The Rigorous Rotational Approach

Even without calculus, treating the cosine and sine simultaneously has great advantages in developing plane analytic geometry and the complex number plane, emphasizing rotation, scaling, and reflection symmetries as a complement to horizontal and vertical shifts. Here, we show how the framework of calculus makes it straightforward to make that "precalculus" approach to the rotation formula, trigonometric identities, Pythagorean relation, and geometric interpretation of complex multiplication and the dot product rigorous.

We wish to understand the properties of the circular functions from the starting point that $\cos s$ and $\sin s$ are defined to be the horizontal and vertical components of a point to which the point $(1,0)$ has been rotated by a counterclockwise rigid rotation about the origin, so that the length of the circular arc from $(1,0)$ to $(\cos s, \sin s)$ is s. We do this in turn by defining a rigid rotation of the vector $\mathbf{v}$ by the angle s radians in terms of proceeding for s units of time along a circular path $\mathbf{v}(s)$ whose velocity vector is given by the position vector rotated a quarter circle counterclockwise. By a congruent triangle argument, related to the "negative reciprocal rule" from elementary geometry (which also motivates the "dot product" condition for orthogonality) this velocity vector at $\begin{pmatrix} x \\ y \end{pmatrix}$ is given by $\begin{pmatrix} -y \\ x \end{pmatrix}$.

This seems quite natural, and leads to fairly straightforward connections with other properties of the circular functions (and other transcendental functions, such as $\cosh s, \sinh s, \exp s$, and $\log s$.) The use of derivatives and differential equations (two coupled!) may be seen as a drawback, but we prefer to view this approach as a demonstration of the power of this framework to simplify and enlighten. One need only consider defining the cosine and sine rigorously and usefully as real-valued functions "without calculus", in terms of the adjacent and opposite sides and hypotenii of right triangles, and you will probably agree.

Define the rotation of a vector $\mathbf{v}$ by an angle of s radians, $R(s)\mathbf{v}$, as the solution at time s of the rotational differential equation

$$\frac{d}{ds}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -y \\ x \end{pmatrix} \qquad \begin{pmatrix} x(0) \\ y(0) \end{pmatrix} = \mathbf{v}$$

and define $\cos s$ and $\sin s$ by $\begin{pmatrix} \cos s \\ \sin s \end{pmatrix} = R(s)\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ by s radians. This should make intuitive sense. We do not need to prove that the derivative of $\cos s$ is $-\sin s$ and the derivative of $\sin s$ is $\cos s$ using this approach, they, along with the values at $s = 0$ are the definitions.

The Rotation and Addition Formulas

In this approach, the issue is to show that they correspond with other concepts of cos and sin. We will justify the necessary technicalities, such as existence, uniqueness, and linearity of solutions of these particular linear, constant coefficient differential equations with respect to initial conditions in an appendix. These demonstrations give concrete examples of and motivation for the more general theory.

First we show if $R(s)\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} A \\ B \end{pmatrix}$ then $R(s)\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -B \\ A \end{pmatrix}$. (Capital letters are used to identify the components of unit vectors.) To do this, we only need to turn our head a quarter circle clockwise, so the rotation of $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ into $\begin{pmatrix} A \\ B \end{pmatrix}$ "looks like" the rotation of $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ into $\begin{pmatrix} -B \\ A \end{pmatrix}$. Mathematically then, we define $u(t) = -y(t)$ and $v(t) = x(t)$, so the $u - v$ axes are a quarter circle clockwise from the $x - y$ axes. Then $u' = -v$ $v' = u$ $u(0) = 0$ $v(0) = 1$ leads to Then $x' = v' = u = -y$ $y' = -u' = v = x$ $x(0) = 1$ $y(0) = 0$, as expected. Since we know $x(s) = A$ $y(s) = B$, we also know that $u(s) = -B$ $v(s) = A$ which we were trying to prove.

Then, to see that the rotation which takes $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ to $\begin{pmatrix} A \\ B \end{pmatrix}$ takes $\begin{pmatrix} x \\ y \end{pmatrix} = x\begin{pmatrix} 1 \\ 0 \end{pmatrix} + y\begin{pmatrix} 0 \\ 1 \end{pmatrix} = x\begin{pmatrix} A \\ B \end{pmatrix} + y\begin{pmatrix} -B \\ A \end{pmatrix} = \begin{pmatrix} Ax - By \\ Bx + Ay \end{pmatrix}$ we check that $\begin{pmatrix} u \\ v \end{pmatrix} = x\begin{pmatrix} \cos s \\ \sin s \end{pmatrix} + y\begin{pmatrix} -\sin s \\ \cos s \end{pmatrix}$ is a solution of $\frac{d}{ds}\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -v \\ u \end{pmatrix}$, since $\frac{d}{ds}u = x(-\sin s) + y(-\cos s) = -v$ and $\frac{d}{ds}v = x(\cos s) + y(-\sin s) = u$. Also, $\begin{pmatrix} u(0) \\ v(0) \end{pmatrix} = x\begin{pmatrix} 1 \\ 0 \end{pmatrix} + y\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$ This proves the rotation formula. (This is a concrete demonstration of the linearity of solutions of linear differential equations with respect to their initial conditions.)

We set $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} = R(t)\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and see that $R(s)R(t)\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos s\cos t - \sin s\sin t \\ \cos s\sin t + \sin s\cos t \end{pmatrix}$.

Finally, we check that that $R(t)R(s)\mathbf{v} = R(s+t)\mathbf{v} = R(s)R(t)\mathbf{v}$, i.e., rotation by s then by t is the same as rotation by $s+t$, which is the same as rotation by t then by s. (This is another property of solutions of differential equations, based in turn upon uniqueness of solutions with the value prescribed at one point.) Then, $R(s+t)\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos(s+t) \\ \sin(s+t) \end{pmatrix} = \begin{pmatrix} \cos s\cos t - \sin s\sin t \\ \cos s\sin t + \sin s\cos t \end{pmatrix}$. These are the trigonometric addition formulas.

The Pythagorean Relation

We begin with a proof corresponding to the "precalculus" demonstration based upon the rotation formula and reflection.

By the rotation formula, if $R(s)\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} A \\ B \end{pmatrix}$ then $R(s)\begin{pmatrix} A \\ -B \end{pmatrix} = \begin{pmatrix} A^2+B^2 \\ 0 \end{pmatrix}$.

By time reversal, If $R(s)\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} A \\ B \end{pmatrix}$ then $R(-s)\begin{pmatrix} A \\ -B \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

Proof: $R((-s)+s)\mathbf{v} = R(0)\mathbf{v}$. So $R(-s)R(s)\begin{pmatrix} 1 \\ 0 \end{pmatrix} = R(0)\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

Finally, if $R(s)\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} A \\ B \end{pmatrix}$, then $R(s)\begin{pmatrix} A \\ -B \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

We use a reflection change of variables similar to the quarter circle rotation change of variables used to prove the rotation formula. Assume $u' = -v$ $v' = u$ $u(0) = A$ $v(0) = -B$ Let $x(t) = u(-t)$ and $y(t) = -v(-t)$. Then $x'(t) = -u'(-t) = v(-t) = -y(t)$ and $y'(t) = v'(-t) = u(-t) = x(t)$. Also $x(0) = A$ $y(0) = B$, so by the time reversal step, $x(-s) = 1$ $y(-s) = 0$, which says So $u(s) = 1$ $v(s) = 0$.

Taken together, these say if a rotation takes $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ to $\begin{pmatrix} A \\ B \end{pmatrix}$, then it takes $\begin{pmatrix} A \\ -B \end{pmatrix}$ to $\begin{pmatrix} A^2+B^2 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. The first component gives the Pythagorean relation in the form $\cos(s-s) = 1$. It gives the relationship between the horizontal and vertical coordinates of a point to which the point $(1,0)$ may be rotated by a rigid rotation about the origin. This implicit form of the unit circle is useful for testing whether a point is on the unit circle, while the explicit form, $\begin{pmatrix} \cos s \\ \sin s \end{pmatrix}$, is useful for generating points on the unit circle.

Another more direct demonstration that $(\cos s)^2 + (\sin s)^2 = 1$ does not appear to have a precalculus analogue.

$$\begin{aligned} \frac{d}{ds}\frac{1}{2}((\cos s)^2 + (\sin s)^2) &= \cos s\frac{d}{ds}\cos s + \sin s\frac{d}{ds}\sin s \\ &= \cos s(-\sin s) + \sin s\cos s = 0. \end{aligned}$$

Since the derivative is 0, the quantity is constant. Since $\cos(0)^2 + \sin(0)^2 = 1^2 + 0^2 = 1$, then $(\cos s)^2 + (\sin s)^2 = 1$ for all s. In vector terms, this says that the velocity is orthogonal to the position (by definition) and thus the length squared, the dot product of the position with itself, is constant.

Corollary: The components of the velocity vector $\begin{pmatrix} -\sin s \\ \cos s \end{pmatrix}$ also satify the Pythagorean relation, since

$$(-\sin s)^2 + (\cos s)^2 = 1.$$

These identities confirm our intuition that the point $(\cos s, \sin s)$ travels counterclockwise on the conventionally defined unit circle

$$x^2 + y^2 = 1$$

with unit speed, starting from $(1, 0)$ at $s = 0$.

By incorporating scaling, $S(r)$, into the discussion, such that $S(r)\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} r \\ 0 \end{pmatrix}$, and showing that $R(s)S(r) = S(r)R(s)$, we may extend these results to show $x^2 + y^2 = r^2$ and the rotation formula to a "rotation-scaling" formula for the complex number plane: If a rotation and scaling takes $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ to $\begin{pmatrix} A \\ B \end{pmatrix}$, then it takes $\begin{pmatrix} x \\ y \end{pmatrix}$ to $\begin{pmatrix} Ax - By \\ Bx + Ay \end{pmatrix}$. This will be done elsewhere.

More significant than the tendency to use the opposite letters from those representing rotation and scaling for the independent variables of their respective transformations, s and r (for arclength and radial distance), is the distinction in their origin. The arclength is based upon the independent variable for the differential equation defining rotation (although unit speed makes it the same as the distance travelled by the dependent variable.) To treat the radial distance analogously, we would define $S(r)$ "logarithmically" so that $S(r)\mathbf{v}$ is the solution at r of the corresponding "scaling differential equation" with intial value $\mathbf{v}$, so that again $S(0)$ is the identity. But there are limits to the challenges even we are prepared to make to convention!

The Geometric Interpretation of the Dot Product

The dot product of two vectors may be motivated as a generalization of the "negative reciprocal rule" for the slopes of perpendicular lines. The rule is based upon a similar triangle argument. If we apply the rule to the slopes of the lines determined by the components of perpendicular vectors $\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ and $\begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$, $\frac{v2}{v1}\frac{w2}{w1} = -1$ When one slope is zero and the other is infinite, the rule fails, but if we clear the denominators we obtain $\mathbf{v} \cdot \mathbf{w} \equiv v_1 w_1 + v_2 w_2 = 0$ which holds for any two perpendicular vectors. This expression is called the dot product of the vectors $\mathbf{v}$ and $\mathbf{w}$.

To see the significance of the dot product when it is not necessarily zero, we observe that if two vectors $\mathbf{v}$ and $\mathbf{w}$ are rotated by the same rotation, $R(s)$ then the dot product of the resulting vectors, $R(s)\mathbf{v} \cdot R(s)\mathbf{w}$, is the same as the dot product of the original vectors, $\mathbf{v} \cdot \mathbf{w}$. We compute:

$$\begin{aligned} R(s)\mathbf{v} \cdot R(s)\mathbf{w} &= \begin{pmatrix} v_1 \cos s - v_2 \sin s \\ v_1 \sin s + v_2 \cos s \end{pmatrix} \cdot \begin{pmatrix} w_1 \cos s - w_2 \sin s \\ w_1 \sin s + w_2 \cos s \end{pmatrix} . \\ &= v_1 v_2((\cos s)^2 + (\sin s)^2) + w_1 w_2((\cos s)^2 + (\sin s)^2) \\ &= v_1 v_2 + w_1 w_2 = \mathbf{v} \cdot \mathbf{w} \end{aligned}$$

using the Pythagorean relation.

If we perform a rotation on both $\mathbf{v}$ and $\mathbf{w}$ which makes the first component of $\mathbf{v}$ positive and the second component equal to zero, i.e., rotates it to the positive horizontal axis, then the components of $\mathbf{v}$ may be written

$$\begin{pmatrix} r_1 \\ 0 \end{pmatrix}$$

and the components of $\mathbf{w}$ may be written

$$\begin{pmatrix} r_2 \cos s \\ r_2 \sin s \end{pmatrix}$$

for some nonnegative numbers r_1, r_2, and s, the angle between $\mathbf{v}$ and $\mathbf{w}$.

The dot product $r_1 r_2 \cos s$ of these vectors is the same as the dot product of the original $\mathbf{v}$ and $\mathbf{w}$ and represents the product of their magnitudes times the cosine of the angle between them. This expression and its interpretation have far reaching consequences and generalizations which we will explore in more depth elsewhere.

The method we have used to give the geometric interpretation of the dot product can also be applied to give the geometric intepretation of the cross product in two and three dimensions, and the dot product $\mathbf{v} \cdot \mathbf{w} = \sum_1^n v_j w_j$ in any number of dimensions as follows. In three dimensions, we first perform a $y-z$ rotation to move the first vector to the $x-y$ plane by making the resulting

$z = 0$. Next we perform an $x - y$ rotation to make $y = 0$ (while z remains zero.) Finally, we perform another $y - z$ rotation (which leaves $y = z = 0$ of our first vector undisturbed) to make the z component of the second vector (which has already undergone the same rotations as the first) equal to zero. We may take any $k \leq n$ vectors in $\mathbf{R^n}$ and perform a sequence of coordinate plane rotations upon all of them (preserving all pairs of dot products, since it preserves the contribution from each pair of components) so that the image of the first vector has all but its first component zero, the image of the second has all but its first two components equal to zero, and so on. This shows that our method generalizes to the QR decomposition of a matrix by Givens' rotations, a computationally important method for orthogonalizing vectors.

Series Expansion

The series representations for the natural cosine and sine functions follow immediately from differentiating the defining derivative relationships:

The jth derivative of $\cos s$, is

$$\cos^{(j)}(s) = \left(\begin{array}{rl} \cos s & j = 0 \bmod 4 \\ -\sin s & j = 1 \bmod 4 \\ -\cos s & j = 2 \bmod 4 \\ \sin s & j = 3 \bmod 4 \end{array} \right.$$

and the jth derivative of $\sin s$, is

$$\sin^{(j)}(s) = \left(\begin{array}{rl} \sin s & j = 0 \bmod 4 \\ \cos s & j = 1 \bmod 4 \\ -\sin s & j = 2 \bmod 4 \\ -\cos s & j = 3 \bmod 4 \end{array} \right.$$

Using the initial conditions, also from the definition, at $s = 0$ the jth derivative of $\cos s$ is

$$\cos^{(j)}(0) = \left(\begin{array}{rl} 1 & j = 0 \bmod 4 \\ 0 & j = 1 \bmod 4 \\ -1 & j = 2 \bmod 4 \\ 0 & j = 3 \bmod 4 \end{array} \right.$$

and the jth derivative of $\sin s$ is

$$\sin^{(j)}(0) = \left(\begin{array}{rl} 0 & j = 0 \bmod 4 \\ 1 & j = 1 \bmod 4 \\ 0 & j = 2 \bmod 4 \\ -1 & j = 3 \bmod 4 \end{array} \right.$$

By considering the preferable "shift" form of the differentiation rule for polynomials, $a\frac{d}{dx}\frac{x^j}{j!} = a\frac{x^{j-1}}{j-1!}$. in which taking derivatives simply shifts the coefficients to of a polynomial to the left, it is easy to see that the polynomial $p(x)$ of smallest degree having prescribed derivatives $p^{(j)}(c) = a_j$, $j = 0, \ldots, k$ is

$$p(x) = \sum_{j=0}^{k} a_j \frac{x^j}{j!}$$

.

Putting these pieces together, we confirm that the polynomial $p(x)$ of degree k having the same jth derivatives as $\cos x$ $p^{(j)}(0) = \cos^{(j)}(0)$, $j = 0, \ldots, k$ is

$$p(x) = \sum_{j=0}^{2j \leq k} -1^j \frac{x^{2j}}{(2j)!}$$

or, written out other form:

$$p(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots + \frac{x^2 l}{(2l)!}$$

where $2l$ is the largest even number less than or equal to k.

The polynomial $p(x)$ of degree k having the same jth derivatives as $\sin x$ $p^{(j)}(0) = \sin^{(j)}(0)$, $j = 0, \ldots, k$ is

$$p(x) = \sum_{j=0}^{2j+1 \leq k} -1^j \frac{x^{2j+1}}{(2j+1)!}$$

or, written out other form:

$$p(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots + \frac{x^2 l + 1}{(2l+1)!}$$

where $2l + 1$ is the largest odd number less than or equal to k.

These formulas are the essence of the Taylor-Maclaurin series representations of the natural cosine and sine functions.

To demonstrate convergence of these polynomial approximations, as well as the existence and uniqueness of solutions of the rotational differential equation, we use the estimate for the remainder (difference) between an infinitely differentiable function f and its kth degree polynomial approximation given above, $R_{k+1}(s) = f^{(k+1)}(\xi)\frac{s^{k+1}}{(k+1)!}$ for some $\xi \in [0, s]$. Since $|\cos^{(k+1)}(s)| \leq 1$ and $|\sin^{(k+1)}(s)| \leq 1$ by applying the Pythagorean relation to the higher derivative formulas, and $\frac{s^{k+1}}{(k+1)!}$ as $k \to \infty$ for all $s \in [-\infty, \infty]$, the series is convergent. We will show elsewhere that the convergence of this series, which is also arises in the "Picard fixed point iteration" for the rotational differential equation also demonstrates the existence and uniqueness of its solutions.

This approach highlights the power of the fundamental framework of differential equations, such as existence, uniqueness, linear dependence on initial conditions, etc.