Math 141: Lecture 8

Sperner's lemma and the Brouwer Fixed Point Theorem

Bob Hough

October 3, 2016

Upper bounds and sups

Review:

- An *upper bound* of a non-empty set is a number at least as large as every member of the set: B is an upper bound of set S if

$$\forall s \in S, B \geq s.$$

- For instance, the numbers $5, 5.5, 6, 10$, and 10^{20} are all upper bounds of the set $\{1, 2, 3, 4, 5\}$.
- If an upper bound for a set exists, the set is said to be bounded above.

Upper bounds and sups

Review:

- Given a non-empty set of real numbers which is bounded above, the *supremum* of the set is the least upper bound of the set.
- The supremum need not be a member of the set. For instance, the set $S = \{x \in \mathbb{R} : x < 5\}$ has supremum 5.
- An important property of the real numbers is that any non-empty set of real numbers which is bounded above has a supremum.

Upper bounds and sups

Review:

- Whereas a set which is bounded above has infinitely many upper bounds, it has only one supremum.
- To prove that B is the supremum of a set S, it is necessary to prove that B is an upper bound for S, and that if C is another upper bound for S, then $B \leq C$.
- To check that 5 is the supremum of $S = \{x \in \mathbb{R} : x < 5\}$ note that for all $x \in S$, $x < 5$ so 5 is an upper bound. If $C < 5$ then $C < (C + 5)/2 < 5$ so C is not an upper bound for S, thus 5 is the l.u.b.

Continuous functions

Review:

- A function f defined on a set S is *continuous* at $p \in S$ if, for each $\epsilon > 0$ there is a $\delta > 0$ such that $x \in S$ and $|x - p| < \delta$ implies $|f(x) - f(p)| < \epsilon$.
- A function f is said to be continuous on S if it is continuous at each point $p \in S$.

Continuous functions

Review:

- We check from the definition that $f(x) = x^2$ is continuous on $[0, 1]$.
- Let $p \in [0, 1]$. Given $\epsilon > 0$, let $\delta = \frac{\epsilon}{2} > 0$. For $x \in [0, 1]$ such that $|x - p| < \delta$,
$$|x^2 - p^2| = |x + p||x - p| < 2\delta = \epsilon$$
which verifies the condition of continuity at p.

Uniformly continuous functions

Review:

- A function f is *uniformly continuous* on S if, for any $\epsilon > 0$ there is $\delta > 0$ such that if $x, y \in S$ and $|x - y| < \delta$, then $|f(x) - f(y)| < \epsilon$.
- We check from the definition that $f(x) = x^2$ is uniformly continuous on $[0, 1]$. Given $\epsilon > 0$, choose $\delta = \frac{\epsilon}{2}$. Then for $x, y \in [0, 1]$ such that $|x - y| < \delta$,
$$|x^2 - y^2| = |x + y||x - y| < 2\delta = \epsilon.$$
- We proved in Lecture 7 that any function f which is continuous on a closed interval $[a, b]$ is uniformly continuous there.

The vector space $\mathbb{R}^n$

- $\mathbb{R}^n$ consists of *n*-tuples of real numbers $\underline{x} = (x_1, x_2, ..., x_n)$ where $x_1, ..., x_n \in \mathbb{R}$.
- Pairs of elements of $\mathbb{R}^n$ are added and subtracted component-wise, that is, $\underline{x} + \underline{y} = (x_1 + y_1, x_2 + y_2, ..., x_n + y_n)$.
- If $a \in \mathbb{R}$, $a \cdot \underline{x} = (ax_1, ax_2, ..., ax_n)$. This is called *scalar multiplication*.
- These definitions make $\mathbb{R}^n$ into a *real vector space*.

The Euclidean norm on $\mathbb{R}^n$

- The *Euclidean norm* of a vector $\underline{x} \in \mathbb{R}^n$ is
$$\|\underline{x}\|_2 = \sqrt{x_1^2 + x_2^2 + ... + x_n^2}.$$
- This satisfies $\|\underline{x}\|_2 = 0$ if and only if $\underline{x} = 0 = (0, 0, ..., 0)$ and, for $a \in \mathbb{R}$, $\|a \cdot \underline{x}\|_2 = |a|\|\underline{x}\|_2$.
- We checked in Lecture 4 that $\|\underline{x} + \underline{y}\|_2 \leq \|\underline{x}\|_2 + \|\underline{y}\|_2$.
- When $n = 1$, $\|\cdot\|_2$ reduces to the absolute value $|\cdot|$.

The Euclidean distance on $\mathbb{R}^n$

- The *Euclidean distance* between two vectors $\underline{x}$ and $\underline{y}$ in $\mathbb{R}^n$ is defined to be
$$d(\underline{x}, \underline{y}) = \|\underline{x} - \underline{y}\|_2.$$
- The distance $d(\underline{x}, \underline{y})$ satisfies
  1. For all $\underline{x}, \underline{y} \in \mathbb{R}^n$, $d(\underline{x}, \underline{y}) = 0$ if and only if $\underline{x} = \underline{y}$.
  2. For all $\underline{x}, \underline{y} \in \mathbb{R}^n$, $d(\underline{x}, \underline{y}) = d(\underline{y}, \underline{x})$.
  3. The triangle inequality holds: For all $\underline{x}, \underline{y}, \underline{z} \in \mathbb{R}^n$,
  $$d(\underline{x}, \underline{z}) \leq d(\underline{x}, \underline{y}) + d(\underline{y}, \underline{z}).$$
- A distance which satisfies the three properties listed is called a *metric*.

Continuity in Euclidean space

Definition

Let $m, n \geq 1$. Let $S \subset \mathbb{R}^m$ and let $f : S \to \mathbb{R}^n$. The function f is *continuous at a point* $\underline{p} \in S$ if, for each $\epsilon > 0$ there exists $\delta > 0$ such that if $\underline{x} \in S$ then

$$d(\underline{x}, \underline{p}) < \delta \qquad \Rightarrow \qquad d(f(\underline{x}), f(\underline{p})) < \epsilon.$$

f is said to be continuous on S if it is continuous at each point $\underline{p} \in S$.

Continuity in Euclidean space

- Given $S \subset \mathbb{R}^m$, a function $f : S \to \mathbb{R}^n$ has the form $f(\underline{x}) = (f_1(\underline{x}), ..., f_n(\underline{x}))$ where $f_1(\underline{x}), ..., f_n(\underline{x})$ are *component functions* each mapping $S \to \mathbb{R}$.
- f is continuous at $\underline{p} \in \mathbb{R}^n$ if and only if all of the component functions $f_1, ..., f_n$ are continuous at $\underline{p}$.
- For example $f : \mathbb{R}^2 \setminus \{(x, y) : y \neq 0\} \to \mathbb{R}^2$, defined by $f(x, y) = \left(xy, \frac{x}{y}\right)$ is continuous.

For a proof of these facts, see Homework 6.

Sequences and subsequences

- A *sequence* taking values in a set S is a function $f : \mathbb{N} \to S$.
- Instead of writing $f(0)$, $f(1)$, $f(2)$, ... we often write $x_0, x_1, x_2, \ldots$ or $a_0, a_1, a_2, \ldots$ etc. Other common notation includes $\{x_n\}_{n=0}^{\infty}$.
- A sequence $\{b_n\}_{n=0}^{\infty}$ is a *subsequence* of a sequence $\{a_n\}_{n=0}^{\infty}$ if there is a strictly increasing function $f : \mathbb{N} \to \mathbb{N}$ such that $b_n = a_{f(n)}$. Intuitively, $\{b_n\}_{n=0}^{\infty}$ may 'skip over' some terms of $\{a_n\}_{n=0}^{\infty}$.

Examples of sequences

Here are some examples of sequences:

- The constant sequence $a_n = 1$ for all n: 1, 1, 1, 1, ...
- The Fibonacci sequence is defined by $a_0 = a_1 = 1$, for $n \geq 2$, $a_n = a_{n-1} + a_{n-2}$. Its first few terms are

$$1, 1, 2, 3, 5, 8, \ldots.$$

- The sequence $a_n = 2^{-n}$, with first few terms

$$1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots.$$

- a_n is the length n truncation of the binary representation of 1,

$$0, 0.1, 0.11, 0.111, 0.1111, \ldots.$$

Examples of sequences and subsequences

- Let $\{a_n = 2^n\}_{n=0}^{\infty}$ be the sequence of powers of 2, with terms

$$1, 2, 4, 8, 16, 32, 64, 128, 256, 512, 1024, 2048, ...$$

- The sequences $\{b_n = 4^n\}_{n=0}^{\infty}$ and $\{c_n = 8^n\}_{n=0}^{\infty}$ are both subsequences

$$b_n : 1, 4, 16, 64, 256, 1024, ...$$
$$c_n : 1, 8, 64, 512, 4096, ...$$

found by taking from a_n the even index terms, respectively those terms with index divisible by 3.

Examples of sequences and subsequences

A sequence could take its values in higher dimensional Euclidean space, for instance, $\{a_n = (n, n^2)\}_{n=0}^{\infty}$ is a sequence in $\mathbb{R}^2$ with first few terms

$$a_n : (0, 0), (1, 1), (2, 4), (3, 9), (4, 16), (5, 25), (6, 36), \ldots.$$

Subsequence of a subsequence

Theorem

Let $\{y_n\}_{n=0}^{\infty}$ be a subsequence of sequence $\{x_n\}_{n=0}^{\infty}$, and let $\{z_n\}_{n=0}^{\infty}$ be a subsequence of $\{y_n\}_{n=0}^{\infty}$. Then $\{z_n\}_{n=0}^{\infty}$ is a subsequence of $\{x_n\}_{n=0}^{\infty}$.

Proof.

Let $f : \mathbb{N} \to \mathbb{N}$ and $g : \mathbb{N} \to \mathbb{N}$ be strictly increasing functions such that $y_n = x_{f(n)}$ and $z_n = y_{g(n)}$. Then $z_n = x_{f(g(n))}$. The function $f \circ g$ is the composition of strictly increasing functions, hence strictly increasing. Thus $\{z_n\}_{n=0}^{\infty}$ is a subsequence of $\{x_n\}_{n=0}^{\infty}$. $\square$

The limit of a sequence

Definition

A sequence $\{a_n\}_{n=0}^{\infty}$ contained in Euclidean space $\mathbb{R}^n$ has a limit A if, for each $\epsilon > 0$ there exists $N \geq 0$ such that

$$n > N \qquad \Rightarrow \qquad d(a_n, A) < \epsilon.$$

A sequence which has a limit is said to be *convergent*.

Examples of limits

- The constant sequence $a_n = 1$ has limit 1, written $\lim_{n\to\infty} a_n = 1$. Given $\epsilon > 0$, $N = 0$ suffices to obtain the required accuracy.
- The sequence of binary approximations to 1, given by $a_n = 1 - 2^{-n}$ has limit 1. Given $\epsilon > 0$, any $N > \log_2 \frac{1}{\epsilon}$ will suffice.
- The sequence $1, 0, 1, 0, 1, 0, 1, 0, \ldots$ which alternates between 1 and 0 does not have a limit. It has as subsequences the constant sequence 1 with limit 1, and the constant sequence 0 with limit 0.

Images of limits

Theorem

Let $\{\underline{a}_n\}_{n=0}^{\infty}$ be a sequence in Euclidean space $\mathbb{R}^n$, with limit $\lim_{n\to\infty} \underline{a}_n = \underline{x}$. Then any subsequence of $\{\underline{a}_n\}_{n=0}^{\infty}$ has limit $\underline{x}$.

See Homework 6.

Images of limits

Theorem

Let $\{\underline{a}_n\}_{n=0}^{\infty}$ be a sequence in the Euclidean space $\mathbb{R}^m$, $m \geq 1$, with $\lim_{n\to\infty} \underline{a}_n = \underline{x}$. Let f be a function defined on $\mathbb{R}^m$, which is continuous at $\underline{x}$. Then

$$\lim_{n\to\infty} f(\underline{a}_n) = f(\underline{x}).$$

Proof.

Given $\epsilon > 0$, by the continuity of f at $\underline{x}$ there exists $\delta > 0$, such that if $d(\underline{y}, \underline{x}) < \delta$ then $d(f(\underline{y}), f(\underline{x})) < \epsilon$. Now choose N such that $n > N$ implies $d(\underline{a}_n, \underline{x}) < \delta$. It follows that for $n > N$, $d(f(\underline{a}_n), f(\underline{x})) < \epsilon$, which proves

$$\lim_{n\to\infty} f(\underline{a}_n) = f(\underline{x}).$$

□

Sequential compactness

Definition

A set S is said to be *sequentially compact* if any sequence contained in S has a subsequence converging to a limit in S.

Sequential compactness of a closed interval

Theorem

Let $a < b$ be real numbers. The interval $[a, b]$ is sequentially compact.

Proof.

The proof is by the method of bisection. Let $\{x_n\}_{n=0}^{\infty}$ be a sequence with values in $[a, b]$.

- Let $[a_0, b_0] = [a, b]$. For $i \geq 0$, choose $[a_{i+1}, b_{i+1}]$ to be a half of $[a_i, b_i]$ which contains infinitely many terms of the sequence $\{x_n\}$.
- Define a subsequence $\{y_n = x_{f(n)}\}_{n=0}^{\infty}$ of $\{x_n\}_{n=0}^{\infty}$ by defining $f(0) = 0$, and, for $n \geq 1$, $f(n)$ is the first index after $f(n-1)$ such that $x_{f(n)} \in [a_n, b_n]$.
- Let $\alpha = \sup\{a_n : n \geq 0\}$. We have $\alpha \in [a, b]$. Given $\epsilon > 0$, choose N sufficiently large such that $n \geq N$ implies $[a_n, b_n] \subset [\alpha - \epsilon, \alpha + \epsilon]$. Thus, for $n > N$, $|y_n - \alpha| < \epsilon$, so $\lim_{n\to\infty} y_n = \alpha$.

$\square$

Sequential compactness of a closed rectangle

Theorem

Let $a < b$ and $c < d$ be real numbers. The closed rectangle $[a, b] \times [c, d] \subset \mathbb{R}^2$ is sequentially compact.

Proof.

- Let $\{\underline{x}_n = (x_{n,1}, x_{n,2})\}_{n=0}^{\infty}$ be a sequence in $[a, b] \times [c, d]$.
- Apply the previous theorem to find a subsequence $\{\underline{y}_n = (y_{n,1}, y_{n,2})\}_{n=0}^{\infty}$ of $\{\underline{x}_n\}_{n=0}^{\infty}$ such that $y_{n,1}$ converges to limit x_1.
- Now find a subsequence $\{\underline{z}_n = (z_{n,1}, z_{n,2})\}_{n=0}^{\infty}$ of $\{\underline{y}_n\}_{n=0}^{\infty}$ such that $z_{n,2}$ converges to limit x_2. $z_{n,1}$ still converges to x_1.
- To prove $\lim_{n\to\infty} \underline{z}_n = (x_1, x_2)$, given $\epsilon > 0$ choose N sufficiently large such that $n > N$ implies $|z_{n,1} - x_1| < \frac{\epsilon}{2}$ and $|z_{n,2} - x_2| < \frac{\epsilon}{2}$. Then

$$d(\underline{z}_n, (x_1, x_2)) \leq d(\underline{z}_n, (x_1, z_{n,2})) + d((x_1, z_{n,2}), (x_1, x_2)) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

□

Sperner's Lemma in 1d

Lemma (1d Sperner's lemma)

Suppose an interval $[a, b]$ is partitioned into finitely many subintervals by points $a = x_0 < x_1 < x_2 < ... < x_n = b$. Color each point either red or blue, and color a red and b blue. Then there is a segment (x_i, x_{i+1}) which has endpoints of opposite color.

Proof.

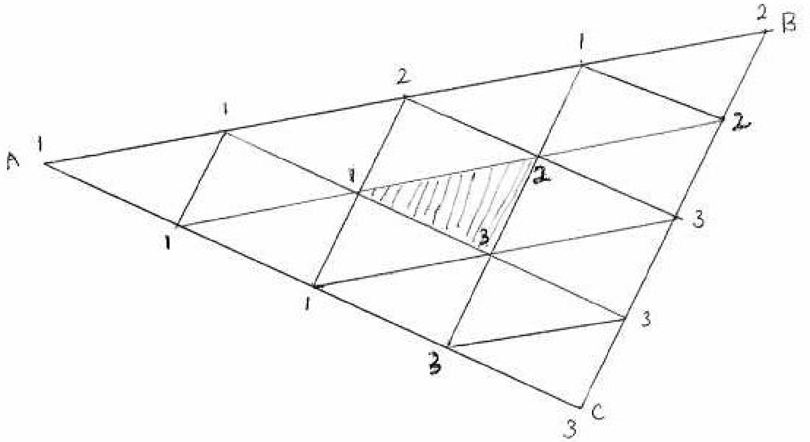
Since a and b receive opposite colors, the number of color changes passing from x_0 to x_n is odd, hence non-zero. $\square$

Sperner's Lemma in 2d

Let $[ABC]$ be a triangle.

- A *proper subdivision* of $[ABC]$ is a partition of $[ABC]$ into sub-triangles such that any two adjacent sub-triangles have a full edge in common.
- Given a proper subdivision of $[ABC]$, a *proper coloring* of the subdivision is an assignment of colors 1, 2, 3 to the vertices of the subdivision such that
  1. Vertices A, B, C are colored 1, 2, 3
  2. Any vertex lying on an edge of $[ABC]$ receives one of the colors of the two endpoints of the edge, e.g. a vertex on $[AB]$ is colored either 1 or 2.

A proper subdivision and coloring



Sperner's Lemma in 2d

Lemma (2d Sperner's Lemma)

Given a proper coloring of a proper subdivision of a triangle $[ABC]$, there is a sub-triangle whose vertices receive all three colors 1, 2, 3.

Proof from Jacob Fox's notes.

- Let Q denote the number of sub-triangles with colors $(1,1,2)$ or $(1,2,2)$ and R denote the number of sub-triangles with colors $(1,2,3)$
- Let X denote the number of boundary edges colored $(1,2)$ and Y the number of interior edges colored $(1,2)$.
- Let N denote the number of pairs (T,E) where T is a sub-triangle, and E is an edge colored $(1,2)$.
- $N = 2Q + R = X + 2Y$. Since X is odd by the 1d Sperner's lemma, R is odd, so $R > 0$.

□

Brouwer's fixed point theorem in 1d

Theorem (1d Brouwer's fixed point theorem)

Let $a < b$ and let $f : [a, b] \to [a, b]$ be continuous. The fixed point equation $f(x) = x$ has a solution.

Proof.

Consider $g(x) = f(x) - x$. A fixed point of f is a zero of g. One has $g(a) \geq 0$ and $g(b) \leq 0$. Thus, either an endpoint is a zero, or by the intermediate value theorem, there exists $a < c < b$ such that $g(c) = 0$. □

Brouwer's fixed point theorem in 2d

Theorem (2d Brouwer's fixed point theorem)

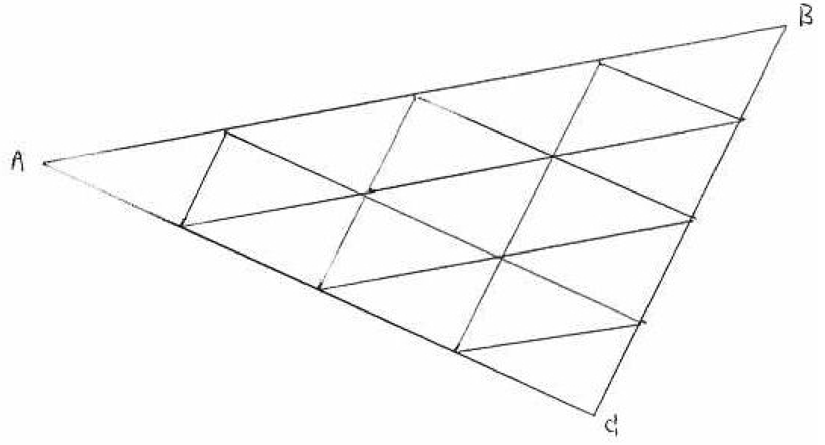
Let $[ABC] \subset \mathbb{R}^2$ *be a triangle (including the interior), and let* $f : [ABC] \to [ABC]$ *be continuous. The fixed point equation* $f(\underline{x}) = \underline{x}$ *has a solution* $\underline{x}_0 \in [ABC]$.

Proof.

- Given a triangle $[A_0B_0C_0]$, define its standard level 1 subdivision to be the subdivision into 4 sub-triangles obtained by connecting the midpoints of the sides.
- Define the standard level n subdivision to be the subdivision obtained by applying a standard level 1 subdivision to each sub-triangle in the standard level $n-1$ subdivision.
- Each sub-triangle in the standard level n subdivision is similar to $[A_0B_0C_0]$ and has been rescaled by $\frac{1}{2^n}$.

□

A standard level 2 subdivison



Brouwer's fixed point theorem in 2d

Proof.

- Treating A, B, C as vectors/points in $\mathbb{R}^2$,

$$[ABC] = \{x_1 A + x_2 B + x_3 C : 0 \leq x_1, x_2, x_3, x_1 + x_2 + x_3 = 1\}$$

  This identifies $[ABC]$ with the standard simplex

$$\Delta_2 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1, x_2, x_3 \geq 0, x_1 + x_2 + x_3 = 1\}.$$

- For instance, A corresponds to $(1, 0, 0)$, B to $(0, 1, 0)$ and C to $(0, 0, 1)$, and the segment $[A, B]$ to $\{(x, 1 - x, 0) : 0 \leq x \leq 1\}$.

$\square$

Brouwer's fixed point theorem in 2d

Proof.

- Treat the function $f : [ABC] \to [ABC]$ as a function $\tilde{f} : \Delta_2 \to \Delta_2$ via the identification.
- Given the standard level n subdivision of $[ABC]$ define a coloring of the vertices of the subdivision by assigning to point $\underline{x} = (x_1, x_2, x_3)$ an index $i \in \{1, 2, 3\}$ such that $x_i > f_i(\underline{x})$.
- Note that A receives color 1, B color 2 and C color 3, since $x_2 = x_3 = 0$ at A, etc. and any point of $[AB]$ has $x_3 = 0$, hence receives color 1 or 2, etc. This verifies that the coloring is proper.
- By Sperner's Lemma, at each level n there is a sub-triangle $[A_nB_nC_n]$ with vertices colored $1, 2, 3$.

□

Brouwer's fixed point theorem in 2d

Proof.

- By sequential compactness in rectangles of $\mathbb{R}^2$, there is a subsequence $\{A_{g(n)}\}$ of the sequence $\{A_n\}$ which converges to a point $\underline{x} \in \mathbb{R}^2$. It's possible to check that $\underline{x} \in [ABC]$.
- Given $\epsilon > 0$ let $0 < \delta < \epsilon$ be sufficiently small that $d(\underline{y}, \underline{x}) < \delta$ implies $d(f(\underline{y}), f(\underline{x})) < \epsilon$.
- There exists $N > 0$ such that for $n > N$, $\max(d(A_{g(n)}, \underline{x}), d(B_{g(n)}, \underline{x}), d(C_{g(n)}, \underline{x})) < \delta$.
- Since $[A_{g(n)}B_{g(n)}C_{g(n)}]$ is colored 1,2,3,

$$
\begin{aligned}
f_1(\underline{x}) &< f_1(A_{g(n)}) + \epsilon < A_{g(n),1} + \epsilon < x_1 + \delta + \epsilon \\
f_2(\underline{x}) &< f_2(B_{g(n)}) + \epsilon < B_{g(n),2} + \epsilon < x_2 + \delta + \epsilon \\
f_3(\underline{x}) &< f_3(C_{g(n)}) + \epsilon < C_{g(n),3} + \epsilon < x_3 + \delta + \epsilon.
\end{aligned}
$$

□

Brouwer's fixed point theorem in 2d

Proof.

- Recall $0 < \delta < \epsilon$, and we've checked,

$$f_1(\underline{x}) < x_1 + \delta + \epsilon$$
$$f_2(\underline{x}) < x_2 + \delta + \epsilon$$
$$f_3(\underline{x}) < x_3 + \delta + \epsilon.$$

- But $f_1(\underline{x}) + f_2(\underline{x}) + f_3(\underline{x}) = 1 = x_1 + x_3 + x_3$. Thus

$$f_1(\underline{x}) > x_1 - 2\delta - 2\epsilon, \; f_2(\underline{x}) > x_2 - 2\delta - 2\epsilon, \; f_3(\underline{x}) > x_3 - 2\delta - 2\epsilon.$$

Letting $\epsilon \to 0$, $f(\underline{x}) = \underline{x}$.

$\square$

Brouwer's fixed point theorem in other domains

Theorem

Let $D \subset \mathbb{R}^2$ and $F : D \to [ABC]$ a continuous bijection with continuous inverse. Then any continuous function $f : D \to D$ has a fixed point.

Proof.

The map $F \circ f \circ F^{-1} : [ABC] \to [ABC]$ is continuous, and hence has a fixed point x. It follows that $f(F^{-1}(x)) = F^{-1}(x)$. $\square$

Ex: A continuous map from the ball $B^2 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$ to itself has a fixed point, see Homework 6.

Applications of Brouwer's fixed point theorem

Brouwer's fixed point theorem finds applications in various fields, from partial differential equations, to economics.