NOTES ON METRIC SPACES

JUAN PABLO XANDRI

1. Introduction

Let X be an arbitrary set, which could consist of vectors in $\mathbb{R}^n$, functions, sequences, matrices, etc. We want to endow this set with a **metric**; i.e a way to measure distances between elements of X. A **distance** or **metric** is a function $d : X \times X \to \mathbb{R}$ such that if we take two elements $x, y \in X$ the number $d(x, y)$ gives us the distance between them. However, not just any function may be considered a metric: as we will see in the formal definition, a distance needs to satisfy certain properties.

Definition 1.1 (**Metric Spaces**). Given a set X and a function $d : X \times X \to \mathbb{R}$, we say that the pair $\mathcal{M} = (X, d)$ is a **metric space** if and only if $d(\cdot)$ satisfies the following properties:

(1) (Non-negativeness) For all $x, y \in X$, $d(x, y) \geq 0$
(2) (Identification) For all $x, y \in X$ we have that $d(x, y) = 0 \iff x = y$
(3) (Symmetry) For all $x, y \in X$, $d(x, y) = d(y, x)$
(4) (Triangular inequality) For all $x, y, z \in X$ we have that

$$d(x, z) \leq d(x, y) + d(y, z) \tag{1.1}$$

Property (1) just states that a distance is always a non-negative number. Property (2) tells us that the distance identifies points; i.e. if the distance between x and y is zero, it is because we are considering the same point. Property (3) states that a metric must measure distances symmetrically; i.e. it does not matter where we start measuring it. Finally, the triangular inequality is a generalization of the famous result that holds for the euclidean distance in the plane

2. Examples of Metric Spaces

2.1. **Norms in vector spaces.** Let $X = \mathbb{R}^n$. The typical distance used is the **euclidean distance**, defined as

$$d_2(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

For this metric, all properties except (1.1) are trivially shown. The triangular inequality is a well known result from linear algebra, known as the Cauchy-Schwartz inequality. However, this is not the only distance we could define over $\mathbb{R}^n$. Consider the distance

$$d_1(x,y) \equiv \sum_{i=1}^{n} |x_i - y_i|$$

which is known as the "taxicab distance". Another one we could use is the so called "max distance", defined as

$$d_\infty(x,y) \equiv \max_{i\in[1:n]} |x_i - y_i|$$

where $[1:n] \equiv \{1,2,\cdots,n\}$. Notice that all this distances can be written as $d(x,y) = \mathbf{N}(x-y)$, for some function $\mathbf{N}:\mathbb{R}^n \to \mathbb{R}$. In general, one could define a lot of distances, based on different functions $N(\cdot)$ that one can come up with. In particular, sometimes we are interested in a subset of functions called **norms.** We will define them in a way that it applies to any **vector space** X (i.e. any space in which you can calculate $x+y$ and αx with $\alpha \in \mathbb{R}$ with the usual rules).

Definition 2.1 (**Norms**)**.** Let X be a vector space (e.g. $X = \mathbb{R}^n$) and $\mathbf{N}: X \to \mathbb{R}$. We say $\mathbf{N}(\cdot)$ is a **norm** if the following 4 conditions hold:

- **(i):** : $\mathbf{N}(x) \geq 0$ for all $x \in X$
- **(ii):** : $\mathbf{N}(x) = 0 \iff x = 0$ for all $x \in X$
- **(iii):** : $\mathbf{N}(\alpha x) = |\alpha| \mathbf{N}(x)$ for all $\alpha \in \mathbb{R}$, $x \in X$
- **(iv):** : $\mathbf{N}(a+b) \leq \mathbf{N}(a) + \mathbf{N}(b)$ for all $a, b \in X$

Exercise 2.1. Show that the functions $\mathbf{N}_2(x) = \sqrt{x_i^2}$, $\mathbf{N}_1(x) = \sum_i |x_i|$ and $\mathbf{N}_\infty(x) = |x_i|$ are norms

Proposition 2.1. *Let X be a vector space (e.g. $X = \mathbb{R}^n$) and define $d_{\mathbf{N}}: X \times X \to \mathbb{R}$ as*

$$d_{\mathbf{N}}(x,y) \equiv \mathbf{N}(x-y)$$

Then, $\mathcal{M}_{\mathbf{N}} = (X, d_{\mathbf{N}})$ is a metric space

Proof. We need to prove each of the properties of a distance (from 1 to 4).

Non negativity (1) : Easy, since $\mathbf{N}(\cdot) \geq 0$ always

Identification (2) : Follows from

$$d_{\mathbf{N}}(x,y) = 0 \iff \mathbf{N}(x-y) = 0 \underbrace{\iff}_{\text{(ii)}} x - y = 0 \iff x = y$$

Symmetry (3) : Follows from

$$d_{\mathbf{N}}(x,y) = \mathbf{N}(x-y) = \mathbf{N}((-1)(y-x)) \underbrace{=}_{\text{(iii) } \alpha=-1}$$

$$= |-1| \mathbf{N}(y-x) = d_{\mathbf{N}}(y,x)$$

Triangle inequality (4) : Take $x, y, z \in X$:

$$d_{\mathbf{N}}(x,z) = \mathbf{N}(x-y) = \mathbf{N}(x-z+z-y) \underbrace{\leq}_{(\mathbf{iv})} \mathbf{N}\left(\underbrace{x-y}_{=a}\right) + \mathbf{N}\left(\underbrace{y-z}_{=b}\right) = d_{\mathbf{N}}(x,y) + d_{\mathbf{N}}(y,z)$$

□

2.2. **Functional Spaces.** Probably the most important new concept will be the space X that consists of functions instead of vectors. The most important one is the so-called "**sup-norm metric**" space: pick a norm $\mathbf{N} : \mathbb{R}^m \to \mathbb{R}$ and a set $A \subseteq \mathbb{R}^n$ and define

$$X = \{f : A \to B \text{ such that } \exists K_f > 0 \text{such that } \mathbf{N}[f(x)] \leq K_f \text{for all } x \in A\} \tag{2.1}$$

$$d_\infty(f,g) \equiv \sup_{x \in A} \mathbf{N}(f(x) - g(x)) \tag{2.2}$$

We define the metric space $\mathcal{M} \equiv \mathcal{B}(A, \mathbb{R}^n) = (X, d_\infty)$ as the **set of bounded functions from A to $\mathbb{R}^m$**. Note that X is a vector space, defining the sum of functions as the point-wise sum; i.e.

$$(f+g)(x) = f(x) + g(x) \text{ for all } x \in A$$

since the sum of bounded functions is also bounded. Hence, if we show that the function

$$\mathbf{N}^*(f) \equiv \sup_{x \in A} |\mathbf{N}[f(x)]| \tag{2.3}$$

is itself a norm for X, we can apply Proposition 2.1 to show $\mathcal{B}(A, \mathbb{R}^m)$ is indeed a metric space. First, we need to show that $\mathbf{N}^*(f) < \infty$ for all $f \in X$ (i.e. it is a well defined object). This can be done since we are only taking functions that are bounded, and hence the supremum always exist (the so-called "Axiom of completeness" of the real numbers). Now, we show the properties of a norm.

(i) : $\mathbf{N}^*(f) \geq 0$ *for all* $f \in X$*.* This follows from N being itself a norm

(ii) : $\mathbf{N}^*(f) = 0 \iff f(x) = 0$*for all* $x \in A$*.*

$$\mathbf{N}^*(f) = 0 \iff \sup_{x \in A} \mathbf{N}[f(x)] = 0 \underbrace{\iff}_{(a)} \mathbf{N}[f(x)] = 0 \text{ for all } x \in A \underbrace{\iff}_{(b)}$$

$$f(x) = 0 \text{ for all } x \in A$$

(a) follows from the fact that if $\mathbf{N}[f(\hat{x})] > 0$ for some $\hat{x} \in A$, then the sup would have to also be strictly positive. (b) follows from $\mathbf{N}(\cdot)$ being a norm (Property (**ii**))

(iii) : $\mathbf{N}^*(\alpha f) = |\alpha| \mathbf{N}^*(f)$*.* This comes from the homogeneity of the sup operator:

$$\mathbf{N}^*(\alpha f) = \sup_{x \in A} N[\alpha f(x)] \underbrace{=}_{(c)} \sup_{x \in A} |\alpha| N[f(x)] = |\alpha| \sup_{x \in A} N[f(x)] = |\alpha| \mathbf{N}^*(f)$$

where (c) follows from the fact that $\mathbf{N}$ is a norm.

(iv) : *Triangular inequality.*

$$\mathbf{N}^*(f+g)=\sup_{x\in A}\mathbf{N}[f(x)+g(x)]\leq\sup_{x\in A}\mathbf{N}[f(x)]+\mathbf{N}[g(x)]=\sup_{x,y\in A:x=y}\mathbf{N}[f(x)]+\mathbf{N}[g(y)]\leq$$

$$\sup_{x\in A}\mathbf{N}[f(x)]+\sup_{y\in A}\mathbf{N}[g(y)]=\mathbf{N}^*(f)+\mathbf{N}^*(g)$$

and hence, $\mathcal{B}(A,\mathbb{R}^m)$ is indeed a metric space.

Exercise 2.2. Let $A\subseteq\mathbb{R}^n$. Show that the following are metric spaces: $\mathcal{L}_1(A,\mathbb{R})=(X_1,d_1)$ where

$$X_1=\left\{f:A\to\mathbb{R}\text{ such that }\int_{x\in A}|f(x)|\,dx<\infty\right\}$$

$$d_1(f,g)\equiv\int_{x\in A}|f(x)-g(x)|\,dx$$

and $\mathcal{L}_2(A,\mathbb{R})=(X,d_2)$ where

$$X_1=\left\{f:A\to\mathbb{R}\text{ such that }\int_{x\in A}f^2(x)\,dx<\infty\right\}$$

$$d_2(f,g)\equiv\sqrt{\int_{x\in A}[f(x)-g(x)]^2\,dx}$$

In general, show that given a norm $\mathbf{N}:\mathbb{R}^m\to\mathbb{R}$ the pair $\mathcal{L}_{\mathbf{N}}(A,\mathbb{R})=(X_{\mathbf{N}},d_{\mathbf{N}})$ defined as

$$X_{\mathbf{N}}=\left\{f:A\to\mathbb{R}\text{ such that }\int_{x\in A}\mathbf{N}[f(x)]\,dx<\infty\right\}$$

$$d_{\mathbf{N}}(f,g)\equiv\int_{x\in A}\mathbf{N}[f(x)-g(x)]\,dx$$

$$\mathbf{N}^*(f)=\left\{\int_{x\in A}\mathbf{N}[f(x)]^\rho\,dx\right\}^{\frac{1}{\rho}}$$

is a metric space as well, for any $\rho>0$.

3. Metric Topology

In this section we will be studying the concept of "neighborhood" or closeness in generic metric spaces. This will be useful when generalizing concepts like "open sets", "continuous functions", "compact sets", etc.

3.1. **Open and closed balls.**

Definition 3.1 (Open Ball). Given a metric space $\mathcal{M}=(X,d)$, $a\in X$ and $r>0$ we define the open ball of center a and radius r as the set

$$B(a,r)=\{x\in X:d(a,x)<r\}$$

Definition 3.2 (**Open Ball**)**.** Given a metric space $\mathcal{M} = (X, d)$, $a \in X$ and $r > 0$ we define the closed ball of center a and radius r as the set

$$\overline{B}(a, r) = \{x \in X : d(a, x) \leq r\}$$

For example, in $\mathcal{B}([-1, 1], \mathbb{R})$ when we take the center $f(x) = x^3$ and radius r, is easy to see that

$$g \in B(f, r) \iff g(x) \in (f(x) - r, f(x) + r)$$

3.2. **Open and closed sets.** In our first calculus courses, we saw that an "open set" was one that did not include its "border", or more formally, its "frontier". However, in generic metric spaces this cannot be graphically checked, so we need to have the formal definition of this concept. The basic idea is that for a set to be open (and not include its border), every time we pick an element $x \in A$ we must be able to find an open ball around it that it is also completely included on the same set A.

Definition 3.3 (**Interior of a set**)**.** Let $\mathcal{M} = (X, d)$ be a metric space and $A \subseteq X$. We say that $x \in A$ is interior of $A \iff \exists r_x > 0$ such that $B(x, r_x) \subseteq A$. The set of all interior points of A is called the interior of A, and is written as $\mathring{A}$

Definition 3.4 (**Open set**)**.** A set $A \subseteq X$ is open $\iff A = \mathring{A}$

The first example of open set is in fact, the open balls themselves:

Proposition 3.1 (**Open balls are open**)**.** *Given $\mathcal{M} = (X, d)$ a metric space, $x \in X$ and $r > 0$, the set $A \equiv B(x, r)$ is an open set.*

Proof. To prove A is open we need to show that for any $z \in B(x, r) \iff d(x, z) < r$ we need to find a radius $r_z > 0$ such that $B(z, r_z) \subseteq A = B(x, r)$. This equivalent to prove the following statement:

$$\exists r_z > 0 : \text{ for all } y \in X, \text{ if } d(y, z) < r_z \Longrightarrow d(x, y) < r \tag{3.1}$$

A natural candidate would be $r_z = r - d(x, z)$. In fact, when $r_z = r - d(x, z)$, then for any $y \in B(z, r_z)$:

$$d(x, y) \leq d(x, z) + d(z, y) \underbrace{<}_{d(y,z) > r_z} d(x, z) + [r - d(x, z)] = r$$

and hence $d(x, y) < r$ (i.e. $y \in B(x, r)$) as we wanted to show. $\square$

In basic calculus, we also thought of "closed sets" as those sets that would include its boundary, unlike open sets. Intuitively, a set is closed if it includes all the points that are "pasted to it" in a sense. This concept is that of closure points.

Definition 3.5 (**Closure Set**)**.** Given a m.s $\mathcal{M} = (X, d)$ and a set $A \subseteq X$, we say that $x \in X$ is a closure point of $A \iff$ the following rule holds:

$$\forall r > 0 \text{ we have } B(x, r) \cap A \neq \emptyset$$

i.e. no matter how close you get to x, there is always a point in A which is even closer to x. The set of all closure points of A is called the closure of A and is denoted by $\overline{A}$

Definition 3.6 (**Closed set**)**.** A set $A \subseteq X$ is closed $\iff A = \overline{A}$

The first result we get, for any given set, is that A is "in between" its interior and its closure

Proposition 3.2. *For any subset $A \subseteq X$, we have*

$$\mathring{A} \subseteq A \subseteq \overline{A}$$

Proof. That $\mathring{A} \subseteq A$ is obvious, since if $x \in \mathring{A} \Longrightarrow \exists B(x, r) \subseteq A$ and $x \in B(x, r)$. For the closure, is just suffices to note that for any $x \in A$ and any $r > 0$ we have that $x \in B(x, r) \cap A \Longrightarrow x$ is a closure point. □

Another important property is the one that relates closed an open sets: an closed set is, by definition, the complement of an open set.

Theorem 3.1. *Let $\mathcal{M} = (X, d)$ be a metric space. Then the following rule holds:*

$$A \subseteq X \text{ is open } \iff A^c \equiv \{x \in X : x \notin A\} \text{ is closed}$$

and vice versa

Proof. (Direct $\Longrightarrow$)Suppose A is open. We need to show that A^c is closed. Suppose, by contradiction, that it is not: i.e. there exists a point $\hat{x} \in \overline{A^c}$ such that $\hat{x} \notin A^c \iff \hat{x} \in A$. Since A is open exists $\hat{r} > 0$ such that $B(\hat{x}, \hat{r}) \subseteq A$

$$\exists \hat{r} > 0 : B(\hat{x}, \hat{r}) \subseteq A \Longrightarrow$$

$$\exists \hat{r} > 0 : B(\hat{x}, \hat{r}) \cap A^c = \emptyset \tag{3.2}$$

But (3.2) contradicts the definition of closure point of A^c for $x = \hat{x}$, and hence $\hat{x} \notin \overline{A^c}$, a contradiction

(Converse $\Longleftarrow$) Suppose A^c is closed, and we have to show A is open. Suppose, by contradiction, that it was not: then, there exist a point $\hat{x} \in A$ that is not interior: i.e.

$$\forall r > 0 \text{ exists } y \in B(\widehat{x, r}) \text{ such that } y \notin A \iff$$

$$\forall r > 0 \text{ we have } B(\hat{x}, r) \cap A^c \neq \emptyset \tag{3.3}$$

Condition (3.3) then implies that $\hat{x}$ is a closure point, and since A^c is closed, we must have $\hat{x} \in A^c \iff x \notin A$. But this contradicts the initial assumption that $\hat{x} \in A$, proving the desired result. □

3.3. **Operations between open and closed sets.**

Theorem 3.2 (**Metric Topology**). *For any given metric space $\mathcal{M} = (X, d)$:*

1. $A = X$ *and* $A = \emptyset$ *are open sets*
2. *Let $\mathcal{O}$ be a family of open sets of (X, d). Then the set*
$$A^* \equiv \bigcup_{A \in \mathcal{O}} A$$
is also open
3. *Let $\{A_1, A_2, \dots, A_k\}$ be a finite family of open sets. Then the set*
$$A_* \equiv \bigcap_{i=1}^{i=k} A_i$$
is also open

Proof. Condition (1) is trivial. Take a family of open sets $\mathcal{F}$ and take $x \in A^*$. Now

$$x \in A^* \iff \exists A_x \in \mathcal{F} : x \in A_x \underbrace{\implies}_{A_x \text{ open}} \exists r_x > 0 : B(x, r_x) \subseteq A_x \subseteq \textstyle\bigcup_{A \in \mathcal{F}} A$$
$$\implies \exists r_x : B(x, r_x) \subseteq A$$

and hence A^* is open. For (3), we know that

$$x \in A_* \iff x \in A_i \text{ for all } i \in [1:k] \underbrace{\implies}_{A_i \text{ open}} \exists r_x^i > 0 : B\left(x_i, r_x^i\right) \subseteq A_i \text{ for all } i \in [1:k]$$

Then, by choosing $r_x \equiv \min_i r_x^i$ we have that

$$B(x, r_x) \subseteq B\left(x, r_x^i\right) \subseteq A_i \text{ for all } i \in [1:k] \implies B(x, r_x) \subseteq \bigcap_{i=1}^{i=k} A_i = A_*$$

as we wanted to show. □

Corollary 3.1. *Given a metric space $\mathcal{M} = (X, d)$ the following properties hold:*

1. $A = X$ *and* $A = \emptyset$ *are closed*
2. *Let $\mathcal{C}$ be a family of closed sets. Then*
$$C_* \equiv \bigcap_{C \in \mathcal{C}} C$$
is closed
3. *If $\{C_1, C_2, \dots, C_k\}$ are closed sets, then*
$$C^* \equiv \bigcup_{i=1}^{k} C_i$$

is closed too.

Proof. Use previous theorem, jointly with DeMorgan rules: for any family of sets $\mathcal{F} \subseteq \{A : A \subseteq X\}$ we have

$$\left(\bigcap_{A\in\mathcal{F}} A\right)^c = \bigcup_{A\in\mathcal{F}} A^c$$
$$\left(\bigcup_{A\in\mathcal{F}} A\right)^c = \bigcap_{A\in\mathcal{F}} A^c$$

□

Remark: Theorem 3.2 is very important, since it relates to the theory of topological spaces. These are spaces of the form $(X, \mathcal{O})$ where $\mathcal{O}$ is a family of subsets of X that satisfy properties (1) to (3) of Theorem 3.2. Such a family $\mathcal{O}$ is referred to as a topology for X. That is, in topological spaces, instead of deriving the notion of open and closed sets from a predefined metric, we start right from Theorem 3.2, treating the "list" of open sets $\mathcal{O}$ as a primitive, and not a result, of the space we study.

4. Continuous Functions

In our elementary calculus courses, he have learned that a continuous function is one that takes "close" points from its domain X and takes them to also "close" points of the image Y. We formalize this idea in the definition below.

Definition 4.1 (**Continuity**)**.** Given metric spaces $\mathcal{M}_X = (X, d_X)$ and $\mathcal{M}_Y = (X, d_Y)$ and $x \in X$, we say that a function $f : X \to Y$ is continuous at x $\iff$ the following rule holds: for all $\epsilon > 0$ we can find $\delta_{\epsilon,x} > 0$ such that

$$\text{if } y \text{is such that } d_X(x,y) < \delta_{\epsilon,x} \Longrightarrow d_Y(f(x), f(y)) < \epsilon.$$

or, equivalently: if $B_X(\cdot)$ and $B_Y(\cdot)$ denote open balls in spaces X and Y respectively, f is continuous at f $\iff$

$$\exists \delta_{\epsilon,x} > 0 : f[B_X(x, \delta_{\epsilon,x})] \subseteq B_Y[f(x), \epsilon]$$

when $f(\cdot)$ is continuous for all $x \in X$ we say "f is continuous".

4.1. **Examples.** **Example (1) :** Let $(X, d_X) = (Y, d_Y) = (\mathbb{R}, |x-y|)$ and $f(x) = 3x$. We want to show that f is continuous. Take $x \in \mathbb{R}$ and $\epsilon > 0$, and we need to find $\delta_{\epsilon.x} > 0$ such that if $|x-y| < \delta_{\epsilon,x} \Longrightarrow$ we also have $|f(x) - f(y)| < \epsilon$. Now, see that

$$|f(x) - f(y)| = 3|x-y|$$

so, if we take $\delta_{\epsilon,x} \equiv \frac{1}{3}\epsilon$, we have that if $y : |x-y| < \frac{1}{3}\epsilon$ then

$$|f(x) - f(y)| = 3|x-y| < 3\frac{1}{3}\epsilon = \epsilon$$

proving the desired result.

Example (2) : Same distance as before, but now $X = \mathbb{R}_+$ and $f(x) = \sqrt{x}$. Note that:

$$\left(\sqrt{x} - \sqrt{y}\right) = \frac{x-y}{\sqrt{x}+\sqrt{y}} \text{ for any } x > 0, y \geq 0$$

Therefore, by taking $\delta_{\epsilon,x} \equiv \epsilon\sqrt{x}$ we have that if $|x-y| < \epsilon\sqrt{x}$:

$$\left|\sqrt{x} - \sqrt{y}\right| = \frac{|x-y|}{\sqrt{x}+\sqrt{y}} \leq \frac{|x-y|}{\sqrt{x}} < \frac{\epsilon\sqrt{x}}{\sqrt{x}} = \epsilon$$

as we wanted to show.

Example (3) : Suppose $A \subseteq \mathbb{R}^n$ is a bounded Borel set (i.e. we can write an integral on it), and take $(X, d_X) = \mathcal{B}(A, \mathbb{R}) \cap \{f \text{ integrable}\}$, $(Y, d_Y) = (\mathbb{R}, |x-y|)$ and $T : \mathcal{B}(A, \mathbb{R}) \to \mathbb{R}$ defined as

$$T(f) = \int_A f(x)\,dx$$

i.e. T is the calculation of the integral, thought of as a function from the set of all functions to the real numbers. We will show that T is continuous (for all f integrable). See that for any pair of functions f, g :

$$|T(f) - T(g)| = \left|\int_A f(x)\,dx - \int_A g(x)\,dx\right| = \left|\int_A [f(x) - g(x)]\right| \leq \int_A |f(x) - g(x)|\,dx$$

$$\leq \int_A \left[\sup_{z \in A} |f(z) - g(z)|\right] dx = \int_A d_\infty(f,g)\,dx = \mu(A)\, d_\infty(f,g)$$

where $\mu(A) = \int_A 1dx$ is the volume of set A. Therefore, given $f \in X$, by taking $\delta_{\epsilon,f} = \frac{1}{\mu(A)}\epsilon$ we have that if $g : d_\infty(f,g) < \frac{\epsilon}{\mu(A)}$ then

$$|T(f) - T(g)| \leq \mu(A)\, d_\infty(f,g) < \mu(A)\frac{\epsilon}{\mu(A)} = \epsilon$$

and hence T is a continuous function.

4.2. **Continuity and Topology.**

Theorem 4.1. *Let (X, d_X) and (Y, d_Y) be two metric spaces, and $f : X \to Y$. Then f is a continuous function $\iff$ the following rule holds:*

$$\text{for any open set } A \subseteq Y \text{ we have that } f^{-1}(A) \subseteq X \text{ is also open} \tag{4.1}$$

i.e. the pre-image of open sets is always open.

Proof. (Direct $\Longrightarrow$) Suppose f is continuous, and we want to prove this rule. Take $A \subseteq Y$ an open set: we want to then show that $U = f^{-1}(A)$ is also an open set in (X, d_X): i.e. for any given $x \in f^{-1}(A)$ we need to find $r_x > 0$ such that $B_X(x, r_x) \subseteq f^{-1}(A)$ as well. Since A is open, we know there exists $r_{f(x)} > 0$ such that $B_Y\left(f(x), r_{f(x)}\right) \subseteq A$. Since f is continuous, take $\epsilon = r_{f(x)}$ and use it in the definition of continuity: this gives as a

$\delta_{\epsilon,x} = \delta_{r_{f(x)},x} > 0$ such that

$$f\left[B_X\left(x, \delta_{r_{f(x)},x}\right)\right] \subseteq B_Y\left[f(x), r_{f(x)}\right] \subseteq A$$

and hence

$$B_X\left(x, \delta_{r_{f(x)},x}\right) \subseteq f^{-1}(A)$$

Therefore, we finish the proof by taking $r_x \equiv \delta_{r_{f(x)},x}$ (which as you see, depends only on the point taken x)

(Converse $\Longleftarrow$) If the rule holds, we need to show f is continuous: i.e. given $x \in X$ and $\epsilon > 0$ we need to find $\delta_{\epsilon,x} > 0$ such that $f[B_X(x, \delta_{\epsilon,x})] \subseteq B_Y\left[f(x), r_{f(x)}\right]$. Let $A_{\epsilon,x} = B_Y(f(x), \epsilon)$. We know that open balls are open, and hence $A_{\epsilon,x}$ is a open subset of Y. Since the rule in (4.1) is assumed to hold, we know that the set $f^{-1}(A_{\epsilon,x}) = f^{-1}[B_Y(f(x), \epsilon)]$ is open as well, and since $x \in A_{\epsilon,x} \Longrightarrow$ there exist some radius $r_{\epsilon,x} > 0$ such that

$$B_X(x, r_{\epsilon,x}) \subseteq f^{-1}(A_{\epsilon,x}) \iff$$

$$\text{for all } y \in X \;:\; d_X(x,y) < r_{\epsilon,x} \Longrightarrow y \in f^{-1}(A_{\epsilon,x}) \iff f(y) \in B_Y(f(x), \epsilon)$$

i.e. $d_Y(f(x), f(y)) < \epsilon$. Hence, we prove the desired result by choosing $\delta_{\epsilon,x} = r_{\epsilon,x}$ (i.e. the radius involved with the fact that x is interior to the pre-image of $B_Y(f(x), \epsilon)$) □

In the theory of topological spaces, this theorem is in fact the definition of continuous function, since it is defined completely in terms of open sets alone. This theorem is also called the sign conservation theorem, when $Y = \mathbb{R}$, as we show in the following Corollary:

Corollary 4.1. *Let (X, d_X) a metric space and $f : X \to \mathbb{R}$ a continuous function. Then, the set*

$$A = \{x \in X \;:\; f(x) > 0\}$$

is an open set. Moreover, the set

$$C = \{x \in X \;:\; f(x) \geq 0\}$$

is closed

Proof. Simply take $U = (0, +\infty)$ (an open set) and note that $A = f^{-1}(U)$, and hence open. For C, see that we can write it as

$$C = \left[f^{-1}(-\infty, 0)\right]^c$$

the complement of an open set as well, and hence closed. □

A lot of important results of continuous functions can be easily shown using Theorem 4.1.

Theorem 4.2 (**Composite of continuous functions**). *Let $(X, d_X), (Y, d_Y), (Z, d_Z)$ three metric spaces, and $f : X \to Y$ and $g : Y \to Z$ two continuous functions. Then, the composite* *function $h : X \to Z$ defined as*

$$h(x) \equiv (gf)(x) = g[f(x)]$$

is also continuous.

Proof. Take an open set $A \subseteq Z$. We need to show that $h^{-1}(A) \subseteq X$ is also open. For that, see that

$$h^{-1}(A) = (gf)^{-1}(A) = f^{-1}\left[g^{-1}(A)\right]$$

Now, $g^{-1}(A) \subseteq Y$ is an open set since $g(\cdot)$ is continuous, and also is $f^{-1}\left[g^{-1}(A)\right]$, from the continuity of f, finishing the proof. □

5. Sequences

A **sequence** is simply a function $x : \mathbb{N} :\to X$. We usually write $x_n = x(n) \in X$ and $\{x_n\}_{n\in\mathbb{N}} = x(\cdot)$ instead, since we try to distinguish between sequences and other functions from more generic domains. One of the most important concepts related to sequences is that of **convergence**. We say that a sequence converges to an element x if the tail of the sequence is, eventually, arbitrarily close to x, and hence $x_n \approx x$ when n is "large". We formalize this idea in the definition below

Definition 5.1 (**Convergence**)**.** For a given metric space $\mathcal{M} = (X, d)$ and a sequence $\{x_n\}_{n\in\mathbb{N}}$ in X, we say that x_n converges to $x \in X \iff$ The following rule holds:

$$\text{for any } \epsilon > 0 \exists N_\epsilon \in \mathbb{N}\text{such that for all } n \geq N_\epsilon \text{ we have } d(x_n, x) < \epsilon$$

i.e. for any given "closeness" to x, we can find an index N_ϵ such that after $n = N_\epsilon$, the whole tail of the sequence is within ϵ of x. We usually write

$$x_n \to x \text{ or } \lim_{n\to\infty} x_n = x$$

5.1. **Examples.** **(1) :** Suppose $(X, d) = (\mathbb{R}, |x - y|)$. We show that $x_n = \frac{1}{n} \to 0$. For this, take $\epsilon > 0$ and we must find $N_\epsilon$such that if $n \geq N_\epsilon$ we must have $\left|\frac{1}{n} - 0\right| = \frac{1}{n} < \epsilon$. Hence, by taking

$$N_\epsilon = \left\lceil \frac{1}{\epsilon} \right\rceil \equiv \min\left\{n \in \mathbb{N} : n > \frac{1}{\epsilon}\right\}$$

we have that if $n \geq N_\epsilon \geq \frac{1}{\epsilon} \implies \frac{1}{n} < \epsilon$ as we wanted to show.

(2) : Suppose we take the same metric and sequence as before, but now $X = (0, 1)$. See that now, it is no longer true that $x_n \to 0$, since $0 \notin X$ (even though it is the intuitive limit of x_n). In some sense, x_n does not converge in the metric space $\mathcal{M} = ((0, 1), |x - y|)$ because X is "incomplete": it is missing some of the limits of sequences that should converge. This will relate to the concept of **complete metric spaces** later on.

(3) : Suppose now that $X = \mathcal{B}([0, 1], \mathbb{R})$ and $d = d_\infty$ and the sequence

$$f_n(x) = \frac{1}{n}x + \frac{2}{n} \text{ for all } x \in [0, 1]$$

we will show that $f_n \to \mathbf{0}(x) = 0$ for all x. See that

$$d_\infty\left(f_n(\cdot), \mathbf{0}(\cdot)\right) = \sup_{x\in[0,1]}\left[\frac{1}{n}x + \frac{2}{n} - 0\right] = \frac{1}{n}\left(\sup_{x\in[0,1]} x\right) + \frac{2}{n} = \frac{3}{n} \to 0$$

so we have shown the desired result: see that $N_\epsilon \equiv \left\lceil \frac{3}{\epsilon} \right\rceil$ satisfies the definition of convergence.

(4) : Same metric space as in (3), only that now the sequence we consider is

$$f_n(x) = x^n$$

What is the natural "candidate" for the limit of this sequence? One would be the so called "point wise limit" : i.e. the limit of $f_n(x)$ for every single x. See that

$$\text{for given } x \in [0,1],\ \lim_{n\to\infty} f_n(x) = \begin{cases} 0 & \text{if } x < 1 \\ 1 & \text{if } x = 1 \end{cases} \equiv f^*(x)$$

However, the fact that f converges to the limit function f^* for any given x does not imply that $f_n \to f^*$ as a sequence in $\mathcal{B}([0,1], \mathbb{R})$. To understand what we mean by this, let's calculate the distance between the sequence and f^*:

$$d_\infty\left(f_n(\cdot), f^*(\cdot)\right) = \sup_{x\in[0,1]} |x^n - f^*(x)| = \sup_{x\in[0,1]} \begin{cases} x^n & \text{if } x < 1 \\ 0 & \text{if } x = 1 \end{cases} = 1$$

But then, $d_\infty\left(f_n(\cdot), f^*(\cdot)\right) \nrightarrow 0$, which is the required definition for convergence.

What went wrong here? The problem is that there is a difference between convergence in the sup-norm d_∞ (usually referred to as uniform convergence) and convergence for all x (called point wise convergence). While is true that uniform convergence implies point wise convergence, this example shows that the converse is not true.

Proposition 5.1. *Let f_n be a sequence of bounded functions in $\mathcal{B}(A, \mathbb{R}^m)$ with $A \subseteq \mathbb{R}^n$, and $f \in \mathcal{B}(A, \mathbb{R}^m)$ such that*

$$f_n \to f \iff \sup_{x\in A} \mathbf{N}\left[f_n(x) - f(x)\right] \to 0 \text{ as } n \to \infty \tag{5.1}$$

i.e. f_n converges uniformly to f. Then, for all $x \in A$, the sequence $f_n(x) \in A$ converges to $f(x)$ (i.e. f_n converges point wise to f)

Proof. It simply follows from the fact that, for any $x \in A$:

$$0 \le \mathbf{N}\left[f_n(x) - f(x)\right] \le \sup_{y\in A} \mathbf{N}\left[f_n(y) - f(y)\right] \to 0 \tag{5.2}$$

where the last part comes from condition 5.1. Therefore $\mathbf{N}\left[f_n(x) - f(x)\right] \to 0 \iff f_n(x) \to f(x)$ for any $x \in A$, as we wanted to show. □

5.2. **Topology and Sequences.** Why do we care about sequences? It turns out they provide an easier characterization of most topological concepts: particularly relating to closed sets, and how we can check a set is closed without referring to the original definition. This is the purpose of Theorem 5.1

Theorem 5.1 (Sequential Definition of Closed sets)**.** *Let $\mathcal{M} = (X, d)$ be a metric space and $C \subseteq X$. Then, the set C is closed $\iff$ the following rule holds: if a sequence $\{x_n\}_{n\in\mathbb{N}}$ is such that (1) $x_n \in C$ for all n and (2) $\exists x : x_n \to x$,* ***then*** *we must also have that $x \in C$*

Theorem 5.1 tells us that showing that a set C is closed is equivalent to showing a "theorem", that states that whenever we have a convergent sequence of points inside the set, we must be able to show that the limit is inside the set as well.

Proof. ($\Longrightarrow$ Direct) Suppose C is closed, and that we have a sequence $x_n \in C$ for all n such that $x_n \to x$. We need to show that if that is the case, then we must have $x \in C$ as well. We will prove this **by contradiction:** suppose the rule is not true, so that $x \notin C \iff x \in C^c$, which is an open set, since C is closed by assumption. Therefore, there exist $r > 0$ such that $\forall y : d(x, y) < r \Longrightarrow y \notin C$. But since $x_n \to x$ we know that if we take $\epsilon = r$ we know that for any $n \geq N_{\epsilon=r}$ we have $d(x_n, x) < \epsilon = r$ and hence $x_n \notin C$ for all $n \geq N_r$. But this contradicts the assumption that $x_n \in C$ **for all** n, and hence x must be in C

($\Longleftarrow$ Reciprocal) Suppose that the rule is true, and we need to show that C is closed. Again, the proof will be **by contradiction:** suppose C is not closed, and hence $\exists \hat{x} \in \overline{C}$ such that $\hat{x} \notin C$. Now,

$$\hat{x} \in \overline{C} \iff \text{for all } \epsilon > 0,\ B(\hat{x}, \epsilon) \cap C \neq \emptyset$$

$$\iff \text{for all } \epsilon > 0 \text{ there exist } z_\epsilon \in C\text{: } d(\hat{x}, z_\epsilon) < \epsilon \tag{5.3}$$

What we will do now is create a sequence x_n such that $x_n \in C$ for all n and $x_n \to \hat{x}$: simply substitute in 5.3 $\epsilon = \frac{1}{n}$ and define $x_n \equiv z_{\epsilon=\frac{1}{n}}$. See that by definition of $z_{\frac{1}{n}}$ we must have $x_n \in C$. Moreover,

$$0 \leq d(x_n, \hat{x}) = d\left(z_{\frac{1}{n}}, \hat{x}\right) < \frac{1}{n} \to 0$$

and hence $d(x_n, \hat{x}) \to 0 \iff x_n \to \hat{x}$. And here lies the contradiction: if the sequence rule is true, then since $x_n \in C$ for all n and moreover, $x_n \to \hat{x}$, we must therefore have $\hat{x} \in C$. But we assumed that $\hat{x} \notin C$, reaching the desired contradiction. $\square$

Example 5.1 (**Set of non-decreasing functions**)**.** Take $(X, d) = \mathcal{B}(A, \mathbb{R})$ and the subset defined as

$$C = \{f \in X : \text{if } x \geq y \Longrightarrow f(x) \geq f(y) \text{ for all } x, y \in A\}$$

i.e. C is the set of non-decreasing functions. We will show that C is a closed set of functions, and we will use Theorem 5.1 to show it: i.e. we need to show that the following "Theorem" is true:

$$\text{if } \{f_n\}_{n\in\mathbb{N}} \text{ is such that } f_n \in C \text{for all } n \text{ and } f_n \to f \Longrightarrow f \in C \text{ as well.}$$

Suppose, by contradiction, that it was not true: hence, there would exist a sequence f_n of non-decreasing functions, converging to a function f that has the following property: $\exists \hat{x}, \hat{y} \in A$ with $\hat{x} > \hat{y}$ but $f(\hat{x}) < f(\hat{y})$. If this was the case, given this pair of points, define the sequence

$$a_n \equiv f_n(\hat{x}) - f_n(\hat{y})$$

Since $f_n \in C$ by assumption, and $\hat{x} > \hat{y}$ we must have $a_n \geq 0$ for all n. Moreover, from Proposition 5.1 we know that

$$a_n \to \lim_{n\to\infty} [f_n(\hat{x}) - f_n(\hat{y})] = f(\hat{x}) - f(\hat{y}) < 0$$

which is not possible, since $a$which is a closed set, and hence a should be in it, and not in $(-\infty, 0)$, reaching hence a contradiction.

Example 5.2 (Set of continuous functions). Same space as in Example 5.1, but now

$$\mathcal{C}(A, \mathbb{R}^m) \equiv \{f \in X : f \text{ is continuous for all } x \in A\}$$

i.e. the set of all bounded, continuous functions. We want to show that this set is closed: for that, take a sequence $f_n \in \mathcal{C}(A, \mathbb{R}^m)$ of continuous functions, and take the limit $f \in X$. We need to show that f is continuous. For that, take any $x \in A$ and $\epsilon > 0$, and we need to find $\delta_{\epsilon,x} > 0$ such that if $y : \mathbf{N}(x-y) < \delta_{\epsilon,x} \implies \mathbf{N}[f(x) - f(y)] < \epsilon$. See that for any $n \in \mathbb{N}$:

$$\mathbf{N}[f(x) - f(y)] \leq \mathbf{N}[f(x) - f_n(x) + f_n(x) - f_n(y) + f_n(y) - f(y)] \leq$$

(5.4)
$$\mathbf{N}[f_n(x) - f(x)] + \mathbf{N}[f_n(x) - f_n(y)] + \mathbf{N}[f_n(y) - f(y)] \leq \mathbf{N}[f_n(x) - f_n(y)] + 2\sup_{z\in A} \mathbf{N}[f_n(z) - f(z)]$$

Now, since $f_n \to f$ we know that there exists $n = n_{\frac{\epsilon}{4}}$ such that if $n \geq n_{\frac{\epsilon}{4}} \implies d_\infty(f_n, f) = \sup_{z\in A} \mathbf{N}[f_n(z) - f(z)] < \frac{\epsilon}{2}$. Since inequality 5.4 is true for any n, is particularly true for $n = n_{\frac{\epsilon}{4}}$, and hence

$$\mathbf{N}[f(x) - f(y)] < \mathbf{N}\left[f_{n_{\frac{\epsilon}{4}}}(x) - f_{n_{\frac{\epsilon}{4}}}(y)\right] + 2\frac{\epsilon}{4} = \mathbf{N}\left[f_{n_{\frac{\epsilon}{4}}}(x) - f_{n_{\frac{\epsilon}{4}}}(y)\right] + \frac{\epsilon}{2}$$

If $y : \mathbf{N}(x-y) < \hat{\delta}_{\frac{\epsilon}{2},x} \implies \mathbf{N}\left[f_{n_{\frac{\epsilon}{4}}}(x) - f_{n_{\frac{\epsilon}{4}}}(y)\right] < \frac{\epsilon}{2}$. Hence,

$$\mathbf{N}[f(x) - f(y)] < \mathbf{N}\left[f_{n_{\frac{\epsilon}{4}}}(x) - f_{n_{\frac{\epsilon}{4}}}(y)\right] + \frac{\epsilon}{2} < \epsilon$$

And since the function $f_{n_{\frac{\epsilon}{4}}}(\cdot)$ is also continuous, if we take $\delta_{\epsilon,x} \equiv \hat{\delta}_{\frac{\epsilon}{2},x}$ where $\hat{\delta}$ is the one that comes from the continuity of $f_{n_{\frac{\epsilon}{4}}}(\cdot)$. Then,

so $\delta_{\epsilon,x} = \hat{\delta}_{\frac{\epsilon}{2},x}$ works to show that f is continuous at x, as we wanted to show

5.3. **Sequences and continuous functions.** Probably one of the most useful theorems of calculus is that continuous functions commute limits: that is, if f is continuous

$$\lim_{n\to\infty} f(x_n) = f\left(\lim_{n\to\infty} x_n\right)$$

Turns out that this is in fact a characterization of continuity in general metric spaces, as we see in the following theorem.

Theorem 5.2. *Let $\mathcal{M}_X = (X, d_X)$ and $\mathcal{M}_Y = (Y, d_Y)$ two metric spaces, and $f : X \to Y$. Then*

$$f \text{ is continuous at } x \in X \iff \text{the following rule holds: if } x_n \to x \text{then } f(x_n) \to f(x)$$

Proof. ($\Longrightarrow$ Direct) Suppose f is continuous: we want to show that if $x_n \to x$ then $f(x_n) \to f(x)$. For that, we need to find N_ϵ^f such that if $n \geq N_\epsilon^f$ then $d_Y(f(x_n), f(x)) < \epsilon$. Since f is a continuous function, we know that $\exists \delta_{\epsilon,x} > 0 : d_X(x, y) < \delta_{\epsilon,x}$ implies $d(f(x), f(y)) < \epsilon$. Since $x_n \to x$, simply take $N_\epsilon^f = N_{\epsilon=\delta_{\epsilon,x}}$ from the definition of convergence of x_n, and continuity proves the rest.

($\Longleftarrow$Converse) By contradiction: suppose the rule is true, but f is not continuous. That is:

$$\exists \bar{\epsilon} > 0 : \forall \delta > 0 \text{ there exist } y_\delta \in B(x, \delta) \text{ such that } d(x, y_\delta) \geq \bar{\epsilon} \tag{5.5}$$

Now, we will define a sequence $x_n \to x$ such that $f(x_n) \not\to f(x)$, which would contradict the fact that f was not continuous. Define

$$x_n \equiv y_{\delta=\frac{1}{n}}$$

i.e. take $\delta = \frac{1}{n}$ and let x_n be the $y_\delta = y_{\frac{1}{n}}$ of condition 5.5. Now, that same condition implies that

$$x_n = y_{\frac{1}{n}} \in B\left(x, \frac{1}{n}\right) \iff 0 \leq d(x_n, x) < \frac{1}{n} \to 0$$

and hence $x_n \to x$. However, we have

$$d(f(x_n), x) \geq \bar{\epsilon} > 0 \text{ for all } n \in \mathbb{N}$$

violating the proposed rule. □

6. Compact Sets

In $\mathbb{R}^n$, we use the definition:

Compact set =Closed and bounded

It is a concept of great importance, since most of the most used theorems in economics involve compact sets, most notably the **Weierstrass theorem**, that states that any continuous function defined on a compact set has maximum and minimum values. Turns out that the extension of the concept of compactness for generic metric spaces is NOT that one of a set that is closed and bounded, but is rather a little more involved.

Definition 6.1 (**Open Cover**)**.** Given a metric space (X, d) and a set $C \subseteq X$, we say that a family $\mathcal{O} \subseteq \{A : A \subseteq X\}$ is an open cover of $C \iff$ (1) For all $A \in \mathcal{O}$, A is open. (2) $C \subseteq \bigcup_{A \in \mathcal{O}} A$

Example 6.1. Let $X = \mathbb{R}$ and d the usual metric. Take the set $C = (0, +\infty)$. One possible open cover is $\mathcal{O}_1 = \{\mathbb{R}\}$, since it is an open set. Also is the family

$$\mathcal{O}_2 = \{A_n = (0, n) \text{ for } n \in \mathbb{N}\}$$

since it is clear that $C = \bigcup_{n=1}^{\infty} A_n =$.

Definition 6.2 (**Compact set**)**.** Let (X, d) be a metric space and $C \subseteq X$. We say that C is a **compact set** $\iff$ for all open cover $\mathcal{O}$ of C, there exist a **finite number of sets** $\{A_1, A_2, ..., A_k\} \subseteq \mathcal{O}$ such that

$$C \subseteq \bigcup_{i=1}^{i=k} A_i$$

i.e. for any open cover, we only need a finite number of elements of it to cover the whole set.

Note that this definition asks as that this rule has to be true for ANY open cover. For example, for the set C, we know that $\mathcal{O}_1$ is an open cover that satisfies this property (since it only consist of one set). However, $\mathcal{O}_2$ does not satisfy this property, since for any finite number of sets $\{A_{n_1}, A_{n_2}, ..., A_{n_k}\} \subset \mathcal{O}$ we have

$$\bigcup_{i=1}^{i=k} A_{n_i} = \bigcup_{i=1}^{i=k} (0, n_i) = \left(0, \max_{i \in [1:k]} n_i\right) \supset C$$

and hence $\mathcal{O}_2$ violates this property, making C not compact.

Notice also that the definition of compactness only talks about open sets, and does not mention distances at all. This therefore generalizes to the study of topological spaces, as we mentioned above. As you might have imagined, checking that a compact set is so with the given definition is hard, since one has to consider a huge class of objects to satisfy this property. Luckily, there is a simpler way to check compactness in general.

Definition 6.3 (Subsequences)**.** Given a sequence $\{x_n\}_{n \in \mathbb{N}}$ we say that $\{y_n\}_{n \in \mathbb{N}}$ is a subsequence of $\{x_n\}_{n \in \mathbb{N}} \iff$ there exist an increasing sequence $k_n \in \mathbb{N}$ such that $y_n = x_{k_n}$ for all $n \in \mathbb{N}$

Example 6.2. Let $x_n = \frac{1}{n} + (-1)^n$. The sequence $y_n = 1 + \frac{1}{2n}$ for all n is a subsequence of $\{x_n\}$, since

$$x_{2n} = \frac{1}{2n} + (-1)^{2n} = 1 + \frac{1}{2n} = y_n$$

i.e. y_n is the composition of x_n with $k_n = 2n$. Note that even if x_n does not have a limit, **we can find subsequences of it that do have limits**: for example

$$y_n \equiv x_{2n} = 1 + \frac{1}{2n} \to 1$$

and

$$z_n \equiv x_{2n+1} = \frac{1}{2n+1} - 1 \to -1$$

It turns out that a compact set is a set such that for any sequence we have inside it, we can always find a subsequence that convergence somewhere inside.

Theorem 6.1. *Let $\mathcal{M} = (X, d)$ be a metric space and $C \subseteq X$. Then, C is compact $\iff$ the following rule holds:*

$$\textit{for any}\, \{x_n\}_{n\in\mathbb{N}} \quad : x_n \in C \;\textit{there exists}\; y_n = \{x_{k_n}\} \textit{and}\; y \in C \textit{such that}\; y_n \to y$$

i.e. any sequence in C has a convergent subsequence x_{k_n}.

Exercise 6.1. Show that if $C \subseteq X$ is compact, then it is also closed. Show also that if $D \subseteq C$ is closed and C is compact, then D is compact as well.

In Problem set 2 you will show that when $(X, d) = (\mathbb{R}, |x - y|)$ then we have

$$C \text{ is compact } \iff \text{is bounded and closed}$$

The following theorem is of extreme importance in the study of topological spaces: it says that the image of compact sets through continuous functions is compact as well

Theorem 6.2. *Let (X, d_X) and (Y, d_Y) two metric spaces, and $C \subseteq X$ a compact subset. If a function $f : X \to Y$ is continuous, then $f(C) = \{y \in Y : \exists x \in X \text{ with } f(x) = y\} \subseteq Y$ is a compact set as well.*

Proof. Take an open cover $\mathcal{O}$ of $f(X)$. We need to find a finite number of sets $\{A_1, A_2, ..., A_k\} \subseteq \mathcal{O}$ such that $f(X) \subseteq \bigcup_{i=1}^{i=k} A_i$. Now, if $\mathcal{O}$ is an open cover of $f(C)$, we have that

$$f(C) \subseteq \bigcup_{A\in\mathcal{O}} A \implies f^{-1}[f(C)] \subseteq f^{-1}\left[\bigcup_{A\in\mathcal{O}} A\right] \implies$$

$$C \subseteq \bigcup_{A\in\mathcal{O}} f^{-1}(A)$$

Since f is continuous and $A \in \mathcal{O}$ is open for all elements of $\mathcal{O}$, we have that the family

$$\hat{\mathcal{O}} \equiv \{B \subseteq X : B = f^{-1}(A) \text{ for some } A \in \mathcal{O}\}$$

is an open cover of C. Since C is compact, we know that there exist a finite number of sets $\{B_1, B_2, ..., B_k\} = \{f^{-1}(A_1), f^{-1}(A_2), ..., f^{-1}(A_k)\}$ such that

$$C \subseteq \bigcup_{i=1}^{i=k} f^{-1}(A_i) \implies f(C) \subseteq \textstyle\bigcup_{i=1}^{i=k} A_i$$

as we wanted to show. ☐

Armed with this, we can easily shown the famous Weierstrass theorem

Theorem 6.3 (**Weierstrass**). *Let (X, d) be a metric space, $C \subseteq X$ a compact set and $f : X \to \mathbb{R}$ a continuous function. Then,*

$$\exists \overline{x} \in \underset{x \in C}{argmax}\ f(x) and\ \underline{x} \in \underset{x \in C}{argmin} f(x)$$

i.e. the maximum and minimum of f over C are well defined objects.

Proof. Let

$$Z = \{y \in \mathbb{R} : \exists x \in X \text{ such that } f(x) = y\} = f(C)$$

We have $f(C) \subset \mathbb{R}$ is a compact set from the previous theorem, hence it is bounded and closed. Let $\overline{f} = \sup[f(X)]$ (which exists from $f(C)$ being bounded), and we will show that $\exists x \in C : f(x) = \overline{f}$. Using the definition of supremum, we know that for all $\epsilon > 0$ there exist $y_\epsilon \in f(C)$ such that $y > \overline{f} - \epsilon$. Therefore, let $\epsilon = \frac{1}{n}$ and define the sequence $z_n = y_{\frac{1}{n}}$ from the previous condition. We clearly have that $z_n \in f(C)$ for all n, and that $z_n \to \overline{f}$. Since $f(C)$ is compact $\implies f(C)$ is also closed, and hence $\overline{f} \in f(C) \iff \exists \overline{x} \in X : f(\overline{x}) = \overline{f}$. The proof for the minimum is analogous. $\square$