REAL ANALYSIS CRASH COURSE
LECTURE 9: CONNECTEDNESS, CANTOR SET

WILLIAM WU

Reading: Chapter 2 in the book, still. (It's sooo deep.)

1. Agenda

- Warmup
- Homework Solutions
- Homeomorphisms
- Connectedness, and sample problems
- Cantor Set

2. Warmup/Quiz

(1) Give four different definitions of continuity.
(2) Give two different definitions of compactness.
(3) What is uniform continuity?
(4) (Not analysis) Show that AB and BA have the same eigenvalues.

3. Homework Solutions

(1) Prove from the definition that $f(x) = 1/x$ is uniformly continuous over the interval $[1, 6)$.

Solution Choose $\delta = \epsilon$. Then

$$|f(x) - f(y)| = \frac{|y - x|}{|xy|} \leq |y - x| \leq \epsilon.$$

□

(2) Use compactness to prove Rolle's theorem: for a continuous function and differentiable function f over $[a, b]$, if $f(a) = f(b)$, then there exists a $c \in (a, b)$ such that $f'(c) = 0$.

Solution The key is to think about the maximum point, and to recall from calculus the fact that local maxima occur when the derivative is zeroed.

If the function f is identically zero (fancy terminology for the "always zero function"), then there is nothing to prove. Hence, assume f is not identically zero. Since f is a continuous function on a compact set, we know that it achieves its maximum and minimum on that set. Either the maximum or the minimum must be non-zero, or else

Date: July 30, 2008.

the function would be identically equal to zero. Assume for now that the maximum is achieved at c and $f(c) \neq 0$. Since $f(a) = f(b) = c$, we know that $c \in (a, b)$, and therefore f is differentiable at c. Note that $f(x) \leq f(c)$ since $f(c)$ is a maximum.

- if $x < c$ then $\frac{f(x)-f(c)}{x-c} \geq 0$ for all $x < c$.
- if $x > c$ then $\frac{f(x)-f(c)}{x-c} \leq 0$ for all $x > c$.

The first inequality implies that as x approaches c from the left, the limit must be greater than or equal to zero. The second one says that as x approaches c from the right, the limit must be less than or equal to zero. But since f is differentiable at c we know that both right and left handed limits exist and must agree. Therefore, $f'(c) = 0$. The proof is similar in the case where $f(c)$ is a minimum and $f(c) \neq 0$. □

(3) (This is problem 38 on page 119 in Pugh.)
Let $\| \ \|$ be any norm on $\mathbf{R}^m$, and let $B = \{x \in \mathbf{R}^m : \|x\| \leq 1\}$.

(a) Prove that B is convex.
Just use the triangle inequality, and the positive homogeneity or positive scalability property of norms.

$$\begin{aligned}\|\lambda x + (1-\lambda)y\| &\leq \|\lambda x\| + \|(1-\lambda)y\| \\ &\leq \lambda\|x\| + (1-\lambda)\|y\| \\ &\leq \lambda + (1-\lambda) \\ &= 1.\end{aligned}$$

(b) (Tough) Prove that B is compact.
Well I'm still stuck on this. See attached solution I found. I don't get it, but maybe you can explain it to me =(

(4) (This is problem 43 on page 119 in Pugh.) Suppose that (K_n) is a nested sequence of compact non-empty sets, $K_1 \supset K_2 \supset \ldots$, and $K := \bigcap K_n$ Suppose that for some $\mu > 0$, each diam $K_n \geq \mu$. Prove that diam $K \geq \mu$.[1]

Solution An outline:

- Construct sequences $x_i, y_i \in K_i$.
- By compactness of K_1, there exist convergent subsequences of these sequences; we can take the intersection of the subsequence indices to make two subsequences with the same indexing.
- x_{m_i} and y_{m_i} converge to x and y.
- As we proved when looking at your own approach to a previous homework problem, the distance metric $d(a, b)$ is continuous in both arguments. So by the sequential definition of continuity, if $d(x_{m_i}, y_{m_i}) \geq \mu$ for all i, then pulling the limit inside the arguments of the continuous function, $d(x, y) \geq \mu$.

□

4. Uniform Continuity

A function is called uniformly continuous if for every $\epsilon > 0$ there exists a $\delta > 0$ such that $|x - t| > \delta$ implies that $|f(x) - f(t)| < \epsilon$.

- Question: What is the difference between uniform continuity and regular continuity?
- Show that a uniformly continuous function is continuous, but continuity does not imply uniform continuity. (For the latter, prove that $\sin(1/x)$ is continuous over $(0, 1)$, but is not uniformly continuous there.

Why did people invent this concept? One motivation has to do with convergence of functions. (Show example involving $f_n(x) = x^n$ over $[0, 1]$.)

From the example, we see that you can have a sequence of continuous functions $f_n : X \to \mathbf{R}$ such that for any fixed $x_0 \in X$, the sequence $(f_n(x_0))$ is convergent – and yet, the limiting function defined by $f(x) = \lim_{n\to\infty} f_n(x)$ is *not* continuous. This is aesthetically disturbing, to have a sequence of continuous functions converging to a non-continuous function. To assure that the limiting function is continuous as well, uniform continuity is sufficient.

Theorem 4.1. *Every continuous function defined on a compact set is uniformly continuous.*

Proof. Use contradiction. Then there is some $\epsilon > 0$ such that no matter how small δ is, there are points $p, q \in M$ with $d(p, q) < \delta$ but $|fp - fq| \geq \epsilon$. Take $\delta = 1/n$ and construct corresponding sequences p_n and q_n. Then bust out the convergent subsequences bestowed upon us by compactness ... □

You will see this concept again when we get to function spaces.

[1] The diameter of a set S is the supremum of the distances $d(x, y)$ between points in S. See page 79.

5. Homeomorphisms

A homeomorphism is a bi-continuous bijection. Many problems in topology are motivated by showing that two shapes are homeomorphic, meaning I can continuously stretch and mold one shape into the other: e.g., a coffee cup with a handle, and a donut. (To a topologist, a donut and a coffee cup appear the same.)

Corollary 5.1. *If M is compact and M is homeomorphic to N, then N is compact.*

Proof. (easy) □

Application: $[0, 1]$ and $\mathbf{R}$ are not homeomorphic.

Theorem 5.2. *If M is compact, then a continuous bijection $f : M \to N$ is a homeomorphism – its inverse bijection $f^{-1} : N \to M$ is continuous.*

Proof. (see book) □

This is a very useful result because you can save a lot of work in not having to show that f^{-1} is continuous.

6. Connectedness

6.1. Clopen Definition.

- If M has a proper clopen subset, then M is called disconnected.
- M is connected if it is not disconnected.

Why? Recall that the complement of a clopen set is clopen. Thus, if M is disconnected, we can write it as the disjoint union of A and A^c where A is clopen.

6.2. Main Theorems.

Theorem 6.1. *If M is connected, $f : M \to N$ is continuous, and f is onto, then N is connected. The continuous image of a connected set is connected.*

Proof. Use contradiction. Pull back from image space to domain space using the continuous function f and the fact that the preimages will be clopen too. □

Corollary 6.2. *Two spaces homeomorphic to each other are either both connected or both disconnected. (Connectedness is a topological property.)*

Example: The union of two disjoint intervals is not homeomorphic to a single interval. (Why?)

Theorem 6.3. $\mathbf{R}$ *is connected.*

Proof. Suppose not connected. Then we can find a proper nonempty clopen subset $U \subset \mathbf{R}$. Use fact that U is countable disjoint union of open intervals (a theorem we proved some lectures ago). Also add that the endpoints of these intervals do not belong to U (consider a small example for the basic reason why). Now consider (a, b) as one of these intervals. Since U is also closed, we would require $b \in U$, which is a contradiction unless $b = \infty$. Similarly, $a = -\infty$. So $U = \mathbf{R}$. But then its not a proper subset anymore, so we have a contradiction. □

6.3. **The Point Removal Technique.**

- $[a, b]$ is not homeomorphic to the circle S^1. (Removing one point of the interval disconnects it, but the circle remains connected after removing any one point. Explain rigorously)
- Circle not homeomorphic to figure eight. (Remove two points.)
- Circle not homeomorphic to the disc. (Remove two points.)

Notice how we didn't develop a whole lot of machinery in our discussion of connectedness, but already we are getting results that are rather impressive. If you asked someone who hadn't studied analysis to rigorously argue that there does not exist a bijective continuous map from a circle to a disc, he'd probably have no idea. It's not clear that something like that shouldn't exist.

6.4. **Clopen Separation.** Say S is disconnected. Then it has a proper closed subset A. But then if we think of S as the entire metric space, then the complement of A can be written as

$$A^c = B = S \setminus A$$

and then since the complement of a clopen set is clopen, we have that if S is disconnected, then it can be written as the disjoint union of two proper clopen nonempty subsets

$$A \sqcup B = S.$$

6.5. **The Two Set Approach.** The clopen approach is Pugh's approach to connectedness. However, sometimes the following alternative definition of disconnectedness is useful, and actually feels more natural: An open set S is called disconnected if there are two open, non-empty sets U and V such that

(1) $U \cap V = \emptyset$
(2) $U \cup V = S$.

If the set S is not open, then we say S is disconnected if there are two open sets U and V such that

(1) $U \cap S \neq \emptyset$ and $V \cap S \neq \emptyset$
(2) $(U \cap S) \cap (V \cap S) = \emptyset$
(3) $(U \cap S) \cup (V \cap S) = S$.

Some exercises:

(1) Is the set $\{x \in \mathbf{R} : |x| < 1, x \neq 0\}$ connected or disconnected ? What about the set $\{x \in \mathbf{R} : |x|1, x \neq 0\}$
(2) Is the set $[-1, 1]$ connected or disconnected ?
(3) Is the set of rational numbers connected or disconnected ? How about the irrationals ?
(4) Prove this theorem:

Theorem 6.4. *If S is any connected subset of $\mathbf{R}$ then S must be an interval.*

Proof. Suppose not an interval. Then $\exists a, b \in S$ and t between a and b such that t is not in S. Define $U = (-\infty, t)$ and $V = (t, \infty)$. Then ... □

6.6. **Totally Disconnected.** In fact, a set can be disconnected at every point.

Clopen Approach:

Definition A metric space is totally disconnected if each point $p \in S$ has arbitrarily small clopen neighborhoods.

Two-Set approach:

Definition A set S is called totally disconnected if for each distinct $x, y \in S$ there exist disjoint open sets U and V such that $x \in U, y \in V$, and $(U \cap S) \cup (V \cap S) = S$.

6.7. **Path-Connectedness.** You'll explore this topic in the homework.

7. The Infamous Cantor Set

7.1. **Motivation.** You may be wondering why we are studying the Cantor Set. After all, you won't find any Cantor sets in real life. Several reasons:

(1) The Cantor set has many properties which you would not expect a set to be able to simultaneously have. It has measure zero. But it is also uncountable. It is totally disconnected. But it is also compact.
(2) By studying some "pathological" mathematical phenomena like this, you will learn to become more skeptical of your intuition. Intuition is very important and powerful, but it is also necessary to realize that it has limitations. Our intuitions largely come from our experiences in $\mathbf{R}^3$, and the geometry of $\mathbf{R}^3$, with finite numbers of objects. However, once math problems demand us to deal with the infinite, or metric spaces foreign from human experience, we have to be more cautious about what intuition says, and rely a little more on systematic analysis. The Cantor set is one of a long list of examples like this which will temper your intuition and help you appreciate these analytical tools.
(3) Much of the development analysis was motivated in trying to understand seemingly pathological or confusing phenomena (e.g., the rearrangement paradox for conditionally convergent series). So these objects were important to the history of the subject.

(4) Lastly, these pathological things may not be as useless as everyone thinks. The best example is Brownian motion, a random process whose sample paths are continuous and yet nowhere differentiable. (Can you imagine a function that is continuous and yet has no definable tangent anywhere?) No one had any use for this idea until over half a century later, when investment bankers use it to model the evolution of stock options, and amounting to billions of dollars in financial transactions every day. Maybe someday someone like you will find a practical use for the Cantor set.

7.2. **The Properties of the Cantor Set.**

(1) Describe interval construction. (Could refer to text.) Remove open middle thirds from remaining intervals.
(2) Nested sequence of sets. At each iteration we have a union of closed intervals.
(3) How many intervals do we have at the n^{th} stage?
(4) What are their endpoints?
(5) Verify this recursive formula:

$$C^n = \frac{C^{n-1}}{3} \cup (\frac{2}{3} + \frac{C^{n-1}}{3}).$$

(6) Cantor set is defined as the set of points that is not excluded.

$$C = \bigcap C^n$$

.

(7) What is the measure of the set? (Compute measure of intervals that have been subtracted out.)
(8) But is it empty? No. What points are left?
(9) Are there any points left that are not interval endpoints? (Consider 1/4.)
(10) Natural next question: does the Cantor set contain a countable infinity, or an uncountable infinity of points? (We might expect a countable number of points because that would help explain the measure zero result.)
(11) Actually, it turns out to be uncountable. Infinite address string argument:

$$C_0 = [0, 1/3] \qquad C_2 = [2/3, 1]$$

$$C^2 = C_{00} \cup C_{02} \quad \cup \quad C_{20} \cup C_{22}$$

An interval of C^n is thus coded by an address string of n symbols. Then

$$C_{\omega_1} \supset C_{\omega_2} \supset \dots C_{\omega_1 \cdots \omega_n} \supset \dots$$

and the intersection is a point in the Cantor set:

$$p(\omega) = \bigcap_{n \in \mathbf{N}} C_{\omega|n}$$

This is a 1-1 correspondence. Verify for yourself:
(a) Any two Cantor points have different addresses.

(b) Any two addresses map to different Cantor points.

This actually is a ternary encoding of the Cantor set:

The Cantor set consists of all numbers in $[0, 1]$ whose decimal expansions, written in base-3, do not contain a 1. (Example: .0222020202002022 is in the Cantor set.)

How to complete the uncountability argument? (You fill in the rest)

(12) Thus the Cantor set is uncountable *and* it has measure zero!

(13) The Cantor set is compact. (Why?)

(14) The Cantor set contains no intervals. (Why?)

(15) The Cantor set is totally disconnected.

If something is not connected, it can break up into connected components. So, a component is a "maximum" connected set. For a totally disconnected set, every component (or maximum connected subset) is a single point.

Now, on the real line, a set is connected iff it is an interval. Since the Cantor set contains no intervals, it must not have any connected sets impling the maximum connected sets are points.... meaning the Cantor set is totally disconnected.

There's lots more weird stuff about the Cantor set in Section 6*.

8. Foreshadowing

This concludes our introduction to topology, for the most part. If you want to know more, there are still important things in Chapter 2 you can read, namely about relative topology (Inheritance Theorem) and completion of metric spaces.

We will be moving to more concrete things now. It is really great that we have learned the tools we did so far, because now you will be able to understand the proofs behind many powerful applied math theorems. These are proofs that the vast majority of college students, including engineers, never ever see because they wouldn't be able to understand them without having taken real analysis – and they probably couldn't handle real analysis anyways. I've had a list of ideas in my head of things we could study ... some highly proper subset of the following:

- Calculus – now with proofs! (why do all those rules actually work?)
- Function spaces and approximating functions (will lead to Fourier Analysis)
- Introduction to Ordinary Differential Equations, with proofs of certain key theorems
- Dynamic Programming and the J-Operator with policy iteration convergence
- Fubini's Theorem and Term-by-Term Differentiation and other exchanges of operations
- Lagrange Multipliers
- General Stokes Theorem (this would be new for me too)
- Inverse Function Theorem and linearization methods in feedback control
- The Lebesgue Integral

Then we'll be done with real analysis. Then, who knows. Some other topics, in descending order of usefulness to you at this stage in my opinion:

(1) probability theory and statistics
(2) optimization theory + more linear algebra
(3) discrete math, combinatorics, and algorithms
(4) abstract algebra
(5) complex analysis
(6) signal processing and control theory
(7) willywutang nunchaku theory

9. Homework

9.1. **Connectedness and Disconnectedness.** Let (M, d) be a metric space.

(a) Prove that the following two conditions (i. and ii.) on M are equivalent
  i. There is a continuous, surjective map $f : M \to \{0, 1\}$.
  ii. There exist non-empty open sets $U, V \subseteq M$ such that $U \cup V = M$ and $U \cap V = \emptyset$.

(b) Let $A \subseteq M$ be connected. Let B be such that $A \subseteq B \subseteq \overline{A}$. Let C be connected such that $A \cap C \neq \emptyset$.
  i. Show that B is connected.
  ii. Show that $A \cup C$ is connected.

(c) Let (N, δ) be another metric space, $f : M \to N$ continuous. Show that if M is connected, then $f(M)$, the image of M in N under f, is connected. Hint: Prove the contrapositive.

(d) Using the concepts of connectedness and disconnectedness, prove this general version of the Intermediate Value Theorem:

  If there exist $p, q \in M$ with $f(p) < 0$ and $f(q) > 0$ then there is some $r \in M$ with $f(r) = 0$.

9.2. **Path-Connectedness.** Let (M, d) be a metric space. We say that M is *path-connected* if for each $p, q \in M$ there is a continuous map $\alpha : [0, 1] \to M$ with $\alpha(0) = p$ and $\alpha(1) = q$. The map α is called a *path from p to q*.

(a) The formal definition of an interval in $\mathbf{R}$ is that an interval is a set $I \subseteq \mathbf{R}$ with the property that if $t_1 < t_2 < t_3$ with $t_1, t_3 \in I$ then $t_2 \in I$.

  Let I be an interval in $\mathbf{R}$. Let $U, V \subseteq I$ be non-empty, disjoint, open subsets. As both are non-empty, there are $a, b \in I$ with $a \in U$ and $b \in V$. Assume, without loss of generality, that $a < b$.

  Let $C = \{c \in I : [a, c] \subseteq U\}$. Show that C has a supremum in I and that this supremum is neither in U nor in V. Hint: Show that if it lies in U then it is not an upper bound, and if it lies in V, then it is not the least upper bound.

  Deduce that I is connected.

(b) Show that if M is path-connected then it is connected.

(c) The converse of the previous result is not true. The classic counterexample is the set $\{(x, \sin(1/x)) : x \in (0, \infty)\} \cup \{(0, y) : y \in [0, 1]\}$ as a subset of $\mathbf{R}^2$. Another is by considering the closure of the set

$$A = \{(\cos t, \sin t, 1/t) : t > 0\}.$$

Explain why $\overline{A}$ is connected. Describe $\overline{A} \setminus A$.

9.3. **Perfect Metric Spaces.** Read about Perfect Metric Spaces on page 93 of Pugh. This is a quick trick for showing that something is uncountable.

9.4. **Optional: Continuous and Nowhere Differentiable.** Very Challenging Extra Credit Project: You can try thinking about how to construct your own function that is continuous and nowhere differentiable. It would be important not to look in the books though, because some examples of such functions are already there.

A fun read is attached regarding how the mathematical community responded to Weierstrass's discovery of a nowhere differentiable continuous function.