# Continuity for Metric Spaces

IngramMaths

January 2026

## 1 Definitions of Continuity

Before looking at some of the important properties that come with continuity, we will study some alternative definitions that are equivalent for metric spaces.

**Definition 1.1 (Classical Continuity)**
Suppose $(X, d)$ and $(Y, \tilde{d})$ are metric spaces and $f : X \to Y$ a function. We say that $f$ is **continuous** at a point $x_0 \in X$ if and only if $\forall \varepsilon > 0\ \exists \delta > 0$ such that

$$\forall x \in X \quad d(x, x_0) < \delta \implies \tilde{d}(f(x), f(x_0)) < \varepsilon$$

We say that $f$ is (globally) continuous if and only if $f$ is continuous at all $x_0 \in X$. If $A \subseteq X$ and $B \subseteq Y$ then a function $g : A \to B$ is continuous if and only if it is continuous from the subspace $(A, d|_A)$ to the subspace $(B, \tilde{d}|_B)$.

*Remark.* Note that, implicitly, continuity is not taken as a property of $f$ here. In particular, "$f$ is continuous" really means "$f$ is continuous from $(X, d)$ to $(Y, \tilde{d})$." This is hinted at with the definition of continuity for subsets of a metric space, where we define it in terms of continuity from one metric subspace to another metric subspace.

**Theorem 1.2 (Sequential Continuity)**
Suppose $(X, d)$ and $(Y, \tilde{d})$ are metric spaces and $f : X \to Y$ a function. Then, $f$ is continuous from $(X, d)$ to $(Y, \tilde{d})$ if and only if, for every convergent sequence $(x_n)$ in $X$,

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

*Proof.* "$\implies$": Suppose that $f$ is continuous, and take an arbitrary convergent sequence $(x_n)$ in $X$, letting $L$ be its limit. If $\varepsilon > 0$ is given, then by the classical definition of continuity there exists $\delta > 0$ such that

$$\forall x \in X \quad d(x, L) < \delta \implies \tilde{d}(f(x), f(L)) < \varepsilon$$

By the definition of convergence of a sequence, there exists $N \in \mathbb{N}$ such that

$$\forall n \in \mathbb{N} \quad n > N \implies d(x_n, L) < \delta$$

Hence, for all $n \in \mathbb{N}$, if $n > N$ then $d(x_n, L) < \delta$ and so $\tilde{d}(f(x_n), f(L)) < \varepsilon$. It follows that $f(x_n) \to f(L)$ as $n \to \infty$, that is,

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

as required.

" $\Longleftarrow$ ": For this direction, we use the contrapositive. Suppose that $f$ is not continuous, that is, for some point $x_0 \in X$, there exists $\varepsilon > 0$ such that for all $\delta > 0$ we can find $x \in X$ where $d(x, x_0) < \delta$ but $\tilde{d}(f(x), f(x_0) \geq \varepsilon$. We hence use this to construct a sequence $(x_n)$ such that, for all $n \in \mathbb{N}$,

$$d(x_0, x_n) < \frac{1}{n} \text{ but } \tilde{d}(f(x_0), f(x_n)) \geq \varepsilon$$

Since $\frac{1}{n} \to 0$ as $n \to \infty$, the Sandwich Theorem tells us that $d(x_0, x_n) \to 0$ as $n \to \infty$, and so $x_n \to x_0$ as $n \to \infty$. However, clearly

$$\tilde{d}(f(x_0), f(x_n)) \not\to 0 \text{ as } n \to \infty$$

since the distance is bounded below by $\varepsilon > 0$. This means that $f(x_n) \not\to f(x_0)$ as $n \to \infty$, meaning that the sequential definition of continuity does not hold for this sequence. Thus, contrapositively, it follows that if the sequential definition holds, then the classical definition must also hold. □

**Theorem 1.3 (Topological Continuity)**
Suppose $(X, d)$ and $(Y, \tilde{d})$ are metric spaces and $f : X \to Y$ a function. Then, $f$ is continuous from $(X, d)$ to $(Y, \tilde{d})$ if and only if for every open set $\Omega$ in $(Y, \tilde{d})$ the preimage $f^{-1}(\Omega)$ is open in $(X, d)$.

*Proof.* " $\Longrightarrow$ ": Suppose that $f$ is continuous, and take any open $\Omega \subseteq Y$. Consider an arbitrary point $x_0 \in f^{-1}(\Omega)$. Since $\Omega$ is open and $f(x_0) \in \Omega$, there exists $\varepsilon > 0$ such that

$$B_{\tilde{d}}(f(x_0), \varepsilon) \subseteq \Omega$$

That is, for any $y \in Y$,

$$d(y, f(x_0)) < \varepsilon \implies y \in \Omega$$

Since $f$ is continuous, there exists $\delta > 0$ such that if $d(x, x_0) < \delta$ then

$$\tilde{d}(f(x), f(x_0)) < \varepsilon$$

That is, if $d(x, x_0) < \delta$ then $f(x) \in \Omega$, or equivalently $x \in f^{-1}(\Omega)$. It follows that

$$B_d(x_0, \delta) \subseteq f^{-1}(\Omega)$$

Since $x_0$ was arbitrary, it follows that $f^{-1}(\Omega)$ is open, as required.

" $\Longleftarrow$ ": Suppose that $f$ satisfies the topological definition of continuity, and take any $x_0 \in X$. For any $\varepsilon > 0$, we know that $B_{\tilde{d}}(f(x_0), \varepsilon)$ is open in $(Y, \tilde{d})$.

Hence, $f^{-1}(B_{\tilde{d}}(f(x_0), \varepsilon))$ is open in $(X, d)$. Thus, there exists $\delta > 0$ such that

$$B_d(x_0, \delta) \subseteq f^{-1}(B_{\tilde{d}}(f(x_0), \varepsilon))$$

That is, for any $x \in X$, if $d(x_0, x) < \delta$ then

$$x \in f^{-1}(B_{\tilde{d}}(f(x_0), \varepsilon))$$

and so

$$f(x) \in B_{\tilde{d}}(f(x_0), \varepsilon)$$

finally giving

$$\tilde{d}(f(x), f(x_0)) < \varepsilon$$

It follows that $f$ is continuous at $x_0$. Since $x_0$ was arbitrary, it follows that $f$ is continuous. $\square$

Having now seen all three of the key definitions of continuity, we will observe some examples of results, showing that different definitions are more applicable in different scenarios.

**Proposition 1.4**
Let $(X, d)$ and $(Y, \tilde{d})$ be metric spaces. Then, for any $y_0 \in Y$, the constant function $f : X \to Y : x \mapsto y_0$ is continuous.

*Proof.* Take any open set $\Omega$ in $Y$. If $y_0 \in \Omega$ then $f^{-1}(\Omega) = X$, which is open, meaning that $f$ is continuous (by the topological definition). Otherwise, $f^{-1}(\Omega) = \varnothing$, which is also open, meaning that $f$ is continuous in this case also. Thus, $f$ is continuous in all cases, as required. $\square$

**Proposition 1.5**
Let $(X, d)$ be a metric space. Then, the identity function $\mathrm{Id}_X : X \to X$ mapping $x \mapsto x$ is continuous from $(X, d)$ to $(X, d)$.

*Proof.* For any convergent sequence $(x_n)$ in $(X, d)$, trivially

$$\lim_{n\to\infty} x_n = \lim_{n\to\infty} x_n$$

meaning that $f$ is continuous, as required. $\square$

**Proposition 1.6**
Let $(X, d_X), (Y, d_Y), (Z, d_Z)$ be metric spaces and $f : X \to Y$, $g : Y \to Z$ be continuous functions. Then, $g \circ f$ is continuous.

*Proof.* Let $(x_n)$ be any convergent sequence in $X$. Then, by the continuity of $f$, we have

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

Then, by the continuity of $g$, we have

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

It follows that $g \circ f$ is continuous from $(X, d_X)$ to $(Z, d_Z)$ by the sequential definition as required. $\square$

*Remark.* The proof using the topological definition is also quite simple: letting $\Omega$ be an open set in $(Z, d_Z)$, we get that $g^{-1}(\Omega)$ is open in $(Y, d_Y)$ and so $f^{-1}(g^{-1}(\Omega))$ is open in $(X, d_X)$. That is, $(g \circ f)^{-1}$ is open in $(X, d)$. Since $\Omega$ was arbitrary, it follows that $g \circ f$ is continuous.

# 2 The Contraction Mapping Theorem

**Definition 2.1 (Lipschitz Functions)**
Let $(X, d)$ and $(Y, \tilde{d})$ be metric spaces. A function $f : X \to Y$ is said to be **Lipschitz continuous** (or just Lipschitz) from $(X, d)$ to $(Y, \tilde{d})$ with **Lipschitz constant** $\lambda \geq 0$ if and only if

$$\forall x_1, x_2 \in X \quad \tilde{d}(f(x_1), f(x_2)) \leq \lambda d(x_1, x_2)$$

The **optimal Lipschitz constant** of $f$ is defined as

$$\lambda^* := \inf\{\lambda \in \mathbb{R} \mid f \text{ is Lipschitz with constant } \lambda\}$$

*Remark.* It can be proven that the optimal Lipschitz constant of $f$ is indeed a Lipschitz constant of $f$ itself. This can be done by first considering the case where $x_1 = x_2$, and then otherwise noting that

$$\frac{\tilde{d}(f(x_1), f(x_2))}{d(x_1, x_2)}$$

is a lower bound on the Lipschitz constants of $f$, and therefore must be at most $\lambda^*$, giving $\tilde{d}(f(x_1), f(x_2)) \leq \lambda^* d(x_1, x_2)$.

**Definition 2.2 (Contractions)**
Let $(X, d)$ be a metric space. A function $f : X \to X$ is said to be a **contraction** if and only if it is Lipschitz continuous from $(X, d)$ to $(X, d)$ with Lipschitz constant $0 \leq \lambda < 1$.

**Theorem 2.3**
Let $(X, d)$ and $(Y, \tilde{d})$ be metric spaces and $f : X \to Y$ be Lipschitz continuous with constant $\lambda$. Then, $f$ is continuous.

*Proof.* If $\lambda = 0$ then for all $x_1, x_2 \in X$ we have

$$\tilde{d}(f(x_1), f(x_2)) \leq 0 \cdot d(x_1, x_2) = 0$$

Hence, by identity of indiscernibles, $f(x_1) = f(x_2)$. It follows that $f$ is constant in this case, and therefore continuous by 1.4. Otherwise, we must have $\lambda > 0$. Taking any $x_0 \in X$ and letting $\varepsilon > 0$ be given, we therefore pick $\delta = \frac{\varepsilon}{\lambda}$. For any $x \in X$, if $d(x, x_0) < \delta$ then

$$\begin{aligned}
\tilde{d}(f(x), f(x_0)) &\leq \lambda d(x, x_0) \\
&< \lambda \cdot \frac{\varepsilon}{\lambda} \\
&= \varepsilon
\end{aligned}$$

By the classical definition of continuity, it follows that $f$ is continuous at $x_0$, and therefore (globally) continuous by the arbitrariness of $x_0$. □

**Definition 2.4 (Fixed Points)**
Let $X$ be a non-empty set and $f : X \to X$ a function. Then, a **fixed point** of $f$ is a point $x_0 \in X$ such that $f(x_0) = x_0$.

**Theorem 2.5 (Contraction Mapping Theorem)**
Let $(X, d)$ be a complete metric space and $f : X \to X$ a contraction. Then,

(i) $f$ has a unique fixed point $\hat{x}$.

(ii) If $(x_n)$ is a sequence with the property $x_{n+1} = f(x_n)$ then $x_n \to \hat{x}$ as $n \to \infty$.

*Proof.* Let $\lambda \in [0, 1)$ be a Lipschitz factor of $f$. Let $(x_n)$ be any sequence with the property that

$$\forall n \in \mathbb{N} \quad x_{n+1} = f(x_n)$$

Then, for any $m, n \in \mathbb{N}$, if $m, n > 1$ then

$$d(x_m, x_n) = d(f(x_{m-1}), f(x_{n-1})) \leq \lambda d(x_{m-1}, x_{n-1})$$

Inductively, if $m \geq n$ then it follows that

$$d(x_m, x_n) \leq \lambda^{n-1} d(x_{m-n+1}, x_1) \tag{*}$$

Notably, this also holds for $n = 1$. Also, for any $n \in \mathbb{N}$, if $n > 1$ then

$$d(x_n, x_1) \leq \sum_{i=1}^{n-1} d(x_i, x_{i+1})$$

by the triangle inequality. By (*), this means that

$$d(x_n, x_1) \leq \sum_{i=1}^{n-1} \lambda^{i-1} d(x_1, x_2) = d(x_1, x_2) \cdot \frac{1 - \lambda^{n-1}}{1 - \lambda} \tag{**}$$

Combining (*) and (**) we get

$$\begin{aligned}
d(x_m, x_n) &\leq \lambda^{n-1} \cdot d(x_1, x_2) \cdot \frac{1-\lambda^{m-n}}{1-\lambda} \\
&= d(x_1, x_2) \cdot \frac{\lambda^{n-1} - \lambda^{m-1}}{1-\lambda} \\
&= \frac{d(x_1, x_2)}{\lambda(1-\lambda)}(\lambda^n - \lambda^m) \\
&\leq \frac{d(x_1, x_2)}{\lambda(1-\lambda)}\lambda^n
\end{aligned}$$

Since $0 \leq \lambda^n \to 0$ as $n \to \infty$, letting $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that if $n > N$ then $\lambda^n < \frac{\lambda(1-\lambda)}{d(x_1,x_2)}\varepsilon$. So, if $m, n > N$, assuming again WLoG that $m \geq n$, we get

$$d(x_m, x_n) \leq \frac{d(x_1, x_2)}{\lambda(1-\lambda)} \cdot \frac{\lambda(1-\lambda)}{d(x_1, x_2)}\varepsilon = \varepsilon$$

So, $(x_n)$ is Cauchy. Since $(X, d)$ is complete, $(x_n)$ must be convergent, say to a limit $\hat{x}$. Since $f$ is a contraction, it is continuous (by 2.3) and so

$$\begin{aligned}
f(\hat{x}) &= f\left(\lim_{n\to\infty} x_n\right) \\
&= \lim_{n\to\infty} f(x_n) \\
&= \lim_{n\to\infty} x_{n+1} \\
&= \hat{x}
\end{aligned}$$

So, $\hat{x}$ is a fixed point of $f$. To verify both (i) and (ii), it's therefore enough to prove that $\hat{x}$ is unique. If $f$ has two fixed points $x_1, x_2$ then we must have

$$\begin{aligned}
d(x_1, x_2) &= d(f(x_1), f(x_2)) \\
&\leq \lambda d(x_1, x_2)
\end{aligned}$$

If we assume for contradiction that $d(x_1, x_2) > 0$, then this gives $d(x_1, x_2) < d(x_1, x_2)$, so this cannot be the case. Thus, we must have $d(x_1, x_2) = 0$, that is, $x_1 = x_2$. So, the fixed point $\hat{x}$ must be unique, as required. $\square$

Having proved this immensely powerful result, we can now put it into action and apply it to develop some numerical approximation methods. A key strategy which is commonly applied is using the **mean value theorem**, allowing us to bound the growth of a function within a given interval.

**Example 2.6**

Let's attempt to find and justify a numerical "iteration" method for approximating a solution to

$$\cos(x^2) = 5x$$

For this, we could (for example) define $f : [-1, 1] \to [-1, 1]$ by

$$f(x) = \frac{1}{4}(\cos(x^2) - x)$$

It can be formally verified quite easily that this function is well-defined. Letting $x_0, y_0 \in [-1, 1]$ be given, the mean value theorem tells us that for some $c \in [-1, 1]$ we have

$$\begin{aligned} \frac{f(x_0) - f(y_0)}{x_0 - y_0} &= f'(c) \\ \implies |f(x_0) - f(y_0)| &= f'(c)|x_0 - y_0| \end{aligned}$$

The derivative of $f$ is given by

$$f'(x) = -\frac{1}{4}(2x\sin(x^2) - 1)$$

Thus,

$$\begin{aligned} |f(x_0) - f(y_0)| &= \frac{1}{4}|x_0 - y_0||2x\sin(x^2) - 1| \\ &\leq |x_0 - y_0| \left(\frac{1}{2}|x\sin(x^2)| + \frac{1}{4}\right) \\ &\leq \frac{3}{4}|x_0 - y_0| \end{aligned}$$

Hence, $f$ is a contraction with Lipschitz factor $\frac{3}{4}$. Since $[-1, 1]$ is closed in $\mathbb{R}$ with the standard metric (which is a complete space), it forms a complete metric subspace. Thus, the CMT tells us that $f$ has a unique fixed point $\hat{x}$, such that

$$\hat{x} = \frac{1}{4}(\cos(\hat{x}^2) - \hat{x})$$

that is,

$$5\hat{x} = \cos(\hat{x}^2)$$

as required. To approximate the solution to this equation, we can start with $x_1 = 0$ and iterate $f$ starting at $x_1$. We have

$$\begin{aligned} x_1 &= 0 \\ x_2 &= \frac{1}{4} \\ x_3 &\approx 0.18701187767 \\ x_4 &\approx 0.20309415345 \\ x_5 &\approx 0.19901382498 \\ x_6 &\approx 0.20005048500 \\ x_7 &\approx 0.19978720345 \\ &\vdots \end{aligned}$$

This gives $\hat{x} \approx 0.200$ (3 d.p.), and we see that indeed the equation approximately holds for this estimate of $\hat{x}$.

**Example 2.7**
Suppose we want to use the CMT to approximately solve the differential equation

$$f'(x) = f(x)$$

over $[0, 0.5]$. First, we transform this into an integral equation using the FTC by integrating:

$$f(x) - f(0) = \int_0^x f(t)\ dt$$

Rearranging, we get

$$f(x) = f(0) + \int_0^x f(t)\ dt$$

Hence, we construct a function $\varphi : C[0, 0.5] \to C[0, 0.5]$ by

$$\varphi(f)(x) = f(0) + \int_0^x f(t)\ dt$$

We claim without formal proof that indeed this gives a continuous function, and consider the subset of $C[0, 0.5]$ with the same initial value $C$ at 0, equipped with the metric $d_\infty$ given by

$$d(f, g) = \sup_{t \in [0,0.5]} |f(t) - g(t)|$$

Note that for functions $f, g$ we have

$$\begin{aligned}
|\varphi(f)(x) - \varphi(g)(x)| &= |f(0) - g(0) + \int_0^x f(t) - g(t)\ dt| \\
&\leq \int_0^x |f(t) - g(t)|\ dt \\
&\leq \int_0^x \sup_{t \in [0,0.5]} |f(t) - g(t)|\ dt \\
&\leq \int_0^x d(f, g)\ dt \\
&= x d(f, g) \\
&\leq 0.5 d(f, g)
\end{aligned}$$

Hence, $\varphi$ is a contraction with Lipschitz factor 0.5. The space we're working in can be shown to be complete, and therefore $\varphi$ has a unique fixed point $\hat{f}$, found by iterating $\varphi$. So, if we want to evaluate $\hat{f}$ given $C = 1$, we can do so as follows:

1. $f_1(x) = x + 1$

2. $f_2(x) = 1 + \int_0^x t + 1 \, dt = \frac{1}{2}x^2 + x + 1$
3. $f_3(x) = 1 + \frac{1}{2}\int_0^x t^2 + t + 1 \, dt = \frac{1}{6}x^3 + \frac{1}{2}x^2 + x + 1$

As we would expect, as we iterate, this tends towards

$$\sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x$$

which we know to be the solution to this equation, though it would need to be shown that this pointwise limit is the same as the actual limit with respect to the metric.

# 3 Invariants and Further Examples

In this section we examine some of the topological properties that are preserved under continuous functions. For these proofs, we will use exclusively the language of open sets rather than the metrics where possible, since many of these properties can be shown completely from a topological perspective rather than for metric spaces in particular.

**Definition 3.1 (Connectedness)**
A metric space $(X, d)$ is said to be **disconnected** if and only if there exists a continuous surjection $f : X \to \{0, 1\}$ from $(X, d)$ to the discrete space $(\{0, 1\}, d_0)$. Thus, $(X, d)$ is said to be **connected** if and only if it is not disconnected.

**Theorem 3.2**
Let $(X, d)$ be a connected metric space, $(Y, \tilde{d})$ a connected metric space and $f : X \to Y$ a continuous function. Then, the subspace $(f(X), \tilde{d}|_{f(X)})$ of $(Y, \tilde{d})$ is connected.

*Proof.* Assume for a contradiction that $(f(X), \tilde{d}|_{f(X)})$ is disconnected. Then, there exists a continuous surjection $g : f(X) \to \{0, 1\}$. Since $f(X)$ contains all attained values of $f$, it follows that $g \circ f$ is a surjection. Since $f$ and $g$ are both continuous, $g \circ f$ is continuous by 1.6. Thus, we have a continuous surjection from $(X, d)$ to $(\{0, 1\}, d_0)$ meaning $(X, d)$ is disconnected. This is a contradiction, so $(f(X), \tilde{d}|_{f(X)})$ must have been connected. $\square$

**Example 3.3**
It is known that the connected subsets of $(\mathbb{R}, d_2)$ are precisely the intervals. With this, we are able to show that $(S^1, d_2)$ (that is, the unit circle) is connected. Indeed, we construct a function $f : [0, 2\pi] \to S^1$ given by

$$f(\theta) = (\cos(\theta), \sin(\theta))$$

From real analysis, the components of $f$ are both continuous, and this can be shown to be enough for $f$ to also be continuous (we will see this later). We also know that $f$ is surjective, attaining every point on the unit circle, meaning that $f(X) = S^1$. Thus, indeed $(S^1, d_2)$ is connected.

**Definition 3.4 (Compactness)**
A **cover** of a set $X$ is a collection $\mathcal{C}$ of sets such that

$$X \subseteq \bigcup_{\Omega \in \mathcal{C}} \Omega$$

A **subcover** drawn from $\mathcal{C}$ is a collection $\mathcal{C}' \subseteq \mathcal{C}$ of sets from $\mathcal{C}$ which is also a cover of $X$. If $(X, d)$ is a metric space, then this cover is said to be **open** if and only if $\Omega$ is open for all $\Omega \in \mathcal{C}$.

A metric space $(X, d)$ is said to be **compact** if and only if every open cover of $X$ has a finite subcover.

**Theorem 3.5**
Suppose $(X, d)$ is a compact metric space, $(Y, \tilde{d})$ is a metric space and $f : X \to Y$ is continuous. Then, $(f(X), \tilde{d}|_{f(X)})$ is compact.

*Proof.* Let $\mathcal{C}_Y$ be an open cover of $f(X)$, indexed by a set $\Lambda$ as follows:

$$\mathcal{C}_Y = \{\Omega_\lambda \in \mathcal{P}(Y) \mid \lambda \in \Lambda\}$$

Then,

$$f(X) \subseteq \bigcup_{\lambda \in \Lambda} \Omega_\lambda$$

Taking the preimage of both sides under $f$, we get

$$X \subseteq f^{-1}\left(\bigcup_{\lambda \in \Lambda} \Omega_\lambda\right) = \bigcup_{\lambda \in \Lambda} f^{-1}(\Omega_\lambda)$$

The working above is valid using standard results from set theory. Define

$$\mathcal{C}_X := \{f^{-1}(\Omega_\lambda) \in \mathcal{P}(X) \mid \lambda \in \Lambda\}$$

From the above, since $f$ is continuous, $\mathcal{C}_X$ must be an open cover of $X$ (indeed, each member of $\mathcal{C}_X$ is a preimage of an open set under a continuous function). Thus, since $(X, d)$ is compact, $\mathcal{C}_X$ must have a finite subcover, say

$$\mathcal{C}'_X = \{f^{-1}(\Omega_\lambda) \in \mathcal{P}(X) \mid \lambda \in \Lambda'\}$$

for some finite $\Lambda' \subseteq \Lambda$. Now define

$$\mathcal{C}'_Y := \{\Omega_\lambda \in \mathcal{P}(Y) \mid \lambda \in \Lambda'\} \subseteq \mathcal{C}_Y$$

We know that this is a finite subset of $\mathcal{C}_Y$, due to $\Lambda'$ being finite. In order to show that $\mathcal{C}'_Y$ is a finite subcover drawn from $\mathcal{C}_Y$, we now need to verify that it is a cover of $f(X)$. Again, we can use standard results regarding images and preimages from set theory for this.

We have

$$
\begin{aligned}
X &\subseteq \bigcup_{\lambda\in\Lambda'} f^{-1}(\Omega_\lambda) \\
\implies f(X) &\subseteq f\left(\bigcup_{\lambda\in\Lambda'} f^{-1}(\Omega_\lambda)\right) \\
&= \bigcup_{\lambda\in\Lambda'} f(f^{-1}(\Omega_\lambda)) \\
&= \bigcup_{\lambda\in\Lambda'} \Omega_\lambda
\end{aligned}
$$

Hence, $\mathcal{C}'_Y$ is indeed a finite subcover drawn from $\mathcal{C}_Y$. Since $\mathcal{C}_Y$ was arbitrary, it follows that $(f(X), \tilde{d}|_{f(X)})$ must be compact. $\square$

We will now delve into a few more properties related to continuity which may be of use.

**Theorem 3.6**

Let $(X, d)$ and $(Y_1, \tilde{d}_1), \cdots, (Y_n, \tilde{d}_n)$ be metric spaces. Let $Y = Y_1 \times \cdots \times Y_n$, and suppose $(Y, \tilde{d})$ is a metric space equipped with the Euclidean product metric,

$$
\tilde{d}(\vec{u}, \vec{v}) = \left(\sum_{i=1}^{n} \tilde{d}_i(u_i, v_i)^2\right)^{\frac{1}{2}}
$$

If $f_1 : X \to Y_1, f_2 : X \to Y_2, \cdots, f_n : X \to Y_n$ are all continuous, then $f : X \to Y$ given by

$$
f(x) = (f_1(x), \cdots, f_n(x))
$$

is continuous.

*Proof.* Let $\varepsilon > 0$ be given and fix any $x_0 \in X$. By continuity of each $f_i$, there exists $\delta_i > 0$ such that if $d(x, x_0) < \delta_i$ then $\tilde{d}_i(f_i(x), f_i(x_0)) < \frac{\varepsilon}{\sqrt{n}}$. Defining

$$
\delta = \min_{i=1}^{n} \delta_i > 0
$$

we get that, if $d(x, x_0) < \delta$ then

$$
\begin{aligned}
\tilde{d}(f(x), f(x_0)) &= \left(\sum_{i=1}^{n} \tilde{d}_i(f_i(x), f_i(x_0))^2\right)^{\frac{1}{2}} \\
&< \left(\sum_{i=1}^{n} \frac{\varepsilon^2}{n}\right)^{\frac{1}{2}} \\
&= \sqrt{\varepsilon^2} \\
&= \varepsilon
\end{aligned}
$$

It follows that $f$ is continuous at $x_0$. Since $x_0$ was arbitrary, it follows that $f$ is continuous everywhere. $\square$

**Theorem 3.7**

Let $(X, d_0)$ be a discrete metric space and $(Y, d)$ be an arbitrary metric space. Then, every function $f : X \to Y$ is continuous.

*Proof.* Let $\Omega \subseteq Y$ be an open set in $(Y, d)$. Then, $f^{-1}(\Omega) \subseteq X$ is open in $(X, d_0)$, since every set is open with respect to the discrete metric. It follows immediately that $f$ is continuous. $\square$

**Corollary 3.8**

Every discrete space $(X, d_0)$ where $|X| > 1$ is disconnected.

*Proof.* Take any $x_0 \in X$ and define $f : X \to \{0, 1\}$ by

$$f(x) = \begin{cases} 0 & \text{if } x = x_0 \\ 1 & \text{otherwise} \end{cases}$$

We can see that $f$ is a surjection, since $f(x_0) = 0$ and $f(x) = 1$ for some other $x \in X$. Further, $f$ is continuous by 3.7. Hence, $(X, d_0)$ must be disconnected. $\square$