Lecture 2: Sequences, Continuity

Originally written by Mauricio Caceres Bravo
Revised by Cole Davis
Webpage: cj-davis99.github.io

August 14, 2024

Contents

Notation

- $\forall$ translates to "for all"
- $\exists$ translates to "there exists"
- $\mathbb{N} = \{1, 2, \ldots\}$ is the set of natural numbers
- $\mathbb{Z} = \{\ldots, -1, 0, 1, \ldots\}$ is the set of integers
- $\mathbb{Q} = \{p/q : p \in \mathbb{Z} \text{ and } q \in \mathbb{Z}\backslash\{0\}\}$ is the set of rational numbers
- $\mathbb{R}$ is the set of real numbers
- If S is a set and $n \in \mathbb{N}$, then S^n is the n^{th} order Cartesian product of S. E.g., $S^2 = S \times S$
- For any $\varepsilon > 0$, $B_\varepsilon(x)$ is the Euclidean ball around x with radius ε
- Unless otherwise specified, $d(x, y)$ is a metric on whatever set x, y belong to
- $\text{cl}(S)$ is the closure of the set S

1 Sequences

Formally, a (real-valued) sequence is a function from $\mathbb{N}$ to $\mathbb{R}^N$ where $N \in \mathbb{N}$; however, it's often simpler to think of sequences as a collection of elements of a set indexed by the natural numbers. We will not be terribly precise with notation,[1] and denote sequences with elements in S as $(x_m) \in S$.

Definition 1. Let $(x_m) \in S$ be a sequence:

a) (x_m) is ***increasing*** if $\forall m$ we have $x_m \leq x_{m+1}$; it is *strictly* increasing if $x_m < x_{m+1}$.

b) (x_m) is ***decreasing*** if $\forall m$ we have $x_m \geq x_{m+1}$; it is *strictly* decreasing if $x_m > x_{m+1}$.

c) (x_m) is (strictly) ***monotonic*** if it is (strictly) increasing or decreasing .

d) (y_k) is a ***subsequence*** of (x_m) if $\exists$ some strictly increasing sequence $(n_k) \in \mathbb{N}$ s.t. $y_k = x_{m_k}$.

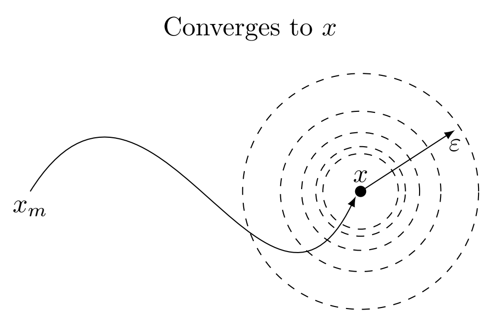
e) (x_m) is bounded above, below, or bounded if $\{x_m\}_{m \in \mathbb{N}}$ is bounded above, below, or bounded (resp).

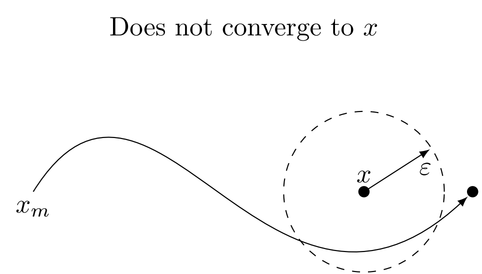
f) $(x_m) \to +\infty$ if $\forall N > 0, \quad \exists M$ s.t. $m \geq M \implies x_m \geq N$. $(x_m) \to -\infty$ if $\forall N < 0, \quad \exists M$ s.t. $m \geq M \implies x_m \leq N$. If either property holds, we say the sequence (x_m) ***diverges to infinity***.

1.1 Convergence

Definition 2. A sequence (x_m) ***converges*** to x if $\forall \varepsilon > 0, \quad \exists M \in \mathbb{N}$ s.t. $d(x_m, x) < \varepsilon$ whenever $m \geq M$. We denote this as $x_m \to x$ or $\lim_{m\to\infty} x_m = x$. Otherwise, we say that the sequence (x_m) ***diverges***.

We can visualize this idea in the figure below: x_m is eventually contained within any ε-ball if x.[2]





Claim 1. *A sequence converges to at most one limit.*

Proof. This is a consequence of the fact $d(x, y) \iff x = y$. Suppose $x_m \to x$ and $x_m \to y$. If $x \neq y$, then $d(x, y) > 0$. Take $\varepsilon = d(x, y)$; we know by the definition of convergence that there is some M_x, M_y s.t. for $m \geq M = \max\{M_x, M_y\}$ we get

$$d(x_m, x) < \varepsilon/2 \quad \text{and} \quad d(x_m, y) < \varepsilon/2$$

Now the triangle inequality gives

$$\varepsilon = d(x, y) \leq d(x_m, x) + d(x_m, y) < \varepsilon/2 + \varepsilon/2 = \varepsilon = d(x, y)$$

[1] We can also think formally of a sequence as an element of the infinite Cartesian product of $\mathbb{R}^N \times \mathbb{R}^N \times \cdots$. While tempting to define a sequence as a countable or finite subset of $\mathbb{R}^N$, sets have no notion of order; further, sequences can have repeated elements.

[2] Recall ε-ball is the equivalent of a neighborhood in Euclidean space, even through here in two dimensions it's technically a ε-circle.

contradiction ($\varepsilon < \varepsilon$). Hence $d(x,y) = 0$, or $x = y$. □

Theorem 1. *S is closed $\iff$ for any $(x_m) \in S$ s.t. $x_m \to x$ for some $x \in \mathbb{R}^N$, we have that $x \in S$.*

This is an equivalent definition of closedness: A set that "contains all its limits."

Proof. See Subsection 1.2 for a visualization (try formally writing the proof yourself!). □

Theorem 2. *Take any set $S \subseteq \mathbb{R}^N$; the following are equivalent:*

a) $x \in cl(S)$.

b) $\forall \varepsilon > 0, B_\varepsilon(x) \cap S \neq \varnothing$.

c) $\exists (x_m) \in S$ *s.t.* $x_m \to x$.

Proof. We prove this equivalence by proving $b \implies c \implies a \implies b$.

$(b \Rightarrow c)$ Take any x such that condition b holds. Define $(x_m) \in S$ such that $\forall m \in \mathbb{N}$

$$x_m \in B_{1/m}(x) \cap S \neq \varnothing$$

Showing that $x_m \to x$ completes the proof. Fix any $\varepsilon > 0$. By the Archimedean Property (see Lecture 1 notes), we know that there exists an $M \in \mathbb{N}$ such that $0 < 1/M < \varepsilon$. Thus,

$$d(x_m, x) < 1/m < \varepsilon$$

whenever $m \geq M$, as desired.

$(c \Rightarrow a)$ Let (x_m) be any convergent sequence in S such that $x_m \to x$ (we know such a sequence exists because we are assuming condition c). We want to show that $x \in \text{cl}(S)$. Since $S \subseteq \text{cl}(S)$, we know that $(x_m) \in \text{cl}(S)$. Moreover, we know that $\text{cl}(S)$ is a closed set. By Theorem 1, it must be that $x \in \text{cl}(S)$.

$(a \Rightarrow b)$ Suppose there exists $\varepsilon > 0$ such that

$$B_\varepsilon(x) \cap S = \varnothing$$

We want to show that $x \notin \text{cl}(S)$. Since $B_\varepsilon(x)$ is open, we know that its complement is closed. Since $\text{cl}(S)$ is a closed set, $\text{cl}(S) \setminus B_\varepsilon(x)$ is a closed set. By definition, $S \subseteq \text{cl}(S)$. Since $B_\varepsilon(x) \cap S = \varnothing$, it must also be that

$$S \subseteq \text{cl}(S) \setminus B_\varepsilon(x)$$

Hence, $\text{cl}(S) \subseteq \text{cl}(S) \setminus B_\varepsilon(x)$ by the definition of closure. Since $x \in B_\varepsilon(x)$, we know that

$$x \notin \text{cl}(S) \setminus B_\varepsilon(x) \implies x \notin \text{cl}(S)$$

as desired.

□

Definition 3. Let (x_m) be any sequence. We define

$$\limsup_{m \to \infty} x_m = \lim_{m \to \infty} \left(\sup_{k \geq m} x_k \right)$$

and

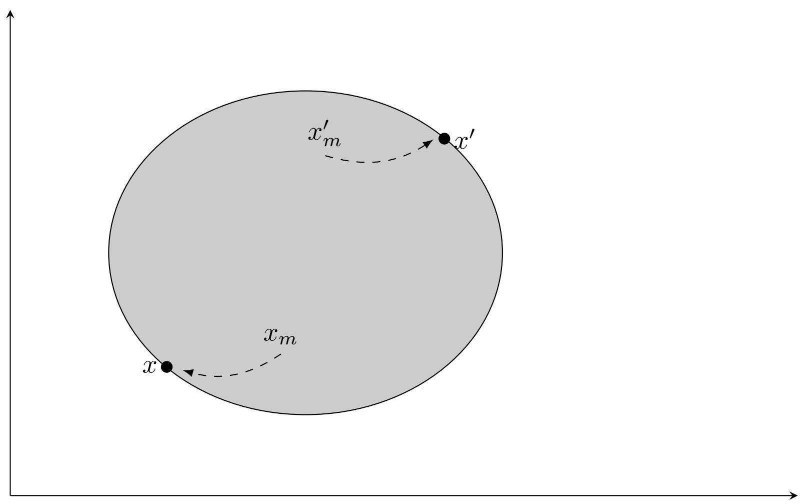
$$\liminf_{m\to\infty} x_m = \lim_{m\to\infty}\left(\inf_{k\geq m} x_k\right)$$

The lim sup and lim inf do not require convergence. Take, for instance, $(x_m) = 0, 1, 0, 1 \ldots$. Clearly $\limsup = 1$ and $\liminf = 0$ but the limit does not exist.

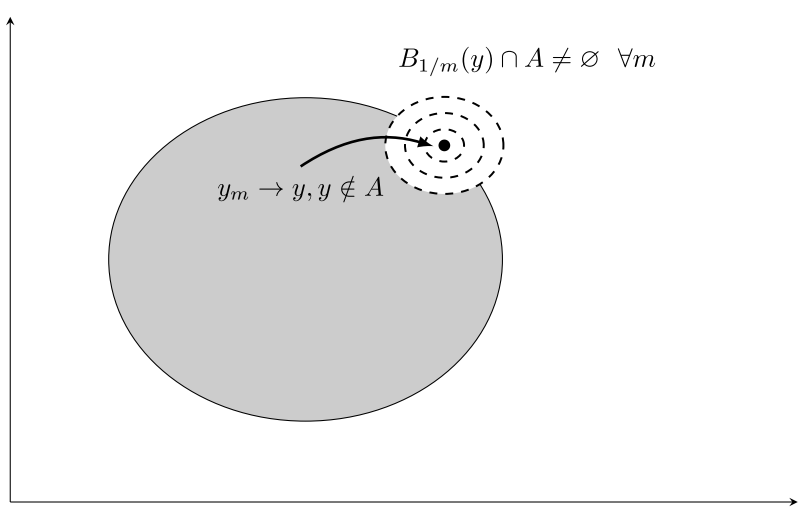
1.2 Sets and Sequences (Visualization)

The mathy sequel to Dungeons & Dragons. In this section we will try to visualize the proof of Theorem 1. This should give some intuition for why convergence is related to sets being closed beyond the formality of the theorem.

- If a sequence x_m converges to x, then it becomes arbitrarily close to a point. If for every $(x_m) \in A$ s.t. $x_m \to x$ we also have $x \in A$, that means that no sequence can ever "escape " outside of A:

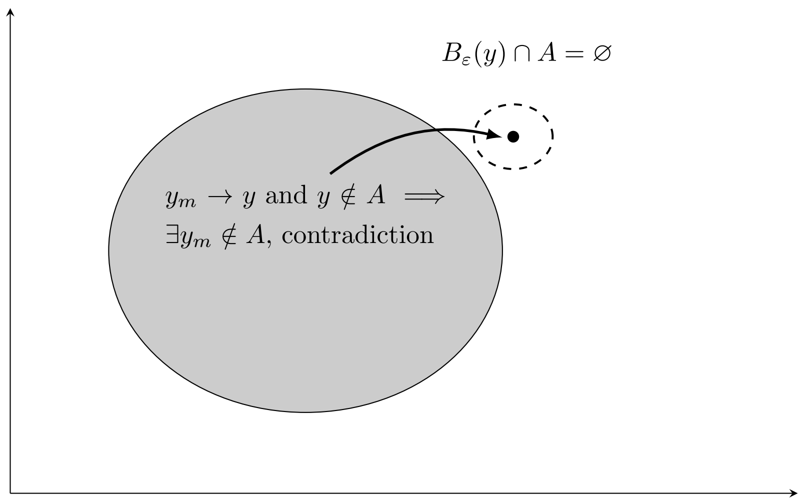


By contrapositive, if A is not closed, its complement is not open, so $\exists y \in \mathbb{R}^N \setminus A$ that cannot be enclosed in an ε-ball inside of $\mathbb{R}^N \setminus A$. In other words, there is some y s.t. for each $\varepsilon_m = 1/m$ we can find a corresponding $y_m \in A \cap B_{1/m}(y)$. The resulting $(y_m) \in A$ converges to $y \notin A$.



In other words, if A is not closed, we can find a sequence that "escapes" A, which by contrapositive proves that if every sequence in A that converges does so to a point in A, the set is closed.

- On the other hand, if A is closed and we have a sequence $(y_m) \in A$ s.t. $y_m \to y$ with $y \notin A$, then we have a sequence that "escaped" A. However, A closed implies $X \setminus A$ is open, and $\exists B_\varepsilon(y) \subseteq \mathbb{R}^N \setminus A$. Since y_m will get arbitrarily close to y, $\exists y_m \in B_\varepsilon(y) \subseteq X \setminus A$. Since $y_m \in A$ by premise, this is a contradiction.



I think in general it's very useful to visualize proofs in $\mathbb{R}^2$ (making drawings, as above):

- $\mathbb{R}$ is not enough: A ton of things will hold in one dimension that won't in general, and one-dimensional intuition can end up being misleading.
- $\mathbb{R}^3$ can be too much: I cannot draw 3D very easily and it is harder to visualie clearly (certainly $\mathbb{R}^N$ would be too many dimensions).
- $\mathbb{R}^2$ is a nice trade-off between rigor an intuition. (I know we discuss properties in more general terms than in $\mathbb{R}^2$, but for intuition I think it's a great benchmark.)

1.3 Properties of Convergent Sequences

What follows are some useful properties of convergent sequences sometimes referred to as "limit laws" in calculus textbooks. We leave the proof of these results as an exercise for the reader.

Theorem 3. *Take any sequence $(x_m) \in \mathbb{R}^N$:*

a) $x_m \to x \implies x_{m_k} \to x$ for all subsequences (x_{m_k}) of (x_m).

b) $x_m \to x \implies (x_m)$ is bounded. (Is the converse true? Can you prove your answer?[3])

c) If $x_m \leq y_m \leq z_m$ and $x_m, z_m \to x$ then $y_m \to x$.

d) If $x_m \to 0$ and (y_m) is bounded then $x_m \cdot y_m \to 0$.

Let $x_m \to x$ and $y_m \to y$:

e) $c \cdot x_m \to c \cdot x$ for any $c \in \mathbb{R}$.

[3] A counterexample is sufficient: $x_m = (-1)^m$ is bounded above by 1 and below by -1 but does not converge.

f) $x_m \pm y_m \to x \pm y$.

g) $x_m \cdot y_m \to x \cdot y$.

h) $x_m / y_m \to x/y$ *if* $y \neq 0$.

1.4 Bolzano-Weierstrass

Theorem 4. *Let* $(x_m) \in \mathbb{R}$. *If* (x_m) *is bounded and monotonic then* (x_m) *converges.*

Proof. See question 6 of problem set 1. *Hint*: If (x_m) is bounded and increasing, then what should (x_m) converge to? Claim 10 in Lecture 1 notes may be useful here... □

Theorem 5 (Nested Intervals Theorem)**.** *Let* $I_m = [a_m, b_m]$ *s.t.* $I_{m+1} \subseteq I_m$.

a) $\cap_{m \in \mathbb{N}} I_m \neq \varnothing$

b) If $b_m - a_m \to 0$ *then* $\cap_{m \in \mathbb{N}}$ *is a singleton.*

Proof. If $I_{m+1} \subseteq I_m$, then $I_m \subseteq I_1$ for all m. Hence

$$a_1 \leq a_m \leq a_{m+1} \leq b_{m+1} \leq b_m \leq b_1$$

for all m. In other words, a_m and b_m are bounded and monotonic, so $a_m \to a$ and $b_m \to b$ for some a, b by Theorem 4. Further, since $a_m \leq b_m$ for all m, we have that $a \leq b$ (do you see why?[4]) and $I_m \to [a, b] \neq \varnothing$. If $b_m - a_m \to 0$, that means $b - a = 0$ so the interval $[a, b]$ is just the singleton $a = b$. □

Theorem 6 (Bolzano-Weierstrass)**.** *Every bounded sequence in* $\mathbb{R}$ *admits a convergent sub-sequence.*

Proof. This follows from Theorem 5 above—the trick is to construct the nested intervals, which we can do for bounded sequences. If (x_m) is bounded, first define

$$I_1 = [L, U] = [a_1, b_1]$$

the interval with endpoints equal to the lower and upper bounds of x_m. Let x_{m_1} the first element of the sub-sequence be any term of $x_m \in I_1$. Now take

$$I_1^- = \left[a_1, \frac{a_1 + b_1}{2}\right] \qquad I_1^+ = \left[\frac{a_1 + b_1}{2}, b_1\right]$$

and let $I_2 = I_1^-$ if $\{x_m : m \in \mathbb{N}\} \cap I_1^+$ is non-finite and $I_2 = I_1^+$ otherwise. Let a_2, b_2 be the endpoints of I_2 and x_{m_2} be any term of $x_m \in I_2$. We iterate this process: In general

$$I_k^- = \left[a_k, \frac{a_k + b_k}{2}\right] \qquad I_k^+ = \left[\frac{a_k + b_k}{2}, b_k\right]$$

[4] If $a > b$ then we will have that $a_m > b_m$ for some m. That is, pick $0 < \varepsilon < a - b$; we can then find M_a, M_b s.t.

$$a - \varepsilon/2 < a_m < a + \varepsilon/2 \qquad b - \varepsilon/2 < b_m < b + \varepsilon/2$$

whenever $m > \max\{M_a, M_b\}$. However, $\varepsilon < a - b$ gives

$$b_m < b + \varepsilon/2 < a - \varepsilon/2 < a_m$$

Since $b_m \geq a_m$ for all m we have a contradiction.

and $I_{k+1} = I_k^-$ if $\{x_m : m \in \mathbb{N}\} \cap I_k^+$ is non-finite and $I_{n+1} = I_k^+$ otherwise, with x_{m_k} any element of $x_m \in I_k$. (What if both halves have a finite intersection?[5]) We have that

$$a_1 \leq a_{k-1} \leq a_k \leq x_{m_k} \leq b_k \leq b_{k-1} \leq b_1$$

Furthermore, $b_n - a_n = \dfrac{b_1 - a_1}{2^{n-1}} \to 0$. Hence we can apply the Theroem 5: $a_n \to a$ and $b_n \to b$ with $a = b$ implies $x_{m_k} \to a = b$. □

1.5 Cauchy Sequences

In this section, we define a generalization of a convergent sequence.

Definition 4. A sequence is ***Cauchy*** if $\forall \varepsilon > 0 \;\; \exists M$ s.t.

$$m, n > M \implies d(x_m, x_n) < \varepsilon$$

Informally, a Cauchy sequence is any sequence whose terms get arbitrarily close together as we look further into the sequence. The following theorem states that every convergent sequence is a Cauchy sequence.

Theorem 7. *If (x_m) converges, then it is Cauchy.*

Proof. Suppose $x_m \to x$; by the triangle inequality

$$d(x_m, x_n) \leq d(x_m, x) + d(x_n, x)$$

Now take any $\varepsilon > 0$; for $\varepsilon/2$ we have that for some M, $m, n > M$ gives

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

Hence

$$d(x_m, x_n) \leq d(x_m, x) + d(x_n, x) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

which shows (x_m) is Cauchy. □

While the converse of the theorem above is also true in $\mathbb{R}^N$, during your math course you will probably encounter the fact that in general metric spaces, Cauchy sequences needn't converge. (The reason is, again, this property of Euclidean space called "completeness.")

Theorem 8. *If (x_m) is a Cauchy sequence in $\mathbb{R}$ then (x_m) converges.*

Proof. We will show this in $\mathbb{R}$: Cauchy sequences are bounded (why?), which means that there exist some sub-sequence x_{m_k} that converges to some x. We show that this is also the limit for the sequence x_m. By the triangle inequality,

$$d(x_m, x) \leq d(x_m, x_{m_k}) + d(x_{m_k}, x)$$

[5]Note that for $k > 1$, it must always be that either I_k^+ or I_k^- have a non-finite intersection with $\{x_m : m \in \mathbb{N}\}$, since we chose I_k to have a non-finite intersection. The only way both halves will have a finite intersection is if I_1 is finite to begin with. However, this means that some M, $x_n = x_m$ whenever $n, m > M$, which means we have a convergent sub-sequence $x_{m_k} = x_{M+1}$ with $m_k = M + k$.

Take any $\varepsilon > 0$. Since $x_{m_k} \to x$, we know there exists a $K \in \mathbb{N}$ such that

$$k > K \implies d(x_{m_k}, x) < \frac{\varepsilon}{2}$$

Since (x_m) is Cauchy, we also know that there exists some $M \in \mathbb{N}$ such that

$$m, k > M \implies d(x_m, x_{m_k}) < \frac{\varepsilon}{2}$$

(note that $m_k \geq k$, so $k > M \implies m_k > M$). Letting $N = \max\{K, M\}$, we have that for any $m, k > N$

$$d(x_m, x) \leq d(x_m, x_{m_k}) + d(x_{m_k}, x) < \varepsilon/2 + \varepsilon/2 = \varepsilon$$

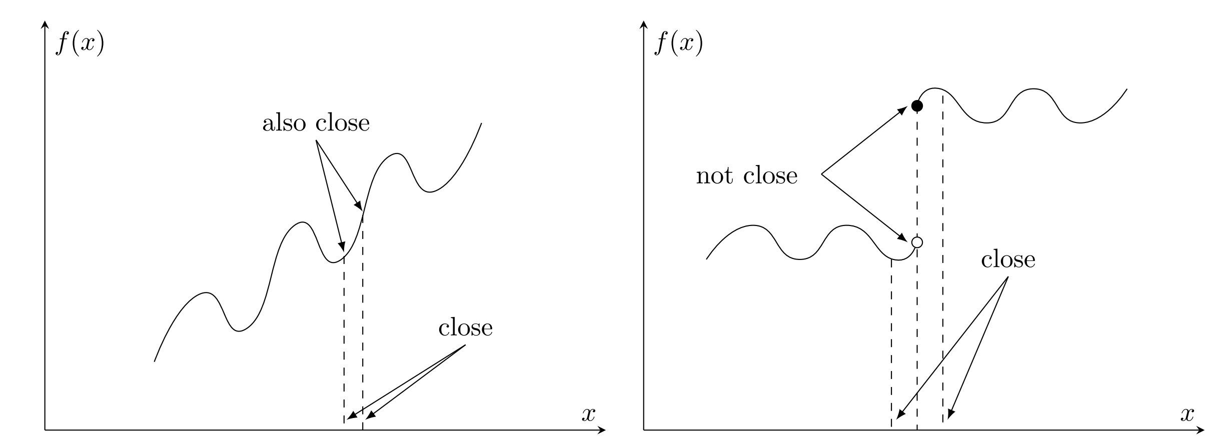
This proof should generalize to $\mathbb{R}^N$ if you argue that along each coordinate, (x_m) is bounded and then use that to find a candidate limit x. Then the identical argument goes through (note it used generic properties of the distance rather than anything specific to $\mathbb{R}$). $\square$

Remark 1. The sticking point about "completeness" being required has to do with the fact the candidate limit x needn't be in the space (e.g. take some sequence in $\mathbb{Q}$ that "converges" to $\pi \notin \mathbb{Q}$).

2 Continuous Functions

The intuition for continuity is that "you can draw the function without picking up your pencil." (What about asymptotes like $1/x$ at 0?[6]) More precisely, if two points in the domain are close, the corresponding points in the co-domain must also be close. Put another way, a small neighborhood in the domain maps to a small neighborhood in the co-comain. (Note that the converse is not true! Take $f(x) = x^2$; $f(-x) = f(x) = x^2$, so the points are close in the image—they are identical—but as x gets large, $d(x, -x)$ gets larger.)

Figure 1: Intuition for Continuity



Definition 5 (continuity)**.** A function $f : X \to Y$ is ***continuous*** at $x \in X$ if for every $\varepsilon > 0$ there exist a

[6]The intuition should still hold because the function is not defined at 0; with infinitely long paper you needn't pick up your pencil.

$\delta > 0$ s.t.

$$d(z, x) < \delta \implies d(f(z), f(x)) < \varepsilon$$

Put another way, $z \in B_\delta(x) \implies f(z) \in B_\varepsilon(f(x))$, or

$$f(B_\delta(x)) \subseteq B_\varepsilon(f(x))$$

If f is continuous at every $x \in X$ we say it is continuous.

Proposition 1. *Let $\varphi : \mathbb{R}^N \to \mathbb{R}$. If φ is continuous then the sets*

$$\left\{x \in \mathbb{R}^N : \varphi(x) \geq \alpha\right\}$$
$$\left\{x \in \mathbb{R}^N : \varphi(x) \leq \alpha\right\}$$

and are closed for all $\alpha \in \mathbb{R}$.

Proof. Let $A = \left\{x \in \mathbb{R}^N : \varphi(x) \geq \alpha\right\}$ and $B = \mathbb{R}^N \setminus A$. If B is open then A is closed. Note

$$B = \left\{x \in \mathbb{R}^N : \varphi(x) < \alpha\right\}$$

Pick any $x \in B$ and let $0 < \varepsilon < \alpha - \varphi(x)$. Then since φ is continuous we know there is some $\delta > 0$ s.t.

$$z \in B_\delta(x) \implies \varphi(x) - \varepsilon < \varphi(z) < \varphi(x) + \varepsilon < \alpha \implies z \in B$$

therefore B is open (for any point in B $\exists\delta$-ball that is entirely contained in B). Thus $\mathbb{R}^N \setminus B = A$ is closed. The proof for the lower sets is analogous. $\square$

Theorem 9. *Let $f : X \to Y$ and $g : f(X) \subseteq Y \to Z$ with $X, Y, Z \subseteq \mathbb{R}^N$. If f, g are continuous then $(g \circ f) : X \to Z$ is continuous.*

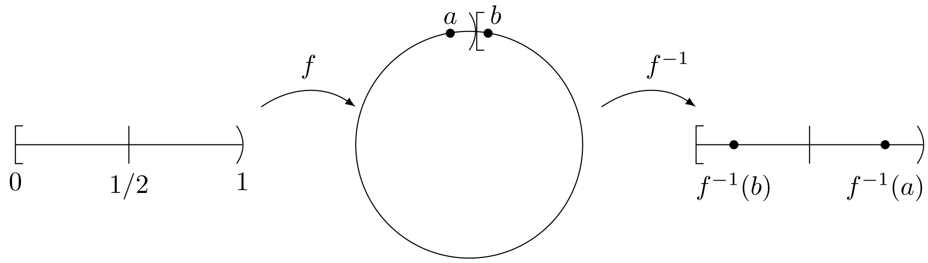
Theorem 10. *Let $f : X \to \mathbb{R}$ and $g : X \to \mathbb{R}$ be continuous functions. Then*

1. *$h(x) = f(x) \pm g(x)$ is continuous.*
2. *$h(x) = f(x) \cdot g(x)$ is continuous.*
3. *$h(x) = f(x)/g(x)$ is continuous whenever $g(x) \neq 0$.*

2.1 Sequential and Open Set Characterizations of Continuity

Remark 2. Continuous functions don't map open sets to open sets! Consider $f(x) = x^2$. The image of $(-2, 2)$ is $[0, 2)$, which is not open. Further, even if all maps of closed sets are closed, the function might not be continuous. For example, Consider a function that is 0 if $x < 0$ and 1 if $x \geq 1$. This has a jump at 1, but $f([a, b])$ is either $\{0\}, \{1\}$, or $\{0, 1\}$, which are closed.

Last, if a function is continuous the inverse image need not be. Consider a function $f : [0, 1) \to \mathbb{R}^2$ that maps the line into a circle:



We can see that a, b are close in the image, but not in the inverse image.

The remarks above all get to the same idea: Continuity states that points in the image are close if the points in the domain are close, not the converse. Hence we have the following characterization in terms of the *inverse image* and *converging sequences*.

Theorem 11. *The following are equivalent for any $f : \mathbb{R}^N \to \mathbb{R}^M$:*

a) f is continuous.

b) If $O \subseteq \mathbb{R}^M$ is open, $f^{-1}(O)$ is open.

c) If $S \subseteq \mathbb{R}^M$ is closed, $f^{-1}(S)$ is closed.

d) For every $(x_m) \in \mathbb{R}^N$ s.t. $x_m \to x$ for some $x \in \mathbb{R}^N$, $f(x_m) \to f(x)$.

We give visual intuition for the proof in Section 2.1.1.

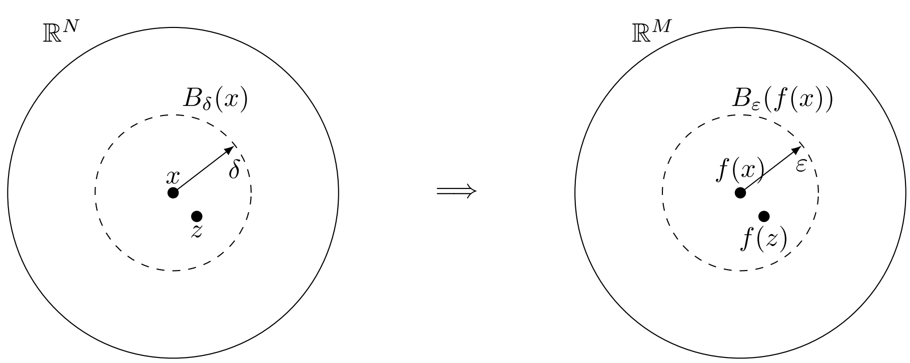
2.1.1 Sequences, Sets, and Continuity (Visualization)

We show Theorem 11 using, again, drawings in $\mathbb{R}^2$.

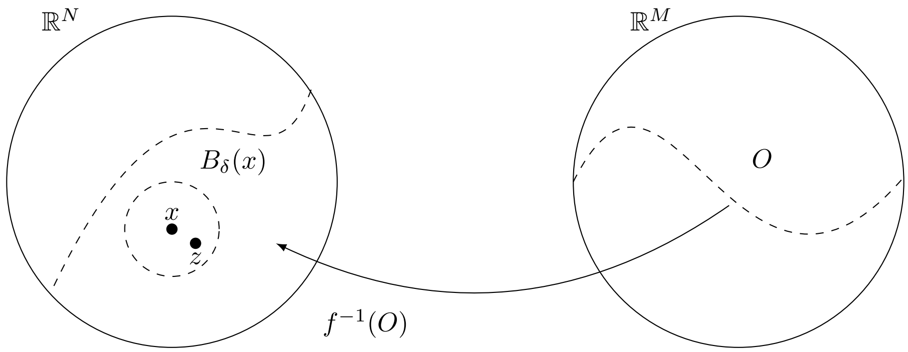
Proof. We cycle through the statements. If we show $a \implies b \iff c \implies d \implies a$ then we have shown they are equivalent.

1. $a \implies b$. First, it is a good idea to write down what you have and what you want to show:

   - Continuity means that $\forall \varepsilon > 0 \;\; \exists \delta > 0$ s.t. $z \in B_\delta(x) \implies f(z) \in B_\varepsilon(f(x))$.

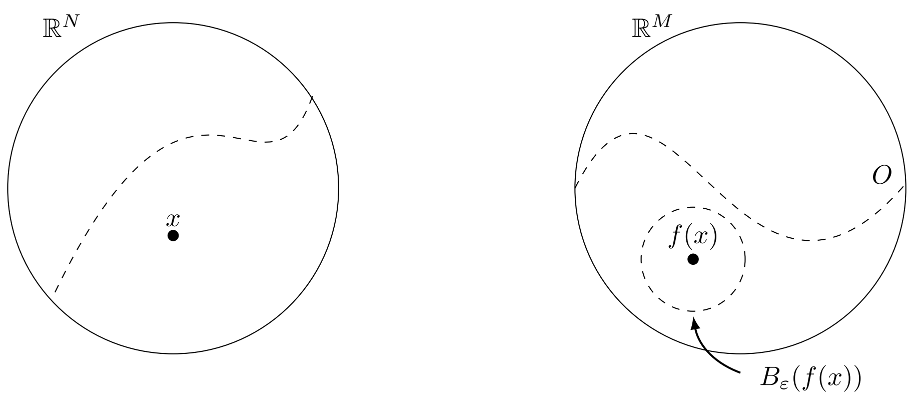


   We want to show that if $O \subseteq \mathbb{R}^M$ is open, then $f^{-1}(O)$ is open. That is, $\forall x \in f^{-1}(O) \;\; \exists \delta > 0$ s.t. $z \in B_\delta(x) \implies z \in f^{-1}(O)$. You will note this is a very similar statement!

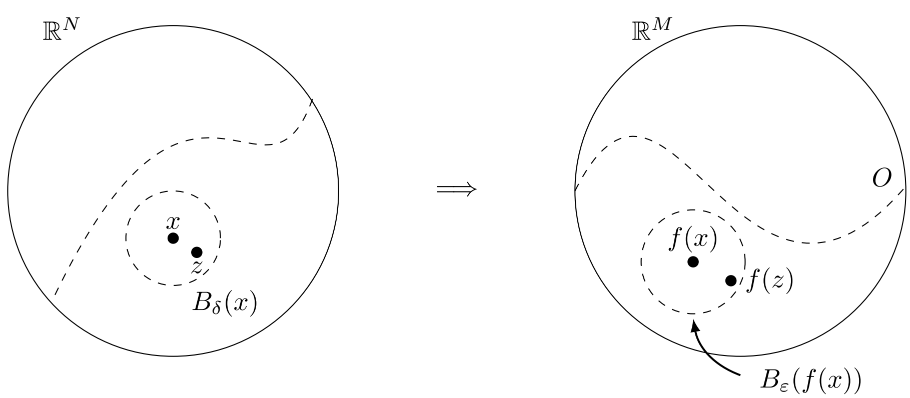


- It's basically the same picture: All we are missing is $B_\varepsilon(f(x))$ and it looks like we're done. How do we get it? We use the fact that O is open in $\mathbb{R}^M$.

Pick any $x \in f^{-1}(O)$ $\exists \varepsilon > 0$ s.t. $B_\varepsilon(f(x)) \subseteq O$



$\exists \delta$ s.t. $z \in B_\delta(x) \implies f(z) \in B_\varepsilon(f(x))$
(by the continuity assumption!)



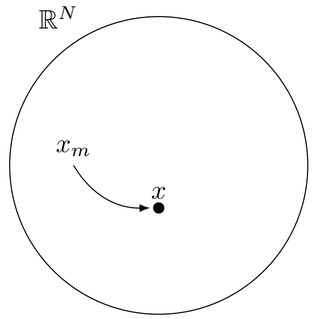
Note $f(z) \in B_\varepsilon(f(x)) \subseteq O$, so $z \in f^{-1}(O)$. This statement is the heart of the proof! It is not obvious that $B_\delta(x)$ will be contained in $f^{-1}(O)$, so we need the link with $B_\varepsilon(f(x))$ we drew above. Only then can we say that for arbitrary x we found $\delta > 0$ s.t. $z \in B_\delta(x) \implies z \in f^{-1}(O)$; by definition that means the set is open.

2. We show $b \iff c$. First, consider any closed set $S \in \mathbb{R}^M$, so $\mathbb{R}^M \setminus S$ is open; by premise, $f^{-1}(\mathbb{R}^M \setminus S)$ is also open, which means $f^{-1}(\mathbb{R}^M) \setminus f^{-1}(\mathbb{R}^M \setminus S) = f^{-1}(\mathbb{R}^M \setminus (\mathbb{R}^M \setminus S)) = f^{-1}(S)$ is closed. (The only sticking point here would be to show that for general subsets A, B of the co-domain of any function f, $f^{-1}(A) \setminus f^{-1}(B) = f^{-1}(A \setminus B)$[7]). Now consider open set $O \in \mathbb{R}^M$, so $\mathbb{R}^M \setminus O$ is open; by premise, $f^{-1}(\mathbb{R}^M \setminus O)$ is also closed, which means $\mathbb{R}^N \setminus f^{-1}(\mathbb{R}^M \setminus O) = f^{-1}(O)$ is open. You will notice this is an entirely analogous argument.

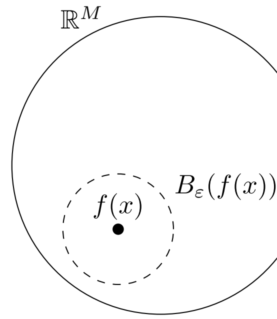
3. For this one it is easier to show that $b \implies d$ (noting we already argued $c \implies b$).
   - $\forall O \subseteq \mathbb{R}^M$, if O open then $f^{-1}(O)$ open. This means that $\forall x \in f^{-1}(O) \;\; \exists \delta$ s.t. $B_\delta(x) \subseteq f^{-1}(O)$.
   - We WTS $x_m \to x \implies f(x_m) \to f(x)$; i.e. $\forall \varepsilon > 0 \;\; \exists M$ s.t. $m \geq M \implies f(x_m) \in B_\varepsilon(f(x))$.

[7] Suppose $x \in f^{-1}(A) \setminus f^{-1}(B)$. If $x \in f^{-1}(B) \Leftrightarrow f(x) \in B$, then $x \notin f^{-1}(A) \setminus f^{-1}(B)$, a contradiction. If $x \notin f^{-1}(A) \Leftrightarrow f(x) \notin A$, then $x \notin f^{-1}(A) \setminus f^{-1}(B)$, a contradiction. Thus, $f(x) \in A \setminus B \Leftrightarrow x \in f^{-1}(A \setminus B)$, implying that $f^{-1}(A) \setminus f^{-1}(B) \subseteq f^{-1}(A \setminus B)$. We can follow a nearly identical argument to show set containment in the other direction (try it!), giving us the desired equality.

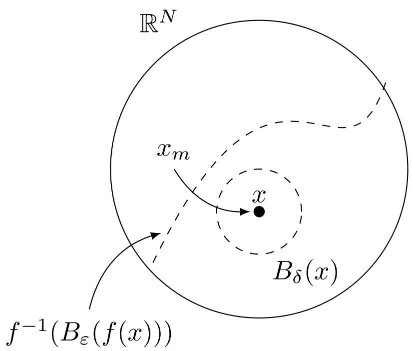
Start with any sequence $x_m \to x$



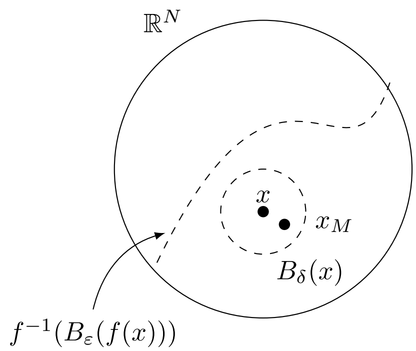
Pick any $\varepsilon > 0$ and note $B_\varepsilon(f(x))$ is open.



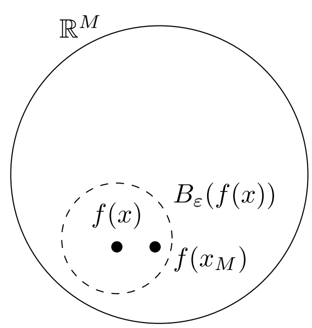
Hence $f^{-1}(B_\varepsilon(f(x)))$ is open, and
$\exists \delta > 0$ s.t. $B_\delta(x) \subseteq f^{-1}(B_\varepsilon(f(x)))$



$x_m \to x$ so $\exists M$ s.t.
$m \geq M \implies x_m \in B_\delta(x)$



$x_m \in B_\delta(x) \subseteq f^{-1}(B_\varepsilon(f(x))) \implies f(x_m) \in B_\varepsilon(f(x))$



Hence for any $x_m \to x$, for any $\varepsilon > 0$ we found M s.t.

$$m \geq M \implies x_m \in f^{-1}(B_\varepsilon(f(x))) \implies f(x_m) \in B_\varepsilon(f(x))$$

which by definition means $f(x_m) \to f(x)$. The tricky step here was that $f^{-1}(B_\varepsilon(f(x)))$ does not need to be a nice set. We need the premise that the inverse image of open sets is open so that we can fit a neighborhood inside of it, and *then* use the fact $x_m \to x$.

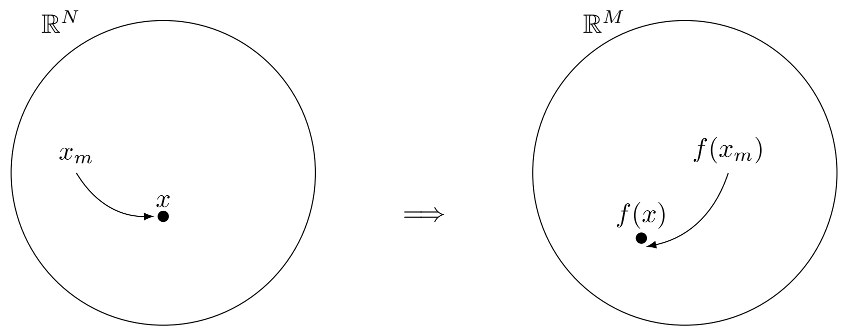
4. Finally, we show that $d \implies a$. We do this by contradiction.

   - It is not clear why contradiction is the way to go; it boils down to the fact I think it's easier, but I don't think that's obvious. In general, if unsure how to start a proof, one strategy is to try to make progress with a direct proof, and if you get stuck, switch to contradiction or contrapositive to see if it helps.

- First, we have $x_m \to x \implies f(x_m) \to f(x)$.

If we had a sequence $x_m \to x$

Then we'd know $f(x_m) \to f(x)$.



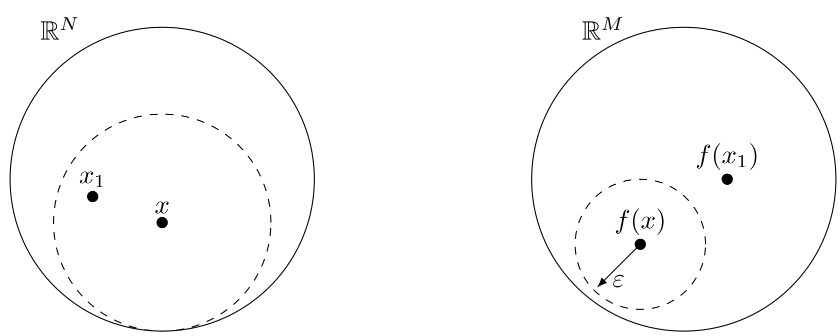
- We want to show that $\forall x \in \mathbb{R}^N \;\; \forall \varepsilon > 0 \;\; \exists \delta > 0$ s.t. $z \in B_\delta(x) \implies f(z) \in B_\varepsilon(f(x))$.
- A great starting point is to construct a sequence in $\mathbb{R}^N$ that converges to x, because the premise here is a statement about sequences. I don't see an obvious way to do this directly, but if we think about doing contradiction, we can negate the previous bullet point:

$$\exists x \in \mathbb{R}^N \;\; \exists \varepsilon > 0 \;\text{ s.t. }\; \forall \delta > 0 \;\; \exists z \in B_\delta(x) \;\text{ and }\; f(z) \notin B_\varepsilon(f(x))$$

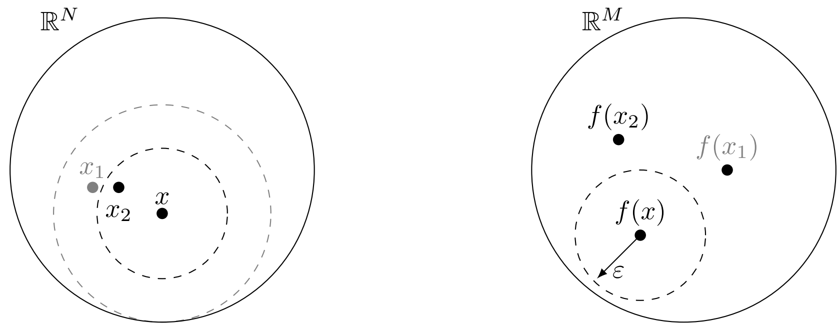
Note that x and ε here are fixed, and that we don't get to choose z—all we know is one such a z exists. However, δ is a free parameter here, because this must be true for any δ.

- If we pick $\delta = 1/m$ then we can construct a sequence $x_m \to x$ s.t. $f(x_m) \notin B_\varepsilon(f(x))$:

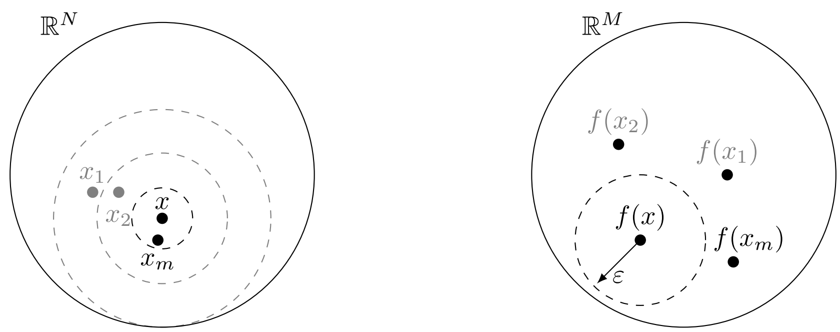
For $m = 1 \;\; \exists x_1 \in B_1(x)$ s.t. $f(x_1) \notin B_\varepsilon(f(x))$



For $m = 2$ $\exists x_2 \in B_{1/2}(x)$ s.t. $f(x_2) \notin B_\varepsilon(f(x))$



$\forall m$ $\exists x_m \in B_{1/m}(x)$ s.t. $f(x_m) \notin B_\varepsilon(f(x))$



We can see as x_m becomes increasingly closer to x, $f(x_m)$ is always at least ε away from $f(x)$. In other words, we have constructed a sequence $x_m \to x$ were $f(x_m) \not\to f(x)$, contradiction.

□

3 Intermediate Value Theorem (IVT)

Theorem 12 (Intermediate Value Theorem (IVT))**.** *If $f : [a, b] \to \mathbb{R}$ is continuous in $[a, b]$ then for any L between $f(a)$ and $f(b)$,*

$$\exists c \in [a, b] \quad s.t. \quad f(c) = L$$

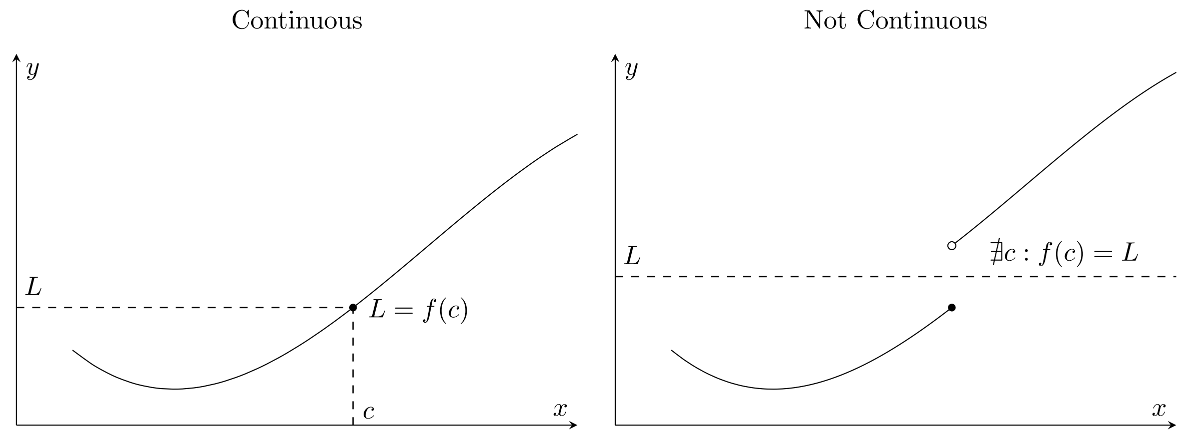


Figure 2: Intermediate Value Theorem (IVT)

Proof. WLOG, suppose $f(a) < f(b)$ (the equality case is trivial). Take any $L \in (f(a), f(b))$ and consider

$$A = \{x \in [a, b] : f(x) \leq L\}$$

A is bounded so the sup exists; let $c = \sup A$. For $\varepsilon_m = 1/m$ take $x_m \in (c - \varepsilon_m, c] \cap A$, so $x_m \to c$ and $f(x_m) \to f(c)$ (by continuity). Since A is closed (see Proposition 1) and $x_m \in A$, $c \in A$ (by Theorem 1) and $f(c) \leq L$. If $f(c) = L$ we are done; if $f(c) < L$ then consider

$$B = \{x \in [c, b] : f(x) \geq L\}$$

Note $B = (c, b]$ because $f(c) < L$, so $c \notin B$, and c is an upper bound for A: For every $x \in (c, b]$ it cannot be $f(x) < L$ because then $x \in A$ and $x > c$, a contradiction, so $f(x) \geq L$. However, this is another contradiction because the upper set of a continuous function is closed, and $B = (c, b]$ is not closed. Hence $f(c) = L$. (If you take $\inf(B)$ you can follow the same logic to get $f(c) > L$; basically we're showing there's a discontinuity at c, which cannot happen for a continuous function.) (Another subtlety here is the case when $c = b$, but then $f(b) = f(c) < L < f(b)$ contradiction.) $\square$

4 Fun Remarks

You can have a fun relaxing minute listening to Tom Lehrer's *There's a Delta for Every Epsilon*:

> There's a delta for every epsilon,
> It's a fact that you can always count upon.
> There's a delta for every epsilon
> And now and again,
> There's also an N.
>
> But one condition I must give:
> The epsilon must be positive
> A lonely life all the others live,

In no theorem
A delta for them.

How sad, how cruel, how tragic,
How pitiful, and other adjec-
Tives that I might mention.
The matter merits our attention.
If an epsilon is a hero,
Just because it is greater than zero,
It must be mighty discouragin'
To lie to the left of the origin.

This rank discrimination is not for us,
We must fight for an enlightened calculus,
Where epsilons all, both minus and plus,
Have deltas
To call their own.

Index