ECE / CS 6501: Geometry of Data **Due Tu 9/13**

Homework 1: Getting Started

Instructions: Submit a single Jupyter notebook (.ipynb) of your work to Collab by 11:59pm on the due date. All code should be written in Python. **Be sure to show all the work involved in deriving your answers! If you just give a final answer without explanation, you may not receive credit for that question.**

You may discuss the concepts with your classmates, but write up the answers entirely on your own. Do not look at another student's answers, do not use answers from the internet or other sources, and do not show your answers to anyone.

1. Consider the set of all **closed** intervals in the real line with non-zero length:
$$\mathcal{B} = \{[a, b] : a < b\}.$$
Is this a valid basis for a topology on $\mathbb{R}$? Why or why not?

2. Given a metric space X with metric $d : X \times X \to \mathbb{R}$ and a subset $A \subseteq X$, show that the metric restricted to A, $d|_{A \times A}$, gives the same topology as the subspace topology on A. **Hint:** you may want to solve Exercise 2 in Section 1 of the notes and use that result.

3. Write a Python function to compute the stereographic projection, $f : S^2 \to \mathbb{R}^2$, from the north pole $(0, 0, 1)$ of the 2-sphere to the $z = 0$ plane. You may want to see the Wikipedia page for the equation:

   `https://en.wikipedia.org/wiki/Stereographic_projection`

   (a) What is the image of a longitudinal line, i.e., a great circle passing through the north and south poles? Use your function to plot several images of logitudinal lines in the plane.

   (b) What is the image of a latitude line, i.e., a circle formed by intersection with a horizontal plane ($z =$ const.)? Plot several images of latitude lines in the plane.

   (c) What is the image of an arbitrary great circle (one that is **not** a longitudinal line or the equator)? Again, plot several in the plane.

4. Write a Python function to sample uniformly random points on the d-sphere, S^d. Do this by first sampling $d + 1$ i.i.d. standard Gaussian random variables, $N(0, 1)$, to get a point $x = (x_1, x_2, \ldots, x_{d+1}) \in \mathbb{R}^{d+1}$. Then project this point to the unit sphere by $x \mapsto \frac{x}{\|x\|}$.
Repeat the following experiments for $d = 2, 4, 10, 100, 1000$

   (a) Sample 10,000 points on S^d and compute their distance from the "equator", or great circle with $x_1 = 0$. Distance to the equator should be arc distance along the sphere. For each value of d, plot a histogram of these distances. What do you notice as d increases? What does this say about uniform random points on the sphere? Would this result be the same or different if you used a different great circle (possibly not axis-aligned)?

   (b) Sample 1,000 pairs of points on S^d and compute the angle between them. Again, for each value of d, plot a histogram of these pairwise angles. What do you notice as d increases and what does it say about the geometry of uniformly random points?