Continuity

Definition 1 (Continuous function) *Let* (X, d_X) *and* (Y, d_Y) *be two metric spaces. Consider* $E \subseteq X$ *and* $P \in E$. *Then* $f : E \to Y$ *is continuous at* p *iff* $\forall \varepsilon > 0$, *there exists* $\delta > 0$ *such that*

$$d_X(x,p) < \delta \Rightarrow d_Y\left(f\left(x\right), f\left(p\right)\right) < \varepsilon.$$

If f *is continuous at* p $\forall p \in E$, *then* f *is continuous.*

Theorem 2 *Let* (X, d_X) *and* (Y, d_Y) *be two metric spaces. A function* $f : E \subseteq X \to Y$ *is continuous at* $p \in E$ *iff for any sequence* $\{x_n\}_{n\in\mathbb{N}} \subseteq E$ *such that* $x_n \overset{d_X}{\to} p$, *it is true that* $f(x_n) \overset{d_Y}{\to} f(p)$.

Proof. "$\Rightarrow$"

Suppose that f is continuous at p, and consider $\varepsilon > 0$. Then

$$d_Y\left(f\left(x\right), f\left(p\right)\right) < \varepsilon$$

if $d_X(x_n, p) < \delta$ for some $\delta > 0$ (because of continuity). Since $x_n \overset{d_X}{\to} p$, there exists some $N \in \mathbb{N}$ such that $\forall n > N$, $d(x_n, p) < \delta$. Then there exists some N such that $\forall n > N$,

$$d_Y\left(f\left(x_n\right), f\left(p\right)\right) < \varepsilon \Rightarrow f(x_n) \overset{d_Y}{\to} f(p).$$

"$\Leftarrow$"

Suppose that every sequence $\{x_n\}_{n\in\mathbb{N}} \subseteq E$ such that $x_n \overset{d_X}{\to} p$ has the characteristic that $f(x_n) \overset{d_Y}{\to} f(p)$, and that f is not continuous. Then there exists some $\varepsilon > 0$ such that $\forall \delta > 0$,

$$d_X(x,p) < \delta \text{ and } d_Y\left(f\left(x\right), f\left(p\right)\right) \geq \varepsilon \text{ for some } x.$$

Consider $\varepsilon > 0$ and set $\delta = 1$. Then there exists x_1 such that $d_X(x_1, p) < 1$ and $d_Y(f\left(x_1\right), f\left(p\right)) \geq \varepsilon$. Now, for $\delta = \frac{1}{2}$, there exists x_2 such that $d_X(x_2, p) < \frac{1}{2}$ and $d_Y(f\left(x_2\right), f\left(p\right)) \geq \varepsilon$. For $\delta = \frac{1}{n}$, there exists x_n such that $d_X(x_n, p) < \frac{1}{n}$ and $d_Y(f\left(x_n\right), f\left(p\right)) \geq \varepsilon$. Then $x_n \overset{d_X}{\to} p$ but $f(x_n)$ does not converge to $f(p)$, which is a contradiction. So f is continuous. ∎

Example 3 *Let* $f : \mathbb{R}^2 \to \mathbb{R}$, $(x,y) \to \begin{cases} \frac{xy}{x^2+y^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$. *Is this function continuous at* $(0,0)$? *Consider* $x_n = \left(\frac{1}{n}, \frac{1}{n}\right)$, *where* $x_n \to (0,0)$. *Also, note that* $f(x_n) = \frac{\frac{1}{n}\cdot\frac{1}{n}}{\frac{1}{n^2}+\frac{1}{n^2}} = \frac{1}{2}$, *so* $f(x_n) \to \frac{1}{2} \neq 0 = f(0,0)$. *Then,* f *is not continuous.*

Remark 4 *Theorem 2 stated that* f *was continuous iff* $x_n \to p \Rightarrow f(x_n) \to f(p)$. *This justifies operations of the type*

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

For example, $\lim_{n\to\infty} 2^{\frac{1}{n}} = 2^{\lim_{n\to\infty}\frac{1}{n}} = 2^0 = 1$.

Theorem 5 *Let* (X, d_X), (Y, d_Y) *and* (Z, d_Z) *be metric spaces and* $f : X \to Y$, $g : Y \to Z$ *continuous. Then* $g \circ f : X \to Z$ *is also continuous.*

Proof. There are two ways to prove this theorem.

(a) Note that:

$$\begin{aligned} d_Z\left(g\circ f\left(x\right), g\circ f\left(p\right)\right) &= d_Z(g(f(x)), g(f(p)) \\ &= d_Z(g(a), g(b)), \end{aligned}$$

where the last step equaled $g\left(f\left(x\right)\right) = g(a)$ and $g\left(f\left(p\right)\right) = g(b)$. Since g is continuous, there exists some δ_1 such that

$$d_Y\left(f\left(x\right), f\left(p\right)\right) < \delta_1 \Rightarrow d_Z\left(g\left(f\left(x\right)\right), g\left(f\left(p\right)\right)\right) < \varepsilon.$$

Since f is continuous, there exists some $\delta_2 > 0$ such that

$$d_X(x,p) < \delta_2 \Rightarrow d_Y(f(x), f(p)) < \delta_1$$

where δ_1 plays the role of ε. Then, considering $\delta = \delta_2$,

$$d_X(x,p) < \delta \Rightarrow d_Z(g(f(x)), g(f(p))) < \varepsilon.$$

(b) We know that $x_n \to p$. Then, by continuity of f, $f(x_n) \to f(p)$. Then, by continuity of g, $g(f(x)) \to g(f(p)) \Rightarrow g \circ f(x_n) \to g \circ f(p)$. ■

Proposition 6 *Consider a sequence $\{x_n\}_{n\in\mathbb{N}}$ in (X,d). The following propositions are equivalent:*
(a) There exists a subsequence $\{x_{k_n}\}_{n\in\mathbb{N}}$ such that $x_{k_n} \to x$, where x is a cluster point.
(b) $\forall \varepsilon > 0$, $\forall n \in \mathbb{N}$, there exists $n > N$ such that $d(x_n, x) < \varepsilon$.

Proof. "(b) $\Rightarrow$ (a)"

It is possible to construct a sequence $\{x_{k_n}\}_{n\in\mathbb{N}}$ which converges to x. Consider $\varepsilon = 1$, $N = 1$. Then there exists some $n > 1$ such that $d(x_n, x) < 1$. Consider $k_1 = n$. Then $d(x_{k_1}, x) < 1$. Now consider $\varepsilon = \frac{1}{2}$, $N = k_1$. Then there exists some $n > k_1$ such that $d(x_n, x) < \frac{1}{2}$. Consider $k_2 = n$. Then $d(x_{k_2}, x) < \frac{1}{2}$. This process can be repeated up to where $x_{k_{n-1}}$ has been chosen. Then consider $\varepsilon = \frac{1}{n}$ and $N = k_{n-1}$. Then there exists some $n > k_{n-1}$ such that $d(x_n, x) < \frac{1}{n}$. Consider $k_n = n$. Then $d(x_n, x) < \frac{1}{n}$ ... It is clear that $\{x_{k_n}\}_{n\in\mathbb{N}}$ that has been constructed is a subsequence of $\{x_n\}_{n\in\mathbb{N}}$, and $d(x_{k_n}, x) < \frac{1}{n}$. Then $x_{k_n} \to x$.

"(a) $\Rightarrow$ (b)"

Consider any $\varepsilon > 0$ and any $N \in \mathbb{N}$. Then there exists some $N_1 \in \mathbb{N}$ such that $\forall n > N_1$, $d(x_{k_n}, x) < \varepsilon$, so for $n > \max\{N, N_1\}$, $d(x_{k_n}, x) < \varepsilon$. If $n > N$ then $k_n > N$ and $d(x_j, x) < \varepsilon$ and $j > N$. ■

Proposition 7 *Let (S,d) be a metric space, $\{x_n\}_{n\in\mathbb{N}}$, $\{y_n\}_{n\in\mathbb{N}}$ Cauchy sequences in S. Then $\{d(x_n, y_n)\}_{n\in\mathbb{N}}$ converges in $(\mathbb{R}, |\cdot|)$.*

Proof. Using the triangular inequality (twice):

$$d(x_n, y_n) \leq d(x_n, x_m) + d(x_m, y_m) + d(y_m, y_n). \tag{1}$$

Then

$$d(x_n, y_n) - d(x_m, y_m) \leq d(x_n, x_m) + d(y_m, y_n). \tag{2}$$

Note that since $\{x_n\}_{n\in\mathbb{N}}$ and $\{y_n\}_{n\in\mathbb{N}}$ are Cauchy, the components in the right hand side are small after some N. From (1) we have that

$$d(x_m, y_m) - d(x_n, y_n) \geq -d(x_n, x_m) - d(y_m, y_n). \tag{3}$$

Switching m and n:

$$d(x_n, y_n) - d(x_m, y_m) \geq -d(x_n, x_m) - d(y_m, y_n), \tag{4}$$

where in the right hand side of (4) nothing happens because of the symmetry of $d(\cdot)$. From (2) and (4):

$$-d(x_n, x_m) - d(y_m, y_n) \leq d(x_n, y_n) - d(x_m, y_m) \leq d(x_n, x_m) + d(y_m, y_n)$$

$$\Leftrightarrow |d(x_n, y_n) - d(x_m, y_m)| \leq d(x_n, x_m) + d(y_m, y_n).$$

Consider $\varepsilon > 0$. Then there exists $N_1 \in \mathbb{N}$ such that $\forall n, m > N_1$,

$$d(x_n, x_m) < \frac{\varepsilon}{2}.$$

Also, there exists $N_2 \in \mathbb{N}$ such that $\forall n, m > N_2$,

$$d(y_n, y_m) < \frac{\varepsilon}{2}.$$

Let $N = \max\{N_1, N_2\}$. Then, $\forall n, m > N$:

$$\begin{aligned} |d(x_n, y_n) - d(x_m, y_m)| &\leq d(x_n, x_m) + d(y_m, y_n) \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon. \end{aligned}$$

Then $\{d(x_n, y_n)\}_{n\in\mathbb{N}}$ is a Cauchy sequence, which converges in $\mathbb{R}$. ■

Theorem 8 (Contraction Mapping Theorem) *Let (S,d) be a complete metric space and $f : S \to S$ such that $d(f(x), f(y)) \leq \beta d(x,y)$ $\forall x, y \in S$ and $0 < \beta < 1$. Then there exists x^* such that $f(x^*) = x^*$. Moreover, x^* is unique.*

Proof. Consider the sequence $\{x_n\}_{n\in\mathbb{N}}$ defined by:

$x_1 = x_1$
$x_2 = f(x_1)$
$x_3 = f(x_2) = f^2(x_1)$
$\vdots$
$x_n = f(x_n) = f^n(x_1)$.

The proof requires the following 4 steps.

(a) Consider

$$d(x_n, x_{n+1}) = d(f(x_{n-1}), f(x_n))$$

But using the fact that $d(f(x), f(y)) \leq \beta d(x,y)$,

$$\begin{aligned} d(x_n, x_{n+1}) = d(f(x_{n-1}), f(x_n)) &\leq \beta d(x_{n-1}, x_n) \\ &= \beta d(f(x_{n-2}), f(x_{n-1})) \\ &\leq \beta^2 d(x_{n-2}, x_{n-1}) \\ &\vdots \\ &\leq \beta^{n-1} d(x_1, x_2). \end{aligned} \tag{5}$$

Consider $m > n$. Then, by the triangular inequality,

$$d(x_n, x_m) \leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + ... + d(x_{m-1}, x_m).$$

Using (5):

$$\begin{aligned} d(x_n, x_m) &\leq \beta^{n-1} d(x_1, x_2) + \beta^n d(x_1, x_2) + ... + \beta^{m-2} d(x_1, x_2). \\ &= \beta^{n-1} d(x_1, x_2)\left[1 + \beta + \beta^2 + ... + \beta^{(m-2)-(n-1)}\right] \\ &\leq \beta^{n-1} d(x_1, x_2)\left[1 + \beta + \beta^2 + ...\right], \end{aligned}$$

since the infinite series is greater than or equal than the finite series. Notice that $\left(1 + \beta + \beta^2 + ...\right) = \frac{1}{1+\beta}$, then:

$$d(x_n, x_m) \leq \beta^{n-1}\left[\frac{d(x_1, x_2)}{1 - \beta}\right]$$

but since the term in $[\cdot]$ is fixed, then there exists some N such that the right hand side goes to zero. Since $\beta^n \to 0$, there exists $N \in \mathbb{N}$ such that

$$|\beta^n| < \frac{\varepsilon(1-\beta)}{d(x_1, x_2)}.$$

Then $\forall m, n > N + 1$,

$$\begin{aligned} d(x_n, x_m) &\leq \beta^{n-1}\left[\frac{d(x_1, x_2)}{1 - \beta}\right] \\ &< \frac{\varepsilon(1-\beta)}{d(x_1, x_2)} \cdot \frac{d(x_1, x_2)}{1 - \beta} \\ &= \varepsilon \end{aligned}$$

then $\{x_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence.

(b) The sequence $\{x_n\}_{n\in\mathbb{N}}$ converges to x^* since (S,d) is a complete metric space.

(c) By the triangular inequality,

$$d(f(x^*), x^*) \leq d(f(x^*), x_n) + d(x_n, x^*)$$

$$= d\left(f\left(x^*\right), f\left(x_{n-1}\right)\right) + d\left(x_n, x^*\right)$$
$$\leq \beta d\left(x^*, x_{n-1}\right) + d\left(x_n, x^*\right)$$
$$< d\left(x^*, x_{n-1}\right) + d\left(x_n, x^*\right)$$

since $0 < \beta < 1$. Now consider $\varepsilon > 0$. Then there exists some $\bar{N}$ such that $\forall n > \bar{N}$,

$$d(x_n, x^*) < \frac{\varepsilon}{2}$$

since by (b), $x_n \to x^*$. Let $N = \bar{N} + 1$, then $\forall n > N$,

$$d(x^*, x_{n-1}) < \frac{\varepsilon}{2}$$

so

$$d\left(f\left(x^*\right), x^*\right) < \varepsilon.$$

Since this is true $\forall \varepsilon > 0$, we can conclude that $d\left(f\left(x^*\right), x^*\right) = 0$ and $f\left(x^*\right) = x^*$.
(d) Consider x^* such that $f\left(x^*\right) = x^*$, and $\bar{x}^*$ such that $f\left(\bar{x}^*\right) = \bar{x}^*$. However, note that:

$$d(x^*, \bar{x}^*) = d\left(f\left(x^*\right), f\left(\bar{x}^*\right)\right)$$
$$\leq \beta d\left(x^*, \bar{x}^*\right), \text{ where } \beta < 1.$$

Then $d(x^*, \bar{x}^*) = 0$ and $x^* = \bar{x}^*$. ■