Metric Spaces and Length Spaces

C. Sormani

Lehman College

Definition: A **metric space** is a set of points, X, and a distance function $d : X \times X \to [0, \infty)$ such that the following three conditions hold:

$$a) \qquad d(x,y) \geq 0 \qquad \forall x,y \in X \text{ and } d(x,y) = 0 \text{ iff } x = y,$$

$$b) \qquad d(x,y) = d(y,x) \qquad \forall x,y \in X,$$

$$c) \qquad d(x,z) \leq d(x,y) + d(y,z) \qquad \forall x,y,z \in X.$$

This third condition is called the triangle inequality.

Note that any space satisfying the axioms of Neutral geometry is a metric space and so is the sphere. Other metric spaces are the taxicab space and the torus. Metric spaces in which the points are functions are used to solve differential equations and other topics in analysis. Lots of information about metric spaces can be found in Marsden's "Elementary Classical Analysis" and I highly recommend this book.

Exercise 1*: The taxicab space is $\mathbb{R}^2$ with the distance function

$$d((x_1, y_1), (x_2, y_2)) = |x_1 - x_2| + |y_1 - y_2|.$$

Verify that this is a metric space.

Exercise 2: The torus is the sset of points

$$[0,1) \times [0,1) = \{(x,y) : x \in [0,1), y \in [0,1)\},$$

and it has the distance function

$$d((x_1, y_1), (x_2, y_2)) = \sqrt{((a)^2 + (b)^2}$$

where

$$a = \min\{|x_1 - x_2|, |x_1 - x_2 + 1|\} \text{ and } b = \min\{|y_1 - y_2|, |y_1 - y_2 + 1|\}$$

Verify that this is a metric space by going through the various cases and using the fact that Euclidean space is a metric space.

One of the most important consequences of being a metric space is that one can define a ball.

Definition: A **ball** of radius r about $x \in X$ is

$$B(p,r) = \{y : d(x,y) < r\}.$$

Note that balls in Euclidean space are just the insides of circles.

Exercise 3*: Describe what balls in the taxicab space look like.

Exercise 4: Describe what balls in a torus look like.

We can also define continuous functions:

Definition: A **continuous** function between metric spaces X and Y is a function $f : X \to Y$ such that for all $x \in X$, for all $\epsilon > 0$ there exists $\delta > 0$ sufficiently small (which depends on x and ϵ) such that $f(B(x,\delta)) \subset B(f(x),\epsilon)$. That is, if $y \in B(x,\delta)$ then $f(y) \in B(f(x),\epsilon)$.

The idea is the same as the one in Calculus I where $f : \mathbb{R} \to \mathbb{R}$ is defined. This is explained in an Appendix of Larson's Calculus for example. This metric space case is explained very nicely in Marsden with graphics.

For our purposes we will focus on continuous curves which are continuous functions from intervals $[a,b] \subset \mathbb{R}$ to a metric space Y. The idea behind a continuous curve is that it is unbroken.

We say that a curve $c : [a,b] \to X$ runs from x to y if $c(a) = x$ and $c(y) = b$. We'd like to measure the lengths of curves. This can be done approximately by taking partitions $t_0 = a < t_1 < t_2 < ... < t_n = b$ of $[a,b]$ and finding the **polygonal length**:

$$PL(c, \{t_0, ...t_n\}) = \sum_{i+1}^{n} d(c(t_i), c(t_{i-1})).$$

In Euclidean space this just gives the length of the curve made up of straight segments running between the point $c(t_i)$.

Exercise 5*: Show that if $\{t_i\}$ is a partition of a curve in a metric space X, and we then create a new partition by adding points s_i between t_i and t_{i+1} for $i = 0..n-1$, then the polygonal length cannot decrease:

$$PL(c, \{t_0, ...t_n\}) \leq PL(c, \{t_0, s_0, t_1, s_1, t_2, ..., s_{n-1}, t_n\}).$$

Hint: use the definition of PL to rewrite both sides and then use the triangle inequality.

Definition: A continuous curve, c, is **rectifiable** if

$$sup\{PL(c,S) : S \text{ is any partition of [a,b]}\} \text{ exists,}$$

in which case this supremum is called the length of the curve, $L(c([a,b]))$.

Note that in $\mathbb{R}^n$, where smoothness is defined, you may have learned about the arclength of a smooth curve which is defined by integration of the length of the first derivative vector. One can prove that this length agrees with the length defined above but it uses a lot of serious calculus.

Note that there are some curves which are not rectifiable because the polygonal lengths just get longer and longer.

Exercise 6*: Show that in Euclidean space or Hyperbolic space X, any line segment is a rectifiable curve and that the length of the curve is the distance between the endpoints. Hint: define the curve c using the function which identifies points on a line with points on $\mathbb{R}$, $c : \mathbb{R} \to X$ and then $c : [a, b] \to X$ is a line segment.

Exercise 7: Find an example of a curve in Euclidean space which isn't rectifiable by searching the library or the web.

Exercise 8*: Show that in Euclidean space $\mathbb{R}^2$ the circle curve (cos(t), sin(t)) is rectifiable by showing that for any partition the length of the partition is ≤ 8. Hint: Put the circle inside a square of side length 2 and project the lengths of the partition pieces to the square (arguing that the lengths increase) and then use the perimeter of the square as an upper bound.

Exercise 9: Explain how the perimeters of n gons with circles inscribed in them can be used to estimate the length of a circle from both sides thus obtaining an estimate for π.

Sometimes the distance function on metric space is defined to match the length of the shortest curve between the points. However, it is possible that there is no shortest curve so instead we use an infimum.

Definition: A **Length Space**, X is a metric space such that for any $x, y \in X$ we have

$$d(x, y) = \inf\{L(c([a, b])) : c \text{ is rectifiable }, c(a) = x, c(b) = y\}.$$

A length space is called a **complete length space** if for all $x, y \in X$ there is a **length minimizing rectifiable curve** c from x to y such that

$$d(x, y) = L(c([a, b])).$$

Exercise 10*: Suppose X is Euclidean or Hyperbolic space, then X is a complete length space. Add justifications to the following proof of this fact:

1) L(c([a,b])) exists. by Exercise 6.
2)For all $t_1 \in (a, b)$, $L(c) \geq PL(c, \{a, t_1, b\})$.
3)So $L(c([a, b])) \geq d(c(a), c(t_1)) + d(c(t_1), c(b)) = d(x, c(t_1)) + d(c(t_1), y)$
4) But $d(x, c(t_1)) + d(c(t_1), y) \geq d(x, y)$.
5) So $L(c([a, b])) \geq d(x, y)$, *by steps 1-4.*
6) Thus $inf(L(c)) \geq d(x, y)$ *where the infimum is taken over all c from* x *to* y.
7) However $d(x, y)$ *equals the length of the line segment from* x *to* y.
8) So the infimum = d(x,y).

9) So by 8, X is a length space.
10) And by step 7 again, X is in fact a complete length space.

It is also fairly easy to prove that the sphere is a length space. Steps 1-6 of the above proof can be used and justified easily. The hard part is showing that there are "geodesic segments" running between the points whose length is the distance between the points. This can be done fairly easily if one uses arclength and differentiation and can be found in Noronha's text for example. However, it is a bit more difficult when done directly using the definition of length as the sup of partition lengths.

Exercise 11*: Suppose X is the taxicab space. Show that X is a complete length space by imitating the above proof and showing that the curve

$$\begin{aligned} c(t) &= (x_1 + t, y_1) \text{ for } t \in (0, x_2 - x_1) \quad (1) \\ &= (x_2, y_1 + (t - x_2 + x_1)) \text{ for } t \in (x_2 - x_1, y_2 - y_1 + x_2 - x_1) \quad (2) \end{aligned}$$

runs from (x_1, y_1) to (x_2, y_2) and is rectifiable and that it's length is $|x_1 - x_2| + |y_1 - y_2|$.

Given any length space X with a distance function d_X, and given any subset Y contained in X, one can define an induced length structure d_Y on Y as follows: $d_Y(x, y)$ is the infimum of the length of any rectifiable curve c running from x to y which remains in Y. Here the length of c is measured using d_X. An example of such a situation is when X is three dimensional Euclidean space d_X the usual three dimensional distance formula and Y is the sphere which is a subset of X. The d_Y turns out to be the standard metric on the sphere and this is in fact the most intuitive way of understanding the sphere as a metric space.

Exercise 12*: Show that when one defines the induced length structure as above then Y with the metric d_Y is always a metric space.

Exercise 13: Show that when one defines the induced length structure as above then Y with the metric d_Y is always a length space.

Exercise 14*: Show that when one defines the induced length structure as above then Y with the metric d_Y is not always a complete length space even if X is. Use the example X is Euclidean space and Y is the whole space with the origin removed. State explicitly what d_Y is and give an explicit pair of points in Y with no length minimizing curve running between them.