and factorization over $\mathbb{C}$

1 Complex Numbers

Recall that for every non-zero real number x, its square $x^2 = x \cdot x$ is always positive. Consequently, $\mathbb{R}$ does not contain the square roots of any negative number. This is a serious problem which rears its head all over the place.

It is a non-trivial fact, however, that any positive number has two square roots in $\mathbb{R}$, one positive and the other negative; the positive one is denoted $\sqrt{x}$. One can show that for any x in $\mathbb{R}$,

$$|x| = \sqrt{x \cdot x}.$$

So if we can somehow have at hand a square root of -1, we can find square roots of any real number.

This motivates us to declare a new entity, denoted i, to satisfy

$$i^2 = -1.$$

One defines the set of *complex numbers* to be

$$\mathbb{C} = \{x + iy \,|\, x, y \in \mathbb{R}\}$$

and defines the basic arithmetical operations in $\mathbb{C}$ as follows:

$$(x + iy) \pm (x' + iy') = (x \pm x') + i(y \pm y'),$$

and

$$(x + iy)(x' + iy') = (xx' - yy') + i(xy' + x'y).$$

There is a natural one-to-one function

$$\mathbb{R} \to \mathbb{C},\ x \to x + i.0,$$

compatible with the arithmetical operations on both sides.

It is an easy exercise to check all the field axioms, except perhaps for the existence of multiplicative inverses for non-zero complex numbers. To this end one defines the *complex conjugate* of any $z = x + iy$ in $\mathbb{C}$ to be

$$\overline{z} = x - iy.$$

Clearly,

$$\mathbb{R} = \{z \in \mathbb{C} \,|\, \overline{z} = z\}.$$

If $z = x + iy$, we have by definition,

$$z\overline{z} = x^2 + y^2.$$

In particular, $z\overline{z}$ is either 0 or a positive real number. Hence we can find a non-negative square root of $z\overline{z}$ in $\mathbb{R}$. Define the *absolute value*, sometimes called *modulus* or *norm*, by

$$|z| = \sqrt{z\overline{z}} = \sqrt{x^2 + y^2}.$$

If $z = x + iy$ is not 0, we will put

$$z^{-1} = \frac{\overline{z}}{z\overline{z}} = \frac{x}{x^2 + y^2} - i\frac{y}{x^2 + y^2}.$$

It is a complex number satisfying

$$z(z^{-1}) = z\frac{\overline{z}}{z\overline{z}} = 1.$$

Done.

It is natural to think of complex numbers $z = x + iy$ as being ordered pairs (x, y) of real numbers. So one can try to visualize $\mathbb{C}$ as a plane with two perpendicular coordinate directions, namely giving the x and y parts. Note in particular that 0 corresponds to the origin $O = (0, 0)$, 1 to $(1, 0)$ and i with $(0, 1)$. Geometrically, one can think of getting from -1 to 1 (and back) by rotation about an angle π, and similarly, one gets from i to its square -1 by rotating by half that angle, namely $\pi/2$, in the counterclockwise direction.

To get from the other square root of -1, namely $-i$, one rotates by $\pi/2$ in the clockwise direction. (Going counterclockwise is considered to be in the positive direction in Math.)

Addition of complex numbers has then a simple geometric interpretation: If $z = x+iy$, $z' = x'+iy'$ are two complex numbers, represented by the points $P = (x, y)$ and $Q = (x', y')$ on the plane, then one can join the origin O to P and Q, and then draw a parallelogram with the *line segments* OP and OQ as a pair of adjacent sides. If R is the fourth vertex of this parallelogram, it corresponds to $z + z'$. This is called the *parallelogram law*.

Complex conjugation corresponds to *reflection* about the x-axis.

The *absolute value* or *modulus* $|z|$ of a complex number $z = x + iy$ is, by the Pythagorean theorem applied to the triangle with vertices $O, P = (x, y)$ and $R = (x, 0)$, simply the *length*, often denoted by r, $\sqrt{x^2 + y^2}$ of the line OP.

The *angle* θ between the line segments OR and OP is called the *argument* of z. The pair (r, θ) determines the complex number z. Indeed High school trigonometry allows us to show that the coordinates of z are given by

$$x = r\cos\theta \quad \text{and} \quad y = r\sin\theta,$$

where cos (or cosine) and sin (or sine) are the familiar trigonometric functions. Consequently,

$$z = r(\cos\theta + i\sin\theta).$$

Those who know about *exponentials* (to be treated below in section 9.4) will recognize the identity

$$e^{i\theta} = \cos\theta + i\sin\theta.$$

(This can also be taken as a definition of $e^{i\theta}$, for any $\theta \in \mathbb{R}$.)

Note that $e^{i\theta}$ has absolute value 1 and hence lies on the *unit circle* in the plane given by the equation $|z| = 1$.

It is customary for the angle θ to be called the **argument** of z, denoted $\arg(z)$, taken to lie in $[0, 2\pi)$.

2 Cardano's formula

This section is mainly for motivational purposes. Recall the well known **quadratic formula** from the days of old, which asserts that the roots of the quadratic equation

$$ax^2 + bx + c \ = \ 0, \quad \text{with} \quad a, b, c \in \mathbb{R},$$

are given by

$$\alpha_\pm \ = \ \frac{-b \pm \sqrt{D}}{2a},$$

where the **discriminant** D is $b^2 - 4ac$. Note that

$$D > 0 \implies \exists\, 2 \text{ real roots;}$$

$$D = 0 \implies \exists \text{ a unique real root (with multiplicity 2);}$$

$$D < 0 \implies \nexists \text{ real root.}$$

There were several people in the old days (up to the middle of the nineteenth century), some of them even quite well educated, who did not believe in imaginary numbers, such as square-roots of negative numbers. Their reaction to the quadratic formula was to just ignore the case when $D < 0$ and thus not deal with the possibility of non-real roots. They said they were only interested in real roots. Their argument was shattered when one started looking at the cubic equation

$$ax^3 + bx^2 + cx + d \ = \ 0, \quad \text{with} \quad a, b, c, d \in \mathbb{R}.$$

Thanks to a beautiful formula of the Italian mathematician Cardano, the roots are given by

$$\alpha_1 \ = \ S + T - \frac{b}{3a},$$

$$\alpha_2 \ = \ -(S+T)/2 - \frac{b}{3a} + \frac{\sqrt{-3}}{2}(S-T),$$

$$\alpha_3 \ = \ -(S+T)/2 - \frac{b}{3a} - \frac{\sqrt{-3}}{2}(S-T),$$

with

$$S = (R + \sqrt{D})^{1/3}, \ T = (R - \sqrt{D})^{1/3},$$

where the **discriminant** D is $Q^3 + R^2$, and

$$R = \frac{9abc - 27a^2d - 2b^3}{54a^3}, \ Q = \frac{3ac - b^2}{9a^2}.$$

One has

$$D > 0 \implies \exists \text{ a unique real root;}$$

$$D = 0 \implies \exists\, 3 \text{ real roots with 2 of them equal;}$$

$$D < 0 \implies \exists\, 3 \text{ distinct real roots.}$$

This presented a problem for the Naysayers. One is for sure interested in the case when there are three real roots, but Cardano's formula for the roots goes through an auxiliary computation, namely that of the square-root of D, which involves imaginary numbers!

3 Complex sequences and series

As with real sequences, given a sequence $\{z_n\}$ of complex numbers z_n, we say that it converges to a limit L, say, in $\mathbb{C}$ iff we have, for every $\varepsilon > 0$, we can find an integer $N > 0$ such that

$$n \geq N \implies |L - z_n| < \varepsilon.$$

Proposition 1 *(i) If $\{a_n\}$ is a convergent sequence with limit L, then for any scalar c, the sequence $\{ca_n\}$ is convergent with limit cL;*

(ii) If $\{a_n\}$, $\{b_n\}$ are convergent sequences with respective limits L_1, L_2, then their sum $\{a_n + b_n\}$ and their product $\{a_n b_n\}$ are convergent with respective limits $L_1 + L_2$ and $L_1 L_2$.

The proof is again a simple application of the properties of absolute values. The following Corollary allows the convergence questions for complex sequences to b reduced to real ones.

Corollary 3.1 *Let $\{z_n = x_n + iy_n\}$ be a sequence of complex numbers, with x_n, y_n real for each n. Then $\{z_n\}$ converges iff the real sequences $\{x_n\}$ and $\{y_n\}$ are both convergent.*

Proof. Suppose $\{x_n\}, \{y_n\}$ are both convergent, with respective limits u, v. We claim that $\{z_n\}$ then converges to $w = u + iv$. Indeed, by the Proposition above, $\{iy_n\}$ is convergent with limit iv, and so is $\{x_n + iy_n\}$, with limit w. Conversely, suppose that $\{z_n\}$ converges, say to w. We may write w as $u + iv$, with u, v real. For any complex number $z = x + iy$, $|x|$ and $|y|$ are both bounded by $\leq \sqrt{x^2 + y^2}$, i.e., by $|z|$. Since $w - z_n = (u - x_n) + i(v - y_n)$, we get

$$|u - x_n| \leq |w - z_n| \quad \text{and} \quad |v - y_n| \leq |w - z_n|.$$

For any $\epsilon > 0$, pick $N > 0$ such that for all $n > N$, $|w - z_n|$ is $< \epsilon$. Then we also have $|u - x_n| < \epsilon$ and $|v - y_n|$ is $< \epsilon$ for all $n > N$, establishing the convergence of $\{x_n\}$ and $\{y_n\}$ with respective limits u and v. $\square$

One can define Cauchy sequences as in the real case, and it is immediate that $\{z_n = x_n + iy_n\}$ is Cauchy iff $\{x_n\}$ and $\{y_n\}$ are Cauchy. We have

Theorem 3.2 *A complex sequence $\{z_n\}$ converges iff it is Cauchy.*

Hence $\mathbb{C}$ is also a complete field like $\mathbb{R}$.

An infinite **series** $\sum_{n=n_0}^{\infty} z_n$ of complex numbers is said to be **convergent** iff the sequence of partial sums $\{\sum_{m=n_0}^{n} z_m\}$ is convergent. (Here n_0 is any integer, usually 0 or 1.)

We will say that $\sum_n z_n$ is **absolutely convergent** iff the series of its absolute values, namely $\sum_n |z_n|$ converges.

Note that the question of absolute convergence of a complex series, one is reduced to a real series, since $|z_n|$ is real, even non-negative.

Check that if a complex series $\sum_n z_n$ is absolutely convergent, then it is convergent.

4 The complex exponential function, and logarithm

For any complex number z, we will define its exponential to be given by

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}.$$

This series is absolutely convergent at any z, because the real sequence $\sum_n \frac{r^n}{n!}$ is convergent, with $r = |z|$.

The exponential function has some nice properties, which we state without proof:

$$e^0 = 1,\ e^z = e^x e^{iy},\ e^{z+z'} = e^z e^{z'},$$

for all $z = x + iy, z' \in \mathbb{C}$.

Lemma 4.1 *$e^{i\theta} = \cos\theta + i\sin\theta$, for any real number θ. In particular, $e^{i\theta}$ is periodic of period 2π like the trig functions, and moreover,*

$$|e^{i\theta}| = 1,\ e^{i\pi} = -1.$$

Proof By definition,

$$e^{i\theta} \;=\; 1 + (i\theta) + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \dots,$$

where the even power terms are real and the odd ones are purely imaginary, since $i^{2n} = (-1)^n$ and $i^{2n+1} = (-1)^n i$. Since the series is absolutely convergent, we may rearrange and express it as a sum of two series as follows:

$$e^{i\theta} \;=\; \left(1 - \frac{\theta^2}{2!} + \dots\right) + i\left(\theta - \frac{\theta^3}{3!} + \dots\right).$$

From the Taylor series expansions for the sine and cosine functions, we then see that the right hand side is the sum of $\cos\theta$ and i times $\sin\theta$, as asserted.

The periodicity relative to 2π is now clear. Moreover,

$$|e^{i\theta}| \;=\; \sqrt{\cos^2\theta + \sin^2\theta} \;=\; 1.$$

Finally,

$$\cos(\pi) = -1, \sin(\pi) = 0 \implies e^{i\pi} \;=\; -1.$$

□

Now let $z = x + iy \in \mathbb{C}$. Note that, since $|e^{iy}| = 1$,

$$|e^z| \;=\; e^x.$$

Here we have used the fact that the real exponential is always positive, so $|e^x| = e^x$.

Furthermore, by the periodicity of e^{iy},

$$e^{z+2i\pi} \;=\; e^x e^{i(y+2\pi)} \;=\; e^x e^{iy} \;=\; e^z.$$

So, the complex exponential function is **not one-to-one**, and is in fact periodic of period $2i\pi$. This presents a problem for us, since we would like to define the logarithm as its *inverse*. However, note that e^z is one-to-one if we restrict $z = x + iy$ to lie in the rectangular strip Φ in the complex plane defined by $0 \leq y < 2\pi$.

The **complex logarithm** is defined, for $z \neq 0$, to be

$$\log(z) \;=\; \log|z| + i\arg(z),$$

where $\arg(z)$ is taken to lie (as usual) in $[0, 2\pi)$. Note that since $|z| > 0$ if $z \neq 0$, $\log|z|$ makes sense.

Since $|e^z| = e^x$ (as seen above) and $\arg(e^z) = y$ if $y \in [0, 2\pi)$, we see that

$$\log(e^z) \;=\; \log(e^x) + iy \;=\; x + iy = z,\ \forall z \in \Phi,$$

as desired.

5 Differentiability, Cauchy-Riemann Equations

Let $f(z)$ be a complex valued function of a complex variable $z = x+iy$, with $x, y \in \mathbb{R}$. We will say that f is **differentiable** at a point z_0 in $\mathbb{C}$ iff the following limit exists:

$$\lim_{z\to z_0} \frac{f(z) - f(z_0)}{z - z_0}.$$

when this limit exists, we call it $f'(z_0)$.

It is important to note that **the existence of this limit is a stringent condition**, because, in the complex plane, one can approach a point $z_0 = x_0 + iy_0$ from infinitely many directions. In particular, there are the two independent directions given, for $h \in \mathbb{R}$, by the horizontal one $z_0 + h \to z_0$, and the vertical one $z_0 + ih \to z_0$. The former corresponds to having $x_0 + h \to x_0$ with the y-coordinate fixed, and the latter $y_0 + h \to y_0$ with the x-coordinate fixed. So we must have

$$f'(z_0) \;=\; \lim_{h\to 0} \frac{f(x_0 + h + iy_0) - f(x_0 + iy_0)}{h} \;=\; \lim_{h\to 0} \frac{f(x_0 + i(y_0 + h)) - f(x_0 + iy_0)}{ih}.$$

The two limits on the right define the *partial derivatives* of f, denoted respectively by $\frac{\partial f}{\partial x}(z_0)$ and $-i\frac{\partial f}{\partial y}(z_0)$.

Clearly, given any function φ, real or complex, depending on x, y, we can define the partial derivatives $\partial\varphi/\partial x$ and $\partial\varphi/\partial y$. In any case, we get the equation (when f is differentiable at z_0)

$$\frac{\partial f}{\partial x}(z_0) \;=\; -i\frac{\partial f}{\partial y}(z_0),$$

which is sometimes written as

$$\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right) f(z_0) \;=\; 0.$$

It is also customary to write

$$f(z) \;=\; u(x, y) + iv(x, y),$$

where u, v are real-valued functions of x, y, and taking the real and imaginary parts of the equation above becomes a pair of differential equations, called the *Cauchy-Riemann equations*, at $z_0 = x_0 + iy_0$:

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}.$$

We will now state the following two (amazing) theorems about complex functions without proof:

Theorem 5.1 *Let f be a differentiable function on an open circular disk $D_c(z_0)$ defined by $|z - z_0| < c$, for some $c > 0$. Then f is infinitely differentiable on $D_c(z_0)$. In fact, it is analytic there, meaning that it is represented by its Taylor series in $z - z_0$.*

This is a tremendous contrast from the real situation.

Theorem 5.2 *Let f be a differentiable complex function on all of $\mathbb{C}$. Suppose f is also bounded. Then it must be a constant function.*

Note that this is false in the real case. Indeed, the real function $f(x) = \frac{1}{1+x^2}$ is analytic and bounded, but is not a constant.

6 Factorization over $\mathbb{C}$

The most important result over $\mathbb{C}$, which is the reason people are so interested in working with complex numbers, is the following:

Theorem 6.1 *(***The Fundamental Theorem of Algebra***) Every non-constant polynomial with coefficients in $\mathbb{C}$ admits a root in $\mathbb{C}$.*

We will not prove this result here. But one should become aware of its existence if it is not already the case! We will now give an important consequence.

Corollary 6.2 *Let f be a polynomial of degree $n \geq 1$ with $\mathbb{C}$-coefficients. Then there exist complex numbers $\alpha_1, \ldots, \alpha_r$, with $\alpha_i \neq \alpha_j$ if $i \neq j$, positive integers $m_1, \ldots, m_r$, and a scalar c, such that*

$$f(x) \,=\, c\prod_{j=1}^{r}(x-\alpha_j)^{m_j},$$

and

$$\sum_{j=1}^{r} m_j \,=\, n.$$

In other words, any non-constant polynomial f with $\mathbb{C}$-coefficients **factorizes completely into a product of linear factors**. For each $j \leq r$, the associated positive integer m_j is called the **multiplicity** of α_j as a root of f, which means concretely that m_j is the highest power of $(x-\alpha_j)$ dividing $f(x)$.

Proof of Corollary. Let $n \geq 1$ be the degree of f and let a_n be the non-zero **leading coefficient**, i.e, the coefficient of x^n. Let us set

$$c \,=\, a_n. \tag{6.1}$$

If $n = 1$,

$$f(x) \,=\, a_1x + a_0 \,=\, c(x-\alpha_1) \quad \text{with} \quad \alpha_1 \,=\, -\frac{a_0}{a_1}.$$

So we are done in this case by taking $r = 1$ and $m_1 = 1$.

Now let $n > 1$ and assume by induction that we have proved the assertion for all $m < n$, in particular for $m = n - 1$. By Theorem 9.6, we can find a root, call it β, of f. We may then write

$$f(x) \;=\; (x - \beta)h(x), \tag{6.2}$$

for some polynomial $h(x)$ necessarily of degree $n-1$. The leading coefficients of f are evidently the same. By induction we may write

$$h(x) \;=\; c\prod_{i=1}^{s}(x - \alpha_i)^{k_i},$$

for some roots $\alpha_1, \ldots, \alpha_s$ of h with respective multiplicities $n_1, \ldots, n_s$, so that

$$\sum_{i=1}^{s} k_i \;=\; n - 1.$$

But by (6.2), every root of h is also a root of f, and the assertion of the Corollary follows.

□

7 Factorization over $\mathbb{R}$

The best way to understand polynomials f with real coefficients is to first look at their complex roots and then determine which ones of them could be real. To this end recall first the baby fact that a complex number $z = u + iv$ is real iff z equals its **complex conjugate** $\overline{z} = u - iv$, where $i = \sqrt{-1}$.

Proposition 2 *Let*

$$f(x) \;=\; a_0 + a_1x + \ldots + a_nx^n \quad \text{with} \quad a_j \in \mathbb{R},\ \forall j \leq n, \quad \text{and} \quad a_n \neq 0,$$

for some $n \geq 1$. Suppose α is a **complex root** *of f. Then $\overline{\alpha}$ is also a root of f. In particular, if r denotes the number of real roots of f and s the non-real (complex) roots of f, then we must have*

$$n \;=\; r + 2s.$$

We get the following consequence, which we proved earlier using the *Intermediate value theorem.*

Corollary 7.1 *Let f be a real polynomial of odd degree. Then f must have a real root.*

Proof of Proposition. Let α be a complex root of f. Recall that for all complex numbers z, w,

$$\overline{zw} = \overline{z}\overline{w} \quad \text{and} \quad \overline{z+w} = \overline{z} + \overline{w}. \tag{7.1}$$

Hence for any $j \leq n$,

$$(\overline{\alpha})^j = \overline{\alpha^j}.$$

Moreover, since $a_j \in \mathbb{R}$ $(\forall j)$, $\overline{a}_j = a_j$, and therefore

$$a_j(\overline{\alpha})^j = \overline{a_j\alpha^j}.$$

Consequently, using (7.1) again, we get

$$f(\overline{\alpha}) = \sum_{j=0}^{n} a_j(\overline{\alpha})^j = \overline{f(\alpha)}. \tag{7.2}$$

But α is a root of f (which we have not used so far), $f(\alpha)$ vanishes, as does its complex conjugate $\overline{f(\alpha)}$. So by (7.2), $f(\overline{\alpha})$ is zero, showing that $\overline{\alpha}$ is a root of f.

So the **non-real roots** come in **conjugate pairs**, and this shows that n minus the number r, say, of the **real roots** is even. Done.

□

Given any complex number z, we have

$$z + \overline{z},\ z\overline{z} \in \mathbb{R}. \tag{7.3}$$

This is clear because both the **norm** $z\overline{z}$ and the **trace** $z + \overline{z}$ are unchanged under complex conjugation.

Proposition 3 *Let f be a real polynomial of degree $n \geq 1$ with real roots $\alpha_1, \ldots, \alpha_k$ with multiplicities $n_1, \ldots, n_k$, and non-real roots $\beta_1, \overline{\beta}_1, \ldots, \beta_\ell, \overline{\beta}_\ell$ with multiplicities $m_1, \ldots, m_\ell$ in $\mathbb{C}$. Then we have the factorization*

$$(*) \qquad f(x) \;=\; c \prod_{i=1}^{k} (x - \alpha_i)^{n_i} \cdot \prod_{j=1}^{\ell} (x^2 + b_j x + c_j)^{m_j},$$

where for each $j \leq \ell$,

$$b_j \;=\; -(\beta_j + \overline{\beta}_j) \quad \text{and} \quad c_j \;=\; \beta_j \overline{\beta}_j,$$

Each of the factors occurring in $(*)$ is a real polynomial, and the polynomials $x - \alpha_i$ and $x^2 + b_j x + c_j$ are all irreducible over $\mathbb{R}$.

Proof. In view of Corollary 9.7 and Proposition 2, the only thing we need to prove is that for each $j \leq \ell$, the polynomial

$$h_j(x) \;=\; x^2 + b_j x + c_j$$

is real and irreducible over $\mathbb{R}$. The reality of the coefficients $b_j = -(\beta_j + \overline{\beta}_j)$ and $c_j = \beta_j \overline{\beta}_j$ follows from (7.3). Suppose it is reducible over $\mathbb{R}$. Then we can write

$$h_j(x) \;=\; (x - t_j)(x - t'_j)$$

for some real numbers t_j, t'_j. On the other hand $\beta_j, \overline{\beta}_j$ are roots of h_j. This forces the equality of the sets $\{t_j, t'_j\}$ and $\{\beta_j, \overline{\beta}_j\}$, contradicting the fact that β_j is non-real. So h_j must be irreducible over $\mathbb{R}$.

□