Arithmetic, Geometry and Topology the Complex Plane

In this chapter, we will study the basic properties of the field of complex numbers. We will begin with a brief historic sketch of how the study of complex numbers came to be and then proceed to develop tools needed to study calculus on the complex plane. We will also give several applications of introducing complex numbers to solving classical problems from geometry and trigonometry.

1 Arithmetic of the complex plane

1.1 Why complex numbers?

If you recall from your study of Real Analysis, the introduction of the real numbers as the completion of the field of rational numbers is unavoidable if one wants to develop calculus in an adequate manner. The question arises as to why we need to enlarge $\mathbb{R}$ further and introduce the field $\mathbb{C}$. Complex numbers were first introduced in 1545 by the Italian mathematician Cardano in his *Ars magna* in connection with quadratic equations and he immediately discards them commenting they were "as subtle as they are useless". In fact, complex numbers were often dismissed as "imaginary" or "impossible" even by prominent mathematicians such as Leibniz. So why study them at all?

Recall from high school mathematics that the solution of the quadratic equation $x^2 = mx+c$ is given by the expression

$$x = \frac{1}{2}\left(m \pm \sqrt{m^2+4c}\right). \tag{1.1}$$

The quantity under the square root is called the *discriminant*, denoted D. According to most textbooks, complex numbers were introduced to ensure that quadratic equations always have solutions. This is not only historically inaccurate but also highly misleading.

Geometrically, solving the quadratic equation $x^2 = mx + c$ is same as finding the points of intersection of the parabola P given by the equation $y = x^2$ and the line L given by the equation $y = mx + c$. Three possibilities can arise,

(i) L and P intersects at two points. This corresponds to (1.1) yielding two real solutions. In this case $D > 0$.
(ii) L and P intersects at one point. In this case $D = 0$.
(iii) L and P do not intersect at all. In this case $D < 0$.

Thus, when $D < 0$, the fact that P and L do not intersect is reflected in the fact that (1.1) are complex numbers. This shows that there is absolutely no reason to introduce complex numbers in the study of quadratic equations. Cardano was perfectly justified in dismissing complex numbers as "useless" in connection to solving quadratic equations. That complex numbers were introduced to solve quadratic equations is a lie repeated blindly by many ill-informed textbook authors!

The correct reason for introducing complex numbers is in connection with solving cubics. We begin with the first theorem of the course.

Theorem 1 (Cardano). *The solution of the cubic equation*

$$x^3 = 3px + 2q \tag{1.2}$$

is given by

$$x = \sqrt[3]{q + \sqrt{q^2 - p^3}} + \sqrt[3]{q - \sqrt{q^2 - p^3}}. \tag{1.3}$$

Any arbitrary cubic $x^3 + ax^2 + bx + c$ *can be transformed by a linear change of coordinates to an equation of the form* (1.2).

Proof. It is an exercise for you to prove that an arbitrary cubic can indeed be transformed to an equation of the form (1.2).

To solve (1.2), we first set $x = u + v$. Expanding the LHS of (1.2), we get

$$u^3 + v^3 + 3uv(u + v) = u^3 + v^3 + 3uvx.$$

Equating the above equation with the RHS of (1.2), we see that $p = uv$ and that $u^3 + v^3 = 2q$. Eliminating v, we end up with the equation

$$u^3 + \frac{p^3}{u^3} - 2q = 0.$$

This is a quadratic in u^3 whose solutions are given by

$$u^3 = q \pm \sqrt{q^2 - p^3}.$$

By symmetry, v^3 has the exact same solutions. As $u^3 + v^3 = 2q$, without loss of generality, we can take as the solutions for u^3 and v^3 as

$$u^3 = q + \sqrt{q^2 - p^3}$$
$$v^3 = q - \sqrt{q^2 - p^3}$$

Thus the required solutions are given by

$$x = \sqrt[3]{q + \sqrt{q^2 - p^3}} + \sqrt[3]{q - \sqrt{q^2 - p^3}}.$$

□

Geometrically, solving for the cubic (1.2) is equivalent finding the intersection of the cubic with the line L given by the equation $y = 3px + 2q$. Note that a cubic equation *always* has at least one real root (why?). This means that the formula (1.3) must always yield at least one real number. It was Bombelli who realized that there is something strange about the formula. He considered the cubic $x^3 = 15x + 4$ which has solutions

$$x = \sqrt[3]{2 + 11i} + \sqrt[3]{2 - 11i},$$

here we are freely using complex notation which I am assuming you are already familiar with. The above does not seem to be a real number at all. But, Bombelli had a "wild thought". By guessing, he realized that $x = 4$ solves the cubic. So he assumed that $\sqrt[3]{2 + 11i}$ is an expression of the type $2+ui$ and $\sqrt[3]{2 - 11i}$ is an expression of the form $2-ui$, so that $x = 2+ui+2-ui = 4$. Of course, for this to make sense Bombelli assumed that the ordinary laws of algebra are true for the complex numbers, i.e.,

$$a + ib + a' + ib' = (a + a') + i(b + b').$$

Next to determine u, he needed to evaluate $(2+ui)^3$. To do this, he assumed that multiplication obeys the following obvious rule

$$(a + ib)(a' + ib') = (aa' - bb') + i(a'b + ab'),$$

where we are using $i^2 = -1$. Expanding out $(2 + ui)^3$ using the above rule we get

$$-iu^3 - 6u^2 + 12iu + 8 = 2 + 11i,$$

which readily yields $u = 1$. Similarly, $(2 - i)^3 = 2 - 11i$.

Bombelli's "wild thought" shows that the "useless" complex numbers are unavoidable in the solution of cubics. However, the study of complex numbers remained a mere curiosity and were considered mysterious for almost 250 years.

1.2 Notation and terminology

We will identify the set of complex numbers, denoted $\mathbb{C}$, with $\mathbb{R}^2$. In this identification, the complex number $a + ib$ corresponds to the pair (a, b). The point $1 \in \mathbb{C}$ corresponds to $(1, 0)$ and the point i corresponds to $(0, 1)$. Geometrically, a complex number is nothing but a vector in the so called *Argand–Gauss* complex plane.

The following picture and table summarizes all the relevant notation and terminology related to complex numbers.

Terminology	Meaning	Notation
modulus of z	length r of the vector z	$\lvert z\rvert$
argument of z	angle θ that the vector z makes with the x-axis	$\arg(z)$
real part of z	x coordinate of the vector z	$\mathrm{Re}(z)$
imaginary part of z	y coordinate of the vector z	$\mathrm{Im}(z)$
real axis	the set of real numbers	
imaginary number	a number that is a real multiple of i	
imaginary axis	the set of imaginary numbers	
complex conjugate of z	reflection of z in the real axis	$\overline{z}$

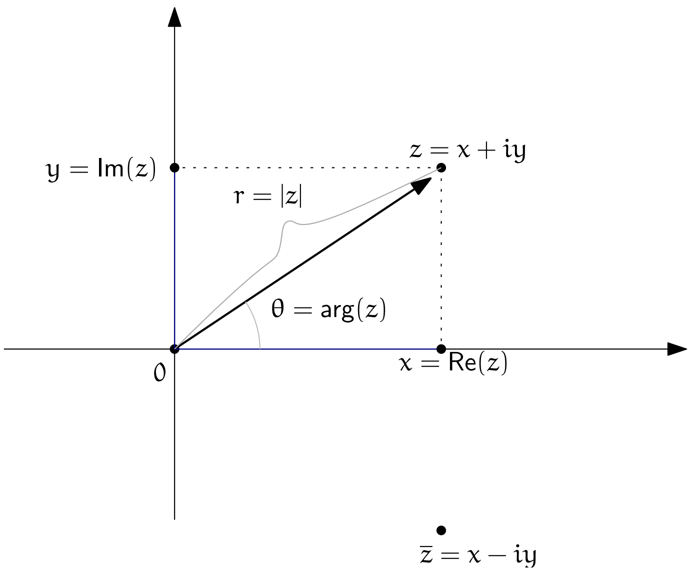

Figure 1: The complex plane

Note that for $z = x + iy, \bar{z} = x - iy$ and

$$\mathrm{Re}(z) = \frac{1}{2}(z + \bar{z}), \quad \mathrm{Im}(z) = \frac{1}{2i}(z - \bar{z}).$$

The sum of two complex numbers z and w can be obtained geometrically using the parallelogram law for addition of vectors.

In order to visualize the product, we need to introduce the polar representation of complex numbers in terms of r and θ. The modulus of $|z|$ is the distance from the origin to the point (x, y). Explicitly, $|z| = \sqrt{x^2 + y^2} = \sqrt{z \cdot \bar{z}}$. The modulus satisfies a number of simple and easy to prove inequalities.

$$\left|\sum_{i=1}^{n} z_i\right| \leqslant \sum_{i=1}^{n} |z_i|,$$
$$||z| - |w|| \leqslant |z \pm w|,$$
$$|z_1 \cdots z_n| \leqslant |z| \cdot |w|.$$

With the sum and product defined as in the previous section, $\mathbb{C}$ is a commutative field. The inverse of the number $z \neq 0$ is given by $z^{-1} = \frac{1}{z} = \frac{\bar{z}}{|z|^2}$.

We now want rigorously define the the argument $\arg(z)$. To do this we will have to first introduce trigonometric functions. The correct way to do this is to use power series which we shall indulge in at a later time. For the time being, we recall the definitions in terms of circular functions that you are no doubt familiar with from calculus.

First of all, we define 2π to be the length of the circumference of the circle of radius 1. Denote the unit circle by $\mathbb{T}$. We now define a map $1_t : \mathbb{R} \to \mathbb{T}$ as follows: for $t \geqslant 0, 1_t$ is the point on $\mathbb{T}$ obtained by starting at the point $1 \in \mathbb{T}$ and moving a distance t counter-clockwise; for $t < 0$ one does the same thing but clockwise. Note that, by definition, 1_t is a 2π-periodic function. We define

$$\sin t := \mathsf{Im} 1_t, \quad \cos t := \mathsf{Re} 1_t.$$

All the familiar identities of trigonometry can easily by derived by using the above definition. The fact that $\sin^2 t + \cos^2 t = 1$ is now obvious as $(\cos t, \sin t)$ lies on the unit circle. Other identities such as the following fundamental identities:

$$\begin{aligned} \cos(t+s) &= \cos t \cos s - \sin t \sin s \\ \sin(t+s) &= \sin t \cos s + \cos t \sin s, \end{aligned} \tag{1.4}$$

can be proved using geometric arguments. These identities prove that

$$1_s \cdot 1_t = 1_{t+s},$$

which says that 1_t is a group homomorphism from the additive group $(\mathbb{R}, +)$ to the multiplicative group $(\mathbb{T}, \cdot)$ with kernel $2\pi\mathbb{Z}$. We immediately get the de Moivre's formula $(\cos t + i \sin t)^n = \cos nt + i \sin nt$. Classical trigonometric identities such as formulas for $\cos 2t, \sin 2t$, etc., easily follow. One also gets the usual differentiation rules for sin and cos easily.

If $0 \neq z \in \mathbb{C}$, then $z/|z| \in \mathbb{T}$, which implies that we can find $\theta \in \mathbb{R}$ such that $1_\theta = z/|z|$.

Definition 2. The *argument* of $0 \neq z \in \mathbb{C}$, denoted $\arg z$, is any number $\theta \in \mathbb{R}$ such that $1_\theta = z/|z|$. Among the all the arguments of z, there is precisely one that belongs to the interval $(-\pi, \pi]$ called the *principal argument* of z and denoted $\operatorname{Arg} z$.

Sometimes we will be sloppy with notation and let $\arg z$ denote the set $\{\theta + 2\pi k, k \in \mathbb{Z}\}$ the set of all possible arguments of z. The argument of z is uniquely determined modulo an integer multiple of 2π. The definition of $\arg z$ combined with (1.4) easily implies that

$$\arg zw = \arg z + \arg w, \quad \arg \frac{1}{z} = \arg \overline{z} = -\arg z.$$

We define the *angle between z and w* when $z, w \neq 0$ as the angle that goes from z to w. i.e., $\arg w - \arg z$. This is the same as $\arg w\overline{z}$ and $\arg \frac{w}{z}$. Note that these are oriented angles and might be negative.

One can determine $\operatorname{Arg} z$ by using the inverse tangent function arctan as follows. Recall that the arctan function is a bijective continuous function from $\mathbb{R}$ to $(-\pi/2, \pi/2)$. If $z = x + iy, x > 0$, then $\operatorname{Arg} z = \arctan \frac{y}{x}$. If $z = iy, y > 0$, then $\operatorname{Arg} z = \pi/2$. In the second quadrant $z = x + iy, x < 0, y > 0$, $\operatorname{Arg} z = \pi + \arctan \frac{y}{x}$ and in the third quadrant $z = x + iy, x < 0, y < 0, \operatorname{Arg} z = \pi - \arctan \frac{y}{x}$. Finally, if $z = iy, y < 0$, then $\operatorname{Arg} z = -\pi/2$.

We are now in a position to give a geometric interpretation of complex multiplication. We represent the complex number z using the *polar representation* as r_θ, where $r = |z|$ and $\theta = \arg z$. Then as $|zw| = |z| \cdot |w|$ and $\arg zw = \arg z + \arg w$, it follows multiplying w z is same as dilation by the real number $|z|$ composed with a rotation by angle θ (the order in which the dilation and rotation is performed is irrelevant). In particular, z itself can be written as $|z| \cdot 1_{\arg z}$. Very soon, we shall give a better way to represent complex numbers in terms of the exp function.

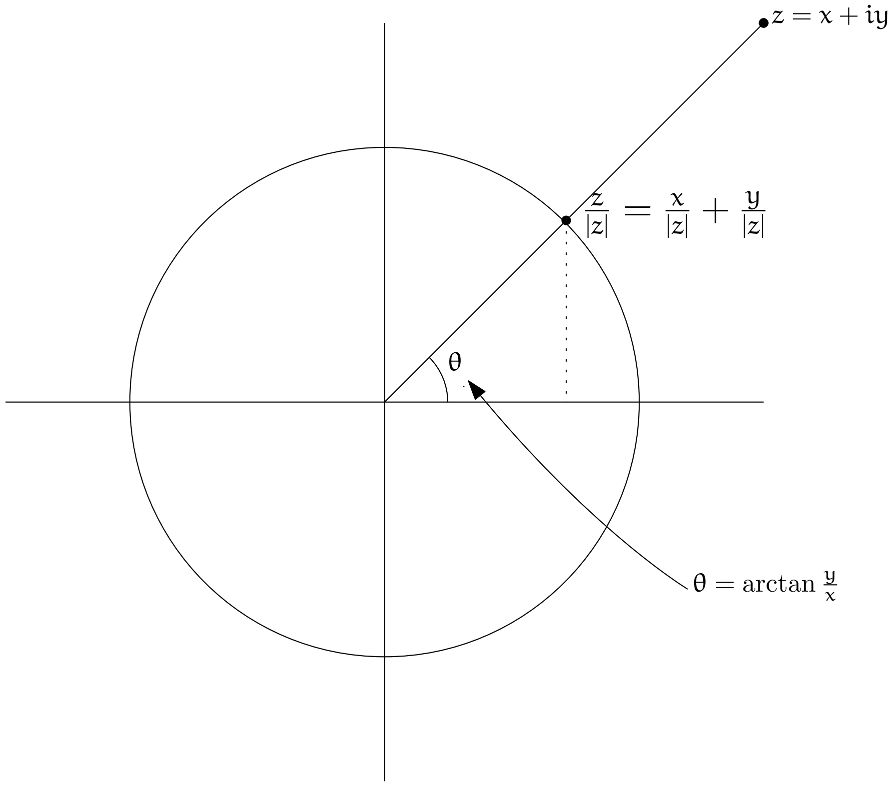

Figure 2: Determining Arg z

1.3 Powers and n-th roots

The historic lie perpetuated in textbooks is that complex numbers were introduced to solve quadratic equations. In fact, the fundamental theorem of algebra asserts that $\mathbb{C}$ is algebraically closed. In particular, given a complex number z_0, the polynomial equation $z^n - z_0 = 0$ must have solutions. We can explicitly obtain the n-roots using the polar representation of z_0. Write $z_0 = r1_\theta$, then $z^n = z_0$ means that

$$|z|^n = r, \quad n \arg z = \theta + 2\pi k, k \in \mathbb{Z}.$$

This means that $|z| = r^{1/n}$ and $\arg z = \frac{\theta}{n} + 2\pi\frac{k}{n}$. Note that two different values of k, say k_1, k_2 will give the same value for $\arg z$ iff $k_1 - k_2$ is an integer multiple of 2π. Thus there are n-distinct solutions, corresponding to $k = 0, 1, \ldots, n-1$.

The n-th roots of unity form a cyclic group of order n generated by the element $1_{2\pi/n}$ which is called the n-th primitive root. For $x > 0$, $\sqrt{x}$ will always denote the positive square root. If $0 \neq z \in \mathbb{C}$, then $\sqrt{z} := \sqrt{|z|}1_{\frac{\operatorname{Arg} z}{2}}$. This is called the *principal branch of the square root.* Similarly, $\sqrt[n]{z} := \sqrt[n]{|z|}1_{\frac{\operatorname{Arg} z}{n}}$.

The integer powers of z are defined in the standard manner:

$$z^n = \underbrace{z \cdots z}_{n\text{-times}}$$
$$z^{-n} = \underbrace{z^{-1} \cdots z^{-1}}_{n\text{-times}}$$

Taking rational powers is more complicated. If $q = \frac{n}{m}$, then $z^q := \left(z^{\frac{1}{m}}\right)^n$. Note that this a set and not a single number. The set does not depend on the particular representation of q. It is easy to see that as sets,

$$\left(z^{\frac{1}{m}}\right)^n = (z^n)^{\frac{1}{m}}.$$

Furthermore, if $q = \frac{n}{m}$ is in its reduced form and $z = |z| 1_\theta$ then in polar representation, z^q is the set of m distinct values

$$\sqrt[m]{|z|^n} 1_{\frac{n}{m}\theta + k\frac{2\pi n}{m}}, \quad k = 0, \ldots, m-1.$$

Note that the law of exponents is not true. For instance,

$$z \neq \left(z^{\frac{n}{m}}\right)^{\frac{m}{n}}$$

simply because the LHS is a number whereas the RHS is a set. Think on how to rephrase the law of exponents to make it true in our situations.

1.4 The field structure on $\mathbb{C}$

As remarked before, $\mathbb{C}$ is a commutative field. It is natural to ask if it is ordered.

Definition 3. Let S be a set. A *total ordering* on S is a relation $\leqslant$ that satisfies

1. **Reflexivity**: $a \leqslant a$ for all a in S.
2. **Antisymmetry**: $a \leqslant b$ and $b \leqslant a$ implies $a = b$.
3. **Transitivity**: $a \leqslant b$ and $b \leqslant c$ implies $a \leqslant c$.
4. **Comparability** (trichotomy law): For any a, b in S, either $a \leqslant b$ or $b \leqslant a$.

If $a \leqslant b$ and $a \neq b$, we often write $a < b$.

Definition 4 (Ordered Field)**.** A field $\mathbb{K}$ with a total ordering $\leqslant$ is said to be an *ordered field* if it satisfies

- if $a \leqslant b$ then $a + c \leqslant a + c$, $\forall c \in \mathbb{K}$,
- if $0 \leqslant a$ and $0 \leqslant b$ then $0 \leqslant a \cdot b$.

Theorem 5. *$\mathbb{C}$ cannot be given the structure of an ordered field.*

Proof. Assume to the contrary that $\leqslant$ makes $\mathbb{C}$ into an ordered field. Then either $i > 0$ or $i < 0$. Suppose $i > 0$. Then $i^2 = -1 > 0$. Adding 1 to both sides, we see that $0 > 1$. On the other hand $-1 > 0$ implies that $(-1)(-1) = 1 > 0$. This is a contradiction. An analogous argument works for $i < 0$. □

Another natural question is the following: view $\mathbb{C}$ as sitting in $\mathbb{R}^3$ as the first two coordinates; Can we give $\mathbb{R}^3$ the structure of a commutative field so that $\mathbb{C}$ is a subfield?

Theorem 6. *$\mathbb{R}^3$ cannot be given the structure of a commutative field such that $\mathbb{C}$ is a subfield.*

Proof. $\mathbb{R}^3$ is obviously a vector space. Assume also that we have defined a multiplication $\cdot$ on $\mathbb{R}^3$ that makes it a commutative field that extends $\mathbb{C}$. Denote the basis vectors $(1,0,0),(0,1,0)$ and $(0,0,1)$ by $1, i, j$, respectively. With this notation, $i^2 = -1$. Let us compute ij. Suppose $ij = a + bi + cj,\ a, b, c \in \mathbb{R}$. Observe first that

$$-j = (i^2)j = i(ij), \quad j^2 = -(ij)^2.$$

So,

$$-j = ai - b + c(ij) = ai - b + c(a + bi + cj).$$

Equating coefficients, we see that $c^2 = -1$ which is absurd. □

2 Geometry and Topology of $\mathbb{C}$

2.1 Conformal and anticonformal linear mappings

In this section, we consider a $\mathbb{R}$-linear mapping $T : \mathbb{C} \to \mathbb{C}$, i.e., $T(\lambda_1 z_1 + \lambda_2 z_2) = \lambda_1 T(z_1) + \lambda_2 T(z_2),\ \forall \lambda_1, \lambda_2 \in \mathbb{R}$. Now, $\mathbb{C}$ is both a vector space over $\mathbb{R}$ of dimension 2 and vector space over $\mathbb{C}$ of dimension 1. It is natural to ask for conditions on T that guarantee that T is $\mathbb{C}$-linear. Two obvious necessary conditions:

1. T is $\mathbb{R}$-linear,
2. T commutes with multiplication by i, i.e., $T(iz) = iT(z) \forall z \in \mathbb{C}$.

It turns out that these two conditions are sufficient to guarantee $\mathbb{C}$-linearity of T. Suppose T is $\mathbb{C}$-linear, then $T(z) = T(1)z$ and therefore T is just multiplication by a scalar. Writing $T(1) = (a, c)$ and $T(i) = (b, d)$, we see that the matrix of T under the standard basis is:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

This means that $T(x, y) = (ax + by, cx + dy)$, setting $x = \frac{z+\bar{z}}{2}$, $y = \frac{z-\bar{z}}{2i}$ and performing a straightforward computation yields that $T(z) = \alpha z + \beta \bar{z}$ where $\alpha = \frac{1}{2}(a + d - ib + ic)$ and $\beta = \frac{1}{2}(a - d + ic + ib)$. Note that given arbitrary $\alpha, \beta \in \mathbb{C}$, we can find a, b, c, d that satisfy the above relations and hence any map of the form $\alpha z + \beta \bar{z}$ is automatically $\mathbb{R}$-linear. Hence, T is $\mathbb{C}$-linear iff $\beta = 0$. This means that $a = d$ and $c = -b$. In matrix form, $\mathbb{C}$-linear maps have a matrix of type

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

Now, if T is $\mathbb{R}$-linear and commutes with i, then $T(iz) = iT(z) = \alpha iz + \beta i \bar{z} = \alpha iz + \beta \overline{iz}$ which means that $i\beta\bar{z} = -\beta i \bar{z}$ and so $\beta = 0$.

Another easy computation shows that

$$\det T = ad - bc = |\alpha|^2 - |\beta|^2,$$

which means that T is invertible iff $|\alpha| \neq |\beta|$. In this case, we can explicitly solve for the inverse T^{-1}:

$$T^{-1}(w) = \frac{\overline{\alpha}w - \beta\overline{w}}{|\alpha|^2 - |\beta|^2}$$

Definition 7. An $\mathbb{R}$-linear and invertible mapping $T : \mathbb{C} \to \mathbb{C}$ is said to be a *conformal linear mapping* if it preserves oriented angles, i.e, if $z, w \in \mathbb{C} \setminus \{0\}$ then

$$\arg Tz - \arg Tw = \arg z - \arg w.$$

Equivalently, $\arg Tz - \arg z = \arg Tw - \arg w$, or in other words, the function $\arg Tz - \arg z$ is constant modulo $2\pi\mathbb{Z}$. This means that $\arg \frac{Tz}{z}$ lies on a ray originating from the 0. However, if $Tz = \alpha z + \beta\overline{z}$, then

$$\frac{Tz}{z} = \alpha + \beta\frac{z}{\overline{z}}.$$

Note that the RHS above traces out a circle centred at α of radius $|\beta|$ and thus can lie on a ray iff $\beta = 0$. We have proved

Proposition 8. *Let $T : \mathbb{C} \to \mathbb{C}$ be an invertible $\mathbb{R}$-linear mapping. Then the following are equivalent:*

1. *T is conformal.*
2. *T is of the form $\alpha z, \alpha \in \mathbb{C}$.*
3. *T is $\mathbb{C}$-linear.*
4. *T is a composition of a rotation followed by a dilation.*

On the other hand, if $\alpha = 0$, then $T = \beta\overline{z}$ and in this case T preserves angles but reverse orientation. We say that T is *anitconformal* or $\mathbb{C}$-*anitlinear*. In this case the matrix of T is of the form

$$\begin{pmatrix} a & -c \\ c & a \end{pmatrix}$$

In addition, if T also preserves distances, then $|T(z)| = |z|$ in which event either $|\alpha|$ or $|\beta| = 1$. It is easy to see that in this case T must be an orthogonal matrix.

2.2 Analytic geometry

We want to express the usual notions from analytic geometry in complex notation. Changing to complex coordinates often greatly simplifies problems of analytic geometry. We equip $\mathbb{C}$ with the usual Hermitian inner-product $\langle z, w\rangle = z \cdot \overline{w}$. It follows that the usual inner-product on $\mathbb{R}^2$ is $\mathrm{Re}\langle z, w\rangle$.

The equation of line in analytic geometry is $Ax + By + C = 0$. Setting $x = \frac{z+\overline{z}}{2}$ and $y = \frac{z-\overline{z}}{2i}$ and computing shows that the general form a line L in complex notation is of the form $\alpha\overline{z} + \overline{\alpha}z + m = 0, \alpha \in \mathbb{C}, m \in \mathbb{R}$. In parametric form, the line joining $z_1, z_2 \in \mathbb{C}$ is of the form $z_1 + tz_2, t \in \mathbb{R}$.

The equation of a circle of radius r centred at $a \in \mathbb{C}$ is given by $|z - a| = r$. In parametric form, $z(t) = a + r1_t$.

2.3 Curves

A substantial portion of our course will involve path integrals. We recall here some basic notions about curves and paths to set the stage for later.

Definition 9. Let $U \subset \mathbb{C}$ be a domain. A *curve* in U is a continuous mapping $\gamma : [a, b] \to U$ where $[a, b]$ is an interval in $\mathbb{R}$. The curve γ is said to be *closed* if $\gamma(a) = \gamma(b)$ and we say it is *simple* if γ is injective.

We will often denote the image of the curve γ called the support of γ by γ^* but often we shall identify both γ and γ^*. We imagine γ^* to be a continuous thread in space but that picture is incorrect as there are curves $\gamma : [0, 1] \to [0, 1] \times [0, 1]$ such that $\gamma^* = [0, 1] \times [0, 1]$! In this context, it is appropriate to state the Hahn–Mazurkiewicz theorem which characterizes all possible continuous images of the unit interval.

Theorem 10 (Hahn–Mazurkiewicz)**.** *A non-empty Hausdorff topological space is a continuous image of the unit interval if and only if it is a compact, connected, locally connected second-countable space.*

We will often call a closed curve that is injective expect at the endpoints a *Jordan curve*. Such a curve has no self-intersections except at the endpoints. Note that that a Jordan curve is homemorphic to the unit circle S^1.

Our main focus in this course will be curves that are "sufficiently nice" to allow us to integrate continuous functions along the curve. To this end, we make the following

Definition 11. A curve $\gamma : [a, b] \to U$ is said to be *differentiable* if γ can be extended to some open set $G \supset [a, b]$ as a differentiable function, i.e., writing $\gamma(t) = x(t) + iy(t)$, both $x(t)$ and $y(t)$ are differentiable functions on (a, b) and the one-sided derivatives of $x(t)$ and $y(t)$ exist at a and b respectively. If moreover, $\gamma'(t) := x'(t) + iy'(t)$ are continuous on $[a, b]$, we say that γ is of class $\mathcal{C}^1$. Furthermore, if the derivative is non-zero at each point, then we say that γ is regular. The adjective *piecewise* is prefixed to indicate that the property is true for all but finitely many points. A *path* is a curve that is piecewise of class $\mathcal{C}^1$.

We will visualize paths as being trajectories of particles moving in the plane. In this visualization, the derivative $\gamma'(t)$ which is geometrically the tangent to γ at the point $\gamma(t)$ is nothing but the velocity of the moving particle at time t.

Example 12. We have already encountered one example of a Jordan curve. The curve $1_t|_{[-\pi,\pi]}$ is Jordan curve that is also regular. This curve is in fact of class $\mathcal{C}^\infty$. Such a curve is called a *smooth curve.* The derivative of this curve

$$1_t'(t) = -\sin t + i\cos t = i\gamma(t) \tag{2.1}$$

Notice that geometrically $\gamma'(t)$ is the tangent to γ at the point $\gamma(t)$. Thus (2.1) says that the tangent to 1_t at the point t is orthogonal to the vector $1_t(t)$. This is because multiplication by i is rotation by $\frac{\pi}{2}$.

It is time to improve the notation 1_t to one that is significantly better. We define

$$e^{it} := 1_t = \cos t + i \sin t. \tag{2.2}$$

This is called Euler's formula. In particular, we have

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

which is widely considered one of the most beautiful equations in all of mathematics.

Recall that e is the unique real number with the property that the function $e^x : \mathbb{R} \to \mathbb{R}$ has derivative itself. It might seem quite mysterious that Euler's constant appears out of thin air in (2.2). Let us motivate the grand entrance of e into complex analysis. The property $(e^x)' = e^x$ characterized the exponential uniquely. If we want to extend the definition to numbers of the form it, it is natural to expect that $(e^{it})' = ie^{it}$. Let us temporarily denote our candidate function for e^{it} by $Z(t)$. Then $Z'(t) = iZ(t)$ and also $Z(0) = 1$. This says that at time 0, the vector $Z(0) = 1$ and the velocity $Z'(0) = i$. A split second later the particle has moved upwards a bit in the direction i. Now, the velocity is again orthogonal to the position $Z(t)$ and thus the curve turns slightly. In this way, we see that the curve $Z(t)$ traces out the unit circle which is exactly what the definition (2.2) is saying.

For convenience, we summarize some facts about e^{it} that you have already seen when you studied the function 1_t:

(i) $e^{it} \cdot e^{is} = e^{i(t+s)}$.
(ii) Any $0 \neq z \in \mathbb{C}$ can be written as $z = |z| \cdot e^{i \arg z}$.
(iii) $z \cdot w = |z| \cdot |w| \cdot e^{i(\arg z + \arg w)}$.
(iv) The reciprocal of e^{it} is the function e^{-it}.
(v) $\cos t = \mathrm{Re}\ e^{it} = \frac{e^{it}+e^{-it}}{2}$.
(vi) $\sin t = \mathrm{Im}\ e^{it} = \frac{e^{it}-e^{-it}}{2i}$.

Often curves are described in polar form as follows. We specify two continuous functions $R : [a, b] \to \mathbb{R}^+$ and $\phi : [a, b] \to \mathbb{R}$ and set

$$\gamma(t) = R(t)e^{i\phi(t)}.$$

γ is clearly a continuous curve in $\mathbb{C} \setminus \{0\}$ and Γ is differentiable iff $\mathbb{R}$ and ϕ are. Furthermore,

$$\gamma'(t) = (R'(t) + iR(t)\phi(t))e^{i\phi(t)}.$$

Question: Can every continuous curve be written this way?

Example 13. A cardioid is a plane curve traced by a point on the perimeter of a circle that is rolling around a fixed circle of the same radius. To illustrate the power of complex notation, let us compute the equation of the Cardiod traced by the point 1 which lies on the circle of radius 1 centred at 2 that is rolling on the unit circle. After the circle has rolled a distance t, its centre is now at $2e^{it}$ and the point that is now touching the unit circle is the point $2e^{it} + e^{i(\pi+t)}$. This means that that the image of the point 1 is $2e^{it} + e^{i(\pi+2t)} = 2e^{it} - e^{i2t}$. This is the complex form of the equation and an easy computations with this expression can be used to obtain the polar form as well as the parametric equations of the cardioid. Please compare this derivations with the usual derivations using classical analytic geometry to understand the power of complex notation.

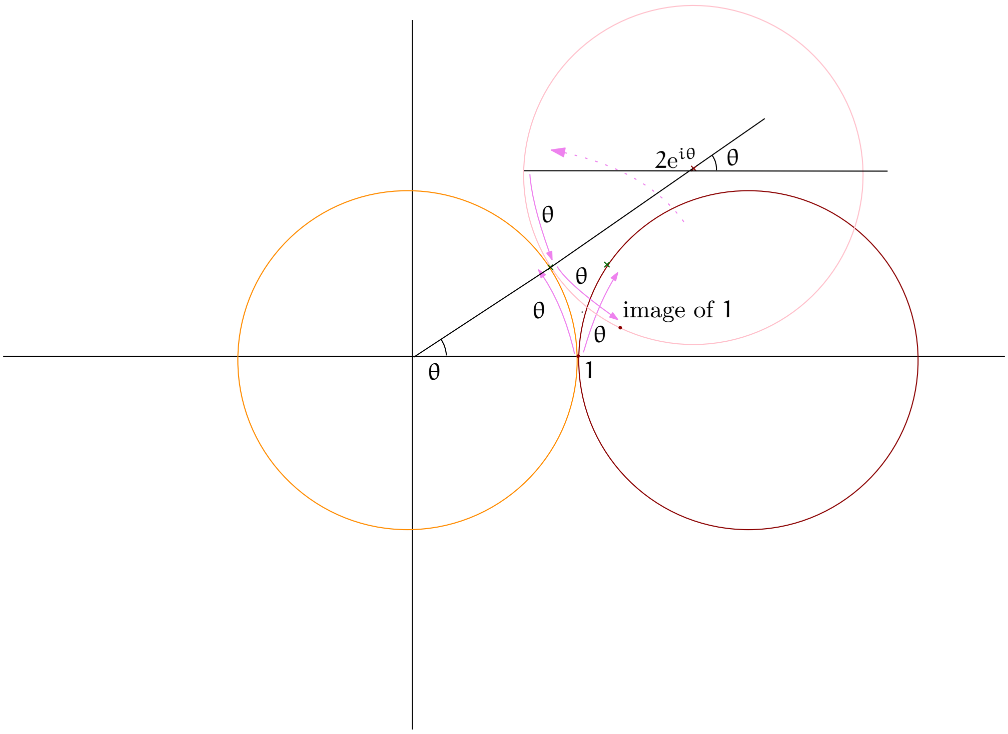

Figure 3: Equation of cardioid

Example 14. Another illustration off the power of complex notation is to find the equation of a spiral. Let $R(t)$ be an increasing continuous positivie function such that $\lim_{t\to-\infty} R(t) = 0$ and $\lim_{t\to\infty} R(t) = \infty$, then it is easy to see that

$$\gamma(t) = R(t)e^{it}$$

traces a spiral which expands as it rotates counter-clockwise.

We will now talk about reparametrization of curves. Given a curve γ, there could be many other curves whose image coincides with γ^*. For, instance the unit circle is the image of both the curves $\gamma_1(t) = e^{2\pi it}$ and $\gamma_2(t) = e^{-2\pi it^2}$. γ_1 traverses the circle counter-clockwise at unit speed whereas γ_2 clockwise with speed 2.

Definition 15. Let $\gamma : [a, b] \to U$ be a curve. We say that $\Gamma : [c, d] \to U$ is a *reparametrization* of γ if

1. $\Gamma^* = \gamma^*$;
2. There exists a continuous strictly monotone map $\Phi : [a, b] \to [c, d]$ such that $\Gamma \circ \Phi = \gamma$.

Remark 16. Such a map Φ is a homeomorphism of intervals. In fact, any homeomorphism of intervals is either strictly increasing or strictly decreasing. This defines an equivalence relation on the curves as follows: two curves γ_1 and γ_2 are equivalent if one is the reparametrization of the other by a strictly increasing homeomorphism.

Given a curve $\gamma : [a,b] \to U$, we define the map $\Phi : [a,b] \to [a,b]$ by $\Phi(t) = b + (a - t)$ and let $\overline{\gamma}$ be the curve $\gamma \circ \Phi$. Intuitively, $\overline{\gamma}$ is the same curve γ traversed in reverse. Any parametrization of γ^* is either equivalent to γ or $\overline{\gamma}$. Thus there are only two equivalenence classes of curves that parametrize γ^* and an *orientation* of γ^* is a choice of one of these classes.

We will now recall notions about the length of curve. We consider a partition $P, a = t_0 < t_1 < \cdots < t_n = b$ of $[a,b]$ and define

$$L(P) := \sum_{i=0}^{n-1} |\gamma(t_{i+1}) - \gamma(t_i)|.$$

P should be thought of as an inscribed polygon on γ with length $L(P)$. The length of the curve

$$L(\gamma) := \sup\{L(P) : P \text{ is a partition of [a,b]}\}.$$

If $L(\gamma) < \infty$ then γ is said to be rectifiable. Note that when if γ is piecewise $\mathcal{C}^1$ then it is rectifiable and

$$L(\gamma) = \int_a^b |\gamma'(t)| dt.$$

This integral is well-defined as $|\gamma'(t)|$ is continuous on $[a,b]$ except for finitely many points. Note also that the integral above is invariant under a change of parametrization.

If γ is regular then defining

$$s(t) := \int_a^t |\gamma'(x)| dx,$$

we see by the fundamental theorem of calculus that $s'(t)$ exists and is equal to $|\gamma'(x)|$. Furthermore as $|\gamma'|$ is a nowhere vanishing function, $s(t)$ is a strictly increasing function. This means that $s : [a,b] \to [0,L]$ is a strictly increasing $\mathcal{C}^1$ function and thus setting $\Psi = s^{-1}$, we see that $\Gamma = \gamma \circ \Psi^{-1}$ is a reparametrization with the property that the length $\Gamma|_{[0,t]} = t$. It is also easy to see that Γ is a unit-speed curve. This parametrization is called the *arc-length parametrization.*

Definition 17. Let $U \subset \mathbb{C}$ be open, $\gamma : [a,b] \to U$ be a path and $h : U \to \mathbb{C}$ be a continuous function. We define the path integral

$$\int_\gamma h := \int_a^b h(\gamma(t)) |\gamma'(t)| dt.$$

This is well-defined and independent of the parametrization of γ.

2.4 Branches of the argument function and the index of a closed curve

As we have mentioned before, the function $\text{Arg} : \mathbb{C} \setminus \{0\} \to (-\pi, \pi]$ is not continuous on any point of the negative real axis. However, it is continuous at any other point in $\mathbb{C} \setminus \mathbb{R}^-$.

Definition 18. Let $E \subset \mathbb{C}$ be a set that does not contain 0. Then any continuous mapping $g : E \to \mathbb{R}$ that satisfies $g(z) \in \arg z$ is called a *continuous branch of the argument in* E.

If $h, g : E \to \mathbb{R}$ are two branches of the argument, then $\frac{h-g}{2\pi}$ is a continuous integer valued function on E and is hence constant on each connected component of E. In particular, if E were connected then any two branches of the argument differ by a constant integer multiple of 2π. On $\mathbb{C} \setminus \mathbb{R}^-$, any branch of the argument is of the form $\operatorname{Arg} z + 2\pi\mathbb{Z}$.

Example 19. On any circle C centred at the origin, there are no continuous branches of the argument. If not, then let $g : C \to \mathbb{R}$ be a branch of the argument and let $a \in \mathbb{R}$ be the point on the negative real axis that intersects C. Then on the connected set $C \setminus \{a\}$, we can find a $k \in \mathbb{Z}$ such that $\operatorname{Arg} z + 2\pi k = g(z)$. As the value of Arg jumps at the point a, we see that g is discontinuous at a, a contradiction.

Example 20. If L is any ray starting from the origin, say $L := \operatorname{Arg} z = \alpha$, all branches of $\arg z$ are given by

$$\arg z = \operatorname{Arg}(z \cdot e^{i}(\pi - \alpha) - \pi + \alpha + 2\pi\mathbb{Z}.$$

This is derived as follows. We first rotate by $\pi - \alpha$ so that the ray L coincides with the negative real axis. Now we find the value of Arg and subtract $\pi - \alpha$ to compensate for our initial rotation.

Example 21. Let $D := D(a, r)$ be a disk that misses the origin, then we can find a continuous branch of arg on D. To see this, consider a (possibly) larger disk $D(a, r')$ such that $0 \in \partial D(a, r')$. Then let L be ray starting at the origin tangent to $\partial D(a, r')$ at 0. Then any branch of arg on $\mathbb{C} \setminus L$ works.

Definition 22. Let X be a connected topological space and let $f : X \to \mathbb{C} \setminus \{0\}$ be a continuous function. A *continuous branch* of the argument of f is any continuous function $h : X \to \mathbb{R}$ such that $h(x) \in \arg f(x)$ $\forall x \in X$.

One important case is when $X = E \subset \mathbb{C}$ and f is the identity function in which case we recover the definition of a continuous branch of arg. We will also consider branches of the n-th root of $f : E \to \mathbb{C} \setminus 0$: any continuous function $h : E \to \mathbb{C}$ such that $h^n = f$. If $g : E \to \mathbb{R}$ is branch of the argument, then $h = |f|^{1/n} e^{ig/n}$ is branch of the n-th root of f. The following example illustrates that the converse is not, in general, true.

Example 23. Let $E = \mathbb{C} \setminus [-1, 1]$ and let $f(z) = z^2 - 1$. Note that the image of $z^2 - 1$ is contained in $(-\infty, 0]$ precisely when $z \in [-1, 1]$ or the imaginary axis. We take $\sqrt{z^2 - 1}$ using the principal square root on the right half-plane and the negative of the principal square root on the left half-plane.

We come back to the question about polar representation of curves. $\gamma : [a, b] \to \mathbb{C}$ is a continuous curve that does not contain the origin. We want to write

$$\gamma(t) = R(t) e^{i\phi(t)}.$$

Of course, $\mathbb{R}(t) = |\gamma(t)|$ which is a continuous function as $\gamma(t) \neq 0$. For each $t \in [a, b]$, we must select $\phi(t) \in \arg \gamma(t)$ in a continuous way. This is of course possible, if we can find a branch of arg on γ^*. But it actually suffices to find a branch of the argument of γ.

Theorem 24. *If $\gamma : [a, b] \to \mathbb{C}$ is a curve that misses the origin, then we can find a continuous branch of the argument of γ, say $\phi(t)$. Any other branch of the argument of γ is of the form $\phi + 2\pi\mathbb{Z}$. Moreover, if γ is differentiable at the point $t \in [a, b]$, then so is ϕ and $\phi'(t) = \mathsf{Im}\frac{\gamma'(t)}{\gamma(t)}$.*

Proof. Observe that on any disk in $\mathbb{C}$ that does not contain the origin, we can define a continuous branch of the argument. Let $r := \text{dist}(\gamma^*, 0)$. As γ^* is compact, $r > 0$. Finitely many sets of the form $\gamma^{-1}(D(t, r)$ cover $[a, b]$ by the compactness of $[a, b]$. Hence, we can find a partition $a = t_0 < t_1 < \cdots < t_n = b$ such that each $\gamma([t_{i-1}, t_i)$ is contained in a disk of the form $D_i := D(t, r)$. Let g_i be a branch of arg on D_i. Then $\phi_i = g_i \circ \gamma$ is a branch of $\arg(\gamma(t),\ t \in [t_{i-1}, t_i]$. At the point t_i, $\phi_i(t_i)$ might not be equal to $\phi_{i+1}(t_i)$. Nevertheless, their difference is an integer multiple of 2π which we add to ϕ_{i+1}. By gluing lemma for continuous functions, this allows to construct a global branch of the argument of γ.

Now, suppose γ is differentiable at the point t. We may, without loss of generality, assume that $\gamma(t)$ lies on the upper half-plane. This in a neighbourhood of t, $\phi(t) = \text{Arg}\,\gamma(t) + 2\pi k$ for some $k \in \mathbb{Z}$. Thus, $\phi'(t)$ exists and is equal to

$$\frac{d}{dt}(\text{Arg}\,\gamma(t)) = \frac{d}{dt}\left(\arctan\frac{y(t)}{x(t)}\right) = \frac{x^2(t)}{x^2(t) + y^2(t)} \cdot \frac{x(t)y'(t) - y(t)x'(t)}{x^2(t)} = \mathsf{Im}\frac{\gamma'(t)}{\gamma(t)},$$

where $\gamma(t) = (x(t), y(t))$.

□

Example 25. The curve $\gamma(t) = e^{it}, t \in [0, 2\pi]$ has a continuous branch of the argument given by $h(t) = t$, but clearly no branch of arg can exist on the circle $\mathbb{T}$.

Now, let $\gamma : [a, b] \to \mathbb{C} \setminus \{0\}$ be a curve that misses 0 and let ϕ be a branch of the argument of γ. Then $\phi(b) - \phi(a)$ does not depend on the choice of ϕ as any two choices differ by an element of $2\pi\mathbb{Z}$. This is called the *variation of the argument* and denoted $\Delta_\gamma \arg$. If, in addition, γ is a path, then

$$\Delta_\gamma \arg = \mathsf{Im} \int_a^b \frac{\gamma'(t)}{\gamma(t)} dt,$$

by the fundamental theorem of calculus and the previous theorem.

2.5 The index or winding number of a closed curve

If $\gamma : [a, b] \to \mathbb{C} \setminus \{0\}$ is a closed curve, then for any branch ϕ of the argument of γ, $\frac{\phi(b) - \phi(b)}{2\pi}$ is an integer as $\gamma(a) = \gamma(b)$.

Definition 26. For a closed curve $\gamma : [a, b] \to \mathbb{C} \setminus \{0\}$, we define the *index or the winding number of γ with respect to* 0, $\text{Ind}(\gamma, 0)$ by the number $\frac{\Delta_\gamma \arg}{2\pi}$.

Intuitively, the winding number measures the number of rotations around 0 a particle moving along γ makes. One can easily define this notion for more general cases.

Definition 27. Let $\gamma : [a, b] \to \mathbb{C} \setminus \{z\}$ be a closed curve. Then, we define the index of γ with respect to z, $\text{Ind}(\gamma, z)$ by

$$\text{Ind}(\gamma, z) = \text{Ind}(\gamma - z, 0).$$

Now, suppose $\gamma = R(t)e^{i\phi(t)}$ is a path that misses the origin. Then $\gamma'(t) = (R'(t) + iR(t)\phi(t))e^{i\phi(t)}$. Thus

$$\frac{\gamma'(t)}{\gamma(t)} = \left(\frac{R'(t)}{R(t)} + i\phi'(t)\right).$$

Evaluating

$$\int_a^b \frac{\gamma'(t)}{\gamma(t)} dt = \int_a^b \mathrm{Re}\frac{\gamma'(t)}{\gamma(t)} dt + i\int_a^b \frac{\gamma'(t)}{\gamma(t)} dt = \log\frac{|\gamma(b)|}{\gamma(a)} + i\Delta_\gamma \arg.$$

This shows that for a path $\gamma : [a, b] \to \mathbb{C} \setminus \{z\}$

$$\operatorname{Ind}(\gamma, z) = \frac{1}{2\pi i}\int_a^b \frac{\gamma'(t)}{\gamma(t) - z} dt.$$

One can rewrite the integral above in the following form

$$\operatorname{Ind}(\gamma, z) = \frac{1}{2\pi i}\int_\gamma \frac{dz}{1 - z}.$$

The above integral expressions clearly show that whenever γ is a path, the function $\operatorname{Ind}(z, \gamma)$ as a function of z on $\mathbb{C} \setminus \gamma^*$ is a continuous function.

Now, let $\gamma : [a, b] \to \mathbb{C}$ be a closed curve. Then $\mathbb{C} \setminus \gamma^*$ is an open set and hence has only countably many connected components. On each of these connected components, the function $\operatorname{Ind}(\cdot, \gamma)$ is a constant function. Moreover, $C \setminus \gamma^*$ has only one connected component that is unbounded. This is because γ^* is compact and is hence contained in some large disk D. Thus the unique unbounded component is the one that contains $\mathbb{C} \setminus D$. We claim that on the unbounded component, $\operatorname{Ind}(\cdot, \gamma)$ is zero. We know that it is constant, so it suffices to pick $z \in \mathbb{C} \setminus D$ and show that $\operatorname{Ind}(z, \gamma) = \operatorname{Ind}(\gamma - z, 0) = 0$. But $\gamma - z \subset D - z$ and $D - z$ is a disk centred at z that misses the origin. This means that we can find a branch of arg on $D - z$ and hence $\Delta_{\gamma - z} \arg = 0$ proving our claim.

2.6 Applications

Our first application will be to find an expression for $\cos^4 t$. The straightforward method is to just compute:

$$\begin{aligned}
2\cos t &= (e^{it} + e^{-it}), \\
16\cos^4 t &= (e^{it} + e^{-it})^4 \\
&= (e^{i4t} + e^{-i4t}) + 4(e^{i2t} + e^{-i2t}) + 6,
\end{aligned}$$

which gives

$$\cos^4 t = \frac{1}{8}(\cos 4t + 4\cos 2t + 3).$$