THE SINE AND COSINE FUNCTIONS

MAT157, WINTER 2021. YAEL KARSHON

As usual, please let me know if you find a typo or anything suspicious or unclear.

We would like $\alpha \mapsto (\cos\alpha, \sin\alpha)$, for $0 \leq \alpha \leq 2\pi$, to be the parametrized curve that starts at the point $(1,0)$ on the positive x-axis, travels along the unit circle in the counterclockwise direction with speed 1, and completes a full turn. For arbitrary $\alpha \in \mathbb{R}$, we'd like to have $\sin(\alpha \pm \pi) = -\sin(\alpha)$ and $\sin(\alpha \pm 2\pi) = \sin(\alpha)$ (do you see why?). The problem is that we have not yet defined the sine and cosine functions. But we did define the length ("arclength") of a curve, and we will use this notion to define the sine and cosine functions.

It will be convenient for us to work with the portion of the unit circle that lies on the right half-plane and corresponds to $-\pi/2 \leq \alpha \leq \pi/2$. We can, and do, use the y coordinate as a parameter for this curve. The x-coordinate is then $\sqrt{1-t^2}$. Thus, the curve is

$$\gamma(t) := (\sqrt{1-t^2}, t) \quad \text{for } -1 \leq t \leq 1.$$

Here is our plan. We will define the sine function in such a way that it will satisfy $y = \sin\alpha$ where α is the angle in radians that corresponds to the point $(\sqrt{1-y^2}, y)$ on this portion of the unit circle. If y is positive, so that we are above the x-axis, this angle is the arclength of the portion of the curve from $t = 0$ to $t = y$. If y is negative, so that we are below the x-axis, we consider the portion of the curve between $t = 0$ and $t = y = -|y|$, and the angle is the *negative* of the length of this portion of the curve. Denoting the angle by $L(y)$, we will show that the resulting function $L\colon [-1,1] \to \mathbb{R}$ is invertible, and on its image we will define the sine function to be the inverse function, L^{-1}. So we will have $\alpha = \sin y$ if and only if $L(y) = \alpha$, which is what we want. We will then extend the definition of the sine function to larger and larger intervals, until eventually we will define it on all of $\mathbb{R}$.

We now carry out this plan.

In the handout "Length of a curve" we showed that the curve $\gamma\colon [-1,1] \to \mathbb{R}^2$ given by $(\sqrt{1-t^2}, t)$ is rectifiable, we defined π to be its length, and we showed that $3 \leq \pi \leq 4$. By "concatenation" (see that handout), the restriction of γ to every sub-interval is also rectifiable, and $\text{length}(\gamma|_{[-1,0]}) + \text{length}(\gamma|_{[0,1]}) = \pi$.

We define $L\colon [-1,1] \to \mathbb{R}$ by

$$L(y) := \begin{cases} \text{length}\,\gamma|_{[0,y]} & 0 < y \leq 1 \\ 0 & y = 0 \\ -\text{length}\,\gamma|_{[y,0]} & -1 \leq y < 0\,. \end{cases}$$

We claim that for all $y \in [-1,1]$ we have $L(-y) = -L(y)$. It is enough to show that, for all $0 < y \leq 1$, we have $\text{length}(\gamma_{[0,y]}) = \text{length}(\gamma_{[-y,0]})$. We now show this. For every partition P of $[0,y]$, given by $0 = t_1 < t_2 < \ldots < t_n = y$, the partition Q of $[-y,0]$ that is given by $-y =$

$-t_n < -t_{n-1} < \ldots < -t_1 < -t_0 = 0$ satisfies $\ell(\gamma|_{[-y,0]}, Q) = \ell(\gamma|_{[0,y]}, P)$. (Exercise: check this.) This implies that the set $\{\ell(\gamma|_{[-y,0]}, Q) \mid Q \text{ a partition of } [-y,0]\}$ contains the set $\{\ell(\gamma|_{[0,y]}, P) \mid P \text{ a partition of } [0,y]\}$. So the suprema of these sets satisfy $\text{length}(\gamma|_{[-y,0]}) \geq \text{length}(\gamma|_{[0,y]})$. A similar argument shows that $\text{length}(\gamma|_{[0,y]}) \geq \text{length}(\gamma|_{[-y,0]})$. (Exercise: check the details.)

In particular, $\text{length}(\gamma|_{[-1,0]}) = \text{length}(\gamma|_{[0,1]})$. Because these lengths are equal to each other and their sum is π, they are both equal to $\pi/2$. So the function L satisfies $L(1) = \pi/2$ and $L(-1) = -\pi/2$. Also, $L|_{[0,1]} \colon [0,1] \to \mathbb{R}$ is continuous and strictly increasing (see the handout "Length of a curve"). It follows that $L \colon [-1,1] \to \mathbb{R}$ is continuous and strictly increasing, and its image is $[-\pi/2, \pi/2]$. (Exercise: fill the details.) So L has an inverse function, $L^{-1} \colon [-\pi/2, \pi/2] \to \mathbb{R}$, which is continuous and strictly increasing too, and which satisfies $L^{-1}(-\pi/2) = -1$ and $L^{-1}(\pi/2) = 1$.

Definition. For $-\pi/2 \leq \alpha \leq \pi/2$, we define $\sin\alpha := L^{-1}(\alpha)$.

Thus, the function $\sin \colon [-\pi/2, \pi/2] \to \mathbb{R}$ is continuous and strictly increasing. Also, it satisfies $\sin(-\pi/2) = -1$, $\sin(0) = 0$, and $\sin(\pi/2) = 1$.

Next, we extend the definition of the sine function to the interval $[-\pi/\pi]$, in such a way that $\sin(\alpha + \pi) = -\sin(\alpha)$:

Definition. For $-\pi \leq \alpha \leq -\pi/2$, we define $\sin(\alpha) := -\sin(\alpha + \pi)$. For $\pi/2 \leq \alpha \leq \pi$, we define $\sin(\alpha) := -\sin(\alpha - \pi)$.

Here, the right hand sides were defined earlier. At $\alpha = -\pi/2$ and at $\alpha = \pi/2$, the new definition is consistent with the old one.

The resulting function $\sin \colon [-\pi, \pi] \to \mathbb{R}$ is continuous on each of the closed subintervals $[-\pi, -\pi/2]$, $[-\pi/2, \pi/2]$, and $[\pi/2, \pi]$, so it is continuous. (Exercise: check the details.) Also, $\sin(-\pi) = \sin(0) = \sin(\pi) = 0$.

Next, we extend the definition of the sine function to the entire real line, in such a way that $\sin(\alpha + 2\pi) = \sin(\alpha)$.

Definition. For each non-zero integer k, for $\alpha \in [-\pi + 2\pi k, \pi + 2\pi k]$, we define $\sin(\alpha) := \sin(\alpha - 2\pi k)$.

Again, the right hand side was defined earlier. When α is an odd multiple of π, we obtain two definitions of $\sin\alpha$, which both give $\sin\alpha = 0$, so they are consistent with each other. The function $\sin \colon \mathbb{R} \to \mathbb{R}$ is continuous on each closed interval $[-\pi + 2\pi k, \pi + 2\pi k]$, so it is continuous.

Exercise. For all α, we have $\sin(\alpha + \pi) = -\sin\alpha$ and $\sin(\alpha + 2\pi) = \sin\alpha$.

Next, we examine the derivative of the sine function.

Claim. On the interval $(-\pi/2, \pi/2)$, the function sin is differentiable, and

$$\sin'(\alpha) = \sqrt{1 - \sin^2\alpha}\,.$$

Proof. On the interval $[0,1)$, the curve γ is C^1. We calculate its speed: the curve is given by $x(t) = \sqrt{1-t^2}$ and $y(t) = t$. So $\dot{x}(t) = \frac{-t}{\sqrt{1-t^2}}$ and $\dot{y}(t) = 1$, and so the speed is $s(t) = \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2} = \sqrt{\frac{t^2}{1-t^2} + 1} = \frac{1}{\sqrt{1-t^2}}$. So, on the interval $[0,1)$, we have

$$L(y) = \int_0^y \frac{1}{\sqrt{1-t^2}} dt,$$

and, by the fundamental theorem of calculus,

$$L'(y) = \frac{1}{\sqrt{1-y^2}}.$$

On the interval $(-1,0]$, because $L(y) = -L(-y)$, we have $L'(y) = L'(-y)$, so again $L'(y) = \frac{1}{\sqrt{1-y^2}}$. Because at the point $y = 0$ the left and right derivatives of L are equal, the function L is differentiable on the entire interval $(-1,1)$, and $L'(y) = \frac{1}{\sqrt{1-y^2}}$ throughout this interval. By the inverse function theorem, the sine function is differentiable on $(-\pi/2, \pi/2)$, and

$$\sin'(\alpha) = \frac{1}{L'(y)\big|_{y=\sin\alpha}} = \sqrt{1-\sin^2\alpha}.$$

$\square$

Because for all α we have $\sin(\alpha + 2\pi) = \sin(\alpha)$, we obtain that for every even integers k the sine function is differentiable on the interval $(-\pi/2 + k\pi, \pi/2 + k\pi)$ and its derivative on this interval is $\sqrt{1-\sin^2(\alpha)}$.

Because for all α we have $\sin(\alpha + \pi) = -\sin(\alpha)$, we further obtain that for every odd integer k the sine function is differentiable on the interval $(-\pi/2 + k\pi, \pi/2 + k\pi)$ and its derivative on this interval is $-\sqrt{1-\sin^2(\alpha)}$.

We now define the cosine function $\cos\colon \mathbb{R} \to \mathbb{R}$ by

$$\cos(\alpha) := \begin{cases} \sqrt{1-\sin^2\alpha} & \alpha \in [-\pi/2 + k\pi, \pi/2 + k\pi] \text{ for an even integer } k \\ -\sqrt{1-\sin^2\alpha} & \alpha \in [-\pi/2 + k\pi, \pi/2 + k\pi] \text{ for an odd integer } k. \end{cases}$$

When α is of the form $\pi/2 + k\pi$ for some integer k, the two definitions of $\cos(\alpha)$ (are both 0, hence) coincide.

Because the cosine function is continuous on each of the intervals $[-\pi/2 + k\pi, \pi/2 + k\pi]$ for $k \in \mathbb{Z}$, it is continuous throughout $\mathbb{R}$.

By what we proved earlier, we have

$$\sin'(\alpha) = \cos(\alpha)$$

on every interval of the form $(-\pi/2 + k\pi, \pi/2 + k\pi)$ for some integer k. So $\sin'(\alpha) = \cos(\alpha)$ holds at every α that is not of the form $\pi/2 + k\pi$. To finish, note that the sine function is continuous at $\pi/2 + k\pi$, its derivative is $\cos(\alpha)$ on a punctured neighbourhood of $\pi/2 + k\pi$, and $\lim_{\alpha \to \pi/2 + k\pi} f'(\alpha)$ exists and is equal to $\cos(\pi/2 + k\pi)$ (because the cosine function is continuous). These conditions imply that $\sin'(\alpha) = \cos(\alpha)$ also at $\alpha = \pi/2 + k\pi$ (See Spivak's Chapter 11, Theorem 7.)