An Algebraic Derivation of De Moivre's Theorem

Chris Clark March 3, 2008

1 Introduction

When I first saw how easy it was to derive trigonometric identities from Euler's theorem, I was inspired to derive the identities starting from axioms, without using geometry. I believe using geometry often hides the fundamental concepts that algebra makes explicit. For example, if you are dealing with the concept of rotation in geometry, you might easily forget about the fact that rotations preserve orientation (orientation is the mathematical term for handedness). When using algebra, this assumption must be explicit or the proof will not work in some cases. Euler's theorem is more powerful than necessary, however, and De Moivre's theorem suffices. In this paper I will derive De Moivre's theorem

$$(\cos(\theta) + i\sin(\theta))^n = \cos(n\theta) + i\sin(n\theta)$$

for the case of rational n, assuming only linear algebra. The reason that we are not just using induction and trigonometric identities to derive the theorem is that then we would not be able to use it to derive those identities without being circular.

2 The Rotation Matrix

In order to prove De Moivre's theorem, we will establish an isomorphism between complex numbers and a class of matrices by

$$z = x + iy \in \mathbb{C} \longleftrightarrow \left(\begin{smallmatrix} x & -y \\ y & x \end{smallmatrix}\right) \in GL(2).$$

So instead of working with the multiplication of complex numbers we will use rotations of vectors in the plane. At the end we can use the isomorphism to convert back to complex numbers.
A rotation $\mathbf{R}$ is a linear transformation on an inner product space that preserves the inner product and orientation (i.e. If u and v are vectors in the inner product space, then $(\mathbf{R}u)\cdot(\mathbf{R}v) = u \cdot v$ and $\det(\mathbf{R}) > 0$). In particular, since rotations preserve all inner product, they also preserve all lengths since the length of a vector since $|v| = \sqrt{v \cdot v}$. Now we specialize to the case of rotations in $\mathbb{R}^2$. Using the matrix form of the transformation in $\mathbb{R}^2$, this constraint gives three equations in the four unknown matrix elements, which leaves us with one free parameter. We call this parameter the angle of rotation and write it as θ.[1] We also write a rotation by an angle θ as R_θ. We define the symbols $\cos(\theta)$ and $\sin(\theta)$ to be the components of the vector $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ after rotation by an angle θ, that is

$$R_\theta \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}.$$

Next we want to determine the rotated form of the other basis vector, so we write

$$R_\theta \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) = \begin{pmatrix} x(\theta) \\ y(\theta) \end{pmatrix},$$

for some unknown functions $x(\theta)$ and $y(\theta)$. Since rotations preserve inner products, we have

$$0 = \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) \cdot \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) = R_\theta \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) \cdot R_\theta \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} \cdot \begin{pmatrix} x(\theta) \\ y(\theta) \end{pmatrix} = x(\theta)\cos(\theta) + y(\theta)\sin(\theta).$$

and

$$1 = \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) \cdot \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) = R_\theta \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) \cdot R_\theta \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) = \begin{pmatrix} x(\theta) \\ y(\theta) \end{pmatrix} \cdot \begin{pmatrix} x(\theta) \\ y(\theta) \end{pmatrix} = x^2(\theta) + y^2(\theta).$$

[1] There is a freedom in the parameterization of a linear transformation i.e. we could transform θ to be $\theta' = a\theta + b$. In order for θ to be the usual angle parameter, we would have to specify the constants so that $\theta = 0$ for no rotation and $\theta = 2\pi$ for a full rotation back to the starting point.

Multiplying this by $\cos^2(\theta)$ gives

$$\cos^2(\theta) = x^2(\theta)\cos^2(\theta) + y^2(\theta)\cos^2(\theta)$$

Using the first relation on the first term on the right hand side, we have

$$\cos^2(\theta) = y^2(\theta)\sin^2(\theta) + y^2(\theta)\cos^2(\theta)$$

$$\cos^2(\theta) = y^2(\theta)$$

$$y(\theta) = \pm\cos(\theta) \qquad \text{and} \qquad x(\theta) = \mp\sin(\theta)$$

Combining the rules for transforming the basis vectors, we can express R_θ as a matrix. Checking the determinate, we find that we must choose the top signs to preserve the orientation according to the definition of rotations. Therefore

$$R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

3 Composition of Rotations

We need to show that

$$R_{\theta_1}R_{\theta_2} = R_{\theta_1+\theta_2}.$$

We will use rotations by an infinitesimal angle to derive this.

$$R_{\delta\theta} = \begin{pmatrix} \cos(\delta\theta) & -\sin(\delta\theta) \\ \sin(\delta\theta) & \cos(\delta\theta) \end{pmatrix}.$$

We know that the identity transformation satisfies the definition of rotation, so $I = R_\theta$ for some value θ. We have not yet defined the zero of the angular parameter, and we will now choose the rotation by an angle of 0 to be the identity. Therefore, since the angle is a continuous parameter, [2]

$$R_{\delta\theta} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \delta\theta \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

for some unspecified constants a and b. Regardless of the value of a, $a\delta\theta$ is negligible compared with 1, so we may drop it to obtain

$$R_{\delta\theta} = \begin{pmatrix} 1 & -b\delta\theta \\ b\delta\theta & 1 \end{pmatrix}.$$

So we can calculate

$$R_{\delta\theta_1}R_{\delta\theta_2} = \begin{pmatrix} 1 & -b\delta\theta_1 \\ b\delta\theta_1 & 1 \end{pmatrix}\begin{pmatrix} 1 & -b\delta\theta_2 \\ b\delta\theta_2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -b(\delta\theta_1+\delta\theta_2) \\ b(\delta\theta_1+\delta\theta_2) & 1 \end{pmatrix} = R_{\delta\theta_1+\delta\theta_2}$$

It remains to show that this holds for finite rotations. This follows from the fact that any finite rotation can be constructed by an infinite composition of infinitesimal rotations. From the last result we have the corollary $(R_{\delta\theta})^n = R_{n\delta\theta}$ for all $n \in \mathbb{Z}$ when $\delta\theta$ is infinitesimal. So

$$\lim_{\delta\theta\to 0}(R_{\delta\theta})^{\theta/\delta\theta} = R_\theta$$

since $\theta/\delta\theta$ becomes infinitely large and can be assumed to be integer without loss of generality. Therefore

$$R_{\theta_1}R_{\theta_2} = \lim_{\delta\theta_1\to 0}\left[(R_{\delta\theta_1})^{\theta_1/\delta\theta_1}\right]\lim_{\delta\theta_2\to 0}\left[(R_{\delta\theta_2})^{\theta_2/\delta\theta_2}\right]$$

Now let $\delta\theta_1 \to \theta_1/M$ and $\delta\theta_2 \to \theta_2/M$.

$$= \lim_{M\to\infty}(R_{\theta_1/M})^M(R_{\theta_2/M})^M$$

[2] We need to use a little bit of analysis here, so maybe this is not a completely algebraic derivation.

Now $R_{\delta\theta_1} R_{\delta\theta_2} = R_{\delta\theta_1+\delta\theta_2}$ implies that infinitesimal rotations commute, so

$$= \lim_{M\to\infty} (R_{\theta_1/M} R_{\theta_2/M})^M$$

$$= \lim_{M\to\infty} (R_{\theta_1/M+\theta_2/M})^M$$

$$= \lim_{M\to\infty} (R_{\theta_1+\theta_2})$$

$$= R_{\theta_1+\theta_2}$$

4 De Moivre's Theorem

In the last two parts we showed that

$$R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

and

$$R_{\theta_1} R_{\theta_2} = R_{\theta_1+\theta_2}$$

The second equation implies that

$$(R_\theta)^n = R_{n\theta} \quad \forall n \in \mathbb{Z}$$

which looks a bit like De Moivre's theorem. In fact, all that remains is to use the first equation to establish the isomorphism with the complex numbers. We choose the isomorphism mentioned earlier

$$z = x + iy \in \mathbb{C} \longleftrightarrow \begin{pmatrix} x & -y \\ y & x \end{pmatrix} \in GL(2).$$

We notice that the matrices on the right hand side are exactly the rotation matrices if we set $x = \cos(\theta)$ and $y = \sin(\theta)$. We must prove that the multiplication rule is preserved through the isomorphism. Let ϕ denote the isomorphism from $\mathbb{C}$ to $GL(2)$.

$$\phi(a+ib)\phi(c+id) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} c & -d \\ d & c \end{pmatrix} = \begin{pmatrix} (ac-bd) & -(ad+bc) \\ (ad+bc) & (ac-bd) \end{pmatrix} = \phi((ac-bd)+i(ad+bc)) = \phi((a+ib)(c+id))$$

Similarly, the map in the other direction also preserves multiplication.
Therefore

$$(R_\theta)^n = R_{n\theta} \Rightarrow (\cos(\theta) + i\sin(\theta))^n = \cos(n\theta) + i\sin(n\theta)$$

by the isomorphism.
Now we need to extend to the case of rational n. First let k be a non-zero integer and let $\theta = k\phi$.

$$(\cos(k\phi) + i\sin(k\phi))^n = \cos(nk\phi) + i\sin(nk\phi) = (\cos(n\phi) + i\sin(n\phi))^k$$

Now exponentiate the leftmost and rightmost sides to the power of $1/k$

$$(\cos(k\phi) + i\sin(k\phi))^{n/k} = \cos(n\phi) + i\sin(n\phi)$$

Finally, replace θ in the equation

$$(\cos(\theta) + i\sin(\theta))^{n/k} = \cos\left(\frac{n}{k}\theta\right) + i\sin\left(\frac{n}{k}\theta\right)$$

This proves De Moivre's theorem for the case of rational exponents.