Corollary 1.71. $|\mathcal{P}(\mathbb{N})| = |\{0,1\}^{\mathbb{N}}| > |\mathbb{N}|$.

Proof. Define $\Phi : \{0,1\}^{\mathbb{N}} \to \mathcal{P}(\mathbb{N})$ by

$$\Phi\big((x_n)_{n\geq 1}\big) = \{\, n \in \mathbb{N} : x_n = 1 \,\}.$$

If $(x_n) \neq (y_n)$ pick k with $x_k \neq y_k$; then k belongs to exactly one of the two images, so $\Phi(x_n) \neq \Phi(y_n)$. Thus Φ is injective. Conversely, for any $A \subseteq \mathbb{N}$ let its characteristic sequence χ_A be given by $\chi_A(n) = 1$ if $n \in A$, 0 otherwise. Then $\Phi(\chi_A) = A$, so Φ is surjective. Hence $|\mathcal{P}(\mathbb{N})| = |\{0,1\}^{\mathbb{N}}|$. $\square$

In fact, this corollary is a special case of a more general result which we will prove later.

Theorem 1.72 (Cantor's Theorem)**.** *For any set S, there is no surjective function from S to its power set $\mathcal{P}(S)$. In particular, $|S| < |\mathcal{P}(S)|$.*

Corollary 1.73. $|[0,1)| = |\{0,1\}^{\mathbb{N}}| > \aleph_0$.

Proof. Define $E \subset \{0,1\}^{\mathbb{N}}$ to be the set of sequences that are not eventually all 1. Define $\Phi : E \to [0,1)$ by

$$\Phi\big((x_n)_{n\geq 1}\big) = \sum_{n=1}^{\infty} \frac{x_n}{2^{\,n}}.$$

This is surjective because for any $y \in [0,1)$, we can write y in binary expansion $y = 0.y_1y_2y_3\ldots$ with $y_i \in \{0,1\}$ and not ending with an infinite sequence of 1's (i.e., when $0.x_1\cdots x_{m-1}1\,0\,0\,0\cdots = 0.x_1\cdots x_{m-1}0\,1\,1\,1\cdots$ happens we choose the expansion that does not end with 1's).

This is also injective: Let $(x_k), (y_k) \in E$ with $(x_k) \neq (y_k)$ and let n be the first index where they differ. WLOG $x_n = 0$ and $y_n = 1$. Then

$$\Phi(y) - \Phi(x) = \sum_{k=1}^{\infty} \frac{y_k - x_k}{2^k} = \frac{1}{2^n} + \sum_{k>n} \frac{y_k - x_k}{2^k}.$$

Since $y_k - x_k \in \{-1, 0, 1\}$,

$$\sum_{k>n} \frac{y_k - x_k}{2^k} \geq -\sum_{k>n} \frac{1}{2^k} = -\frac{1}{2^n},$$

with equality only if $y_k - x_k = -1$ for all $k > n$, i.e. $x_k = 1$ and $y_k = 0$ for every $k > n$, which would make (x_k) eventually 1, contrary to $(x_k) \in E$. Hence the tail sum is strictly greater than $-1/2^n$, so $\Phi(y) - \Phi(x) > 0$ and therefore $\Phi(x) \neq \Phi(y)$.

Therefore, $|[0,1)| = |E|$. Now, we let $D \subset \{0,1\}^{\mathbb{N}}$ be the set of sequences that are eventually 1. $E = \{0,1\} \setminus D$. We use the following two claims to finish the proof.

Claim 1 D is countable.

Claim 2 An uncountable set minus a countable set will maintain its cardinality.

Hence, we have $|E| = |\{0,1\}^{\mathbb{N}}|$. Hence $|[0,1)| = |\{0,1\}^{\mathbb{N}}| > \aleph_0$.

We will leave Claim 1 as an exercise in your homework 2. We prove Claim 2 below. $\square$

Proposition 1.74. *Let A be an infinite set and F a finite set. Then,*

- $|A \cup F| = |A|$.
- $|A \setminus F| = |A|$.

The conclusion still holds if A is uncountable and F is infinitely countable.

Proof. Let $G := F \setminus A$. Then G is finite and disjoint from A. We need to show that $|A \cup G| = |A|$. Since A is infinite, we can choose a sequence of distinct elements $a_1, a_2, \ldots$ in A. If $G = \emptyset$, the result is trivial. If $G \neq \emptyset$, we write $G = \{g_1, g_2, \ldots, g_n\}$, where n is finite. Define $f : A \to A \cup G$ as follows:

1. if $a \notin \{a_1, \ldots\}$, then set $f(a) := a$;
2. if $a = a_i$ for some $i \leq n$, then set $f(a) := g_i$.
3. If $a = a_i$ for some $i > n$, then set $f(a) := a_{i-n}$.

Then, f is a bijection. Thus $|A \cup G| = |A|$.

For the second part, we let $B := A \setminus F$. Then, $|A \setminus F| = |B| = |B \cup (A \cap F)| = |A|$.

Now, assume that F is countably infinite and A is uncountable. Let $G := F \setminus A$. Then, G is countable. WLOG, we assume G is infinite. We then can write $G = \{g_1, g_2, \ldots\}$. Still choose a sequence of distinct elements $a_1, a_2, \ldots$ in A. Define $f : A \to A \cup G$ as follows:

1. if $a \notin \{a_1, \ldots\}$, then set $f(a) := a$;
2. if $a = a_{2n}$, then set $f(a_{2n}) := g_n$;
3. if $a = a_{2n-1}$, then set $f(a_{2n-1}) := a_n$.

Then, f is a bijection. Thus $|A \cup G| = |A|$.

$\square$

Corollary 1.75. $|[0,1)| = |[0,1]| = |(0,1)| = |(0,1]|$.

Proposition 1.76. $|(0,1)| = |\mathbb{R}|$.

Proof. Let $f : (0,1) \to \mathbb{R}$ be defined by

$$f(x) = \tan\big((x - \tfrac{1}{2})\pi\big).$$

Then f is a bijection, hence $|(0,1)| = |\mathbb{R}|$. $\square$

Corollary 1.77. $|\mathbb{R} \setminus \mathbb{Q}| = |\mathbb{R}|$.

1.5.3 Proofs of some general theorems

Theorem 1.78 (Cantor-Schröder-Bernstein)**.** *If there exist injections $f : A \to B$ and $g : B \to A$, then there exists a bijection $h : A \to B$.*

Proof. WLOG, we assume that A and B are disjoint (otherwise replace B by a disjoint copy). For each $a \in A$, we have the following sequence:

$$\dots \overset{f}{\mapsto} b_{-3} \overset{g}{\mapsto} a_{-2} \overset{f}{\mapsto} b_{-1} \overset{g}{\mapsto} a_0 = a \overset{f}{\mapsto} b_0 \overset{g}{\mapsto} a_1 \overset{f}{\mapsto} b_1 \overset{g}{\mapsto} a_2 \overset{f}{\mapsto} \dots$$

There are only four possibilities for this sequence:

1. (B-stopper) it stops at some b_{-k} where $k \geq 1$.
2. (A-stopper) it stops at some a_{-k} where $k \geq 0$.
3. (bi-infinite) it continues indefinitely in both directions.
4. (cyclic) it is a cyclic sequence, i.e., $a_i = a_j$ for some $i < j$.

This also holds for any $b \in B$ and hence these four possibilities give rise to a partition of $A \sqcup B$.

Now, we define $h : A \to B$ as follows:

$$h(a) = \begin{cases} f(a), & \text{if } a \text{ is an } A\text{-stopper;} \\ f(a), & \text{if } a \text{ is in a bi-infinite chain or a cycle;} \\ g^{-1}(a), & \text{if } a \text{ is a } B\text{-stopper.} \end{cases}$$

This is a bijection. □

We proved this using Well-ordering theorem. We can also prove this using Zorn's lemma.

Theorem 1.79. *If A, B are sets then either $|A| \leq |B|$ or $|B| \leq |A|$.*

Proof. Let $\mathcal{M}$ be the set of *matchings* between A and B:

$\mathcal{M} = \{M \subseteq A \times B \mid$ each $a \in A$ appears in at most one pair of M, and each $b \in B$ appears in at most one$\}$.

We partially order $\mathcal{M}$ by inclusion. If $\{M_i\}_{i \in I}$ is a chain in $\mathcal{M}$, then

$$M = \bigcup_{i \in I} M_i$$

is again a matching, since no element of A or B can be matched to more than one partner within a chain. Thus every chain has an upper bound.

By *Zorn's Lemma*, $\mathcal{M}$ has a maximal element M^*.

Suppose there exist $a \in A$ and $b \in B$ both unmatched in M^*. Then

$$M' = M^* \cup \{(a, b)\}$$

is a matching strictly larger than M^*, contradicting maximality. Therefore, at least one side is fully matched: either every $a \in A$ is matched, or every $b \in B$ is matched.

- If every $a \in A$ is matched, define $f : A \to B$ by $f(a) = b$ whenever $(a, b) \in M^*$. This f is injective, so $|A| \leq |B|$.
- Otherwise, every $b \in B$ is matched. Define $g : B \to A$ by $g(b) = a$ whenever $(a, b) \in M^*$. Then g is injective, so $|B| \leq |A|$.

Thus, for any sets A and B, either $|A| \leq |B|$ or $|B| \leq |A|$. □

Theorem 1.80 (Cantor's Theorem)**.** *For any set S, there is no surjective function from S to its power set $\mathcal{P}(S)$. In particular, $|S| < |\mathcal{P}(S)|$.*

Proof. Suppose $f\colon S \to \mathcal{P}(S)$ is any function. Define $B = \{s \in S : s \notin f(s)\}$. If $B \subseteq S$ is in the range of f, then there exists $s_0 \in S$ such that $f(s_0) = B$. The following contradictions arise: one can show that

$$s_0 \in B \implies s_0 \notin f(s_0) = B \quad \text{and} \quad s_0 \notin B = f(s_0) \implies s_0 \in B,$$

which forces a contradiction. Consequently any function $f\colon S \to \mathcal{P}(S)$ must fail to be surjective, and hence $|S| \neq |\mathcal{P}(S)|$. On the other hand, there is an obvious injection $S \to \mathcal{P}(S)$ given by $s \mapsto \{s\}$. Therefore

$$|S| < |\mathcal{P}(S)| < |\mathcal{P}(\mathcal{P}(S))| < \dots.$$

□

2 Metric spaces

For real functions $f : \mathbb{R} \to \mathbb{R}$, we have the notion of continuity using ε-δ definition: a function $f : \mathbb{R} \to \mathbb{R}$ is continuous at a point $x_0 \in \mathbb{R}$ if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$|f(x) - f(x_0)| < \varepsilon \quad \text{whenever } |x - x_0| < \delta.$$

We also have the notion of continuity using sequences: a function $f : \mathbb{R} \to \mathbb{R}$ is continuous at a point $x_0 \in \mathbb{R}$ if for every sequence (x_n) in $\mathbb{R}$ converging to x_0, the sequence $(f(x_n))$ converges to $f(x_0)$.

In an attempt to generalize these notions to more general settings than $\mathbb{R}$, Frechet introduced the notion of metric spaces in 1906. Metric spaces provide a natural setting to generalize the notions of continuity above. Later, Hausdorff introduced the more general notion of topological spaces in 1914. We will first study metric spaces and then topological spaces.

Definition 2.1 (Metric Space)**.** A *metric space* is a pair (X, d) consisting of a set X and a function $d\colon X \times X \to [0, \infty)$ called a *metric* such that

1. $d(x, y) = 0$ if and only if $x = y$;
2. $d(x, y) = d(y, x)$ for all $x, y \in X$ (symmetry);
3. $d(x, z) \leq d(x, y) + d(y, z)$ for all $x, y, z \in X$ (triangle inequality).

The value $d(x, y)$ is called the *distance* between points x and y in the metric space.

If (X, d) satisfies 2 and 3 and 1' below, it is called a *pseudometric space*.

1'. $d(x, x) = 0$ for all $x \in X$.

Example 2.2. 1. $X = \mathbb{R}^n$, $d_p(x, y) = \|x - y\|_p = \left(\sum_{i=1}^n |x_i - y_i|^p\right)^{1/p}$ for $1 \leq p < \infty$ and $d_\infty(x, y) = \|x - y\|_\infty = \max_{1 \leq i \leq n} |x_i - y_i|$ are metrics on $\mathbb{R}^n$.

2. Let X be any set and $f : X \to \mathbb{R}$ be any function. Then, the following defines a pseudometric on X:

$$d(x, y) = |f(x) - f(y)|.$$

3. Let (X, d) be a metric space and $A \subset X$ be a nonempty subset. Then, $(A, d|_{A \times A})$ is a metric space, called a *subspace* of (X, d).

Definition 2.3. Let (X, d_X) and (Y, d_Y) be metric spaces. A function $f\colon X \to Y$ is called an isometric embedding if

$$d_Y(f(x), f(x')) = d_X(x, x') \quad \text{for all } x, x' \in X.$$

The function f is called an isometry if it is an isometric embedding that is also surjective and we say that (X, d_X) and (Y, d_Y) are *isometric* if there exists an isometry between them.

Example 2.4 (Isometric embedding)**.** Let $f : \mathbb{R} \to \mathbb{R}^2$ be given by $f(t) = (t, 0)$. For $s, t \in \mathbb{R}$,

$$\|f(s) - f(t)\|_2 = \sqrt{(s-t)^2 + 0^2} = |s - t|.$$

Thus f preserves distances (an isometric embedding). It is not surjective, so it is not an isometry. Its image is the x-axis in $\mathbb{R}^2$.

Is the map $g(x) := (x, x)$ an isometric embedding?

Definition 2.5. Let (X, d_X) and (Y, d_Y) be metric spaces. A function $f\colon X \to Y$ is called *Lipschitz continuous* if there exists a constant $L \geq 0$ such that

$$d_Y(f(x), f(x')) \leq L \cdot d_X(x, x') \quad \text{for all } x, x' \in X.$$

Definition 2.6. Let (X, d_X) and (Y, d_Y) be metric spaces. A function $f\colon X \to Y$ is called *continuous* at a point $x_0 \in X$ if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$d_Y(f(x), f(x_0)) < \varepsilon \quad \text{whenever } d_X(x, x_0) < \delta.$$

The function f is called *continuous* if it is continuous at every point $x_0 \in X$.

Proposition 2.7. *Any Lipschitz continuous function between metric spaces is continuous. Any isometric embedding is also Lipschitz continuous.*

Proof. Let $f : X \to Y$ be Lipschitz continuous with constant L. Given $\varepsilon > 0$, let $\delta = \varepsilon/L$. Then, for any $x \in X$ with $d_X(x, x_0) < \delta$, we have

$$d_Y(f(x), f(x_0)) \leq L \cdot d_X(x, x_0) < L \cdot \frac{\varepsilon}{L} = \varepsilon.$$

Thus, f is continuous at x_0. Since x_0 is arbitrary, f is continuous.

If f is an isometric embedding, then for any $x, x' \in X$, we have

$$d_Y(f(x), f(x')) = d_X(x, x') \leq 1 \cdot d_X(x, x').$$

Thus, f is Lipschitz continuous with constant $L = 1$. $\square$

2.1 Characterizations of continuity

2.1.1 A sequential characterization of continuity

For real functions, we can define continuity using sequences. This can be generalized to metric spaces as well.

Definition 2.8. A sequence $(x_n)_{n\in\mathbb{N}}$ in a metric space (X, d) is said to converge to a point $x \in X$ if for every $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $d(x_n, x) < \epsilon$ for all $n \geq N$.

One can characterize continuity using sequences:

Proposition 2.9. *A function $f : X \to Y$ is continuous at a point $x_0 \in X$ if and only if for every sequence $(x_n)_{n\in\mathbb{N}}$ in X converging to x_0, the sequence $(f(x_n))_{n\in\mathbb{N}}$ converges to $f(x_0)$.*

Proof. Suppose first that f is continuous at x_0, and let $(x_n)_{n\in\mathbb{N}}$ be a sequence converging to x_0. For any $\varepsilon > 0$, by continuity there exists $\delta > 0$ such that

$$d_X(x, x_0) < \delta \quad \implies \quad d_Y(f(x), f(x_0)) < \varepsilon.$$

Since (x_n) converges to x_0, there exists an $N \in \mathbb{N}$ with $d_X(x_n, x_0) < \delta$ for all $n \geq N$. Hence, for all $n \geq N$, we have

$$d_Y(f(x_n), f(x_0)) < \varepsilon,$$

so $(f(x_n))$ converges to $f(x_0)$.

Conversely, suppose that for every sequence (x_n) converging to x_0, the sequence $(f(x_n))$ converges to $f(x_0)$. Assume for contradiction that f is not continuous at x_0. Then there exists $\varepsilon_0 > 0$ such that for every $\delta > 0$ there exists some $x \in X$ with

$$d_X(x, x_0) < \delta \qquad \text{but} \qquad d_Y(f(x), f(x_0)) \geq \varepsilon_0.$$

For each $n \in \mathbb{N}$, choosing $\delta = \frac{1}{n}$, there exists $x_n \in X$ with

$$d_X(x_n, x_0) < \frac{1}{n} \quad \text{and} \quad d_Y(f(x_n), f(x_0)) \geq \varepsilon_0.$$

Then (x_n) converges to x_0, yet $(f(x_n))$ cannot converge to $f(x_0)$ since the distance remains at least ε_0, which is a contradiction. $\square$

2.1.2 A open set characterization of continuity

Definition 2.10. Let (X, d) be a metric space. For any $x \in X$ and $r > 0$, the *open ball* of radius r centered at x is defined as

$$B_r(x) = \{y \in X : d(x, y) < r\}.$$

Example 2.11 (Shapes of balls for different metrics on $\mathbb{R}^2$)**.** For $r > 0$ the open unit balls (centered at the origin) for the standard metrics on $\mathbb{R}^2$ have the following shapes:

$$\begin{aligned}
B_r^{(2)}(0) &= \{x \in \mathbb{R}^2 : \|x\|_2 < r\} && \text{(Euclidean) — disk,}\\
B_r^{(1)}(0) &= \{x \in \mathbb{R}^2 : |x_1| + |x_2| < r\} && \text{(Taxicab / } \ell^1\text{) — diamond,}\\
B_r^{(\infty)}(0) &= \{x \in \mathbb{R}^2 : \max(|x_1|, |x_2|) < r\} && \text{(Maximum / } \ell^\infty\text{) — square.}
\end{aligned}$$

Below we draw the case $r = 1$.

Then, we can rewrite the definition of continuity using open balls.

Proposition 2.12. *A function $f : X \to Y$ between metric spaces (X, d_X) and (Y, d_Y) is continuous at a point $x_0 \in X$ if and only if for every $r > 0$, there exists $\delta > 0$ such that*

$$f(B_\delta(x_0)) \subseteq B_r(f(x_0)).$$

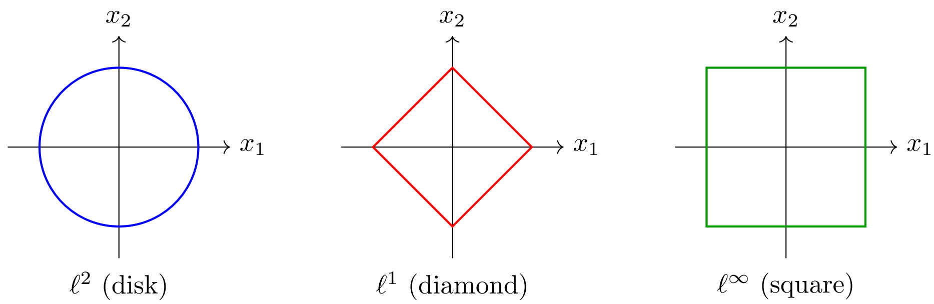


Figure 1: Unit balls $B_1(0)$ in $\mathbb{R}^2$ for ℓ^2, ℓ^1, ℓ^∞ metrics.

The definition of open balls of course depends on the metric. Suppose you want to generalize everything so far to the setting when you don't have metric. Then, let's try to define some language where the notion of metric will be hidden.

Definition 2.13. A subset $U \subseteq X$ is called *open* if for every point $x \in U$ there exists a radius $r > 0$ such that the open ball $B_r(x)$ is contained in U.

Proposition 2.14. *Any open ball $B_r(x)$ is open.*

Proof. Let's prove the last item. Let $x \in B_r(x_0)$ for some $r > 0$. Pick $r' := r - d_X(x, x_0)$. Then, $B_{r'}(x) \subseteq B_r(x_0)$. $\square$

Corollary 2.15. *Let (X, d) be a metric space and $U \subseteq X$ be a subset. Then, U is open iff U is a union of open balls.*

Proof. Assume that U is open. Then, for every $x \in U$ there exists $r_x > 0$ so that $B_{r_x}(x) \subseteq U$. Thus, we can write

$$U = \bigcup_{x \in U} B_{r_x}(x).$$

Conversely, if $U = \bigcup_{i \in I} B_{r_i}(x_i)$ is a union of open balls, then for every $x \in U$, we have that $x \in B_{r_i}(x_i)$ for some $i \in I$. This means that there exists a radius $r > 0$ such that $B_r(x) \subseteq B_{r_i}(x_i) \subseteq U$. Thus, U is open. $\square$

Now we can revisit the definition of continuous functions.

Theorem 2.16. *Let $f : X \to Y$ be a function between metric spaces (X, d_X) and (Y, d_Y). Then f is continuous at a point $x_0 \in X$ if and only if for every open set $V \subseteq Y$ containing $f(x_0)$, there exists an open set $U \subseteq X$ containing x_0 such that $f(U) \subseteq V$.*

Proof. Suppose that f is continuous at x_0. Let $V \subseteq Y$ be an open set containing $f(x_0)$. By openness, there exists $r > 0$ such that $B_r(f(x_0)) \subseteq V$. By continuity, there exists $\delta > 0$ such that $f(B_\delta(x_0)) \subseteq B_r(f(x_0)) \subseteq V$. Thus, we can take $U = B_\delta(x_0)$.

Conversely, suppose that for every open set $V \subseteq Y$ containing $f(x_0)$, there exists an open set $U \subseteq X$ containing x_0 such that $f(U) \subseteq V$. Given $\varepsilon > 0$, let $V = B_\varepsilon(f(x_0))$. By assumption, there exists an open set $U \subseteq X$ containing x_0 such that $f(U) \subseteq V$. Since U is open and contains x_0, there exists $\delta > 0$ such that $B_\delta(x_0) \subseteq U$. Thus,

$$f(B_\delta(x_0)) \subseteq f(U) \subseteq V = B_\varepsilon(f(x_0)).$$

This shows that f is continuous at x_0. $\square$

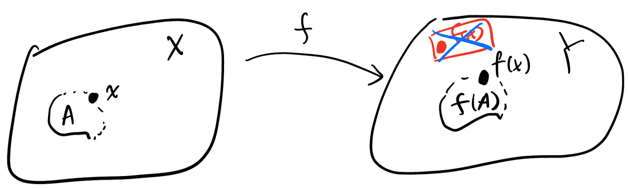
Finally, we will be able to establish the following characterization from the theorem above. We will not give a proof here as we will give a proof in a more general setting (topological spaces) later.

Theorem 2.17. *f is continuous if and only if for every open set $V \subseteq Y$, the preimage $f^{-1}(V)$ is open in X.*

The above theorem reads "preimage of open is open".

2.1.3 A closed set characterization of continuity

However, I never find this characterization natural to understand. We can think of the continuity as not ripping apart the space at any given point. More precisely, if a point is "in touch" with some region in the domain, it should remain in touch with the corresponding region in the codomain.



Definition 2.18. A subset $A \subseteq X$ is called *closed* if its complement $X \setminus A$ is open.

Definition 2.19. For any set $A \subseteq X$, the *closure* of A, denoted $\overline{A}$, is the smallest closed set containing A. Equivalently, it is the intersection of all closed sets containing A.

Theorem 2.20. *Let $f : X \to Y$ be a function between metric spaces (X, d_X) and (Y, d_Y). Then f is continuous if and only if $f(\overline{A}) \subseteq \overline{f(A)}$*

The theorem above reads "image of closure is contained in closure of image", or more intuitively, "points that are in touch with a set remains in touch after the map".

Now we prove the theorem.

Lemma 2.21. *Let $A \subseteq X$ be a set. Then,*

$$\overline{A} = \{x \in X : \textit{there exists a sequence } (x_n)_{n\in\mathbb{N}} \textit{ in } A \textit{ converging to } x\}.$$

Proof. Let $x \in \overline{A}$. Consider any $r > 0$. If $B_r(x) \cap A = \emptyset$, then the complement $X \backslash B_r(x) \supset A$. But this contradicts the fact that $x \in \overline{A}$. Now let $r_n := 1/n$. Then, pick $x_n \in B_{r_n}(x) \cap A$. Then, (x_n) is a sequence in A converging to x.

Conversely, let $x \in X$ be such that there exists a sequence $(x_n)_{n\in\mathbb{N}}$ in A converging to x. Consider a closed set $C \supset A$. If $x \notin C$, then there exists $r > 0$ such that $B_r(x) \cap C = \emptyset$ (why?). This implies that $B_r(x) \cap A = \emptyset$ which contradicts the fact that (x_n) converges to x. Thus, $x \in C$. Since $C \supset A$ is arbitrary, we have $x \in \overline{A}$. □

Now, we prove Theorem 2.20.

Proof of Theorem 2.20. Suppose that f is continuous. Let $x \in \overline{A}$. By the lemma, there exists a sequence (x_n) in A converging to x. By continuity, $(f(x_n))$ converges to $f(x)$. Since each $f(x_n) \in f(A)$, by the lemma again, we have $f(x) \in \overline{f(A)}$. Thus, $f(\overline{A}) \subseteq \overline{f(A)}$.

Conversely, suppose that $f(\overline{A}) \subseteq \overline{f(A)}$ for every set $A \subseteq X$. Let $x_0 \in X$ and let (x_n) be a sequence converging to x_0. Assume on the contrary that $(f(x_n))$ does not converge to $f(x_0)$. Then, there exists $\varepsilon > 0$ such that there is a subsequence $(f(x_{k_n}))$ such that $d_Y(f(x_{k_n}), f(x_0)) \geq \varepsilon$ for all $n \in \mathbb{N}$.

Now we let $A = \{x_{k_n} : n \in \mathbb{N}\}$. Then, $x_0 \in \overline{A}$. Thus, $f(x_0) \in f(\overline{A}) \subseteq \overline{f(A)}$. By the lemma, there exists a sequence (y_n) in $f(A)$ converging to $f(x_0)$. But this is not possible due to the ϵ constraint. Therefore, we conclude that $(f(x_n))$ must converge to $f(x_0)$. □

We now see that many intuitive properties of continuous real functions can be generalized nicely to metric spaces. We will see that the open and closed language still works well in the more general setting of topological spaces, while the sequential language does not. In Willard's book, a whole chapter is devoted to the study of sequences in topological spaces.

2.2 Ultrametric

There is a special class of metric spaces called ultrametric spaces, which satisfy a stronger version of the triangle inequality. Ultrametric spaces have some interesting properties that are not shared by general metric spaces.

Definition 2.22 (Ultrametric space)**.** A metric space (X, d) is an *ultrametric space* if d satisfies the *strong triangle inequality*

$$d(x, z) \leq \max\{d(x, y), d(y, z)\} \qquad \text{for all } x, y, z \in X.$$

We would usually use u instead of d to denote an ultrametric.

Proposition 2.23. *Let (X, u) be an ultrametric space. Then,*

1. *If $r > 0$ and $x, y \in X$ with $u(x,y) < r$, then $B_r(x) = B_r(y)$.*

2. *Any two open balls are either disjoint or one contains the other.*

3. *Every triangle is isosceles, i.e., if $x, y, z \in X$, then at least two of the distances $u(x,y)$, $u(y,z)$, and $u(z,x)$ are equal, and the third is less than or equal to the common value.*

Example 2.24. Recall the definition of the Cantor set $C \subseteq [0,1]$: it is constructed by repeatedly removing the open middle third from each interval. More precisely, set $C_0 = [0,1]$, and for each $k \geq 1$, let C_k be the union of 2^k closed intervals obtained by removing the open middle third from each interval in C_{k-1}. The Cantor set is defined as

$$C = \bigcap_{k=0}^{\infty} C_k.$$

Equivalently, C consists of all numbers in $[0,1]$ whose base-3 expansion contains only the digits 0 and 2. We endow the Cantor set $C \subseteq [0,1]$ with the ultrametric as follows:

$$d_C(x,y) = \begin{cases} 0 & \text{if } x = y, \\ 3^{-n} & \text{if } x \neq y \text{ and } n = \min\{k \geq 0 : x \text{ and } y \text{ lie in different segments in } C_k\}. \end{cases}$$

It is perhaps difficult for you to understand what happens in the example above. It turns out that one can always create a nice visualization of ultrametric spaces using trees. We will use this to help us understand ultrametric spaces better. Although the following definition is for finite sets, it can be generalized to infinite sets as well.