CANTOR SETS

ANUSH TSERUNYAN

1. Definition and topological properties

Definition 1 (Homeomorphism)**.** For metric (topological) spaces X, Y, a function $f : X \to Y$ is called a *homeomorphism* if it is a bijection and both f and f^{-1} are continuous. Call X and Y *homeomorphic* if there is a homeomorphism $X \to Y$.

Observation 2. *Let X, Y be metric (topological) spaces and $f : X \to Y$ a homeomorphism.*

(a) *f and f^{-1} preserve the open sets, i.e. for sets $U \subseteq X, V \subseteq Y$,*

$$U \textit{ is open } (\text{in } X) \iff f(U) \textit{ is open } (\text{in } Y)$$
$$V \textit{ is open } (\text{in } Y) \iff f^{-1}(V) \textit{ is open } (\text{in } X).$$

(b) *X and Y have the same* topological properties*, i.e. properties that are phrased only using open sets, e.g. connectedness, compactness, etc.*

Definition 3 (Cantor sets)**.** For a metric (topological) space X, call a set $C \subseteq X$ a *Cantor set* if C is homeomorphic to the Cantor space $2^{\mathbb{N}}$.

Observation 4. *Cantor sets are compact and have cardinality continuum.*

Proposition 5. *In a connected metric space X* (such as $\mathbb{R}^d$)*, any Cantor set has empty interior and hence is nowhere dense.*

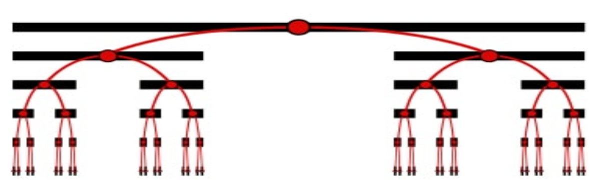
Proof. Let $C \subseteq X$ be a Cantor set and let $f : C \to 2^{\mathbb{N}}$ be a homeomorphism. Suppose towards a contradiction that $U := \mathrm{Int}(C) \neq \emptyset$. Then $f(U)$ is nonempty open in $2^{\mathbb{N}}$ and hence contains a nonempty clopen subset $V \subseteq f(U)$, so V and $f(U) \setminus V$ are both open. Therefore, $V' := f^{-1}V \subseteq U$ is nonempty clopen in C. It remains to show that V' is clopen in X as well. But the closedness of C in X implies that V' is closed in X, and the openness of U in X implies that V' is open in X. □

2. Constructing Cantor sets in $\mathbb{R}$

Recall that $2^{<\mathbb{N}}$ denotes the set of all finite binary sequences. For $s \in 2^{<\mathbb{N}}$ and $i \in 2 := \{0, 1\}$, let $s^\frown i$ denote the extension of s by appending the symbol i at the end of s.

Definition 6. A sequence $(I_s)_{s \in 2^{<\mathbb{N}}}$ of closed intervals in $\mathbb{R}$ is called a *Cantor scheme* if, for each $s \in 2^{<\mathbb{N}}$,

(i) $I_s \neq \emptyset$
(ii) $I_{s^\frown 0}, I_{s^\frown 1} \subseteq I_s$
(iii) $I_{s^\frown 0} \cap I_{s^\frown 1} = \emptyset$
(iv) for every $x \in 2^{\mathbb{N}}$, $|I_{x|_n}| \to 0$ as $n \to \infty$.



Observation 7. *Conditions (i)–(iii) guarantee that the length of each I_s is positive.*

Observation 8. *For any $x \in 2^{\mathbb{N}}$, condition (iv) and the completeness of $\mathbb{R}$ guarantee that $\bigcap_{n \in \mathbb{N}} I_{x|_n}$ is a singleton (i.e. contains exactly one element).*

Date: February 19, 2018.

In the light of this observation, define $f : 2^{\mathbb{N}} \to \mathbb{R}$ by mapping each $x \in 2^{\mathbb{N}}$ to the unique element in $\bigcap_{n\in\mathbb{N}} I_{x|_n}$. Call this f the *function induced by the Cantor scheme* $(I_s)_{s\in 2^{<\mathbb{N}}}$. Note that $f(2^{\mathbb{N}}) = C := \bigcap_{n\in\mathbb{N}} C_n$, where

$$C_n := \bigcup_{s\in 2^n} I_s.$$

We refer to C as the set induced by the Cantor scheme $(I_s)_{s\in 2^{<\mathbb{N}}}$.

Proposition 9. *The function f induced by any Cantor scheme is a homeomorphism $2^{\mathbb{N}} \to f(2^{\mathbb{N}})$. Thus, the set induced by a Cantor scheme is a Cantor set.*

Proof. Condition (iii) guarantees that f is injective. Note that for distinct $s, t \in 2^n$, I_s and I_t are of positive distance from each other. This, together with (iv), implies that both f and f^{-1} are continuous, and we leave the details to the reader. □

One way to build Cantor schemes is by iteratively removing open strict subintervals from the closed intervals. More precisely, say that the set C is *obtained from* $[a, b]$*, by removing an open strict subinterval* if $C = [a, b] \setminus (c, d)$, where $a < c < d < b$. Put $I_\emptyset := [a, b]$ for any $a < b$. Supposing that $I_s := [a_s, b_s]$ is defined for $s \in 2^{<\mathbb{N}}$, we take any $U_s := (c_s, d_s)$ with $a_s < c_s < d_s < b_s$ and put

$$I_{s^\frown 0} := [a_s, c_s] \text{ and } I_{s^\frown 1} := [d_s, b_s].$$

We can always guarantee that condition (iv) holds by choosing the open intervals U_s appropriately; for example, if each U_s contains the midpoint of I_s, then $|I_{s^\frown 0}|, |I_{s^\frown 1}| \leqslant 2^{-1}|I_s|$, so, by induction, $|I_s| \leqslant 2^{-|s|}|I_\emptyset| \to 0$ as $|s| \to \infty$. Thus, we have shown the following.

Proposition 10. *Any interval $[a, b]$ with $a < b$ contains a Cantor set.*

3. Cantor sets and Lebesgue measure

Corollary 11. *For any Cantor set $C \subseteq \mathbb{R}^d$ and any nonempty open $U \subseteq \mathbb{R}^d$, $U \setminus C$ has positive Lebesgue measure.*

Proof. Because C is closed nowhere dense, $U \setminus C$ is a nonempty open set, so it has positive Lebesgue measure. □

Let $(I_s)_{s\in 2^{<\mathbb{N}}}$ be a Cantor scheme obtained by iteratively removing open strict subintervals $(U_s)_{s\in 2^{\mathbb{N}}}$ as described above. Let C be the induced Cantor set and let the sets C_n be defined as above. Let λ denote the Lebesgue measure on $\mathbb{R}$.

Lemma 12. *If for some $\alpha \in (0, 1)$, for each $s \in 2^{<\mathbb{N}}$, $|U_s| \geqslant \alpha |I_s|$, then C is null.*

Proof. Indeed, it is enough to show that $\lambda(C_n) \to 0$ as $n \to \infty$. For each n, because, for each $s \in 2^n$, the removal of U_s leaves at most $(1 - \alpha)$-fraction of the measure of I_s, it is clear $\lambda(C_{n+1}) \leqslant (1 - \alpha)\lambda(C_n)$, so induction gives $\lambda(C_n) \leqslant (1 - \alpha)^n \lambda(C_0) \to 0$ as $n \to \infty$. □

Lemma 13. *If for each $s \in 2^{<\mathbb{N}}$, $|U_s| \leqslant 4^{-|s|}\frac{|I_\emptyset|}{4}$, then $\lambda(C) \geqslant \frac{|I_\emptyset|}{2} > 0$.*

Proof. For each $n \in \mathbb{N}$,

$$\lambda(\bigsqcup_{s\in 2^n} U_s) \leqslant 2^n \cdot 4^{-n} \frac{|I_\emptyset|}{4} = 2^{-n} \frac{|I_\emptyset|}{4}.$$

Thus,

$$\lambda(I_\emptyset \setminus C) = \lambda(\bigcup_{s\in 2^{<\mathbb{N}}} U_s) \leqslant \sum_{n\in\mathbb{N}} 2^{-n} \frac{|I_\emptyset|}{4} = \frac{|I_\emptyset|}{2},$$

so $\lambda(C) \geqslant \frac{|I_\emptyset|}{2}$. □

Corollary 14. *Any nonempty open interval contains a Cantor set of positive measure as well as a null Cantor set.*