MATH 521, WEEKS 11 & 12: Functions, Continuity, Uniform Continuity

1 Functions

In the course so far, we have focused our attention on the details underlying the rational and real number systems, general set structures—especially metric spaces and topology—and a somewhat rigorous treatment of the convergence of sequences in metric spaces.

One thing that should seem notable absent in our list of topics of study is *functions*. This should feel strange! After all, functions have been the basis for our studies in the majority of math courses we have take, from calculus, to differential equations, to algebra. In calculus, in particular, we dove headfirst into such topics as continuity, differentiability, and integration. So what does all this additional rigor allow us to do?

The answer depends on who you ask. Certainly, there is value in rigor for its own sake—after all, we like to know what we are doing is justified—and there is a segment of mathematics that values this endeavor in rigor for its own sake very highly. But we should not lose sight of the big picture, either. In many applications (e.g. imagine/signal processing, study of differential equations, quantum information theory, etc.), generalizing to a more general metric spaces than $(\mathbb{R}, d)$ will prove critical to understanding the problems at play. The rigor we have gone through to generalize the notions we are familiar with in $\mathbb{R}$ to general metric spaces will allow us to apply results to areas of applications which deal with systems other than simple $\mathbb{R}$.

To get started, we re-iterate the basic functional notions with which we are already familiar (in $\mathbb{R}$).

Definition 1.1. *Consider sets X and Y. We say that $f : X \mapsto Y$ is a **function** from X to Y if for every $x \in X$ we may associated a unique element $y \in Y$. We assign this element the notation $y = f(x)$.*

We are most familiar with functions of real numbers (i.e. $X \subseteq \mathbb{R}$, $Y \subseteq \mathbb{R}$). We need not, however, define functions only on this set. For an extreme example, we could define the sets $X = \{\text{square, circle, rectangle}\}$ and $Y = \{\text{red, yellow, green}\}$ and define a function f from X to Y as $f(\text{square}) = \text{red}$, $f(\text{circle}) = \text{yellow}$, and $f(\text{rectangle}) = \text{green}$. In general, these sets really can be *anything*.

We have the following additional structure.

Definition 1.2. *Consider sets X and Y and a function $f : X \mapsto Y$.*

1. *X is called the **domain** of f.*

2. *We define the **range** or **codomain** of f to be the set $R(f) = \{y \in Y \mid y = f(x) \text{ for some } x \in X\}$.*

The domain of a function f is sometimes denoted $D(f)$.

The definitions are as straight-forward and familiar as they seem. In $\mathbb{R}$ we must simply remember to exclude those values from the domain (i.e. the x variable) or the range (i.e. the y variable) which are not associated a value in the other variable through f. For instance, for the function $f(x) = \frac{1}{x}$, we may not associate any y value to $x = 0$, since we may not divide by zero. It follows that the domain is $D(f) = \mathbb{R} \setminus \{0\}$. Similarly, there is no x value so that $f(x) = 0$ so that the range is $R(f) = \mathbb{R} \setminus \{0\}$.

We may also restrict our attention to subsets of the domain and range.

Definition 1.2 (con't)

3. For a subset $A \subseteq X$ we define the **image** of A under f to be the set $f(A) = \{y \in Y \mid y = f(x) \text{ for some } x \in A\}$.

4. For a subset $B \subseteq Y$ we define the **pre-image** of B to be $f^{-1}(B) = \{x \in X \mid f(x) = y \in B\}$.

The image is as straight-forward as it sounds. For example, if we consider the function $f(x) = x^2$ and consider $A = [1, 2]$ we have that $f(A) = [1, 4]$.

It is important to recognize, however, that the **pre-image is not the same as the inverse**. The pre-image captures *all* points which map to a particular $y \in Y$, not just a single one. For instance, for $f(x) = x^2$, we have $f^{-1}(1) = \{-1, 1\}$ and, for $B = [1, 4]$, we have $f^{-1}(B) = [-2, -1] \cup [1, 2]$.

It is also important to note that, for a general pre-image $f^{-1}(B)$, we have $f(f^{-1}(B)) \subseteq B$, not $f(f^{-1}(B)) = B$. For instance, if we consider $f(x) = x^2$ and $B = (-\infty, 1]$, we have $f^{-1}(B) = [-1, 1]$ but $f(f^{-1}(B)) = [0, 1] \subseteq (-\infty, 1]$. (Note, however, that the inclusion can only be strict if there is a point $b \in B$ for which there does not exist an $a \in X$ such that $f(a) = b$, i.e. $f^{-1}(b) = \emptyset$.)

We may use these concepts to classify functions in the following ways.

Definition 1.2 (con't) A function $f : X \mapsto Y$ is said to be:

5. **injective** or **one-to-one** if $f^{-1}(y)$ is a singleton for every $y \in R(f)$.

6. **surjective** or **onto** if $R(f) = Y$.

7. **bijective** if it is both injective and surjective.

In other words, a function is injective if every $x \in X$ maps to exactly one element $y \in Y$, and a function is surjective if every element $y \in Y$ is mapped to by f. An equivalent statement of injectivity is

$$f(x_1) = f(x_2) \quad \Longleftrightarrow \quad x_1 = x_2, \quad \text{for all } \ x_1, x_2 \in X.$$

Injectivity is important for defining inverses. For instance, we can see that the function $f(x) = x^2$ is not injective over the domain $X = \mathbb{R}$ since $f(-1) = f(1) = 1$. However, on the domain $X = [0, \infty)$ we have that each point in the domain corresponds to a unique point in the range $R(f) = [0, \infty)$. We may therefore define the inverse $f^{-1}(x) = \sqrt{x}$. We may, however, define the pre-image $f^{-1}(B)$ for any set $B \subseteq Y$, regardless of whether f is injective or not. Again, we emphasize that the two are not equivalent in general!

The operations we define for functions (e.g. additions, multiplication, etc.) will depend on the chosen domain and range of the function. However, simply interpreting functions as mappings between sets, there is one operation which works in the general setting: *composition*.

Definition 1.3. *Let X, Y, and Z be sets and suppose $f : X \mapsto Y$ and $g : Y \mapsto Z$. Then the* ***composition*** *$(g \circ f) : X \mapsto Z$ is defined as the mapping which assigns to each $x \in X$ the value $g(f(x)) \in Z$.*

Composition is familiar when $X = Y = Z = \mathbb{R}$. For instance, if we have $f(x) = x^2$ and $g(x) = 1 - x$, then we have $g(f(x)) = 1 - x^2$ and $g(f(x)) = (1 - x)^2$. For general sets, we will have to be careful that we apply the composition in the correct order.

2 Continuous Functions

If we are simply speaking about functions between sets, with no additional structure, we may content ourselves with the discussion was have already had. If we are discussing sets together with a metric, however, we have a notion of distance, a notion of set topology, and a notion of convergence. This will allow us to say a great deal more! (And prolong the course by some weeks.)

We should notice before we begin that, since functions represent a mapping from one set into another and that the sets need not be the same, when we extend to mappings from one metric space into another, the *metrics* need not be the same, either. That is to say, for $f : X \mapsto Y$, we allow the distance between points in the domain X to be measured differently than points in the codomain Y. Unless otherwise indicated, we will denote our metric spaces (X, d_X) and (Y, d_Y) so that d_X is the metric on X and d_Y is the metric on Y.

We now extend our discussion of convergence into general metric spaces to convergence of functions.

Definition 2.1. *Suppose (X, d_X) and (Y, d_Y) are metric spaces and $f : X \mapsto Y$. Suppose $a \in X$ is a limit point in the topology of X. We say $f(x)$ has the limit $b \in Y$ as $x \to a$ if, for every $\epsilon > 0$, there is a $\delta > 0$ such that, if $0 < d_X(x, a) < \delta$ for $x \in X$, then $d_Y(f(x), b) < \epsilon$.*

Notationally, we write $\lim_{x \to a} f(x) = b$.

Note that we may have $\lim_{x \to a} f(x) \neq f(\lim_{x \to a} x) = f(a)$. That is to say, the value of the limit need not coincide with the value at the limiting point. Said another way, it is not necessarily true that we may interchange the order of limits and functions. They do not commute in general.

We now define the following.

Definition 2.2. *Suppose (X, d_X) and (Y, d_Y) are metric spaces and $f : X \mapsto Y$. We say f is **continuous** at $a \in X$ if, for every $\epsilon > 0$, there is a $\delta > 0$ so that $d_X(x, a) < \delta$ implies $d_Y(f(x), f(a)) < \epsilon$. f is **continuous** on $A \subseteq X$ if f is continuous at every $a \in A$.*

Notice that a consequence of f being continuous at a is that we may pass the limit operation through the function itself, i.e. we have

$$\lim_{x \to a} f(x) = f\left(\lim_{x \to a} x\right) = f(a).$$

This is a consequence of not requiring that $d_X(x, a) > 0$, as was required in Definition 2.1. Consequently, for continuous points, it is required that *every* sequence $\{x_n\}$ which converges to a satisfies the property that $\lim_{n \to \infty} f(x_n) = f(a)$.

We have the following result regarding compositions of continuous functions.

Theorem 2.1. *Let (X, d_X), (Y, d_Y), and (Z, d_Z) be metric spaces. Suppose that $f : X \mapsto Y$, $g : Y \mapsto Z$, and that f is continuous at a, and g is continuous at $b = f(a)$. Then the function $h : X \mapsto Z$ defined by $h = g \circ f$ is continuous at a.*

Proof. We need to prove that, for every $\epsilon > 0$, there is a $\delta > 0$ so that $d_X(x,a) < \delta$ implies $d_Z(h(x), h(a)) = d_Z(g(f(x)), g(f(a)) < \epsilon$.

Take $\epsilon > 0$ fixed. Since g is continuous at $f(a) \in Y$, and f maps X to Y, we have that there is a $\eta > 0$ such that $d_Y(f(x), f(a)) < \eta$ implies that $d_Z(g(f(x)), g(f(a)) < \epsilon$. Now apply the continuity of $f(x)$ at a. For $\eta > 0$, there is a $\delta > 0$ such that $d_X(x,a) < \delta$ implies $d_Y(f(x), f(a)) < \eta$. Altogether, we have that, for $\epsilon > 0$, there is an $\eta > 0$ and a $\delta > 0$ so that

$$d_X(x,a) < \delta \quad \Longrightarrow \quad d_Y(f(x), f(a)) < \eta \quad \Longrightarrow \quad d_Z(g(f(x)), g(f(a))) < \epsilon.$$

Since this implies $h = g \circ f$ is continuous at a, we are done. ☐

Example 1: Prove that $f(x) = x^2$ is continuous on the whole real number line $\mathbb{R}$.

Proof: We need to show that, at any point $a \in \mathbb{R}$, we can bound $|f(x) - f(a)|$ by an arbitrarily small value, so long as we take $|x - a|$ to be sufficiently small. The trick will be showing that $|f(x) - f(a)|$ may be bounded by $|x - a|$, and then relating this bound to our arbitrarily chosen $\epsilon > 0$.

For this example, we start by taking $\epsilon > 0$. We have

$$|f(x) - f(a)| = |x^2 - a^2| = |x + a| \cdot |x - a|.$$

We notice that we can bound $|x - a|$ by assumption. That is, we are allowed to assume that $|x - a| < \delta$ for some $\delta > 0$. What about $|x + a|$? Well, we might notice that

$$|x + a| = |x - a + 2a| \leq |x - a| + 2|a| < \delta + 2|a|.$$

It follows that we have

$$|x + a| \cdot |x - a| < (\delta + 2|a|)\delta = \epsilon.$$

We can easily determine by the quadratic formula that the value we need is

$$\begin{aligned} \delta &= \frac{-2|a| \pm \sqrt{4|a|^2 + 4\epsilon}}{2} \\ &= -|a| \pm \sqrt{|a|^2 + \epsilon}. \end{aligned}$$

Omitting the negative root, and noting that the parabola opens up, we have, for any $\epsilon > 0$, if we take $\delta < -|a| + \sqrt{|a|^2 + \epsilon}$, then $|x - a| < \delta$ implies that $|x^2 - a^2| < \epsilon$.

Example 2: Prove that $f(x) = x^3$ is continuous on the whole real number line $\mathbb{R}$.

Proof: We will be a little less precise, but the idea will be the same as for x^2. Note that we know the continuous of x (trivial!) and x^2 (just shown) so that the problem is really just a matter of showing the continuity of the product $x^3 = x \cdot x^2$.

Take $\epsilon > 0$. Note that, because x and x^2 are continuous, we have that there are $\delta_1 > 0$ and $\delta_2 > 0$ so that

$$\begin{aligned} |x-a| < \delta_1 \quad &\Longrightarrow \quad |x-a| < \epsilon' \qquad (\text{take } \delta_1 = \epsilon'!) \\ |x-a| < \delta_2 \quad &\Longrightarrow \quad |x^2 - a^2| < \epsilon' \end{aligned}$$

where we will leave the precise relationship between $\epsilon' > 0$ and $\epsilon > 0$ undetered for the moment. We now want to bound $|x^3 - a^3|$. We will use a trick from Theorem 3.3 of Rudin. Notice that, for any functions $g(x)$ and $h(x)$ and values b and c, we have

$$g(x)h(x) - bc = (g(x) - b)(h(x) - c) + b(h(x) - c) + c(g(x) - b).$$

For our case, we want to take $g(x) = x$, $h(x) = x^2$, $b = a$, and $c = a^2$. This gives

$$x^3 - a^3 = (x-a)(x^2 - a^2) + a(x^2 - a^2) + a^2(x - a).$$

We now bound this. We have

$$\begin{aligned} |x^3 - a^3| &= |x-a| \cdot |x^2 - a^2| + |a| \cdot |x^2 - a^2| + a^2 \cdot |x - a| \\ &< (\epsilon')^2 + |a|\epsilon' + a^2\epsilon' \\ &= \epsilon'(\epsilon' + |a| + a^2). \end{aligned}$$

In order bound this by $\epsilon < 0$, we should relate ϵ' to ϵ. This guarantees we can pick ϵ' small enough so that the whole term is bounded. We can do this explicitly by the quadratic formula, but it is messy, so we will instead simply notice that, regardless of the value of $a \in \mathbb{R}$, we have that the first term can be taken close to zero, and the second term is approaches $|a| + a^2$. Clearly, the product goes to zero (although the rate depends on the choice of $a \in \mathbb{R}$!).

Let's suppose that the second term is "close" to its eventual limit, but not quite there. For instance, let's pick the value $1 + |a| + a^2$, which will be true for any $0 < \epsilon' < 1$. This is a worse bound than the eventual limit of

$|a| + a^2$, but it still does not affect the convergence to zero. So let's bound there! We can see that, if we take $\epsilon' < \min\left\{1, \dfrac{\epsilon}{1+|a|+a^2}\right\}$ we have

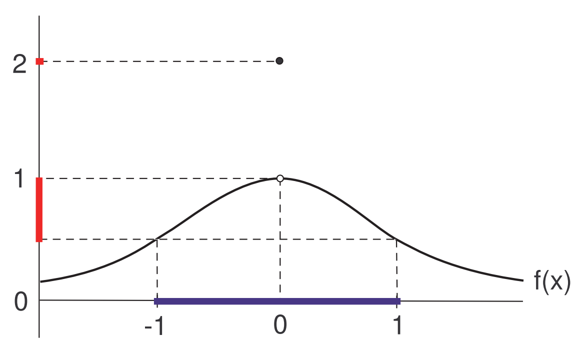
$$|x^3 - a^3| < \epsilon'(\epsilon' + |a| + a^2) < \epsilon'(1 + |a| + a^2) < \epsilon.$$

Since $\epsilon > 0$ and $a \in \mathbb{R}$ were chosen arbitrarily, we have shown $f(x) = x^3$ is continuous at any value $a \in \mathbb{R}$.

3 Continuity and Topology

Since functions associate one metric space with another, a natural question to ask is how the topology in one is transferred to the other through f. For instance, consider the following function ($f : \mathbb{R} \mapsto \mathbb{R}$):

$$f(x) = \begin{cases} \dfrac{1}{1+x^2}, & \text{for } x \neq 0 \\ 2, & \text{for } x = 0. \end{cases}$$



Now consider mapping different sets through f. We can see in Figure 3 that:

1. $A = (-1, 1) \implies f(A) = (\frac{1}{2}, 1) \cup \{2\}$
2. $A = [-1, 1] \implies f(A) = [\frac{1}{2}, 1) \cup \{2\}$

We can see that crucial properties such as sets being open, closed, connected, and compact are not maintained through the mapping f. Perhaps, however, we should consider starting in the second metric space and mapping back to the first through the pre-image. We have that:

3. $B = (\frac{1}{2}, 1) \implies f^{-1}(B) = (-1, 0) \cup (0, 1)$

4. $B = [\frac{1}{2}, 1] \implies f^{-1}(B) = [-1, 0) \cup (0, 1]$

Again, we cannot count on crucial topological properties such as sets being open, closed, connected, or compact being maintained when we consider the pre-image f^{-1}.

Of course, we have performed a sleight of hand, since we have manufactured the function f to have an obvious discontinuity at $x = 0$. It is natural to ask how topological properties are changed through *continuous* mappings f.

The answer is more subtle than it appears. It is not generally true, for instance, that open or closed sets $A \subseteq X$ always yield open or closed sets $f(A) \subseteq Y$ (homework!). We are able to say the following about pre-images, however.

Theorem 3.1. *Let (X, d_X) and (Y, d_Y) be metric spaces and consider a function $f : X \mapsto Y$. Then the following are equivalent:*

(a) f is continuous on X;

(b) If $B \subseteq Y$ is open in (Y, d_Y), then $f^{-1}(B)$ is open in (X, d_X);

(c) If $B \subseteq Y$ is closed in (Y, d_Y), then $f^{-1}(B)$ is closed in (X, d_X).

Proof. We will show that $(a) \iff (b)$ and $(b) \iff (c)$.

$(a) \implies (b)$: Suppose that $f : X \mapsto Y$ is continuous and $B \subseteq Y$ is an open set. Take $a \in X$ and $b \in B$ such that $b = f(a)$. By assumption, we have that b is an interior point of B so that there is an $r > 0$ such that $d(b, y) < r$ implies $y \in B$. However, since f is continuous on X, we have that there is a $\delta > 0$ so that $d_X(a, x) < \delta$ implies $d_Y(f(a), f(x)) = d_Y(b, f(x)) < r$. It follows that $d_X(a, x) < \delta$ implies $x \in f^{-1}(B_r(b)) \subseteq f^{-1}(B)$. Therefore, a is an interior point of $f^{-1}(B)$, and since a was chosen arbitrarily, it follows that $f^{-1}(B)$ is open.

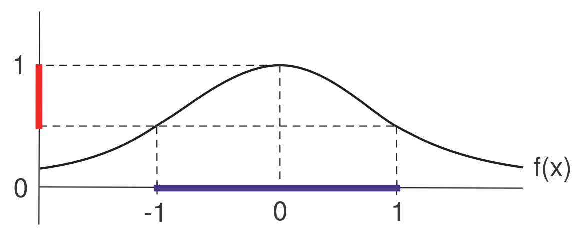
$(b) \implies (a)$: Suppose that $f : X \mapsto Y$ and $B \subseteq Y$ being open implies that $f^{-1}(B) \subseteq X$ is open. Take $\epsilon > 0$, $a \in X$, and $b \in B$ so that $f(a) = b$, and $B_\epsilon(b) \subseteq B$. The ball $B_\epsilon(b)$ is open, which implies by assumption that $f^{-1}(B_\epsilon(b))$ is open. It follows that there is a $\delta > 0$ so that $B_\delta(a) \subseteq f^{-1}(B_\epsilon(b))$. Expanding the definitions of the balls, we have that $d_X(a, x) < \delta$ implies that $d_Y(b, f(x)) < \epsilon$, and we are done.

$(b) \iff (c)$: Suppose B is closed in (Y, d_Y) but $f^{-1}(B)$ is not closed in (X, d_X). It follows that B^c is open in (Y, d_Y) and $[f^{-1}(B)]^c$ is not open in

(X, d_X) (since this would imply $f^{-1}(B)$ is closed). This, however, contradicts (b). The other direction follows from the same argument with "open" in the place of "closed", and vice versa, so that the result is shown. $\square$

It might seem surprising the topological properties of being open or closed do not necessarily transfer in the forward direction through continuous functions f. We should, however, be able to easily convince ourselves that the properties do carry through to the pre-image f^{-1}. For example, consider our previous $f : \mathbb{R} \mapsto \mathbb{R}$ with the discontinuity removed:

$$f(x) = \frac{1}{1+x^2}.$$



We can clearly see that:

1. $B = (\frac{1}{2}, 1) \implies f^{-1}(B) = (-1, 0) \cup (0, 1)$
2. $B = [\frac{1}{2}, 1] \implies f^{-1}(B) = [-1, 1]$

So, at least for this example, open and closed sets transfer from Y to the pre-image space X. (We can check, even with this example, that the other direction does not necessarily hold!)

4 Continuity and Connectedness

It was probably surprising that we could not transfer the properties of subsets being open and closed through the mapping f, even if f was continuous; rather, we are only permitted to work in the other direction.

If we thought about it for a moment, however, we would probably agree that there is one property which is trivially maintained through the mapping f: namely, connectedness. In the previous examples, we were able to map intervals to intervals from through f (since we are working in $\mathbb{R}$). In fact, this is a general property of functions on metric spaces, and the content of the following result.

Theorem 4.1. *Let (X, d_X) and (Y, d_Y) be metric spaces and $f : X \mapsto Y$. Then, if $S \subseteq X$ is connected, $f(S) \subseteq Y$ is connected.*

Proof. Suppose otherwise. That is to say, suppose S is connected but that there are two nonempty subsets $A, B \subseteq Y$ such that $A \cup B = f(S)$, $\overline{A} \cap B = \emptyset$, and $A \cap \overline{B} = \emptyset$.

Define $A^* = f^{-1}(A) \cap S$ and $B^* = f^{-1}(B) \cap S$. In other words, map the disconnected sets which compose $f(S)$ back into sets which compose S. (In general, we may have to remove elements of the pre-images $f^{-1}(A)$ and $f^{-1}(B)$; however, this is not a large concern.)

By construction, we have that $A^* \cup B^* = S$ and $A^* \cap B^* = \emptyset$. Now use the fact that $\overline{A} \cap B = \emptyset$ to show that $\overline{A}^* \cap B^* = \emptyset$. Notice, first of all, that $\overline{A}$ is a closed set and f is continuous so that we have that $f^{-1}(\overline{A})$ is a closed set (from Theorem 3.1). It follows that $\overline{A}^* \subseteq f^{-1}(\overline{A})$ and, consequently, that $f(\overline{A}^*) \subseteq \overline{A}$. Since $f(B^*) = B$, we have that $\overline{A}^* \cap B^* = \emptyset$. A similar argument shows that $A^* \cap \overline{B}^* = \emptyset$, which is a contradiction of the assumption that S is connected (since $A^* \cup B^* = S$). The result follows. □

An immediate application of Theorem 4.1 is the following.

Corollary 4.1 (Intermediate Value Theorem)**.** *Suppose $f : [x, y] \mapsto \mathbb{R}$ is continuous on $[x, y]$ and $f(x) = a$ and $f(y) = b$. Without loss of generality, assume $x < y$ and $a < b$. Then, for every $c \in (a, b)$, there is a $z \in (x, y)$ such that $f(z) = c$.*

Proof. To suppose otherwise is to suppose that there is a $c \in (a, b)$ such that $f(z) \neq c$ for all $z \in (x, y)$. It follows that $f([x, y])$ (image of the interval) is not an interval. Since we know (from Assignment #4) that the only connected components of $\mathbb{R}$ are intervals, this contradicts Theorem 4.1, and we are done. □

5 Continuity and Compactness

In our discussion on continuity, we have yet to consider compact sets. It should come as no surprise that, just as compact sets were often the most desirable sets to have for topological reasons, there are many reasons to prefer compact sets when dealing with continuous functions.

First of all, we have the following result.

Theorem 5.1. *Let (X, d_X) and (Y, d_Y) be metric spaces and $f : X \mapsto Y$ be continuous. Then, if $A \subseteq X$ is compact, $f(A) \subseteq Y$ is compact.*

Proof. Let $\{S_\alpha\}$, $\alpha \in \Lambda$, be an arbitrary open cover of $f(A)$. Since each S_α is open and f is continuous, it follows by Theorem 3.1 that each pre-image $f^{-1}(S_\alpha)$ is open. We also have that $\{f^{-1}(S_\alpha)\}$ covers A (otherwise, there is an $x \in A$ so that $f(x) \in f(A)$ but $f(x) \notin S_\alpha$ for any $\alpha \in \Lambda$). Since A is compact, we have that there is a finite subcover $\{f^{-1}(S_{\alpha_n})\}$, $n = 1, \ldots, N$, of A. Since $f(f^{-1}(S_{\alpha_n})) \subseteq S_{\alpha_n}$ and $\{f^{-1}(S_{\alpha_n}))\}$ covers A, we have that $\{S_{\alpha_n}\}$ covers $f(A)$, and we are done. □

A particularly nice feature of working with functions of compact sets is the following.

Theorem 5.2. *Let (X, d) be a metric spaces and $f : X \mapsto \mathbb{R}$ (we use $(\mathbb{R}, d)$ with $d(x, y) = |x - y|$). Suppose that $A \subseteq X$ is compact and $f : X \mapsto Y$ is continuous on A. Then there exist $p, q \in A$ such that*

$$f(p) = \inf_{x \in A} f(x) \quad \textit{and} \quad f(q) = \sup_{x \in A} f(x).$$

In other words, any continuous mapping from a compact set into the real numbers must attain its maximal and minimal elements. This is a particular useful result when considering multivariate functions $f : \mathbb{R}^n \mapsto \mathbb{R}$ since the Heine-Borel Theorem allows us to characterize compactness in $\mathbb{R}^n$ as those sets which are closed and bounded.

Proof. We know every compact set $A \subseteq X$ is closed (result earlier in the semester) and, since f is continuous, it follows from Theorem 5.1 that $f(A) \subseteq \mathbb{R}$ is compact. It follows from the Heine-Borel Theorem that $f(A)$ is closed and bounded. Since $f(A)$ is bounded, it follows that $\inf(f(A))$ and $\sup(f(A))$ exist in $\mathbb{R}$ (because $\mathbb{R}$ has the least-upper-bound property). Furthermore, since $\inf(f(A)) \in \overline{f(A)}$ and $\sup(f(A)) \in \overline{f(A)}$ and $f(A)$ is closed, it follows that $\inf(f(A)) \in f(A)$ and $\sup(f(A)) \in f(A)$, and the result is shown. □

Example: Prove that the following function $f : \mathbb{R}^2 \mapsto \mathbb{R}$ attains its minimum and maximum on $A = [1, 2] \times [1, 2]$:

$$f(x, y) = \frac{x - y}{x + y}.$$

Solution: Since f maps into $\mathbb{R}$, in order to apply Theorem 5.2, it is sufficient to guarantee that the set A is compact and f is continuous on A. We note in this case that $A \subset \mathbb{R}^2$ is closed and bounded, and therefore compact by the Heine-Borel theorem. Since the only discontinuities of f

occur at $y = -x$, which does not pass through A, we may conclude that f is continuous on A, and therefore that f attains its maximal and minimal values in A. (It can be checked the techniques of Math 234 that the minimal and maximal values are $-\frac{1}{3}$ and $\frac{1}{3}$ and are attained at the points $(1, 2)$ and $(2, 1)$, respectively.)

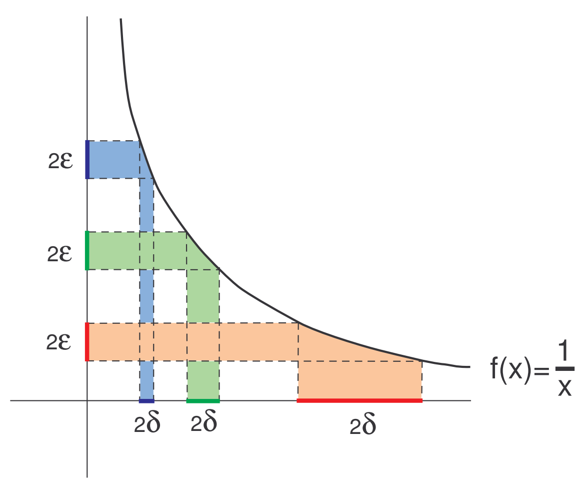
6 Uniform Continuity

In order to more fully understand the subtleties of the definition of continuity, consider the following real-valued function $f : \mathbb{R} \mapsto \mathbb{R}$:

$$f(x) = \frac{1}{x}.$$

We are well-aware, of course, that there is a discontinuity at $x = 0$. What we want to investigate now is the dependence of $\delta > 0$ on $\epsilon > 0$ as we consider points $a \in \mathbb{R}$ closer and closer to the discontinuity (from the right).

Take $\epsilon > 0$ fixed and ask the following question: for a fixed $a > 0$, how small does $\delta > 0$ have to be in order to guarantee that $|x - a| < \delta$ implies $|f(x) - f(a)| < \epsilon$? We can see by consider Figure 6 that the choice depends on sensitively on the point $a > 0$ we are considering. If we choose an a "far" from zero, we may choose δ to be fairly large, while if we choose an a to be "close" to zero, we must choose δ to be fairly small.



Now consider the following question: Is there is a uniform $\delta > 0$ which will work *for all* $a > 0$? This is not a senseless question! After all, the smaller bounds for the points closer to zero will still work for the points which are farther away. We might wonder if we could do away with the dependent of δ on the point a.

A little thought should convince us that this is not possible. Not only are the required δ's smaller as a becomes closer to zero, but in the limit as $a \to 0$ we have that $\delta \to 0$. We may not, therefore, find a smallest $\delta > 0$ which will work for all $a > 0$.

So even though $f(x) = \dfrac{1}{x}$ is continuous at all $a > 0$, it is perhaps not as continuous as we might like. We therefore introduce the following stronger notion of continuity.

Definition 6.1. *Suppose (X, d_X) and (Y, d_Y) are metric spaces and $f : X \mapsto Y$. We say f is* ***uniformly continuous*** *on $A \subseteq X$ if, for every $\epsilon > 0$, there is a $\delta > 0$ so that, for all $x, y \in A$, $d_X(x, y) < \delta$ implies $d_Y(f(x), f(y)) < \epsilon$.*

The key distinction between continuity and uniform continuity is that for uniform continuity we must choose the $\delta > 0$ prior to selecting the point $x \in A$. Another way of stating this is that the chosen $\delta > 0$ may depend on $\epsilon > 0$, but not on the choice of $x \in A$.

By our previous argumentation, we can see that $f(x) = \dfrac{1}{x}$ is not uniformly continuous on $A = \{x \in \mathbb{R} \mid x > 0\}$, even though it is continuous at every point $x \in A$.

We might wonder, however, whether the problem was our choice of A. For instance, imagine doing the same conceptual exercise as before, but on the set $A = \{x \in \mathbb{R} \mid x > m\}$ where $m > 0$ is some lower bound. Since the smallest required $\delta > 0$ occurs for $a = m$, and this δ will work for all other points, we may make out choice of δ prior to our choice of $a \in A$!

There are further subtleties to this construction. In general, however, the following result tells us exactly which kinds of sets we need to operate on in order to guarantee continuity implies uniform continuity.

Theorem 6.1. *Let (X, d_X) and (Y, d_Y) be metric spaces and $f : X \mapsto Y$. Suppose that $A \subseteq X$ is compact and f is continuous on A. Then f is uniformly continuous on A.*

Proof. Take $\epsilon > 0$. Since f is continuous on A, for every $a \in A$ there is a $\delta(a) > 0$ so that $d_X(a, x) < \delta(a)$ implies $d_Y(f(a), f(x)) < \frac{\epsilon}{2}$.

We now construct an open cover of A. We take $r(a) = \frac{1}{2}\delta(a)$ and consider the family $\{B_{r(a)}(a)\}$, $a \in A$ (notice, $B_{r(a)}(a)$ is defined with respect to the metric d_X). This clearly covers A so that, since A is compact, we may select a finite subcover $\{B_{r(a_n)}(a_n)\}$, $n = 1, \dots, N$. We now set

$$\delta = \min_{n=1,\dots,N} \{r(a_n)\} = \frac{1}{2}\left[\min_{n=1,\dots,N} \{\delta(a_n)\}\right] > 0.$$

Now consider two arbitrary points $a, b \in A$ and suppose that $d_X(a, b) < \delta$. Since our family covers A, we have that that $a \in B_{r(a_{n^*})}(a_{n^*})$ for some $n^* \in \{1, \dots N\}$. It follows that $d_X(a, a_{n^*}) < \delta(a_{n^*})$, which implies by the continuity of f that $d_Y(f(a), f(a_{n^*})) < \frac{\epsilon}{2}$. It furthermore follows by the triangle inequality that

$$d_X(b, a_{n^*}) \leq d_X(b, a) + d_X(a, a_{n^*}) < \delta + r(a_{n^*}) \leq 2r(a_{n^*}) = \delta(a_{n^*}).$$

It follows from the continuity of f that $d_Y(f(b), f(a_{n^*})) < \frac{\epsilon}{2}$.

Piecing everything together, we have that, for any $a, b \in A$,

$$\begin{aligned} d_Y(f(a), f(b)) &\leq d_Y(f(a), f(a_{n^*})) + d_Y(f(a_{n^*}), f(b)) \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

It follows that f is uniformly continuous on A. □

Example: Even though $f(x) = \frac{1}{x}$ was not uniformly continuous on $(0, \infty)$, we can now claim that it is uniformly continuous on any compact subset of $(0, \infty)$, i.e. on any closed and bounded sets. For instance, we *can* choose a minimal $\delta > 0$ which will work for all points in the sets $A = [1, 2]$, $B = [2, 10]$, $C = [\alpha, \frac{1}{\alpha}]$ for any fixed $\alpha \in (0, 1)$, etc.