Lecture 23: Complex numbers

Today, we're going to introduce the system of complex numbers. The main motivation for doing this is to establish a somewhat more invariant notion of angle than we have already. Let's recall a little about how angles work in the Cartesian plane.

A brief review of two dimensional analytic geometry

Points in the Cartesian plane are given by pairs of numbers (x, y). Usually when we think of points, we think of them as fixed positions. (Points aren't something you add and the choice of origin is arbitrary.) The set of these points is sometimes referred to as the affine plane. Within this plane, we also have the concept of vector. A vector is often drawn as a line segment with an arrow at the end. It is easy to confuse points and vectors since vectors are also given by ordered pairs, but in fact a vector is the difference of two points in the affine plane. (A change of coordinates could change the origin to some other point, but it couldn't change the zero vector to a vector with magnitude.) It is between vectors that we measure angles.

If $\vec{a} = (a_1, a_2)$, we define the magnitude of $\vec{a}$ written $|\vec{a}|$ by $\sqrt{a_1^2 + a_2^2}$, as suggested by the Pythagorean theorem. Given another vector $\vec{b} = (b_1, b_2)$, we would like to define the angle between $\vec{a}$ and $\vec{b}$. We define the dot product

$$\vec{a} \cdot \vec{b} = a_1 b_1 + a_2 b_2.$$

A quick calculation shows that

$$|\vec{a} - \vec{b}|^2 = |\vec{a}|^2 + |\vec{b}|^2 - 2|\vec{a} \cdot \vec{b}|.$$

Therefore, we can start to define the angle θ between $\vec{a}$ and $\vec{b}$ by

$$\vec{a} \cdot \vec{b} = |\vec{a}||\vec{b}| \cos \theta,$$

inspired by the law of cosines. Note that this only defines the angle θ up to its sign. The angle between $\vec{a}$ and $\vec{b}$ is indistinguishable from the angle between $\vec{b}$ and $\vec{a}$.

There is another product we can define between two dimensional vectors which is the cross product:

$$\vec{a} \times \vec{b} = a_1 b_2 - b_1 a_2.$$

We readily observe that

$$|\vec{a} \times \vec{b}|^2 + |\vec{a} \cdot \vec{b}|^2 = |\vec{a}|^2 |\vec{b}|^2.$$

This leads us to

$$\vec{a} \times \vec{b} = |\vec{a}||\vec{b}| \sin \theta,$$

which gives a choice of sign for the angle θ.

Complex numbers

We now introduce the complex numbers which give us a way of formalizing a two-dimensional vector as a single number, and defining the multiplication of these numbers in a way that involves both of the forms of multiplication that we say before.

We introduce i to be a formal square root of -1. Of course, the number -1 has no square root which is a real number. i is just a symbol, but we will define multiplication using $i^2 = -1$. A complex number is a number of the form

$$a = a_1 + ia_2,$$

where a_1 and a_2 are real numbers. We write

$$Re(a) = a_1$$

and

$$Im(a) = a_2.$$

We can define addition and subtraction of complex numbers. If

$$b = b_1 + ib_2,$$

then we define

$$a + b = (a_1 + b_1) + i(a_2 + b_2),$$

and

$$a - b = (a_1 - b_1) + i(a_2 - b_2).$$

These, of course, exactly agree with addition and subtraction of vectors. The fun begins when we define multiplication. We just define it so that the distributive law holds.

$$ab = a_1b_1 - a_2b_2 + i(a_1b_2 + a_2b_1).$$

We pause for a quick remark. There is something arbitrary about the choice of i. Certainly i is a square root of -1. But so is $-i$. Replacing i by $-i$ changes nothing about our number system. We give this operation a name, complex conjugation. Namely if

$$a = a_1 + ia_2,$$

then the complex conjugate of a is

$$\bar{a} = a_1 - ia_2.$$

Once we have the operation of complex conjugation, we can begin to understand the meaning of complex multiplication. Namely to the complex number a is associated the vector

$$\vec{a} = (a_1, a_2).$$

Similarly to the complex conjugate of b is associated the vector

$$\overrightarrow{\bar{b}} = (b_1, -b_2).$$

Then

$$ab = \vec{a} \cdot \vec{\bar{b}} + i\vec{a} \times \vec{\bar{b}}.$$

To every complex number is associated a magnitude

$$|a| = \sqrt{a_1^2 + a_2^2}.$$

Notice complex conjugation doesn't change this:

$$|a| = |\bar{a}|.$$

To each complex number a is also associated its direction which we temporarily denote as $\theta(a)$, the angle θ that a makes with the x axis. Complex conjugation reflects complex numbers across the x-axis so

$$\theta(\bar{a}) = -\theta(a).$$

Now from our description of multiplication of complex numbers in terms of vectors, we see that

$$ab = |a||b|\cos(\theta(a) + \theta(b)) + i|a||b|\sin(\theta(a) + \theta(b)).$$

Thus

$$|ab| = |a||b|,$$

and

$$\theta(ab) = \theta(a) + \theta(b).$$

This gives a geometrical interpretation to multiplication by a complex number a. It stretches the plane by the magnitude of a and rotates the plane by the angle $\theta(a)$. Note that this always gives us that

$$a\bar{a} = |a|^2.$$

This gives us a way to divide complex numbers:

$$\frac{1}{b} = \frac{\bar{b}}{|b|^2},$$

so that

$$\frac{a}{b} = \frac{a\bar{b}}{|b|^2}.$$

There is no notion of one complex number being bigger than another, so we don't have least upper bounds of sets of complex numbers. But it is easy enough to define limits. If $\{a_n\}$ is a sequence of complex numbers, we say that

$$\lim_{n \longrightarrow \infty} a_n = a,$$

if for every real $\epsilon > 0$, there exists $N > 0$ so that if $n > N$, we have

$$|a - a_n| < \epsilon.$$

You will prove for homework that magnitude of complex numbers satisfies the triangle inequality.

In the same way, we can define limits for complex valued functions. Given a power series

$$\sum_n a_n z^n,$$

it has the same radius of convergence R as the real power series

$$\sum_n |a_n| x^n,$$

and converges absolutely for every z with $|z| < R$.

We can complete our picture of the geometry of complex multiplication by considering

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

This power series converges for all complex z since its radius of convergence is infinite. We restrict our attention to the function

$$f(\theta) = e^{i\theta},$$

with θ real. We may ask what is $|e^{i\theta}|$? We calculate

$$|e^{i\theta}|^2 = e^{i\theta}\bar{e^{i\theta}} = e^{i\theta}e^{-i\theta} = 1.$$

(You will verify the identity $e^{z+w} = e^z e^w$ in your homework.) Thus as θ varies along the real line, we see that $e^{i\theta}$ traces out the unit circle. How fast (and in which direction) does it trace it? We get this by differentiating $f(\theta)$ as a function of θ. We calculate

$$\frac{d}{d\theta} f(\theta) = ie^{i\theta}.$$

In particular, the rate of change of $f(\theta)$ has magnitude 1 and is perpindicular to the position of $f(\theta)$. We see then that f traces the circle by arclength. (That is, θ represents arclength travelled on the circle and from this, we obtain Euler's famous formula

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

By plugging into the definition of e^z and extracting real and imaginary parts, we obtain Taylor series for sin and cos by

$$\sin\theta = \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} + \dots,$$

and

$$\cos\theta = 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} + \dots$$