Before proceeding to the next topic, we’ll prove a completely obvious Fact about sequences and subsequences. (We could call it a Lemma, but that’s glorifying it.) Recall that to show that \((y_n)\) is a subsequence of \((x_n)\), you always have to show the existence of a strictly increasing function \(\tau : \mathbb{N} \to \mathbb{N}\) such that for all \(n \in \mathbb{N}\), \(y_n = x_{\tau(n)}\). Now note:

**Fact**: If \(\tau : \mathbb{N} \to \mathbb{N}\) is a strictly increasing mapping, then for all \(n\), \(\tau(n) \geq n\).

To prove this, we’ll argue by induction.
• **Initial step:** $\tau(1) \geq 1$ (duh).

• **Inductive step:** suppose that $\tau(n) \geq n$; then $\tau(n + 1) \geq n + 1$.

  *Proof of the Inductive step:* Let $\tau(n) = k \in \mathbb{N}$, $k \geq n$. Since $\tau(n + 1) > \tau(n)$ and $\tau(n + 1) \in \mathbb{N}$,

  $$\tau(n + 1) \geq k + 1 \geq n + 1.$$  

1.9. **Continuous Functions**

A function is continuous if it maps nearby points to nearby points. Draw the graph without taking pen off paper. Graph is connected. Formally:

**Definition:** Consider $f : X \to \mathbb{R}^k$. Fix $x_0 \in X$. The function $f$ is **continuous at** $x_0$ if whenever $\{x_m\}_{m=1}^\infty$ is a sequence in $X$ which converges to $x_0$, then $\{f(x_m)\}_{m=1}^\infty$ converges to $f(x_0)$. The function $f$ is **continuous** if it is continuous at $x$, for every $x \in X$.

While this definition seems obvious and intuitive, all is not what it seems to be. We’ve noticed ad nauseum that whether or not a sequence converges depends on the metric and the universe. In this section, we now have *two* different sequences, one in the domain and one in the range; so we have to worry about what kinds of things converge in the domain and what kinds of things converge in the range. So we have roughly double the number of counter-intuitive possibilities. To see what kinds of things can happen, consider the following question:

**Question:** let $f : X \to \mathbb{R}^k$ and let $X$ be endowed with the discrete metric. What can we say about the continuity of $f$?

**Answer:** $f$ is continuous.
In other words, if $X$ is endowed with the discrete metric, then *every* function is continuous! The reason for this is that the discrete metric makes convergence an *extremely* stringent requirement: $(x_n)$ converges to $x$, iff, eventually, the sequence is constant at $x$. But whenever this stringent condition is satisfied, the resulting sequence in the range, $\{f(x_n)\}$ is eventually constant at $f(x)$. That is, convergence in the domain is so difficult to accomplish that whenever it is accomplished, convergence in the range is assured.

Now consider the reverse question:

**Question:** let $f : X \to \mathbb{R}^k$ and let $\mathbb{R}^k$ be endowed with the discrete metric. What can we say about the continuity of $f$.

**Answer:** (first attempt). An obvious first shot at an answer to this question is: $f$ is continuous iff it is constant. The “if” part is right but the “only if” part is very wrong.

(1) It matters what the metric is on $X$. If the metric on $X$ is the discrete metric, then, as we just saw, *every* function is continuous.

(2) Now let’s assume that the metric on $X$ is the Pythagorian and restrict attention to $\mathbb{R}^1$. In this case, there are non-constant functions which are continuous, provided that the domain has “holes” in it. For example, consider the function $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$, defined by

$$f(x) = \begin{cases} 
1 & \text{if } x > 0 \\
-1 & \text{if } x < 0 
\end{cases}$$

In this case, there’s a “hole” in the domain at zero. To see what role the “hole” plays, consider a sequence like $x_n = (-1)^n/n$, which has the apparently discontinuous property that $f(x_n)$ bounces up and down between 1 and $-1$. It looks like this sequence ought to fail the requirement that the definition of continuity imposes, even with the *regular* metric
on \( \mathbb{R} \), and most certainly with the discrete metric on \( \mathbb{R} \). But, since 0 is not the limit of this sequence (0 simply doesn’t exist in this world), the sequence \((x_n)\) doesn’t converge—indeed, it doesn’t even have any subsequences that converge—so the definition of continuity imposes absolutely no restriction on this sequence.

**Answer:** (second (and correct) attempt). If \( f : X \to \mathbb{R}^k \) and \( \mathbb{R}^k \) is endowed with the discrete metric, then \( f \) is continuous iff \( f \) is constant on every connected subset of \( X \). I haven’t defined connected yet, but, intuitively, a set is connected if it is, well, connected. More precisely, a set \( C \subset \mathbb{R} \) is connected if for any \( x, y \in C \), the entire line segment connecting \( x \) to \( y \) is also in \( C \). In the example given by display (1) above, the domain \( X \) has two connected subsets, i.e., \( \mathbb{R}^{++} \) and \( \mathbb{R}^{--} \).

1.9.1. **Continuous functions on compact sets attain extreme values.** For us, the most important result relating to continuity is the following theorem, sometimes know as the “extreme value theorem.”

**Theorem:** (Weierstrass) Consider a function \( f : X \to \mathbb{R}^1 \), where both \( X \) and \( \mathbb{R} \) are endowed with the Pythagorian metric. If \( X \) is a compact set and \( f \) is continuous on \( X \), then \( f \) attains a global maximum and a global minimum on \( X \).

Before proving the theorem, we need the following lemma.

**Lemma:** Consider \( f : X \to \mathbb{R} \), where both \( X \) and \( \mathbb{R} \) are endowed with the Pythagorian metric. If \( f \) is continuous and \( X \) is compact, then \( f \) is bounded.

**Proof of the Lemma:** We’ll just show that the function is bounded above, by proving that if \( X \) is compact and \( f \) isn’t bounded, then \( f \) cannot be continuous. Assume that \( f \) isn’t bounded, i.e., for all \( m \in \mathbb{N} \), \( \exists x_m \) such that \( f(x_m) > m \). Since \( X \) is compact, the sequence \( \{x_n\} \) contains a convergent subsequence. Call this subsequence \( \{y_n\} \) and let \( y \in X \) denote its limit. Define \( \tau : \mathbb{N} \to \mathbb{N} \) by
\[ y_n = x_{\tau(n)} \] for all \( n \). Since \( f \) is defined on \( X \), \( f(y) \in \mathbb{R} \), that is, \( f(y) < N \), for some \( N \in \mathbb{N} \). Now pick \( n \geq N + 1 \) and note that by the Fact above, \( \tau(n) \geq n \geq N + 1 \). Moreover, by assumption, \( f(y_n) = f(x_{\tau(n)}) \geq \tau(n) \geq N + 1 \). Hence for all \( n > N + 1 \), \( f(y_n) - f(y) > 1 \), so that \( f \) is not continuous at \( y \). Similarly, \( f \) is bounded below.

Coming back to the Weierstrass theorem, here’s a sketch of the proof.

- show that the image of the function must be bounded.
- let \( \bar{f} \) denote the supremum of the image of the function.
- pick a sequence \( \{x_n\} \) such that the sequence \( \{f(x_n)\} \) gets closer to the supremum.
- while the sequence \( \{x_n\} \) needn’t converge, it follows from the compactness of \( X \) that there must exist a subsequence \( \{y_n\} \) of \( \{x_n\} \) such that \( \{y_n\} \) converges to \( y \in X \).
- since \( f \) is continuous, the sequence \( \{f(y_n)\} \) must converge to \( f(y) \). But by defn of the supremum, \( \{f(y_n)\} \) converges also to \( \bar{f} \). Hence \( f(y) = \bar{f} \).
- Since \( \bar{f} \) is the supremum of the image of \( X \) under \( f \), then \( f(y) = \bar{f} \geq f(x) \), for all \( x \in X \).

**Proof:** Let \( \bar{f} \) denote the supremum of the image of \( X \) under \( f \), i.e., the set \( \{f(x) : x \in X\} \). By the Lemma above, \( \bar{f} \in \mathbb{R} \). By definition of the supremum, for all \( n \), there exists \( x_n \) such that \( f(x_n) > \bar{f} - 1/n \). Since \( X \) is compact, the sequence \( \{x_n\} \) contains a convergent subsequence. Call this subsequence \( \{y_n\} \) and let \( y \in X \) denote its limit. Since \( f \) is continuous, the sequence \( \{f(y_n)\} \) converges to \( f(y) \in \mathbb{R} \). To complete the proof, we’ll show that \( \{f(y_n)\} \) also converges to \( \bar{f} \). Since a sequence has at most one limit\(^1\), this will imply that \( f(y) = \bar{f} \geq f(x) \), for all \( x \in X \), and hence imply that \( f \) attains a global maximum at \( y \).

Since \( \{y_n\} \) is a subsequence of \( \{x_n\} \), there exists a strictly increasing mapping \( \tau : \mathbb{N} \rightarrow \mathbb{N} \) such that for all \( n \), \( y_n = x_{\tau(n)} \). To prove that \( \{f(y_n)\} \) converges to \( \bar{f} \), we need to show that for all \( \epsilon > 0 \), there

\(^1\) We have noted this before but haven’t proved it. It’s a good exercise to prove it.
exists $N \in \mathbb{N}$ such that for all $n > N$, $f(y_n) \in B(\bar{f}, \epsilon)$. Note first that for all $n$, $f(y_n) \leq \bar{f}$, since $\bar{f}$ is an upper bound for the set $\{f(x) : x \in X\}$. Moreover, it follows from the construction of $\{x_n\}$ that for all $n$, $f(y_n) = f(x_{\tau(n)}) > \bar{f} - 1/\tau(n) \geq \bar{f} - 1/n$ (by the above Fact). It follows therefore that for an arbitrarily chosen $\epsilon > 0$, we can pick $N_\epsilon \in \mathbb{N}, N_\epsilon > 1/\epsilon$, so that $1/\tau(N_\epsilon) \leq 1/N_\epsilon < \epsilon$.

To complete the proof, observe that for all $n > N_\epsilon$, $f(y_n) \in (\bar{f} - 1/\tau(n), \bar{f}] \subset (\bar{f} - 1/\tau(N_\epsilon), \bar{f}] \subset (\bar{f} - \epsilon, \bar{f} + \epsilon) = B(\bar{f}, \epsilon)$

So we have, $(f(y_n)) \to f(y) \in \mathbb{R}$ and $(f(y_n)) \to \bar{f} \in \mathbb{R}$. Since a function can have only one limit, $f(y) = \bar{f}$. Since $\bar{f}$ is an upper bound for $\{f(x) : x \in X\}$, we now have $f(y) \geq f(x)$, for all $x \in X$.

A common source of puzzlement is: since $f(x_n)$ already converges to $\bar{f}$, why do I need to pick a subsequence $(y_n)$ and show that $f(y_n)$ also converges to $\bar{f}$? The reason is that $(x_n)$ doesn’t necessarily converge to some $x \in X$, so I can’t invoke continuity to establish that $f(x_n)$ converges to $f(x)$, for some $x \in X$. Here are two examples that illustrate conclusively (I hope) why you absolutely have to pick the subsequence.

(1) Consider the function $f$ and sequence $(x_n)$ graphed in Fig. 1 below. The example illustrates the point that even though the sequence $(f(x_n))$ converges to $\bar{f}$, the sequence $(x_n)$ doesn’t converge. However, $(x_n)$ has two convergent subsequences, each of which work fine.

(2) Let $X = (0,1)$ and consider $f : X \to \mathbb{R}$, defined by $f(x) = x$. Clearly $f$ doesn’t attain a maximum on $X$. The theorem doesn’t apply because $X = (0,1)$ isn’t compact. Here’s where the proof would break down if we tried to apply it. Using the notation of the proof, $\bar{f} = 1$. Pick $x_n = 1 - 1/(n + 1)$ and note that for all $n$, $f(x_n) > \bar{f} - 1/n$. But we can’t go past this point, because without compactness, we can’t pick a subsequence $(y_n)$ and $y \in X$
Figure 1. Why \((x_n)\) isn’t enough: you need to pick a subsequence \((y_n)\)

such that \(y_n \to y\). The point of the example is that having the sequence \(x_n\) such that the

\(f(x_n)\)'s approach \(\bar{f}\) doesn’t do us much good, without further help.

Our next result establishes a useful alternative definition of continuity.

Definition: Given a mapping \(f : X \to Y\), and \(O \subset Y\), \(f^{-1}(O)\) is the subset of \(X\) that \(f\) maps into \(O\), i.e., \(f^{-1}(O) = \{x \in X : f(x) \in O\}\). \(f^{-1}(O)\) is called the inverse image of \(O\) under \(f\).

The result is that a function is continuous iff the inverse image of every open subset of the range of the function is an open set in the domain.

Theorem: A function \(f : X \to Y\) is continuous iff for every open set \(O \subset Y\), \(f^{-1}(O)\) is an open subset of \(X\).
Before doing the proof, I’ll state the argument in the formal language of logic.

First I define three “statements”

(1) given a function $f$, the statement $C(f)$ is true if $f$ is continuous
(2) given a set $O$, the statement $A(O)$ is true if $O$ is open
(3) given a function-set pair $(f, O)$, the statement $B(f, O)$ is true if $f^{-1}(O)$ is open

Our theorem can now be written as

$$\forall f, \left( C(f) \iff (\forall O, A(O) \Rightarrow B(f, O)) \right)$$

We’ll now make two different kinds of contra-positive arguments.

(1) To prove the $\Rightarrow$ part the theorem, we’ll show: $\forall f$, given $C(f)$, $\forall O, \neg B(f, O) \Rightarrow \neg A(O)$.
(2) To prove that $\Leftarrow$ part the theorem, we’ll show: $\forall f$, given $\neg C(f)$, $\exists O$ s.t. $A(O) \land \neg B(f, O))$

This is a great exercise for getting the hang of formal logical relationships.

Proof:

(1) first prove that continuity implies inverse images of open sets are open. Fix an arbitrary set $S \subset Y$ such that $f^{-1}(S)$ isn’t open in $X$. We’ll argue that $S$ isn’t open in $Y$. This will prove that when $f$ is continuous, $S$ open in $Y$ implies $f^{-1}(S)$ is open in $X$.

- if $f^{-1}(S)$ isn’t open there must exist a point $x$ in $f^{-1}(S)$ (i.e. such that $f(x) \in S$) which is a boundary point of $f^{-1}(S)$.

- i.e., there’s a sequence of points $x^n$ converging to $x$ all of which get mapped to points outside of $S$, that is, for all $n$, $f(x^n) \notin S$. 

• since $f$ is continuous, $f(x^n)$ must converge to $f(x)$.

• but this means that $f(x)$ is a boundary point of $S$.

• conclude that $S$ isn’t open

(2) now prove that inverse images of open sets are open implies continuity, for the case $Y = \mathbb{R}$.

We’ll show that if $f$ is not continuous, then there exists an open set $O \subset Y$ such that $f^{-1}(O)$ isn’t open in $X$.

• if $f$ isn’t continuous, there exists $x \in X$ and a sequence $\{x_n\}$ which converges to $x$ such that $f(x^n)$ doesn’t converge to $f(x)$, i.e., there exists $\epsilon > 0$ and a subsequence $(y^n)$ of $(x^n)$ such that for each $n$, $|f(y^n) - f(x)| > \epsilon$. Let $O = (f(x) - \epsilon, f(x) + \epsilon)$. Clearly $f(x) \in O$ so that $x \in f^{-1}(O)$. We’ll show that $x$ is not an interior point of $f^{-1}(O)$ and conclude that $f^{-1}(O)$ is not open.

• pick an arbitrary open set $W$ containing $x$. Since $\{y_n\}$ converges to $x$, there exist $n$ sufficiently large that $y^n \subset W$. But since by assumption, $f(y^n) \notin O$, it follows that $y^n \notin f^{-1}(O)$. Since $W$ was chosen arbitrarily, we have established that there does not exist an open set which contains $x$ and is itself contained in $f^{-1}(O)$. Conclude that $f^{-1}(O)$ is not open in $X$.

1.10. Upper and Lower Hemi continuous correspondences

Consider a correspondence $\xi$ mapping a metric space $S$ to a metric space $T$. For correspondences, we need to generalize the notion of an inverse image.

**Definition:** Given a set $O \subset T$, the upper inverse image of $O$, denoted $\bar{\xi}^{-1}(O)$, is the set $\{s \in S : \xi(x) \subset O\}$. 
Definition: Given a set $O \subset T$, the **lower inverse image** of $O$, denoted $\bar{\xi}^{-1}(O)$, is the set 
\[ \{ s \in S : \xi(x) \cap O \neq \emptyset \} . \]

Note that the requirement of a non-empty intersection with $O$ is much weaker than the requirement of containment in $O$, so that for every subset $O$ of the range of the correspondence, the upper inverse is contained in the lower inverse. Fig. 2 illustrates the two concepts:

1. the **red** set is the graph of the correspondence $\xi$.
2. now consider the (open) set in the range, denoted by the **blue** interval $O$.
3. the **upper inverse image** of $O$, $\bar{\xi}^{-1}(O)$, is the small **dark green** interval in the domain. Verify that for this example it happens to be an *open* interval.
4. the **lower inverse image** of $O$ is the larger **light green** interval in the domain. Verify that for this example it is also an open interval.

Intuitively (and imprecisely), $\xi$ is **upper hemi continuous** (u.h.c.) if its graph has no “holes”, and **lower hemi continuous** (l.h.c.) if its graph no “jumps.”
More precisely,

**Definition:** \( \xi \) is said to be upper hemi continuous if for every open set \( O \subset T \), the upper inverse image of \( O \), \( \bar{\xi}^{-1}(O) \), is an open set.

**Definition:** \( \xi \) is said to be lower hemi continuous if for every open set \( O \subset T \), the lower inverse image \( \xi^{-1}(O) \), is an open set.

To see the difference between the two definitions, compare the two “mirror image” correspondences, \( \xi^u : \mathbb{R} \to \mathbb{R} \) and \( \xi^\ell : \mathbb{R} \to \mathbb{R} \), defined by

\[
\xi^u = \begin{cases} \mathbb{R} & \text{if } x = 0 \\ \{0\} & \text{if } x \neq 0 \end{cases} \quad \xi^\ell = \begin{cases} \{0\} & \text{if } x = 0 \\ \mathbb{R} & \text{if } x \neq 0 \end{cases}
\]

the graph of \( \xi^u \) has a monstrous “jump” at zero in the domain, and the graph of \( \xi^\ell \) has a monstrous “hole” at zero in the domain.

More precisely,

1. consider the open interval \((1, 2)\), and observe \( \xi^u(1, 2) \) is the closed set \( \{0\} \), establishing that \( \xi^u \) is not lhc;

2. consider the open interval \((-1, 1)\), and observe \( \bar{\xi}^\ell(-1, 1) \) is the closed set \( \{0\} \), establishing that \( \xi^\ell \) is not uhc.

1.10.1. **Alternative defns of upper hemicontinuity.**

**Theorem:** \( \xi \) is upper hemi continuous at \( \bar{x} \in S \) iff \( \xi(\bar{x}) \neq \emptyset \) and for every neighborhood \( U \) of \( \xi(\bar{x}) \) there is a neighborhood \( V \) of \( \bar{x} \) such that \( \xi(x) \in U \) for every \( x \in V \).
Now suppose $\xi$ is compact valued, i.e., for every $x \in S$, $\xi(x)$ is a compact set. In this case, u.h.c. is equivalent to the statement that the graph of the correspondence is closed in the cartesian product $\text{Domain} \times \text{Range}$.

**Theorem:** The correspondence $\xi$ is u.h.c at $\bar{x}$ if and only if $\xi(\bar{x}) \neq \emptyset$ and if for every sequence $(x_n)$ converging to $\bar{x}$ and every sequence $(y_n)$ with $y_n \in \xi(x_n)$, there is a convergent subsequence of $(y_n)$ such that $\lim_n y_n \in \xi(\bar{x})$.

To see the role of compact-valuedness, consider the correspondence $\xi : \mathbb{R} \to \mathbb{R}$ defined by

$$\xi(x) = \begin{cases} 
\{0\} & \text{if } x = 0 \\
\{0, 1/x\} & \text{if } x \neq 0 
\end{cases}$$

Look at the graph of this function: clearly it’s closed in $\mathbb{R}^2$, i.e., it contains all it’s boundary points. (Indeed, in this example, every element of the graph is a boundary point.) But it fails the definition of u.h.c, i.e., there exists a neighborhood of $\xi(0) = \{0\}$—indeed, the following is true for all nbds of $\xi(0)$—with the property that there is no neighborhood of $V$ of $x = 0$ such that $\xi(V) \subset U$.

What goes wrong here? It’s the same idea that we’ve seen many times: (e.g., $\mathbb{R}$ is closed since it contains all of it’s accumulation points. Infinity isn’t an accumulation point, so we don’t have to worry that it isn’t “contained” in $\mathbb{R}$.) Similarly, in the case of $\xi$, sequences in the graph go off into outer space, but these sequences don’t accumulate to anything, so we don’t have to worry about them. But when you make the graph compact-valued, every sequence in the graph has to have a convergent subsequence.

1.10.2. Alternative definitions of lower hemicontinuity.
Theorem: \( \xi \) is lower hemi continuous at \( \bar{x} \in S \) iff \( \xi(\bar{x}) \neq \emptyset \) and for every open set \( G \subset T \) with \( G \cap \xi(\bar{x}) \neq \emptyset \), there exists a neighborhood \( Z \) of \( \bar{x} \) such that \( G \cap \xi(z) \neq \emptyset \), for every \( z \in Z \).

Theorem: The correspondence \( \xi \) is l.h.c at \( \bar{x} \) if \( \xi(\bar{x}) \neq \emptyset \) and if for any \( y \in \xi(\bar{x}) \), and any sequence \( (x_n) \) converging to \( \bar{x} \), there exists a sequence \( (y_n) \) such that \( y_n \in \xi(\bar{x}) \) and \( \lim_n y_n = y \).