The definition of continuity (as stated in Lecture 1 for functions from $\mathbb{R}^2$ to $\mathbb{R}^2$) makes sense for functions from any metric space $(X, d)$ to any other metric space $(Y, d')$.

A function $f: X \to Y$ is continuous at the point $a \in X$ if for every real number $\varepsilon > 0$ there exists a real number $\delta > 0$ such that the following condition holds:

for all $x \in X$, if $d(x, a) < \delta$ then $d'(f(x), f(a)) < \varepsilon$.

Using the concept of “open ball”, this can be rephrased as follows:

A function $f: X \to Y$ is continuous at $a \in X$ if and only if for every open ball $B$ with centre at $f(a)$ there is an open ball $C$ with centre $a$ such that $f(C) \subseteq B$.

Note that the condition $f(C) \subseteq B$ is equivalent to $C \subseteq f^{-1}(B)$. (This is easy to prove: it follows immediately from the definitions of “image” and “preimage”.)

The following proposition generalizes the above statement slightly.

**Proposition.** Let $(X, d)$, $(Y, d')$ be metric spaces and $f: X \to Y$ a function, and let $a \in X$. Then $f$ is continuous at $a$ if and only if for every open subset $U$ of $Y$ with $a \in f^{-1}(U)$ there is an open ball $C$ with centre $a$ such that $C \subseteq f^{-1}(U)$.

**Proof.** Suppose first that $f$ satisfies the stated condition; we shall show that $f$ is continuous at $a$.

Let $\varepsilon > 0$. Then $U = B(f(a), \varepsilon)$ is an open subset of $Y$, and $a \in f^{-1}(U)$ (since $f(a) \in U$). So by the given condition there exists an open ball $C$ centred at $a$ such that $C \subseteq f^{-1}(U)$. Let $\delta$ be the radius of $C$ (so that $C = B(a, \delta)$). Now if $x$ is an arbitrary element of $X$ satisfying $d(x, a) < \delta$, then

$$x \in C \subseteq f^{-1}(U),$$

whence $f(x) \in U = B(f(a), \varepsilon)$, which means that $d'(f(x), f(a)) < \varepsilon$.

Thus we have have shown that for every $\varepsilon > 0$ there exists $\delta > 0$ such that, for all $x \in X$, if $d(x, a) < \delta$ then $d(f(x), f(a)) < \varepsilon$. That is, we have shown that $f$ is continuous at $a$.

Conversely, suppose that $f$ is continuous at $a$, and let $U$ be an open subset of $Y$ such that $a \in f^{-1}(U)$. Since $U$ is open and $f(a) \in U$ there is an $\varepsilon > 0$ such that $B(f(a), \varepsilon) \subseteq U$. Since $f$ is continuous at $a$ there exists $\delta > 0$ such that, for all $x \in X$, if $d(x, a) < \delta$ then $d'(f(x), f(a)) < \varepsilon$. Now put $C = B(a, \delta)$, an open ball centred at $a$. For all $x \in C$ we have $d(x, a) < \delta$, which gives $d'(f(x), f(a)) < \varepsilon$, and hence $f(x) \in B(f(a), \varepsilon) \subseteq U$. So $x \in f^{-1}(U)$ whenever $x \in C$; in other words, $C \subseteq f^{-1}(U)$. Thus we have shown that for every open set $U$ containing $f(a)$ there is an open ball centred at $a$ and contained in $f^{-1}(U)$, as required.

In view of the definition of the interior of a set, we can restate the above result as follows.

**Corollary.** The function $f: X \to Y$ is continuous at $a$ if and only if, for all open subsets $U$ of $Y$, if $a \in f^{-1}(U)$ then $a \in \text{Int}(f^{-1}(U))$.

This enables us to now give a concise characterization of continuous functions.

**Corollary.** If $(X, d)$ and $(Y, d')$ are metric spaces then a function $f: X \to Y$ is continuous if and only if $f^{-1}(U)$ is an open subset of $X$ whenever $U$ is an open subset of $Y$.

**Proof.** To say that $f$ is continuous is to say that it is continuous at all points $a \in X$. By the previous corollary, this holds if and only if for all open $U \subseteq Y$ and all $a \in X$, if
\( a \in f^{-1}(U) \) then \( a \in \text{Int}(f^{-1}(U)) \). That is, for every open \( U \subseteq Y \), all points of \( f^{-1}(U) \) are interior points. But to say that all points of \( f^{-1}(U) \) are interior points is to say that \( f^{-1}(U) \) is open. \( \square \)

**Some inequalities**

Suppose that \( 0 \leq \theta \leq 1 \). If \((x_0, y_0)\) and \((x_1, y_1)\) are points in \( \mathbb{R}^2 \) then the point \((x, y)\) defined by

\[
\begin{align*}
x &= \theta x_0 + (1 - \theta)x_1 \\
y &= \theta y_0 + (1 - \theta)y_1
\end{align*}
\]

lies on the line segment joining \((x_0, y_0)\) and \((x_1, y_1)\). Now the graph of \( y = \ln x \) is concave downwards; so if \((x_0, y_0)\) and \((x_1, y_1)\) are on this graph then \((x, y)\) will be below it; that is, \( y \leq \ln x \). In other words, if \( a, b > 0 \) and we define

\[
\begin{align*}
x_0 &= a \\
y_0 &= \ln a \\
x_1 &= b \\
y_1 &= \ln b
\end{align*}
\]

so that

\[
\begin{align*}
x &= \theta a + (1 - \theta)b \\
y &= \theta (\ln a) + (1 - \theta)(\ln b)
\end{align*}
\]

then it follows that

\[
\theta (\ln a) + (1 - \theta)(\ln b) \leq \ln(\theta a + (1 - \theta)b).
\]

Taking exponentials of both sides, using the fact that \( e^x \) is monotone increasing, it follows that \( e^{\theta \ln a + (1 - \theta)\ln b} \leq \theta a + (1 - \theta)b \).

But \( e^{\theta \ln a + (1 - \theta)\ln b} = e^{\theta \ln a} e^{(1 - \theta)\ln b} = a^\theta b^{1 - \theta} \); so we have shown that

\[
\theta a^\theta b^{1 - \theta} \leq \theta a + (1 - \theta)b.
\]

for all \( a, b > 0 \). The same in fact holds for \( a, b \geq 0 \), since if either \( a \) or \( b \) is zero then the left hand side is zero, while the right hand side remains nonnegative.

**Hölder’s Inequality.** Let \( p > 1 \) and put \( q = p/(p - 1) \) (so that \( q > 1 \) and \( \frac{1}{p} + \frac{1}{q} = 1 \)). Let \( a_k, b_k \) be arbitrary complex numbers, where \( k \) runs from \( 1 \) to \( n \). Then

\[
\sum_{k=1}^{n} |a_k b_k| \leq \left( \sum_{k=1}^{n} |a_k|^p \right)^{1/p} \left( \sum_{k=1}^{n} |b_k|^q \right)^{1/q}.
\]

**Proof.** Let \( c_k = |a_k|^p \) and \( d_k = |b_k|^q \), and put \( C = \sum_{k=1}^{n} c_k \) and \( D = \sum_{k=1}^{n} d_k \). Put \( \theta = 1/p \), so that \( 1 - \theta = 1/q \), and apply \((*)\) with \( c_k/C \) in place of \( a \) and \( d_k/D \) in place of \( b \). We obtain

\[
(c_k/C)^{1/p}(d_k/D)^{1/q} \leq (1/p)(c_k/C) + (1/q)(d_k/D).
\]
Summing from \( k = 1 \) to \( n \) gives
\[
\sum_{k=1}^{n} \frac{c_k^{1/p} d_k^{1/q}}{C^{1/p} D^{1/q}} \leq \frac{1}{pC} \sum_{k=1}^{n} c_k + \frac{1}{qD} \sum_{k=1}^{n} d_k
\]
\[
= \frac{1}{p} + \frac{1}{q} = 1.
\]
Hence \( \sum_{k=1}^{n} c_k^{1/p} d_k^{1/q} \leq C^{1/p} D^{1/q} \); that is,
\[
\sum_{k=1}^{n} |a_k b_k| \leq \left( \sum_{k=1}^{n} |a_k|^p \right)^{1/p} \left( \sum_{k=1}^{n} |b_k|^q \right)^{1/q}
\]
as required. \( \square \)

The special case of Hölder’s Inequality in which \( p = q = 2 \) is known as Cauchy’s Inequality.

**Minkowski’s Inequality.** Let \( p \geq 1 \), and let \( a_k, b_k \in C \) be arbitrary. Then
\[
\left( \sum_{k=1}^{n} |a_k + b_k|^p \right)^{1/p} \leq \left( \sum_{k=1}^{n} |a_k|^p \right)^{1/p} + \left( \sum_{k=1}^{n} |b_k|^p \right)^{1/p}.
\]

**Proof.** Since \( |a + b| \leq |a| + |b| \) for all complex numbers \( a \) and \( b \), it is clear that the result holds for \( p = 1 \). So we assume that \( p > 1 \). Put \( q = p/(p-1) \).

For all \( k \) from \( 1 \) to \( n \) we have
\[
(a_k + b_k)^p = a_k(a_k + b_k)\,^{p-1} + b_k(a_k + b_k)^{p-1}
\]
and so using standard properties of the modulus function for complex numbers (namely \( |ab| = |a||b| \) and \( |a + b| \leq |a| + |b| \) for all \( a, b \in \mathbb{C} \), and \( |a^t| = |a|^t \) for all \( a \in \mathbb{C} \) and \( t \in \mathbb{R} \)) we deduce that
\[
|a_k + b_k|^p \leq |a_k|(|a_k + b_k|)^{p-1} + |b_k|(|a_k + b_k|)^{p-1}
\]
for all \( k \). Summing from \( k = 1 \) to \( n \), and then applying Hölder’s Inequality to each of the sums on the right hand side gives
\[
\sum_{k=1}^{n} |a_k + b_k|^p \leq \sum_{k=1}^{n} |a_k|(|a_k + b_k|)^{p-1} + \sum_{k=1}^{n} |b_k|(|a_k + b_k|)^{p-1}
\]
\[
\leq \left( \sum_{k=1}^{n} |a_k|^p \right)^{1/p} \left( \sum_{k=1}^{n} (|a_k + b_k|)^{(p-1)q} \right)^{1/q} + \left( \sum_{k=1}^{n} |b_k|^p \right)^{1/p} \left( \sum_{k=1}^{n} (|a_k + b_k|)^{(p-1)q} \right)^{1/q}
\]
\[
= \left( \left( \sum_{k=1}^{n} |a_k|^p \right)^{1/p} + \left( \sum_{k=1}^{n} |b_k|^p \right)^{1/p} \right) \left( \sum_{k=1}^{n} (|a_k + b_k|)^{p-1} \right)^{1/q},
\]
where in the last line we have used \( (p-1)q = p \). Dividing through by the second factor on the right hand side gives
\[
\left( \sum_{k=1}^{n} |a_k + b_k|^p \right)^{1-(1/q)} \leq \left( \sum_{k=1}^{n} |a_k|^p \right)^{1/p} + \left( \sum_{k=1}^{n} |b_k|^p \right)^{1/p},
\]
which is the required result, since \( 1 - \frac{1}{q} = \frac{1}{p} \). \( \square \)