Lecture 6: Metric Spaces and Fractals

October 7, 2023

Definition 1.1

If X is a set, then $d : X \times X \to \mathbb{R}$ is said to be a *metric* or *distance function* if

(1) $d(x, y) = d(y, x)$,

(2) $d(x, y) = 0$ if and only if $x = y$,

(3) $d(x, y) + d(y, z) \geq d(x, z)$.

Putting $z = x$ in (3) and using (1) and (2) shows that $d(x, y) \geq 0$. If d is a metric, then (X, d) is said to be a *metric space*.

Definition 1.2

The set $B(x,r) := \{y \in X | d(x,y) < r\}$ is called the *(open) r-ball* around x. A sequence $(x_n)_{n\in\mathbb{N}}$ in X is said to *converge* to $x \in X$ if for all $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that for every $n \geq N$ we have $d(x_n, x) < \epsilon$.

Definition 1.3

A sequence $(x_i)_{i\in\mathbb{N}}$ is said to be a *Cauchy sequence* if for all $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that $d(x_i, x_j) < \epsilon$ whenever $i, j \geq N$. A metric space X is said to be *complete* if every Cauchy sequence converges.

Definition 1.4

Let (X, d), (Y, d') be metric spaces. A map $f : X \to Y$ is said to be *continuous* at $x \in X$ if for every $\epsilon > 0$ there exists a $\delta > 0$ such that $d(x, y) < \delta$ implies $d'(f(x), f(y)) < \epsilon$. A continuous bijection (one-to-one and onto map) with continuous inverse is said to be a *homomorphism*.

Definition 1.5

Let (X, d), (Y, d') be metric spaces. A map $f : X \to Y$ is said to be an isometry if $d'(f(x), f(y)) = d(x, y)$ for all $x, y \in X$.

Definition 1.6

Let (X, d), (Y, d') be metric spaces. A map $f : X \to Y$ is said to be *Lipschitz-continuous* (or *Lipschitz*) with Lipschitz constant C, or C-Lipschitz, if $d'(f(x), f(y)) \leq Cd(x, y)$. A map is said to be a *contraction* (or, more specifically, a λ-*contraction*) if it is Lipschitz-continuous with Lipschitz constant $\lambda < 1$.

Definition 1.7

We say that two metrics are isometric if the identity establishes an isometry between them. Two metrics are said to be *uniformly equivalent* (sometimes just *equivalent*) if the identity and its inverse are Lipschitz maps between the two metric spaces.

- The unit circle $S^1 = \{x \in \mathbb{R}^2 | \|x\| = 1\}$ in the plane can also be described as the set of complex numbers of modulus 1.
- On the circle one can in natural way introduce several metrics.
- The first choice that comes to mind is to measure the distance of two points of S^1 using the Euclidean metric of $\mathbb{R}^2$. We refer to this metric as the Euclidean metric d.
- On the other hand, one may decide that the distance between two points of S^1 should be the distance traveled when moving from one point to the other along the circle, that is, the length of the shorter arc connecting the two points. This we call the length metric d_l.

Lemma 2.1

d and d_l are uniformly equivalent.

Proof.

$d(x, y) = 2\sin(d_l(x, y)/2)$, $d_l(x, y) \in [0, \pi/2]$, and $2t/\pi \leq 2\sin(t/2) \leq t$ for $t \in [0, \pi/2]$. Thus the identity map from the circle (S^1, d) with the Euclidean metric to the circle (S^1, d_l) with the length metric is Lipschitz-continuous with Lipschitz constant $\pi/2$. Its inverse (also the identity, but "in the other direction") is Lipschitz-continuous with Lipschitz constant 1. Therefore these two metrics are uniformly equivalent. □

- Consider the real line $\mathbb{R}$ and define the equivalence relation $\sim$ by setting $x \sim y$ if $x - y \in \mathbb{Z}$, that is, we define points to be equivalent if they differ by an integer.
- We define the *equivalence class* of $x \in \mathbb{R}$ by $[x] := \{y \in \mathbb{R} | y \sim x\}$.
- The equivalence class of 0 is just $\mathbb{Z}$ itself, and every equivalence class is a translate of $\mathbb{Z}$ by a member of the class, that is, $[x] = x + \mathbb{Z}$.
- To define a new metric space we consider the set $X = \mathbb{R}/\mathbb{Z} := \{[x] | x \in \mathbb{R}\}$ of all equivalence classes.

Proposition 2.2

$d(x, y) := \min\{|b-a| \big| a \in [x], b \in [y]\}$ defines a metric on $X = \mathbb{R}/\mathbb{Z}$.

Proof.

d is clearly symmetric. To check $d(x, y) = 0 \Rightarrow x = y$, note first that the metric does not change if instead we take the minimum over $a \in [x]$ only for a fixed $b \in [y]$, because the least distance from any integer translate of b to elements of x. But obviously $\min\{|b - a| \big| a \in x\}$ is actually attained, and hence is only zero if $b \in x$ and hence $x = y$.

To prove the triangle inequality take $x, y, z \in \mathbb{R}/\mathbb{Z}$ and $a \in [x],\ b \in [y]$ such that $d(x, y) = |b - a|$. Then for any $c \in [z]$ we have

$$d(x, z) \leq |c - a| \leq |c - b| + |b - a| = |c - b| + d(x, y).$$

Taking the minimum over $c \in [z]$ then shows that $d(x, z) \leq d(y, z) + d(x, y)$. □

Example 2.3

$d([\pi], [3/2]) = 7/2 - \pi = 0.5 - 0.14159265\cdots = 0.3584073\cdots$ and $d([0.9], [0]) = 0.1$.

Lemma 2.4

(1) If $a, b \in [0,1)$ with $|a-b| \leq 1/2$, then $d([a],[b]) = |a-b|$.

(2) If $|a-b| \geq 1/2$, then $d([a],[b]) = 1 - |a-b|$.

Proof.

(1) $d([a],[b]) \leq |a-b|$ by definition, but the inequality cannot be strict because every integer translate of b is further from a than b itself.

(2) $d([a],[b]) = 1 - |a-b|$ because this is the smaller of $|a-(b-1)|$ and $|a-(b+1)|$. □

- For example, the distance between the classes $[1-\epsilon]$ and $[0]$ is ϵ, if $\epsilon < 1/2$.
- Therefore, this construction intuitively corresponds to taking the interval $[0, 1]$ and "attracting" the open end to 0.
- Or, referring to the identification on the entire line $\mathbb{R}$, the construction amounts to "rolling up" the entire line on to a circle of circumference 1, so that integer translates of the same number all end up on the same point of the circle.
- Conversely, going from the circle to the line is like rolling a bicycle wheel along and leaving periodic tire prints.

- The cylinder is a space naturally visualized as a tube or pipe. There are several ways of defining it from more basic ingredients.
- One of these is motivated by a natural parametrization of a cylinder as follows:

$$(\cos 2\pi t, \sin 2\pi t, z) \text{ for } t \in \mathbb{R} \text{ and } -1 \leq z \leq 1.$$

- Of course, taking $0 \leq t \leq 1$ suffices to get the whole cylinder, and by periodicity of the trigonometric functions the points $(0, z)$ and $(1, z)$ are mapped to the same point in $\mathbb{R}^3$.
- Therefor this parametrization can be visualized as taking a unit square and rolling it up into a tube.

- The torus is the surface usually visualized as the surface of a doughnut.
- One can think of this surface as obtained by taking a circle in the xz-plane of $\mathbb{R}^3$ that does not intersect the z-axis and sweeping out a surface by revolving it around the z-axis.
- That is, by moving its center around a circle in the xy-plane. Doing this with the circle parametrized by

$$(R + r\cos 2\pi\theta, \sin 2\pi\theta)$$

gives the parametrization

$$((R + r\cos 2\pi\theta)\cos 2\pi\phi, (R + r\cos 2\pi\theta)\sin 2\pi\phi, \sin 2\pi\theta)$$

of the torus.

- Once we view the torus as $\mathbb{T}^2 = S^1 \times S^1$, however, we can utilize the description of S^1 as $\mathbb{R}/\mathbb{Z}$ just given and describe $\mathbb{T}^2$ directly as $\mathbb{R}^2/\mathbb{Z}^2$ by considering equivalence classes of points $(x_1, x_2) \in \mathbb{R}^2$ under translation by integer vectors $(k_1, k_2) \in \mathbb{Z}^2$, that is, $[(x, y)] = ([x], [y])$.
- As before, the Euclidean metric on $\mathbb{R}^2$ induces a metric on $\mathbb{T}^2$, which is the same as the product metric

$$d((x_1, x_2), (y_1, y_2)) = \sqrt{(d(x_1, y_1))^2 + (d(x_2, y_2))^2}.$$

- Continuing the rolling-up construction of the cylinder one more step (to roll the z-interval up into a circle as well we obtain a description of $\mathbb{T}^2$ by taking the unit square $[0, 1) \times [0, 1)$ and gluing the right and left edges together as well to get the torus.
- Likewise, we can construct and describe tori $\mathbb{T}^n$ of any dimension as n-fold products of the circles or as $\mathbb{R}^n/\mathbb{Z}^n$.

Proposition 4.1

Let X be a complete metric space. Under the action of iterates of a contraction $f : X \to X$, all points converge with exponential speed to the unique fixed point of f.

Definition 4.2

A map f of a metric space is said to be *eventually contracting* if there are constants $C > 0$, $\lambda \in (0, 1)$ such that

$$d(f^n(x), f^n(y)) \leq C\lambda^n d(x, y)$$

for all $n \in \mathbb{N}$.

Proposition 4.3

If $f : X \to X$ is a map of a metric space and there are C, $\lambda > 0$ such that $d(f^n(x), f^n(y)) \leq C\lambda^n d(x, y)$ for all $x, y \in X$, $n \in \mathbb{N}_0$, then for every $\mu > \lambda$ there exists a metric d_μ uniformly equivalent to d such that $d_\mu(f(x), f(y)) \leq \mu d(x, y)$ for all $x, y \in X$.

Proof.

Take $n \in \mathbb{N}$ such that $C(\lambda/\mu)^n < 1$ and set

$$d_\mu(x,y) := \sum_{i=0}^{n-1} d(f^i(x), f^i(y))/\mu^i.$$

This is called an *adapted* or *Lyapunov metric* for f. The two metrics are uniformly equivalent:

$$d(x,y) = d_\mu(x,y) \le \sum_{i=0}^{n-1} C(\lambda/\mu)^i d(x,y) \le \frac{C}{1-(\lambda/\mu)} d(x,y).$$

Note now that

$$\begin{aligned} d_\mu(f(x),f(y)) &= \sum_{i=1}^{n} \frac{d(f^i(x),f^i(y))}{\mu^{i-1}} = \mu(d_\mu(x,y) + \frac{d(f^n(x),f^n(y))}{\mu^n} - d(x,y)) \\ &\le \mu d_\mu(x,y) - \big(1 - C(\lambda/\mu)^n\big) d(x,y) \le \mu d_\mu(x,y). \end{aligned}$$

□

Corollary 4.4

Let X be a compact metric space and $f : X \to X$ an eventually contracting map. Then under the iterates of f, all points converge to the unique fixed point of f with exponential speed.

- Let us point out one of the major strengths of the notion of an eventually contracting map.
- As we just found, whether or not a map is a contraction can depend on the metric.
- This is not the case for eventually contracting maps.
- If a map f satisfies $d(f^n(x), f^n(y)) \leq C\lambda^n d(x, y)$ and d' is a metric uniformly equivalent to d, specifically $md'(x, y) \leq d(x, y) \leq Md'(x, y)$, then

$$d'(f^n(x), f^n(y)) \leq Md(f^n(x), f^n(y)) \leq MC\lambda^n d(x, y) \leq \frac{MC}{m}\lambda^n d'(x, y).$$

- In other words, only the constant C depends on the metric, not the existence of such a constant.

Proposition 4.5

If X, Y are metric spaces with X complete, $f : X \times Y \to X$ a continuous map such that $f_y := f(\cdot, y)$ is λ-contraction for all $y \in Y$, then the fixed point $g(y)$ of f_y depends continuously on y.

- The *ternary Cantor set* or *middle-third Cantor set* is described as follows.
- Consider the unit interval $C_0 = [0, 1]$ and remove from it the open middle third $(1/3, 2/3)$ to retain two intervals of length $1/3$ whose union we denote by C_1.
- Apply the same prescription to these intervals, that is, remove their middle thirds. The remaining set C_2 consists of four intervals of length $1/9$ from each of which we remove the middle third.
- Continuing inductively we obtain nested sets C_n consisting of 2^n intervals of length 3^{-n} [for a total length of $(2/3)^n \to 0$].
- The intersection C of all of these sets is nonempty and closed and bounded because all C_n are.
- It is called the *middle-third* or *ternary Cantor set*.

Lemma 5.1

C is the collection of numbers in $[0, 1]$ that can be written in ternary expansion (that is, written with respect to base 3 as apposed to base 10) without using 1 as a digit.

Proof.

The open middle third $(1/3, 2/3)$ is exactly the set of numbers that must have a 1 as the first digit after the (ternary) point, that is, that cannot be written in base 3 as $0.0\ldots$ or $0.2\ldots$ (Note that $1/3$ can be written as $0.02222\ldots$ and $2/3$ as $0.20000\ldots$) Correspondingly, the middle thirds of the remaining intervals are exactly those remaining numbers whose second digit after the point must be 1, and so on. □

Definition 5.2

A metric space X is said to be *connected* if it contains no two disjoint nonempty open sets. A *totally disconnected* space is a space X where for every two points $x_1, x_2 \in X$ there exist disjoint open sets $O_1, O_2 \subset X$ containing x_1, x_2, respectively, whose union is X.

Lemma 5.3

The ternary Cantor set is totally disconnected.

Proof.

Any two points of C are in different components of some C_n. Taking a sufficiently small open neighborhood of one of these together with the interior of its complement gives two disjoint open sets whose union contains C and each contains one of the points in question. □

Lemma 5.4

The ternary Cantor set is uncountable.

Proof.

Mapping each point $x = 0.\alpha_1\alpha_2\alpha_3\cdots = \sum_{i=1}^{\infty}(\alpha_i/3^i) \in C(\alpha_i \neq 1)$ to the number $f(x) := \sum_{i=1}^{\infty}(\alpha_i/2/2^i) = \sum_{i=1}^{\infty} \alpha_i 2^{-i-1} \in [0,1]$ defines a surjective map because all binary expansions indeed occur here. The fact that the image is uncountable implies that C is uncountable. □

Definition 5.5

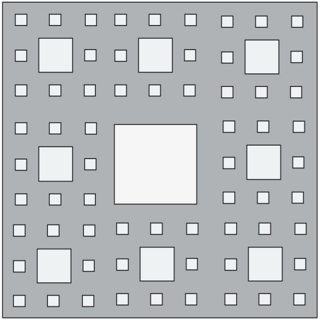
A set homeomorphic to the ternary Cantor set will be referred to as a *Cantor set*.

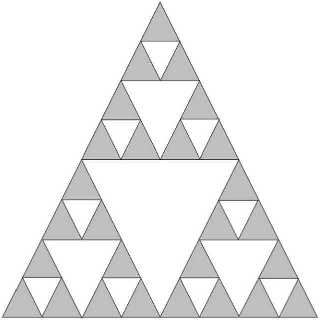
Proposition 5.6

All sets in $\mathbb{R}$ that are bounded, perfect, and nowhere dense are homeomorphic to the ternary Cantor set.

- There is an interesting example of a contraction on the middle-third Cantor set.
- Namely, $f : [0,1] \to [0,1]$, $f(x) = x/3$.
- Since f is a contraction, it is also a contraction on every invariant subset, and in particular on the Cantor set.
- The unique fixed point is obviously 0.
- This property of invariance under a linear contraction is often referred to as *self-similarity* or *rescaling property*.
- The micro structure of the Cantor set is exactly the same as its global structure; it does not become any simpler at any smaller scale.

The square and Sierpinski carpet





The von Koch snowflake